

UNA DOCUMENTACIÓN ACÚSTICA
DE LA LENGUA SHIPIBO-CONIBO (PANO)
(CON UN BOSQUEJO FONOLÓGICO)

José Elías Ulloa

UNA DOCUMENTACIÓN ACÚSTICA
DE LA LENGUA SHIPIBO-CONIBO
(PANO)

(Con un bosquejo fonológico)



**FONDO
EDITORIAL**

PONTIFICIA **UNIVERSIDAD CATÓLICA** DEL PERÚ

Una documentación acústica de la lengua shipibo-conibo (Pano)
(Con un bosquejo fonológico)

José Elías Ulloa

© José Elías Ulloa, 2011

De esta edición:

© Fondo Editorial de la Pontificia Universidad Católica del Perú, 2011

Av. Universitaria 1801, Lima 32, Perú

Teléfono: (51 1) 626-2650

Fax: (51 1) 626-2913

feditor@pucp.edu.pe

www.pucp.edu.pe/publicaciones

Cuidado de la edición, diseño de cubierta y diagramación de interiores:

Fondo Editorial PUCP

Primera edición: enero de 2011

Tiraje: 200 ejemplares

Prohibida la reproducción de este libro por cualquier medio, total o parcialmente,
sin permiso expreso de los editores.

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N°

ISBN:

Registro del Proyecto Editorial: 31501361100815

Impreso en Tarea Asociación Gráfica Educativa

Pasaje María Auxiliadora 156, Lima 5, Perú

Contenido

Sumilla	11
Agradecimientos	13
Lista de abreviaturas	15
1. Introducción	17
1.1 Los shipibos, el Ucayali y la migración a centros urbanos	22
1.2 Los hablantes shipibos que participaron en el estudio	23
1.3 Recajo de datos	24
1.3.1 El protocolo de las entrevistas	26
1.3.2 El equipo de grabación y el programa de análisis acústico	28
1.4 La organización del libro	29
2. Vocales orales	31
2.1 Los formantes F1 y F2	31
2.2 La centralización de las vocales en sílabas no-iniciales inacentuadas	35
2.3 La duración de las vocales cortas	40
2.4 La duración de las vocales largas	42
2.4.1 Algunas ilustraciones de las vocales largas	47
2.5 Vocales en contacto con oclusivas glotales epentéticas y subyacentes	49
2.6 El comportamiento de los formantes en vocales adyacentes: asimilación y coalescencia	55
2.7 Resumen	57
3. Las oclusivas	59
3.1 Introducción	59
3.2 Las oclusivas	60
3.2.1 Mediciones de duración de las oclusivas	61
3.2.2 Análisis espectral de la fase de explosión de las consonantes oclusivas	65
3.2.3 Las oclusivas y las transiciones de los formantes de las vocales adyacentes	72
3.2.4 Sobre la fonética de la labialización causada por [ɪ]	77
3.2.4.1 Una caracterización acústica de la labialización del shipibo	77
3.3 Resumen	83
4. Fricativas y africadas	85
4.1 Fricativas sibilantes sordas	85

4.1.1	Análisis espectral de las sibilantes	88
4.1.2	Las sibilantes y las transiciones de formantes	93
4.2	La fricativa laríngea	96
4.2.1	Análisis espectral de [h]	100
4.3	Las africadas sordas	101
4.3.1	Análisis espectral de la fase de fricción de las africadas sordas	104
4.3.2	Las africadas sordas y la transición de formantes	107
4.4	La africada sonora bilabial [b ^β]	108
4.4.1	La duración de [b ^β]	118
4.4.2	Análisis espectral de la fase de fricción de la africada bilabial [b ^β]	118
4.4.3	La africada bilabial sonora en contextos después de pausa	120
4.4.4	[b ^β] y [b]: Variabilidad entre hablantes	123
4.5	La africada retrofleja sonora	125
4.5.1	Análisis espectral de la fase de fricción de la africada [dʒ _r]	129
4.5.2	Transición de formantes y la africada retrofleja sonora [dʒ _r]	132
4.6	Resumen	132
5.	Nasales	135
5.1	Las consonantes nasales como arranques de sílaba	135
5.1.1	Análisis espectral de las consonantes nasales	137
5.1.2	Las transiciones de formantes y las nasales	139
5.2	Vocales nasalizadas y las codas nasales	141
5.2.1	La duración de las vocales nasalizadas y de las codas nasales	143
5.2.2	La estructura de formantes de las vocales nasalizadas	145
5.2.2.1	El comportamiento de F1 y las asimetrías entre las vocales acentuadas e inacentuadas	150
5.3	Sobre el punto de articulación de las codas nasales	152
5.4	Resumen	158
6.	Glides	161
6.1	Los formantes F1 y F2 de las glides	161
6.2	Sobre la fonética de las vocales altas y las glides	163
6.2.1	Fuera de la fonología: las vocales /i/ y /u/ como semi-vocales	164
6.2.2	Las vocales /i/ y /u/ como glides en posición inicial de palabra	166
6.3.2	Las vocales /i/ y /u/ como semi-consonantes en contextos intervocálicos	169
6.3	Resumen	171
7.	Las sílabas y las propiedades duracionales de las rimas	173
7.1	Estructura de la sílaba y sus propiedades duracionales	173
7.1.1	Duración de las codas	173
7.1.2	duración de las vocales cortas en sílabas cerradas	177
7.1.3	Duración de las rimas de las sílabas	178
7.2	Codas complejas	181
7.3	Resumen	184

8. La frecuencia fundamental, el tono perceptual y el acento	187
8.1 El acento primario en shipibo	189
8.1.1 Casos de acento léxico en raíces	193
8.1.2 El acento primario y los segmentos latentes	194
8.2 El F0 de los tonos altos en sílabas con acento primario	197
8.2.1 Tono alto-nivelado	198
8.2.2 Tono alto-ascendente	200
8.2.3 Tono alto-descendente	201
8.3 El F0 del tono bajo de sílabas iniciales inacentuadas	202
8.3.1 Tono bajo-nivelado en sílabas iniciales inacentuadas	203
8.3.2 Tono bajo-descendente en sílabas iniciales inacentuadas	204
8.3.3 Tono bajo-ascendente en sílabas iniciales inacentuadas	206
8.4 El género como factor condicionante del tono de las sílabas inacentuadas segundas	208
8.5 Tono perceptual, duración e intensidad como claves para el acento principal	212
8.6 El F0 fuera de la ventana del acento primario	213
8.6.1 El F0 de las sílabas cerradas	216
8.7 El F0 en sufijos con acento léxico	222
8.8 F0 y la posición de los acentos secundarios en compuestos	226
8.8.1 La realización de los acentos secundarios en compuestos	227
8.8.2 Choques acentuales y desacentuación	230
8.9 Resumen	236
9. Un Bosquejo de la fonología del shipibo	239
9.1 Una fonología de las vocales del shipibo	240
9.1.1 Los rasgos de las vocales	240
9.1.2 La longitud vocálica	241
9.1.2.1 Un análisis alternativo para las vocales largas	245
9.1.3 Cambios de la duración vocálica determinados por el contexto lingüístico	250
9.2 Una fonología de las consonantes del shipibo	256
9.2.1 Los rasgos de las consonantes	256
9.2.1.1 Sobre el estatus subyacente de las semi-consonantes /j, w/	258
9.2.1.2 Las restricciones fonotácticas sobre la adyacencia de segmentos	261
9.2.2 Los desafíos de un análisis fonológico de la labialización del shipibo	267
9.2.3 Segmentos latentes	271
9.2.3.1 Los segmentos latentes y la asignación del acento	274
9.2.3.2 El punto de articulación por defecto de los segmentos latentes	275
9.2.3.3 Los segmentos latentes dorsales	278
9.2.3.4 El punto de articulación de las nasales finales de raíces	282
9.2.4 Problemas fonológicos que rodean a las nasales	285
9.3 La interacción entre sílaba y acento	287
9.3.1 La estructura de la sílaba del shipibo	287
9.3.1.1 Las glides y la silabificación	289

9.3.1.2 Las vocales altas y la reestructuración fonética de las sílabas	291
9.3.2 El acento primario	293
9.3.3 El acento secundario	296
9.3.4 Casos de acento léxico	299
9.3.5 Choques acentuales y la preservación del tono alto del acento primario en compuestos	303
9.3.6 El género y la adquisición de la distribución de los tonos altos	305
9.4 Resumen	308
10. Conclusiones	311
Bibliografía	319

SUMILLA

Una de las riquezas más notables del Perú es su diversidad lingüística. Además del castellano, en la actualidad, se hablan en el Perú más de cuarenta lenguas indígenas diferentes. A pesar de esta riqueza, sabemos muy poco de la mayoría de esas lenguas y muchas de ellas están en serio peligro de volverse extintas. Además de ser la primera documentación acústica de una lengua indígena peruana y una contribución importante al campo de la lingüística amazónica, este libro innova el paradigma tradicional de documentación de lenguas al ofrecer un estudio inédito de las propiedades acústicas del sistema de sonidos y de los contornos tonales de la lengua shipibo-conibo (pano). El estudio describe en detalle los procedimientos utilizados para cada tipo de análisis acústico de modo que otros trabajos de documentación puedan tomarlos o adaptarlos al estudio de lenguas indígenas. El libro además incluye un bosquejo de la fonología del shipibo-conibo, el cual conecta los resultados acústicos con los patrones de los sonidos de la lengua y los compara con fenómenos lingüísticos similares encontrados en otras lenguas del mundo.

AGRADECIMIENTOS

Agradezco a la comunidad shipiba de Cantagallo. En particular, me gustaría expresar mi profunda gratitud al Sr. Ascencio Franco, Sr. Benito García, Sra. Nimia García, Sr. Jonás Franco, Sra. Ana García, Sr. Cesar Maynas, Sra. Danitza Torres, Sr. Dionel Canayo, Sra. Viginia Bardales, Sra. Teresa Cayruna, Sra. Matilde Baneo, Sr. Emerson Mejía, y a la Sra. Florentina. Doy un agradecimiento muy especial a la Sra. Adelina Ahuanari, a la Sra. Luz Franco, y al Sr. Pepe Silvano por su apoyo y amistad durante todos estos años. Sin ellos, este trabajo no habría sido posible.

Agradezco a Daniel Valle, Armando Cuba, y Jan Carrillo, quienes se desempeñaron como asistentes de investigación. Su compromiso con la documentación de lenguas amazónicas, su profesionalismo, sus incontables horas de trabajo arduo durante las entrevistas así como también en la organización de los archivos de audio hicieron posible que este estudio saliera adelante. Les estaré eternamente agradecido.

También agradezco a Jorge Pérez, Daniel Valle, y a Seungun Lee por sus sugerencias, comentarios, reacciones y discusión de las varias versiones que tomó este libro desde el primer borrador. Un agradecimiento muy especial va para Carolina González por sus comentarios muy detallados, los cuales me ayudaron mucho en obtener la versión final de este libro. Cualquier error que quede es completamente mío. Finalmente, le estoy como profundamente agradecido a mi querido Phillip quien como siempre me ofrece su cariño, su apoyo y la energía para seguir adelante.

Este proyecto no hubiese sido posible de realizar sin las becas de investigación otorgadas por la Universidad de Stony Brook (Stony Brook University Fine Arts, Humanities and Social Science Initiative, 2007 y 2008) y el apoyo del departamento de Hispanic Languages and Literature de la Universidad de Stony Brook, el cual facilitó la publicación de este libro.

LISTA DE ABREVIATURAS¹

CON.3PL	Concordancia verbal de tercera persona plural /-kan/
AGTZ	Agentivizador /-mis/
AND1	Andativo Singular Intransitivo /-kain/
AND2	Andativo No-singular, singular transitivo /-b ^β ain/
AUX	Auxiliar /ʔiki/
C	Consonante
CAUS	Causativo /-ma/
CMPL	Aspecto completivo /-ki/
dB	Decibel
DES	Desiderativo /-kas/
EV	Evidencial directo /-dʒa/
F0	Frecuencia fundamental
F1	Primer formante
F2	Segundo formante
F _n 1	Primer formante nasal
FFT	Fast fourier transform
FRUSTR	Frustrativo /-kian/
FUT1	Futuro Mañana /-jat/
REP	Reportativo /-dʒʊnki/
Hz	Hercios
IMP	Imperativo /- ^l wi/
INF	Infinitivo /-ti/

¹ He adoptado la mayoría de los nombres y abreviaturas usados para los sufijos del shipibo de Valenzuela 2003b.

INTENS	Intensificador /-idz̥a/
INTVZ	Intransitivizador /-t/
FL	Forma Larga de Sustantivos y Adjetivos /-(a)n/, /-nin/
LPC	Linear Prediction Coding
ms	Milisegundos
NMLZ	Nominalizador /-ti/
NEG	Negación /-jama/
PL	Plural /-b ^β ʊ/
FLPL	Forma Larga del Plural /-b ^β aʊn/
PP1	Participio Incompleto /-ai/
PP2	Participio Completo /-a/
PRIV	Privativo /-jaʊma/
PSD1	Pasado Temprano Hoy Día /-wan/
PSD2	Pasado Ayer /-ʔib ^β at/
REM1	Pasado Remoto /-'ni/
REM2	Incompleto en el Pasado Remoto /-paʊni/
ESPC	Especulativo /-main/, /-m̥in/
CR	Cambio de Referencia /-i/, /-ʂʊn/
ST	Semitono
TRVZ	Transitivizador /-n/
V	Vocal
VOT	Tiempo de Inicio de la Voz (<i>Voice onset time</i>)
VZ1	Verbalizador Transitivo /-ʔati/
VZ2	Verbalizador Intransitivo /-ʔiti/

1. INTRODUCCIÓN

La primera pregunta que se viene a la mente al ver el título de este libro es ¿por qué está dedicado a la documentación del shipibo? En shipibo, hay dos sonidos que han sido tradicionalmente descritos como una fricativa bilabial sonora y como una vibrante simple, ortográficamente representados por las letras ‘b’ y ‘r’. Este estudio, sin embargo, mostrará a través de un examen acústico que estos sonidos en realidad son respectivamente una africada bilabial sonora y una africada retrofleja sonora, dos tipos de sonidos raros de encontrar en las lenguas del mundo y que este trabajo de investigación, a través del estudio del shipibo, está contribuyendo a su descripción. Sin una documentación acústica, la descripción tradicional de esos sonidos habría quedado intacta. Pero por supuesto, esa no es la justificación principal de este libro.

Mi interés en el shipibo comenzó cuando era un estudiante de lingüística de la Universidad Nacional Mayor de San Marcos, en un trabajo de campo organizado en la comunidad shipiba de Callería, en el Ucayali, probablemente alrededor de 1996. Lo que inmediatamente se volvió obvio para mí, aún como un joven aspirante a lingüista, era que sabíamos tan poco de la manera cómo funcionaba la lengua shipiba, y qué tan diferente del castellano eran sus sonidos. En particular, estaba fascinado, y aún lo estoy, con sus patrones acentúales y con la alofonía de ciertos sufijos, la cual está condicionada por el número de sílabas de las palabras a las cuales se añaden (Elías-Ulloa 2000, 2006). Así, como otros fenómenos lingüísticos notables, un aspecto que se volvió cada vez más importante para mí, mientras con los años aprendía más del shipibo, era la necesidad de documentar la lengua. El grado de prejuicio que existe en Perú entre los hablantes de castellano monolingües contra peruanos cuya lengua nativa no es el castellano se combina con las presiones sociales y económicas que sufren estas comunidades indígenas.

Pero aún así, el pueblo shipibo tiene una tradición orgullosa de proteger su lengua y cultura, uno de los tesoros más grandes de su patrimonio. Este estudio, entonces, intenta documentar una pequeña parte de ese patrimonio.

Este libro no solo representa la primera documentación acústica del shipibo, una lengua amazónica de la familia lingüística pano, sino que también es la primera caracterización acústica completa del sistema de sonidos de una lengua indígena peruana. Mi principal objetivo es lograr una documentación detallada de las propiedades acústicas de las consonantes, vocales y patrones acentuales del shipibo. Mientras ya existen trabajos de documentación del shipibo en términos de la publicación de vocabularios y diccionarios (von den Steinen 1904; Alemany 1906; Lorient, Lauriault *et al.* 1993), de estudios fonológicos y morfo-sintácticos (Lauriault 1948; Lorient and Hollenbach 1970; Black 1992; García-Rivera 1994; Elias-Ulloa 2000a, b, 2001; Valenzuela, Márquez-Pinedo *et al.* 2001; Valenzuela 2002b, a; Gonzalez 2003; Valenzuela 2003b, a; Camacho and Elias-Ulloa 2004; Elias-Ulloa 2005; Gonzalez 2005; Elias-Ulloa 2006; Camacho 2007; Gonzalez 2007; Sánchez, Camacho *et al.* 2009; Camacho and Elias-Ulloa en prensa; Camacho en prensa-a, b, entre otros), aún no existía un trabajo enfocado en la fonética acústica de la lengua.¹ Este libro busca llenar este vacío.

La segunda pregunta que el título de este libro nos hace es ¿por qué es la tarea de documentar el shipibo realizada desde una perspectiva acústica? Esta pregunta podría colocarse aún en términos más generales: ¿por qué es importante la documentación acústica de una lengua? Esencialmente, el uso de la fonética acústica en la documentación de lenguas incrementa la confiabilidad y exactitud de los patrones descritos. Al reportar los resultados, la fonética acústica provee una representación más exacta de qué representan las transcripciones que el trabajador de campo usa al describir una lengua. Las transcripciones, ya sean estas fonéticas u ortográficas, son un ingrediente importante en la documentación de lenguas y, a su turno, cuando se usan como una fuente primaria de datos sobre los patrones de una lengua, se convierten en las piezas principales sobre las cuales se proponen o rechazan teorías lingüísticas. Mientras las transcripciones juegan un rol importante en lingüística, la manera cómo son usadas es aún bastante informal. Generalmente se les toman como fieles caracterizaciones de una lengua o fenómeno lingüístico. Sin embargo, en muchos casos, las transcripciones son representaciones de la *impresión subjetiva* del lingüista de campo. Reflejan lo que el lingüista de campo oye mientras trabaja con una lengua. Estas impresiones son

¹ Existe un artículo publicado de cinco páginas (Valenzuela, Márquez-Pinedo y Maddieson (2001)), que bosqueja las realizaciones fonéticas de las consonantes y vocales del shipibo pero no brinda ningún tipo de información acústica.

afectadas inevitablemente por la fonología del transcriptor y por los prepuestos teóricos implícitos o explícitos que pueda tener el lingüista de campo. Inclusive las transcripciones fonéticas ostentosamente más neutrales en términos teóricos asumen la posición que el continuo del habla puede ser dividido en segmentos, lo cual ya es una posición teórica (en oposición, por ejemplo, a una teoría que asume que el habla debe representarse como la sobreposición de gestos articulatorios como ha sido propuesto por Browman and Goldstein 1986, 1989, 1990, 1995). Ya que la transcripción del trabajador de campo está afectada por estos factores externos, la pregunta que debemos hacernos es: ¿qué representa una transcripción? ¿Qué estamos haciendo cuando usamos esas transcripciones como datos que van a servir como base a análisis y teorías lingüísticas? Éste es, por supuesto, no un problema que enfrenta exclusivamente los trabajos lingüísticos del Shipibo pero un problema mucho más general que la lingüística como campo debe enfrentar.

Los estudios de acústica ayudan a atenuar la subjetividad que manifiestan las transcripciones. En la documentación de lenguas, la fonética acústica nos brinda los medios de contrastar nuestras interpretaciones personales de los patrones de la lengua que estamos estudiando con análisis acústicos más objetivos. Ambos elementos juntos, aunque no siempre están de acuerdo, nos proveen una descripción más confiable del sistema de sonidos de una lengua. En este sentido, la fonética acústica se convierte en una herramienta indispensable para la documentación de lenguas.

Uno podría argüir que las propiedades fonéticas de los sonidos de las lenguas se pueden recuperar de las descripciones articulatorias asociadas con cada símbolo fonético del AFI (Alfabeto Fonético Internacional). Desafortunadamente, esta asunción no es correcta. Las descripciones articulatorias del AFI no pueden capturar todas las propiedades fonéticas de los segmentos, las cuales pueden ser cruciales al entendimiento de los patrones fonológicos (Bradlow 1995; Ladefoged and Maddieson 2001). Por ejemplo, Disner 1983, 1984 encontró que a pesar que el italiano y el yoruba tienen el mismo inventario vocálico, representado en el AFI por los símbolos [i, e, ε, a, ɔ, o, u], los detalles acústicos de estas vocales difieren según la lengua. En particular, las vocales [e] y [o] en yoruba aparecen consistentemente más cercanas a las vocales [i] y [u] que a las vocales [ε] y [ɔ], mientras que en italiano, las vocales están uniformemente distribuidas en el espacio vocálico. Es decir, hay una diferencia acústica consistente en el valor del Primer Formante (F1) de las vocales [e] y [o] entre ambas lenguas. De este modo, la documentación acústica de una lengua complementa y da más exactitud a la descripción de los símbolos fonéticos usados en las transcripciones.

La documentación acústica de lenguas nos permite llevar a cabo comparaciones tipológicas más rigurosas entre lenguas. Tomemos, por ejemplo, los

casos de las consonantes oclusivas sonoras del inglés y del castellano, [b, d, g], y asumamos por un momento que nunca se ha llevado a cabo un estudio acústico de estos segmentos en esas dos lenguas. Un investigador interesado en hacer comparaciones entre estas dos lenguas y solo teniendo como base las transcripciones fonéticas y las respectivas descripciones articulatorias convencionalmente asociadas con esos símbolos fonéticos erróneamente concluiría que las oclusivas sonoras del inglés y del castellano son los mismos sonidos. Pero realmente no lo son. Aunque los segmentos sean clasificados como sonoros en ambas lenguas, los estudios acústicos han mostrado que en lenguas como castellano, la vibración de las cuerdas vocales, las cuales se encargan de hacer el sonido sonoro, comienza tan pronto como el aire es obstruido en la cavidad oral. En contraste, en lenguas como el inglés, la vibración de las cuerdas vocales durante la articulación de [b, d, g] comienza mucho más tarde dentro de la fase de cerrazón. Es decir, las oclusivas sonoras [b, d, g] del inglés no son completamente sonoras y de hecho, para algunos hablantes de inglés, tienden a no serlo en lo absoluto. La diferencia en el tiempo del inicio de la voz durante la fase de cerrazón de [b, d, g] no es un detalle fonético irrelevante, imperceptible. Esta diferencia puede ciertamente causar que un hablante nativo del castellano escuche estas consonantes (en decir, /b, d, g/ en inglés) como si fuesen sordas, es decir, como si fuesen [p, t, k].

Afortunadamente para los lingüistas interesados en castellano e inglés, hay una literatura abundante que caracteriza las propiedades acústicas de la sonoridad de estas consonantes para estas dos lenguas. Esto, desafortunadamente, no es el caso para todas las lenguas, y en especial no para las lenguas indígenas peruanas. De este modo, en general, en este libro arguyo que no es suficiente decir que una lengua dada tiene este o ese segmento y representarlo con un símbolo fonético en la esperanza que las descripciones articulatorias estándares de los símbolos fonéticos puedan elucidar todas las características fonéticas de los sonidos de una lengua específica. Las transcripciones *deben* ser complementadas con descripciones acústicas detalladas.

Es obviamente importante asegurarnos que la mayor cantidad de estudios de documentaciones acústicas estén disponibles para tantas lenguas como sea posible a fin de mejorar la comparaciones entre lenguas, de dar un contenido más preciso a los resultados de trabajos documentación, de lograr que el uso de las transcripciones sea menos un acto de fe y más una herramienta eficaz que pueda representar con fidelidad los fenómenos lingüísticos de las lenguas que estudiamos. Los resultados de los estudios de documentación acústica no solo se pueden usar en la construcción de teorías lingüísticas sino que también en tareas con fines más prácticos, como por ejemplo, en la propuesta y elaboración de alfabetos, en la creación de materiales de enseñanza, en la lucha contra la discriminación

lingüística como, por ejemplo, lo están haciendo Pérez Silva, Acurio Palma, y Bendezú Araujo (2008).

Este libro presenta al lingüista de campo peruano un modelo que podría usarse para examinar las propiedades acústicas de otras lenguas. El libro muestra de forma detallada los procedimientos de cómo se llevaron a cabo los análisis acústicos para el shipibo de modo que pueden ser replicados e inclusive mejorados en la documentación de otras lenguas. He intentado que estas explicaciones sean tan accesibles como sea posible a un público diverso. No asumo que el lector tenga ningún conocimiento profundo sobre fonética acústica. Para aquellos lectores interesados en conocer más sobre la fonética acústica, recomiendo los siguientes trabajos, los cuales han sido invaluable también para este estudio: Fry 1979; Olive, Greenwood *et al.* 1993; Ladefoged 1996; Ladefoged y Maddieson 1996; Johnson 2003; Ladefoged 2003, 2005; Hewlett y Beck 2006; Ladefoged 2006; Silverman 2006; Raphael, Borden *et al.* 2007; Lodge 2009.

Aunque el principal objetivo es presentar una documentación acústica del shipibo, el lector debe saber que dedico un capítulo completo a bosquejar la fonología del shipibo y hacer las conexiones necesarias entre los resultados acústicos y los principales patrones fonológicos encontrados en la lengua. En ese bosquejo, me he concentrado en develar los desafíos que ponen las propiedades acústicas del shipibo a los análisis fonológicos. En el bosquejo de la fonología del shipibo, he tratado de ser tan neutral como es posible en términos de teoría fonológica de manera que el libro pueda ser de interés a lingüistas que vienen de diferentes corrientes teóricas.

Intento que este libro pueda servir como referencia a investigadores en lenguas amerindias, a lingüistas interesados en la documentación de lenguas y en comparaciones tipológicas, a aquellos interesados en metodologías lingüísticas de trabajo de campo y a lingüistas que en general trabajan en las áreas de fonética y fonología. También espero que anime a trabajadores de campo e investigadores de lenguas indígenas a documentar en detalle no solo las propiedades morfológicas y sintácticas de una lengua pero también las propiedades fonológicas y acústicas de las lenguas que estudian.

Las siguientes secciones presentan información general acerca de la población shipiba, del bagaje socio-lingüístico de los colaboradores shipibos que participaron en el estudio, de cómo se recogieron y analizaron los datos así como también se describe el equipo que se usó.

1.1 LOS SHIPIBOS, EL UCAYALI Y LA MIGRACIÓN A CENTROS URBANOS

La mayoría de investigadores concuerdan que la lengua shipiba (también conocida como Shipibo-Konibo) tiene hoy en día alrededor de 25000 hablantes (Pozzi-Escot 1998; Valenzuela 2003b; Lewis 2009). Los shipibos han logrado conservar su lengua viva, a pesar del contacto prolongado con el castellano. Han retenido su modo de vida, la mayoría de sus tradiciones, su identidad étnica, cultural y lingüística.

Los shipibos viven en cientos de pequeñas comunidades a lo largo de las riberas del río Ucayali y sus tributarios, algunas de ellas cerca de centros urbanos donde el castellano es predominante, otras más distantes de esos centros. La mayoría de las comunidades shipibas están localizadas en los departamentos de Ucayali y Loreto. Debido a la migración a las áreas urbanas, se puede encontrar también comunidades shipibas viviendo en ciudades como Pucallpa y en la capital, Lima.

Ya que es posible encontrar hablantes de todas las edades que tiene el shipibo como su primera lengua, éste ha sido clasificado como *una lengua potencialmente en peligro* (Brenzinger 2007, véase también Solís Fonseca 2003). Esto quiere decir que los niños se muestran menos y menos inclinados a aprender y usar la lengua (Moseley 2007:166). Esta tendencia ya puede observarse en los niños shipibos, especialmente aquellos que viven en áreas urbanas o en comunidades que están cerca de centros urbanos en los cuales el castellano es predominante. Este tipo de migración hacia las ciudades se debe a la expansión urbana que está experimentando la región del Ucayali. Los shipibos migran de sus comunidades en búsqueda de mejores oportunidades económicas y educativas en las ciudades. Los centros urbanos que más los han atraído son la ciudad de Pucallpa (localizada en las riberas del río Ucayali) y el distrito cercano de Yarinacocha. Además de migrar a centros urbanos dentro de la zona del Ucayali, algunos shipibos han migrado fuera del área amazónica, a Lima. Se estima que hay aproximadamente setenta familias shipibas (alrededor de trecientas personas) viviendo en Cantagallo (en el distrito del Rímac, Lima). De aquí en adelante, me referiré a esta comunidad como la *Comunidad Shipiba de Cantagallo*. La primera ola de migración a Lima se produjo a principios de los años noventas. Una segunda ola más marcada de migración a Lima ocurrió en el año 2000.

La comunidad shipiba de Cantagallo crece cada año por la llegada de nuevos migrantes shipibos. Los miembros de la comunidad, aunque están tan distantes de la región amazónica, mantienen su lengua e identidad étnica. Estas se ven reforzadas por los lazos comunitarios que los shipibos han formado: la mayoría de ellos viven en la misma zona, siendo muchos de ellos vecinos, planean actividades juntos y enfrentan sus problemas unidos. Conducen sus actividades diarias en

shipibo y usan castellano principalmente para comunicarse con personas fuera de su grupo, vender sus trabajos artesanales y para actividades laborales fuera de su comunidad. Además, muchos de ellos están constantemente yendo y viniendo de sus comunidades. Aunque algunos expresan su deseo que preferirían que sus hijos sean monolingües en castellano porque piensan que así sus vidas serán más fáciles en la ciudad, admiten que esto es difícil de lograr ya que dentro de la comunidad en Cantagallo, sus niños aprenden shipibo de los más ancianos que los cuidan mientras ellos van a laborar.

1.2 LOS HABLANTES SHIPIBOS QUE PARTICIPARON EN EL ESTUDIO

Los análisis acústicos que este libro ofrece fueron hechos sobre la base de los datos recogidos de ocho hablantes nativos de shipibo (cuatro hombres y cuatro mujeres). Todos los hablantes entrevistados pertenecen a la comunidad de Cantagallo. El género, edad y nivel de educación así como también el nombre de las comunidades indígenas donde crecieron y vivieron antes de migrar a Lima se indican en la Tabla 1-1. Todos hablan castellano como segunda lengua pero muestran notable diferencias con respecto a su nivel de bilingüismo. En general, contra más joven el colaborador, más alto tiende a ser su grado de bilingüismo. Los hablantes que participaron en el estudio vienen originalmente de las siguientes comunidades shipibas: Amaquiria, Betel, Caco, Nuevo Nazareth, Nuevo Samaria, y Sempaya. Todas localizadas a lo largo del río Ucayali. Estas comunidades indígenas se encuentran en el departamento de Ucayali. Con la excepción de Sempaya que se localiza en la provincia de Atalaya, todas las otras comunidades pertenecen a la provincia de Coronel Portillo. Todos los participantes migraron a Lima por la primera vez entre 1993 y 1998.

Los hablantes shipibos de este estudio eran mayores de 30 años al momento de las entrevistas, las cuales se llevaron a cabo en el 2008. El colaborador más joven, HBL1 en la Tabla 1-1, tenía 31 años y migró a Lima en 1998. Esto significa que tenía 21 años en ese momento, una edad que garantiza su dominio nativo de la lengua shipiba ya para ese entonces. Solo HBL2 era más joven que HBL1 cuando llegó a Lima. HBL2 tenía 18 años en 1994. Nuevamente, esa es una edad que también garantiza un dominio nativo de la lengua shipiba al momento de migrar a Lima.

El factor edad divide a los participantes del estudio en dos grupos diferentes. Cuatro de ellos están dentro de sus treinta años y los otros cuatro son mayores (entre cuarentas y cincuentas). Obsérvese que hay una correlación entre estos dos grupos de edades y el nivel de educación alcanzado. Esta correlación coincide con la impresión que uno tiene de la comunidad en general. A diferencia de los

hablantes mayores, los jóvenes usualmente han completado su educación básica (educación primaria y secundaria) y muchos de ellos tienen algún tipo de educación superior en institutos, aunque en la mayoría de los casos no la han terminado.

Tabla 1-1: Información sobre los participantes del estudio

	Género	Edad	Nivel de Educación	Comunidad Indígena
HBL1	Femenino	31	Instituto (incompleto)	Betel y Amaquiria
HBL2	Femenino	33	Secundaria (completa)	Nuevo Nazareth
HBL3	Masculino	35	Instituto (completo)	Betel
HBL4	Femenino	35	Instituto (incompleto)	Betel
HBL5	Masculino	54	Primaria (incompleta)	Caco y Nuevo Samaria
HBL6	Masculino	58	Primaria (completa)	Sempaya y Nuevo Nazareth
HBL7	Femenino	60	Primaria (completa)	Betel
HBL8	Masculino	64	Primaria (completa)	Betel

1.3 RECOJO DE DATOS

Las entrevistas se llevaron a cabo diariamente por el lapso de dos meses. La Tabla 1-1 indica correctamente que se entrevistaron a ocho hablantes nativos de shipibo. Sin embargo, los resultados acústicos que se presentan en este libro para cada propiedad fonética estudiada siempre se basan en los datos de seis hablantes. La discrepancia se debe a que por problemas familiares y de salud, los hablantes HBL7 y HBL8 no pudieron completar las entrevistas. Ambos colaboradores dejaron de participar al final del primer mes. A fin de poder mantener el número de hablantes en seis personas, HBL1 y HBL3 tomaron los lugares de HBL7 y HBL8. Desafortunadamente, por restricciones de tiempo, no pudo ser posible encontrar a dos colaboradores de las mismas edades que HBL7 y HBL8.

Además de los ocho colaboradores shipibos que fueron entrevistados, el estudio se benefició de los resultados previamente obtenidos en un proyecto piloto llevado a cabo en el 2007, el cual involucró el análisis de datos de dos hablantes shipibos (un hombre y una mujer). Al momento de las entrevistas, tenían 43 y 46 años respectivamente. Ambos fueron de la comunidad shipiba de Caimito y Betel (ambas comunidades están ubicadas en Masisea, Coronal Portillo, Uca-yali). Durante el proyecto piloto del 2007, comencé la exploración sistemática de las propiedades acústicas del shipibo. Los resultados obtenidos durante esta fase guiaron el diseño final del estudio y me permitieron modificar y mejorar

tanto en el protocolo usado para recoger los datos así como también en el tipo de datos necesarios para lograr una caracterización detallada de las propiedades acústicas del shipibo.

Las entrevistas se llevaron a cabo en Lima. Cada entrevista duró alrededor de dos horas y consistió en la elicitación de palabras dentro de una frase. Para el recojo de datos, se preparó diecinueve listas de palabras, cada una de ellas diseñada para examinar una propiedad acústica en particular (por ejemplo, duración de consonantes/vocales, la estructura de formantes en las vocales, la transición entre vocales y consonantes, el tiempo de inicio de la voz, nasalización, frecuencia fundamental, etc). Las palabras de cada lista fueron cuidadosamente seleccionadas de acuerdo al tipo de datos que se deseaba obtener.

Se prestó especial atención al contexto segmental que rodeaba la propiedad acústica que se iba a estudiar. Por ejemplo, en el estudio sobre duración vocálica, las vocales aparecieron rodeadas de consonantes oclusivas, ya que estas consonantes facilitan la segmentación de las vocales (es decir, la identificación de los momentos donde se inician y terminan las vocales). Para el estudio de las transiciones entre vocales y consonantes, se controló cuidadosamente el tipo de vocal y al punto de articulación de la consonante involucrada.

Además del contexto segmental, al crear la lista de palabras también se tomó en cuenta otros factores: el tamaño de las palabras, si el segmento que se iba a estudiar aparecía o no dentro de una sílaba acentuada o una sílaba inicial, si ocupaba la posición de arranque o coda, y si estaba en el límite con una frontera morfológica. Un factor extra-lingüístico importante que se consideró fue el género del hablante.

El tamaño de la palabra es particularmente importante al estudiar la duración de los segmentos y los acentos tonales (pitch) de una lengua. Con el fin de neutralizar cualquier cambio en la duración de segmentos debido al tamaño de las palabras que los contienen, las listas usadas tendieron a incluir palabras trisilábicas. En la mayoría de los casos, ya que las raíces shipibas tienden a ser bisilábicas, se agregó un sufijo para formar la tercera sílaba. A menos que se indique de otro modo, la propiedad acústica estudiada siempre ocurre en segmentos en la primera o en la segunda sílaba de estas palabras trisilábicas. Para otros estudios específicos, como por ejemplo, el caso de las vocales largas, se incluyó además palabras monosilábicas y bisilábicas con el fin de detectar cambios en su duración. Para otros estudios, como el del acento, se consideraron palabras de más de tres sílabas.

La realización de segmentos puede verse afectada por la posición del acento primario y la aparición del segmento en la sílaba inicial de la palabra. Estos efectos aparecen en términos de cambios en la duración, y en el valor de las frecuencias de los formantes. De este modo, el estudio fue cuidadoso de si las consonantes

que se analizaron aparecían en posiciones de arranque o coda de la sílaba ya que estas posiciones muestran importante efectos en la duración de los segmentos.

Con respecto al factor género, el estudio incluyó igual número de participantes hombres y mujeres. El género es sumamente importante para el estudio de los formantes y las características espectrales de los segmentos. Las diferencias en el tamaño y longitud de las cuerdas vocales así como también las diferencias en el tamaño de las cavidades orales en hombres y mujeres tienen un impacto directo en cómo la energía acústica se distribuye a lo largo de diferentes frecuencias. En consecuencia, este estudio reporta por separado los resultados relacionados a formantes y espectros.

Los resultados acústicos presentados en este libro se basan en los análisis de palabras objetivos recogidas dentro de frases. Sin embargo, a fin de complementar la información que ellas proporcionaron, el estudio también recogió narraciones. Los datos recogidos a través de narraciones se usaron para hacer comparaciones impresionistas de qué ocurre con las propiedades acústicas de los segmentos en un tipo más espontáneo de habla (comparado con la producción más controlada de las palabras usadas para los análisis acústicos).

Dado que en las narraciones, uno no puede garantizar que los hablantes van a usar las mismas palabras y frases o inclusive producir los mismos segmentos en contextos segmentales o prosódicos similares, los datos que se obtuvieron a través de las narraciones no son comparables y en consecuencia, no se usaron para llevar a cabo los análisis acústicos. En contraste, las palabras que usaron en los análisis, todas ocurrieron dentro de la misma frase, la cual se encargó de proporcionar un contexto prosódico y segmental idéntico, y ya que todos los entrevistados produjeron el mismo conjunto de palabras, se pudo garantizar la comparabilidad de los datos entre diferentes hablantes.

1.3.1 El protocolo de las entrevistas

Las entrevistas se llevaron a cabo en un ambiente cómodo con la ayuda de asistentes de investigación. Los asistentes de investigación recibieron entrenamiento en los lineamientos éticos que se siguen en proyectos de investigación que involucran seres humanos. Del mismo modo, siguiendo tales lineamientos, se les explicó a los colaboradores shipibos que el objetivo del estudio era estudiar los sonidos de su lengua, que la relación entre su identidad y la información lingüística y socio-lingüística que dieran se mantendrían anónima y que podrían detener su participación en el proyecto cuando lo desearan.

Los asistentes de investigación también recibieron entrenamiento en cómo elicitar los datos de acuerdo a los objetivos del estudio y el protocolo que se

describe en esta sección. Durante las primeras entrevistas, los entrevistadores y entrevistados sostuvieron charlas informales para conocerse y hacer las entrevistas menos formales. El nivel de comodidad y confianza que los colaboradores shipibos obtuvieron con los entrevistadores fue sumamente bueno al punto que uno de ellos en algún momento expresó que le gustaba más trabajar en el proyecto que ir a su trabajo regular. Este nivel de comodidad es necesario para garantizar que el tipo de habla recogida sea lo más natural posible.

Durante las primeras entrevistas, los entrevistadores se encargaron de explicar y practicar con los entrevistados la rutina de cómo se iban a elicitar las palabras. Para cada palabra, se le preguntó en castellano al entrevistado ¿Cómo se dice la palabra [X] en Shipibo? En el espacio de la [X] aparecía la traducción castellana de la palabra shipiba que se deseaba obtener. Si la palabra shipiba que se obtenía no correspondía con la palabra esperada, entonces el entrevistador preguntaba por otras posibles palabras relacionadas hasta que se obtenía la palabra deseada: (¿Conoce alguna otra palabra para decir [X] en Shipibo?) Si aún así la palabra objetivo no podía ser elicitada, entonces el entrevistador usaba como último recurso pronunciar la palabra shipiba en la pregunta: ¿Conoce la palabra [X]? Las palabras elicitadas de este modo se les indicaba apropiadamente en la lista de palabras. En la mayoría de los casos, las palabras fueron elicitadas rápidamente ya sea por la primera pregunta o preguntando por palabras relacionadas. Los casos de palabras que tuvieron que ser pronunciados por el entrevistador fueron comparativamente raros.

Una vez que el entrevistado identificaba la palabra objetivo, se le pedía que la pronunciase dentro de la frase que se muestra en (1):

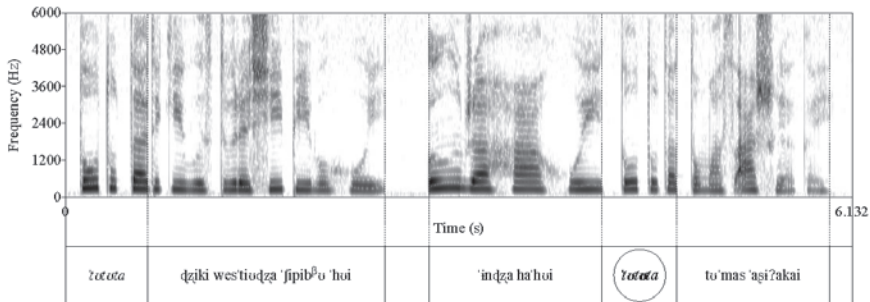
- (1) [[X] dʒiki wiʂ'tiudʒa 'ʃipib^hu 'hwi. 'indʒa ha'hwi [X] tu'mas 'aʃi(?)akai]
 [X] es una shipibo palabra yo la palabra [X] Tomás enseño
 “[X] es una palabra shipiba. Le enseño la palabra [X] a Tomás”

En el sistema ortográfico del shipibo, la frase en (1) se representa como “[X] riqui westíora Shipibo joi. Enra ja joi [X] Tomás aʃheacai”. Esta frase muestra dos lugares, indicados por la [X], donde ocurre la palabra objetivo. Primero, aparece al inicio de la frase y luego entre [ha.'hwi] ‘la palabra’ y [tu.'mas] ‘Tomas’. El entrevistado entonces repetía la frase en (1) tres veces por cada palabra objetivo.

Los resultados que se presentan se basan en el análisis de la segunda ocurrencia de la palabra objetivo dentro de la frase en (1). La ocurrencia inicial de la palabra objetivo tiene como función atraer efectos asociados con prominencia discursiva (información nueva). Obsérvese que la segunda ocurrencia de la palabra objetivo aparece precedida por una palabra que termina en una vocal inacentuada y seguida

por una palabra que empieza con una consonante oclusiva sorda dentro de una sílaba inacentuada. Esta configuración hace más fácil la identificación de los puntos donde empieza y termina la palabra que se está elicitando y al mismo tiempo, evita los efectos de acentos vecinos sobre la palabra objetivo. El espectrograma de la Figura 1-1 ilustra la ocurrencia de la palabra objetivo [ʔutʊtʊ] ‘amarillarse (por ejemplo, una banana)’ (PP2) dentro de la frase en (1). La segunda ocurrencia de la palabra, la cual se usó en los análisis acústicos, aparece encerrada en un círculo. Además, el hecho que la palabra objetivo aparece flanqueada por palabras dentro de una frase neutraliza los efectos segmentales (tales como, glotalización, ensordecimiento o alargamiento de los segmentos) que ocurren a los bordes de los constituyentes prosódicos. Las palabras que se pronuncian aisladamente sufren también estos efectos. Se comportan como unidades prosódicas del nivel de frase y por lo tanto, están también sujetas a los patrones de entonación de las frases (Himmelman and Ladd 2008).

Figura 1-1: [ʔutʊtʊ] ‘amarillarse’ (PP2) como palabra objetivo



Antes de que empezase el recojo de los datos, el entrevistador le dio al entrevistado el tiempo apropiado para aprender la frase en (1). La frase se repitió tres veces por cada palabra. Se les pidió a los colaboradores shipibos que en todo momento usasen una velocidad de habla normal.

1.3.2 El equipo de grabación y el programa de análisis acústico

Las grabaciones de audio fueron hechas usando tres grabadoras digitales: una grabadora portátil Marantz PMD 671 de estado sólido, y dos grabadoras Zoom H4. Las tres grabadoras permiten hacer grabaciones PCM y de este modo, el audio se obtuvo en formato no comprimido WAV, mono, digitalizado a 44.100 Hz (*sampling rate*) y 16-bit (*quantization*). Estos parámetros garantizan la obtención de grabaciones de audio de alta calidad (Ladefoged 2003; Decker and Carrell 2004).

Para las entrevistas, cada grabadora se conectó a un micrófono externo Shure WH30XLR (un micrófono de condensador cardioide con conector XLR que se sujeta a la cabeza del entrevistado). Este tipo de micrófono es unidireccional con una frecuencia mínima de respuesta de 40 Hz, que es bastante cercano a 20 Hz, el promedio de la frecuencia mínima de respuesta de la audición en seres humanos. Dado que el micrófono iba montado sobre la cabeza del entrevistado, la calidad de las grabaciones se mantuvo la misma sin importar los movimientos que el entrevistado hiciese. El micrófono se mantuvo a una distancia de 3 a 4 cm de la esquina derecha de la boca del entrevistado. Las características unidireccionales del micrófono externo garantizaron que las señales acústicas que no viniesen de la boca del hablante shipibo fuesen mayormente ignoradas.

Durante las entrevistas, uno de los micrófonos Shure WH30XLR funcionó mal. Se le descartó y se usó uno de los micrófonos internos que vienen con la grabadora Zoom H4 para continuar con las grabaciones. Los micrófonos internos de esta grabadora también son unidireccionales y tienen una frecuencia mínima de respuesta de 30 Hz.

El estudio utilizó el programa Praat (Boersma y Weenink 2008, versiones 5.0.06 - 5.1.05) para realizar los análisis acústicos. Además de estar bien documentado, este programa es continuamente actualizado. Para este trabajo de investigación, los análisis acústicos realizados con Praat fueron semi-automatizados con la ayuda de scripts y la creación de archivos TextGrids (un tipo de archivo que utiliza Praat y que sirve para segmentar y anotar archivos de audio por medio de la inserción de fronteras y texto). La creación de los archivos TextGrids se llevó a cabo con la ayuda de los asistentes de investigación quienes fueron previamente entrenados en segmentar y anotar el habla contenida en los archivos de audio. Su trabajo fue continuamente monitoreado y revisado para asegurar que se siguieran los procedimientos de segmentación y anotación del estudio.

Una vez que se terminó de hacer los archivos TextGrid, los archivos de audio fueron analizados en Praat. Cuando los análisis involucraban mediciones de formantes, espectros o acentos tonales, se instruyó a los scripts que hicieran pausas de modo que los parámetros para el análisis acústico pudieran ser modificados si fuese necesario a fin de prevenir errores debido, por ejemplo, a factores como género o cambio en la calidad de la voz.

1.4 LA ORGANIZACIÓN DEL LIBRO

El libro está estructurado con el objetivo de servir de referencia al investigador que necesita hallar información específica sobre un determinado tipo de segmento o patrón acentual del shipibo. En el caso de los segmentos, primero presento

un capítulo sobre las vocales y luego en cada capítulo subsiguiente, un tipo de consonante tomando los modos de articulación como criterio agrupador de las consonantes en cada capítulo. Aunque he tratado que cada capítulo dependa en lo mínimo posible de otros capítulos, he colocado una serie de referencias a secciones específicas en otros capítulos de modo que el lector pueda ampliar o relacionar la información que le interesa con la información sobre interacción de esa propiedad con otros segmentos o posiciones prosódicas. La información contenida en este libro, entonces, no necesita ser leída de una manera secuencial. El lector, es por ejemplo, bienvenido a empezar a leer el libro por el capítulo del bosquejo fonológico y conforme encuentre algún patrón que le interese o que está investigando, puede ir al capítulo o a la sección de un capítulo donde se presenta la evidencia fonética que discute el análisis fonológico.

El capítulo 2 está dedicado a la caracterización acústica de las vocales cortas y largas del shipibo. El capítulo las describe de acuerdo a sus patrones de duración y su estructura de formantes. Los capítulos 3 al 6 se enfocan en las consonantes del Shipibo. Estos capítulos brindan una representación detallada de la duración consonántica, de las propiedades espectrales y del comportamiento de los formantes en las transiciones entre vocales y consonantes. El capítulo 3 estudia las características acústicas de las oclusivas. El capítulo 4 se encarga tanto de las consonantes fricativas como africadas. El capítulo 5 documenta las consonantes nasales y las vocales nasalizadas y el capítulo 6 se dedica a las propiedades acústicas de los segmentos semi-consonánticos.

El capítulo 7 ofrece un estudio de las propiedades de duración de los segmentos dentro del dominio de la sílaba y en particular, dentro de la rima (vocales y consonantes en posición de coda). El capítulo 8 presenta una caracterización detallada de los patrones que muestran los acentos tonales en shipibo dentro del dominio de la palabra. Finalmente, el capítulo 9 presenta un bosquejo fonológico del shipibo, el cual no solo conecta los patrones fonológicos de la lengua a los resultados acústicos presentados en los capítulos anteriores sino que también compara los fenómenos fonológicos del shipibo a fenómenos similares encontrados en una diversidad de lenguas. Inmediatamente después el libro termina ofreciendo las conclusiones donde reflexiono con el lector acerca de los beneficios y la necesidad de los trabajos de documentación acústica de lenguas indígenas.

2. VOCALES ORALES

Shipibo tiene cuatro vocales cortas básicas: [i, ɨ, a, u]. Ortográficamente, siguiendo los trabajos de Faust 1973; Lorient, Lauriault *et al.* 1993, la vocal [ɨ] se representa por la letra ‘e’ y la vocal [u], por la letra ‘o’. La mayoría de diccionarios y artículos publicados sobre shipibo siguen esta convención ortográfica. Los datos en (2), tomados de Lorient, Lauriault *et al.* 1993, muestran algunos pares mínimos en los cuales las vocales son fundamentales a las distinciones de significado.

(2) Pares mínimos basados en la calidad vocálica

- a. [ˈta.pi.ti] ‘escalera’
- b. [ˈtɨ.pi.ti] ‘almohada’
- c. [ˈti.pi.ti] ‘estar echados (para cosas o animales)’
- d. [ˈtu.pi.ti] ‘ir juntando o recogiendo’

Shipibo también posee vocales largas ([iː, ɨː, aː, uː]) y vocales nasalizadas ([ĩ, ɨ̃, ã, õ]). Las propiedades fonéticas de las vocales largas las presentaré en la sección 2.4. Las propiedades fonéticas de las vocales nasalizadas las discutiré en el capítulo 5, sección 5.2. El estatus fonológico de estos dos tipos de vocales será discutido en el capítulo 9 (sección 9.1.2 y sección 9.2.4, respectivamente).

2.1 LOS FORMANTES F1 Y F2

Las características articulatorias de las vocales se puede correlacionar desde el punto de vista acústico con la frecuencia de sus formantes (es decir, los picos más altos en un espectro). En particular, se sabe que el *primer formante* y el *Segundo formante* (en adelante, me referiré a ellos como F1 y F2) se correlacionan la altura

vocálica y el grado de posterioridad vocálica, respectivamente (Ladefoged 2005, 2006, Johnson 2003). En general, contra más alto sea el valor de F1, la vocal es más baja. Sin embargo, en el caso de F2, contra más alto sea el valor, la vocal se localiza más al frente en el espacio vocálico. Debido a las diferencias usuales en la forma y tamaño de la cavidad oral entre hombres, mujeres y niños, el promedio de los valores de F1 y F2 son ligeramente diferentes para cada grupo. La Tabla 2-1 muestra el promedio de la frecuencia de los formantes F1 y F2, en Hercios (Hz), para hombres y mujeres.

Los resultados presentados en la Tabla 2-1 se basan en el análisis acústico de aproximadamente 8000 vocales, pronunciadas por tres hablantes hombres y tres hablantes mujeres.¹ Todas estas vocales ocurrieron en palabras trisilábicas (la mayoría de ellas formadas por una raíz bisilábica y un sufijo). Para obtener los valores promedios de F1 y F2, solo se consideraron las vocales de la primera y segunda sílaba de las palabras de modo que se pudiese excluir cualquier efecto prosódico sobre la sílaba final que podría afectar la duración o los formantes de la vocal.

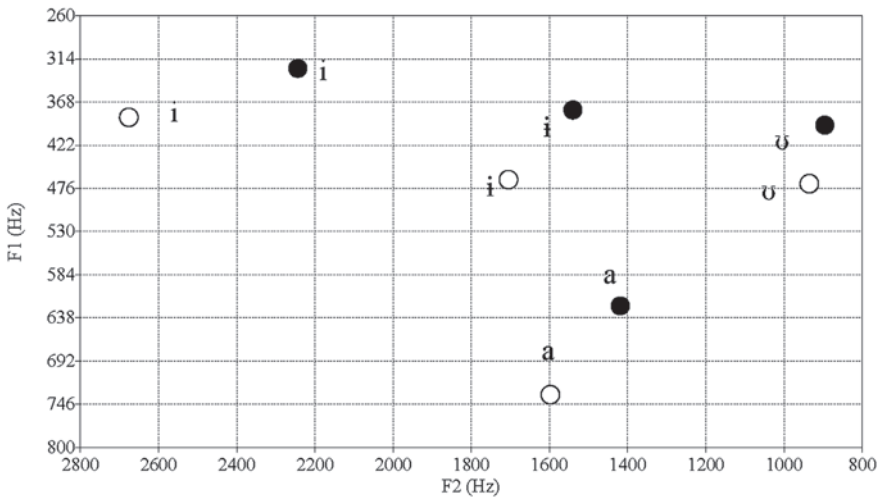
Tabla 2-1: Promedios de las frecuencias de los formantes de las vocales del shipibo

		F1	F2
[i]	Hombres	326 Hz	2244 Hz
	Mujeres	387 Hz	2676 Hz
[ɨ]	Hombres	378 Hz	1539 Hz
	Mujeres	465 Hz	1704 Hz
[u]	Hombres	397 Hz	896 Hz
	Mujeres	470 Hz	935 Hz
[a]	Hombres	623 Hz	1418 Hz
	Mujeres	734 Hz	1597 Hz

¹ Una advertencia importante sobre los promedios para los formantes de las vocales presentados en la Tabla 2-1 es que no discriminan entre vocales en sílabas acentuadas e inacentuadas o entre vocales en sílaba iniciales y no iniciales. Estos son dos factores fonológicos que sí afectan la frecuencia de los formantes en las vocales. Sin embargo, presento los promedios de la Tabla 2-1 porque es útil tener un solo promedio idealizado de F1 y F2 para cada vocal. De este modo, los promedios de la Tabla 2-1 son importantes porque los usaré en los siguientes capítulos para discutir los movimientos de los formantes que ocurren en las transiciones entre vocales y consonantes. Las siguientes secciones de este capítulo mostrarán además varias tablas con los promedios de F1 y F2 teniendo en cuenta los efectos del acento y de la posición de la sílaba.

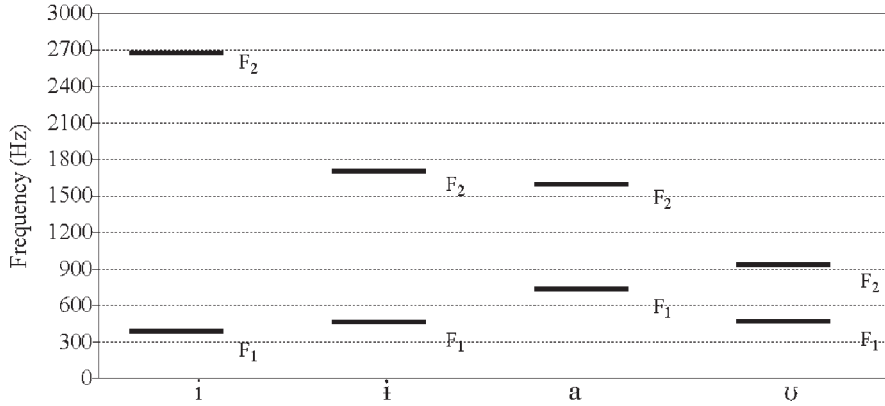
La inspección de los promedios de la Tabla 2-1 revela que la vocal [i] es ligeramente más alta que las vocales [ɨ] y [u] tanto en hombres como en mujeres y que la vocal [ɨ] aparece ligeramente más hacia el frente que la vocal [a]. De hecho, aunque no se muestra en la Tabla 2-1, esto se encontró en todos los hablantes que participaron en el estudio. Podemos obtener una manera más familiar de ver estas relaciones si utilizamos un diagrama de puntos de F1 (en el eje vertical) y F2 (en el eje horizontal). En la Figura 2-1, las vocales de los hombres aparecen representadas por círculos negros y las vocales de las mujeres por círculos blancos. Los valores usados en el diagrama de puntos corresponden a los promedios de la Tabla 2-1.

Figura 2-1: Posición acústica de las vocales orales del shipibo



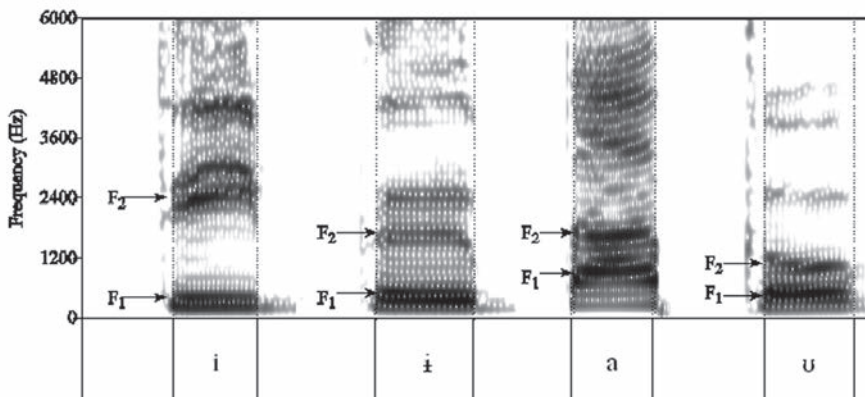
La Figura 2-2 nos permite ver los mismo datos pero de una manera diferente. En esta figura, solo considero los promedios de los formantes de las vocales de las hablantes mujeres. En los capítulos sobre las consonantes, presentaré qué pasa con los formantes a los bordes de las vocales. En el caso de la vocal [i], vemos que el F1 y F2 están muy distantes mientras que ambos formantes se acercan máximamente para la vocal [u]. El F1 y F2 de las vocales [a] y [ɨ] muestran una distancia intermedia. No están tan distantes como en el caso de la [i] pero tampoco tan cercanos como en el caso de la [u]. La diferencia más importante entre las vocales [a] y [ɨ] es el F1, el cual es más alto para la vocal [a].

Figura 2-2: Relaciones entre el F1 y F2 de las vocales del shipibo (mujeres)



La distancia mostrada entre F1 y F2 en la Figura 2-2 no solo es válida para las vocales de las hablantes mujeres. Aunque los varones muestran valores diferentes para los formantes, mantienen las proporciones entre F1 y F2. Esto se muestra en la Figura 2-3, la cual nos permite ver un ejemplo real de los formantes de las cuatro vocales del shipibo de un hablante hombre (HBL6). Todos los espectrogramas corresponden a la vocal acentuada de la sílaba inicial de las palabras ['ti.pi.ti] 'estar echados (para cosas o animales)', ['ti.pi.ti] 'almohada', ['ta.pi.ti] 'escalera' y ['tu.pi.ti] 'ir juntando o recogiendo'. Como se puede observar, las frecuencias de los formantes F1 y F2 en los ejemplos de la Figura 2-3 se acercan bastante a los valores idealizados que se mostraron en la Figura 2-2.

Figura 2-3: Ejemplos de F1 y F2 de las vocales del shipibo (HBL6)



2.2. LA CENTRALIZACIÓN DE LAS VOCALES EN SÍLABAS NO-INICIALES INACENTUADAS

Los resultados presentados en la Tabla 2-1 indican los valores de los formantes F1 y F2 de las vocales. Estos valores corresponden a promedios sin tener en consideración si la vocal es acentuada o no. La presencia o ausencia del acento primario muestra efectos consistentes en los formantes. Sin embargo, en shipibo, el acento primario no es el único factor que puede afectar los formantes de las vocales. La pertenencia a la sílaba inicial de una palabra también es un factor que debe considerarse. En general, las vocales en sílabas débiles tienden a centralizarse mientras que en sílabas fuertes, resisten ser centralizadas. De este modo, las vocales de sílabas no-iniciales inacentuadas tienden a sufrir un mayor grado de centralización que aquellas que ocurren en sílabas iniciales acentuadas.

Las siguientes tablas (Tabla 2-2 a Tabla 2-5) ofrecen los valores promedios de F1 y F2 para cada vocal tomando en cuenta esta vez tanto si la vocal es acentuada y si pertenece a la sílaba inicial de la palabra. Estos valores se obtuvieron de una ventana de tiempo localizada en el medio de cada vocal. La ventana fue equivalente a 30% de la duración total de la vocal. Aunque solo extrajo los valores de las frecuencias de los formantes F1 y F2, al calcular la estructura de formantes de cada vocal, se tomó en cuenta un máximo de cinco formantes dentro del rango 0-5000 Hz para hombres y 0-5500 Hz para mujeres (Ladefoged 2003). Luego, se separó cada vocal de acuerdo al género del hablante, a la posición del acento y a la posición de la sílaba que contiene la vocal dentro de la palabra. Finalmente, se calculó el promedio aritmético y la desviación estándar (la desviación estándar aparece entre paréntesis en todas las tablas del libro). Sobre la base de estas cuatro tablas, se calculó los promedios que se presentaron en la Tabla 2-1.

Tabla 2-2: Vocales orales acentuadas en sílabas iniciales

		F1	F2
[i]	Hombres	341 Hz ⁽²⁶⁾	2271 Hz ⁽⁹⁵⁾
	Mujeres	372 Hz ⁽³⁸⁾	2703 Hz ⁽¹⁰⁸⁾
[i̥]	Hombres	394 Hz ⁽²⁷⁾	1479 Hz ⁽¹⁵⁶⁾
	Mujeres	483 Hz ⁽⁵⁸⁾	1687 Hz ⁽²⁰⁷⁾
[u]	Hombres	414 Hz ⁽³³⁾	862 Hz ⁽⁸⁴⁾
	Mujeres	493 Hz ⁽⁵⁴⁾	887 Hz ⁽¹²²⁾
[a]	Hombres	671 Hz ⁽³⁶⁾	1367 Hz ⁽⁸³⁾
	Mujeres	782 Hz ⁽⁸²⁾	1622 Hz ⁽¹⁰³⁾

Tabla 2-3: Vocales orales acentuadas en segundas sílabas

		F1	F2
[i]	Hombres	347 Hz ⁽²⁷⁾	2281 Hz ⁽¹⁵²⁾
	Mujeres	391 Hz ⁽³²⁾	2658 Hz ⁽¹⁴⁹⁾
[i̥]	Hombres	379 Hz ⁽²⁶⁾	1558 Hz ⁽¹⁶⁴⁾
	Mujeres	472 Hz ⁽⁴⁸⁾	1731 Hz ⁽²³¹⁾
[u]	Hombres	398 Hz ⁽³⁰⁾	870 Hz ⁽⁹⁵⁾
	Mujeres	475 Hz ⁽⁵³⁾	910 Hz ⁽¹⁵³⁾
[a]	Hombres	644 Hz ⁽³⁶⁾	1419 Hz ⁽⁸⁰⁾
	Mujeres	773 Hz ⁽⁸²⁾	1628 Hz ⁽¹⁰⁴⁾

Tabla 2-4: Vocales orales inacentuadas en sílabas iniciales

		F1	F2
[i]	Hombres	294 Hz ⁽²¹⁾	2238 Hz ⁽¹¹⁶⁾
	Mujeres	378 Hz ⁽²⁷⁾	2679 Hz ⁽¹¹⁴⁾
[i̥]	Hombres	372 Hz ⁽³³⁾	1511 Hz ⁽¹⁸²⁾
	Mujeres	440 Hz ⁽³⁸⁾	1695 Hz ⁽²³⁰⁾
[u]	Hombres	396 Hz ⁽³⁶⁾	872 Hz ⁽¹⁰⁸⁾
	Mujeres	448 Hz ⁽⁴³⁾	904 Hz ⁽¹²²⁾
[a]	Hombres	623 Hz ⁽³⁵⁾	1374 Hz ⁽¹¹³⁾
	Mujeres	701 Hz ⁽⁸⁴⁾	1590 Hz ⁽¹¹⁵⁾

Tabla 2-5: Vocales orales inacentuadas en segundas sílabas

		F1	F2
[i]	Hombres	320 Hz ⁽³²⁾	2184 Hz ⁽¹¹⁷⁾
	Mujeres	405 Hz ⁽³²⁾	2603 Hz ⁽¹⁰⁷⁾
[i̥]	Hombres	365 Hz ⁽³⁰⁾	1607 Hz ⁽¹⁹¹⁾
	Mujeres	463 Hz ⁽³⁷⁾	1706 Hz ⁽²³⁵⁾
[u]	Hombres	379 Hz ⁽²⁸⁾	980 Hz ⁽¹⁰¹⁾
	Mujeres	465 Hz ⁽⁴⁷⁾	1039 Hz ⁽⁴⁷⁾
[a]	Hombres	554 Hz ⁽³⁴⁾	1511 Hz ⁽¹³¹⁾
	Mujeres	681 Hz ⁽⁷⁰⁾	1549 Hz ⁽¹²⁴⁾

Los resultados de la Tabla 2-2 a la Tabla 2-5 muestran que las vocales se mueven hacia el centro del espacio vocálico conforme van apareciendo en contextos prosódicos débiles. Contra más débil es la posición prosódica, la centralización de las vocales se hace más fuerte. Si una vocal pertenece a una sílaba que tiene algún tipo de protección ya sea porque está acentuada o porque pertenece a una sílaba inicial, sea ésta última acentuada o no, la centralización de las vocales no es tan dramática como cuando la misma vocal aparece en una sílaba inacentuada no-inicial. Es importante indicar que la centralización vocálica del shipibo es de naturaleza fonética. Esto contrasta, por ejemplo, con la centralización o reducción vocálica de lenguas como el inglés donde una vocal inacentuada se vuelve laxa (por ejemplo, *progress* como el sustantivo ‘progreso’ se pronuncia [ˈprɒgrəs] pero como el verbo ‘progresar’ se pronuncia [prəˈgrɛs]). Sin embargo, en inglés, este cambio no es fonético sino fonológico dado que las vocales laxas [ɪ, ə, ʊ] son categorías fonémicas en la lengua. Esto no ocurre en shipibo, donde por ejemplo la vocal /i/ en una segunda sílaba inacentuada podría alejarse de la posición que tomaría en el espacio vocálico si estuviese acentuada pero esto no implica un cambio en la categoría fonémica.

Los resultados de la Tabla 2-2 a la Tabla 2-5 pueden entenderse mucho más fácilmente en un cuadro vocálico F1-F2. La Figura 2-4 y la Figura 2-5 muestran el comportamiento de los formantes en vocales que pertenecen a sílabas con diferentes propiedades prosódicas. La Figura 2-4 corresponde a las vocales de los hombres y la Figura 2-5, a las vocales de las mujeres. Los círculos negros indican vocales acentuadas en sílabas iniciales; los círculos grises oscuros, vocales acentuadas en segundas sílabas; los círculos grises claros, vocales inacentuadas en sílabas iniciales y los círculos blancos, vocales inacentuadas en segundas sílabas.

La distancia más notable entre los valores de los formantes en términos de la centralización de las vocales se da entre las vocales acentuadas de sílabas iniciales (círculos negros) y las de segundas sílabas inacentuadas (círculos blancos). Obsérvese que las sílabas inacentuadas en sílabas iniciales muestran algún grado de resistencia a ser centralizadas. Éstas tienden a mostrar un comportamiento bastante cercano a las sílabas acentuadas.

Se puede capturar mejor el contraste entre vocales acentuadas de sílabas iniciales e inacentuadas de segundas sílabas, los dos extremos de la centralización fonética de las vocales del shipibo, por medio de un diagrama de puntos. Éste se muestra en la Figura 2-6. Las vocales acentuadas aparecen en negro y las inacentuadas en blanco. He omitido la vocal [ɨ] del diagrama debido a que su realización fonética coincide en parte con los bordes de las otras vocales. El diagrama de puntos correspondiente a la vocal [ɨ] se muestra en la Figura 2-8.

Figura 2-4: Vocales de acuerdo al acento y el número de sílabas (hombres)

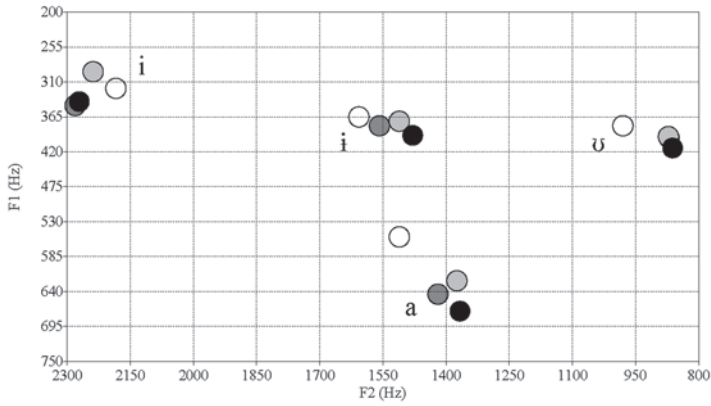
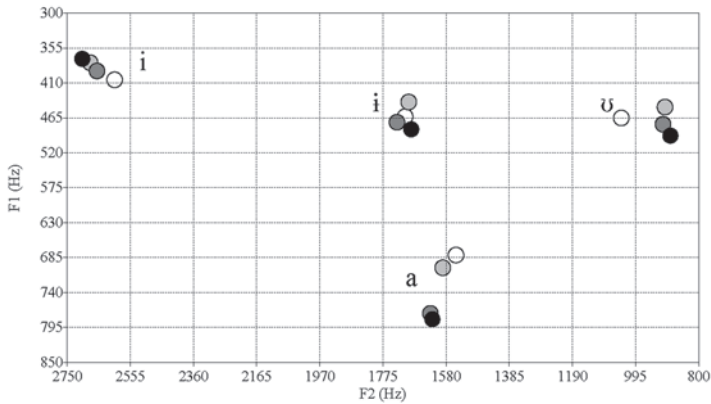
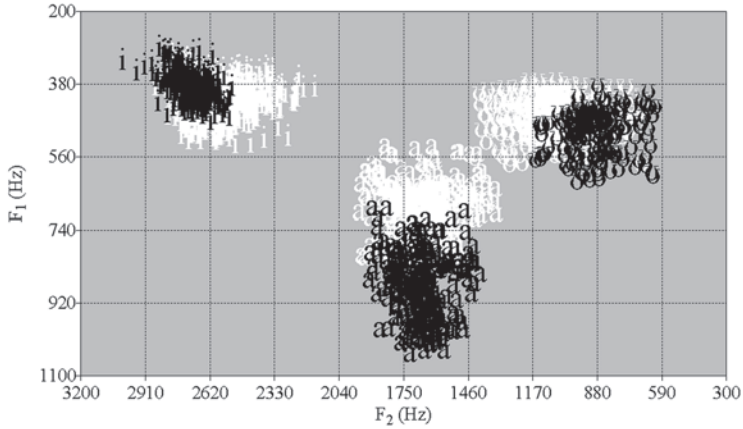


Figura 2-5: Vocales de acuerdo al acento y el número de sílabas (mujeres)



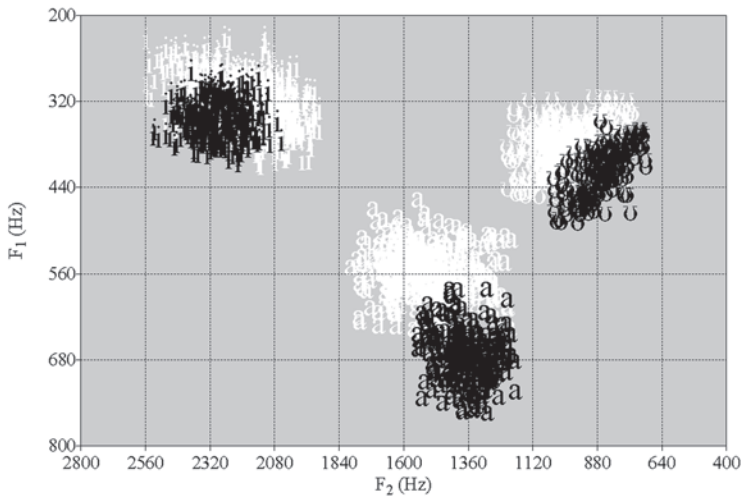
El diagrama en la Figura 2-6 demuestra que aunque las vocales acentuadas e inacentuadas tengan una zona de intersección, las vocales inacentuadas claramente se mueven hacia el centro del espacio vocálico. La vocal [a] se centraliza disminuyendo la frecuencia de su F1 y la vocal [u] incrementando el valor de su F2 mientras que la vocal [i] se mueve hacia el centro disminuyendo su F2.

Figura 2-6: Centralización vocálica en mujeres



Como en el caso de las hablantes mujeres, las vocales de los hombres también sufren centralización en sílabas débiles. La Figura 2-7 permite observar que las inacentuadas vocales [i], [u] y [a] claramente se mueven hacia el centro del espacio vocálico. La vocal [a] muestra el caso más dramático de centralización.

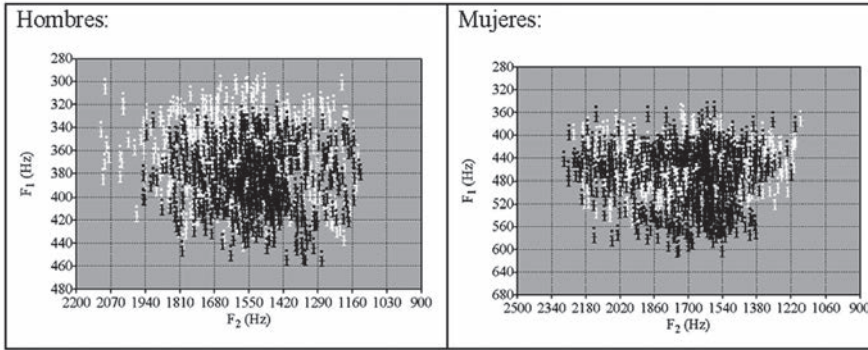
Figura 2-7: Centralización vocálica en hombres



La vocal central [ɨ] no muestra ningún cambio notable en sus formantes F1 y F2 cuando ocurre en sílabas débiles. La Figura 2-8 representa el espacio ocupado por esta vocal, tanto en hombres como en mujeres. A diferencia de

las otras vocales, tanto las versiones acentuadas e inacentuadas de la vocal [i] se sobrepone completamente.

Figura 2-8: La vocal [i] acentuada e inacentuada



2.3 LA DURACIÓN DE LAS VOCALES CORTAS

En shipibo, la duración de las vocales cortas es afectada por el acento y, como en el caso de los formantes, por la posición de la sílaba en la palabra. La Tabla 2-6 muestra los promedios de duración de las vocales cortas en segundos. Discutiré estos resultados después de indicar las condiciones y procedimientos que se utilizaron para obtener las mediciones de la duración vocálica.

Tabla 2-6: Promedios de la duración de las vocales cortas (en segundos)

	Primera sílaba acentuada	Segunda sílaba acentuada	Sílaba inacentuada
[i]	0,094 ^(0,021)	0,109 ^(0,027)	0,066 ^(0,022)
[ĩ]	0,098 ^(0,020)	0,112 ^(0,031)	0,068 ^(0,026)
[u]	0,101 ^(0,023)	0,113 ^(0,031)	0,068 ^(0,022)
[a]	0,111 ^(0,024)	0,127 ^(0,033)	0,080 ^(0,023)

Los resultados provienen del análisis del mismo conjunto de vocales usadas para determinar el valor de los formantes F1 y F2. Los hitos usados para determinar las fronteras de las vocales fueron el comienzo y el final de los formantes más bajos de cada vocal (es decir, F1 y F2). Todas las vocales incluidas en las

mediciones ocurrían entre consonantes oclusivas sordas ([p], [t], [k]) o entre una consonante oclusiva sorda y una africada sonora retrofleja [dʒ].² Las vocales medidas pertenecían ya sea a una sílaba inicial o a una segunda sílaba y todas las sílabas eran abiertas (es decir, sílabas CV). Después que se determinó las fronteras de las vocales, el análisis separó cada tipo de vocal, [i, ī, a, u], de acuerdo a si llevaban acento primario, si pertenecían a una sílaba inicial o a una segunda sílaba y si la vocal fue pronunciada por un hombre o por una mujer.

La Tabla 2-6 no incluye información sobre género debido a que no se encontró diferencias en la duración de las vocales según el género del hablante. En el caso de la duración de las vocales inacentuadas, tampoco se encontró diferencias entre las sílabas iniciales y no-iniciales, su duraciones son virtualmente las mismas. Sin embargo, las vocales acentuadas en segundas sílabas sí tienden a ser ligeramente más largas que aquellas que ocurren en sílabas iniciales.

La Tabla 2-6 muestra que la vocal [a] es la que mayor duración tiene comparada a las otras vocales en todos los tres contextos prosódicos presentados en la tabla: sílabas iniciales acentuadas, segundas sílabas acentuadas y sílabas inacentuadas. No hay una distinción clara en términos de duración que pueda ser observada entre las vocales [ī], [u] y [i].

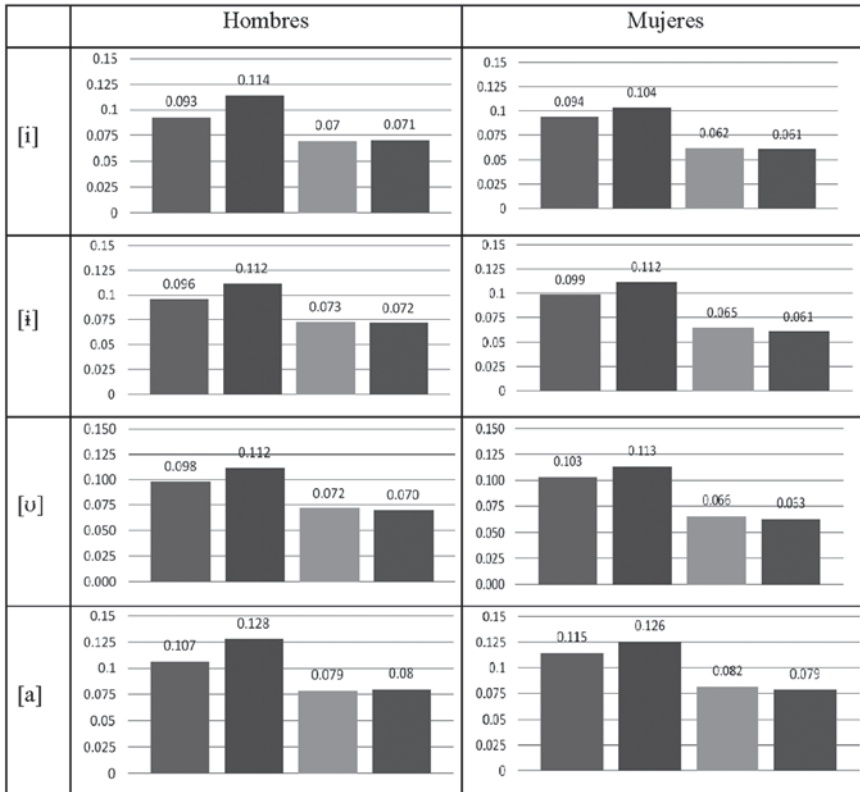
El resultado más notable mostrado en la Tabla 2-6 es la diferencia de duración entre las vocales acentuadas e inacentuadas. Mientras las vocales cortas inacentuadas duran en promedio aproximadamente 70 ms, las vocales cortas que llevan acento primario las sobrepasan por alrededor de 30 a 40 ms.

La Tabla 2-6 también indica que las vocales acentuadas en segundas sílabas tienden a ser ligeramente más largas que las acentuadas en sílabas iniciales. Esto se ilustra en los gráficos de barra de la Figura 2-9. Las vocales con acento primario en segunda sílabas son consistentemente más largas que las vocales acentuadas de sílabas iniciales (por alrededor de 15 ms). La barra de más a la izquierda en cada gráfico representa la duración promedio de la vocal en una sílaba acentuada inicial; la siguiente barra, la duración de la vocal en una segunda sílaba acentuada; la tercera barra, la duración en una sílaba inicial inacentuada y la barra de más a la derecha representa la duración de la vocal en una segunda sílaba inacentuada.

² Este último caso corresponde a algunas vocales que pertenecen a una segunda sílaba en sustantivos. Recuérdese todas las palabras usadas para este estudio son trisilábicas. Sin embargo, debido a que el shipibo no tiene muchas raíces trisilábicas, la mayoría de las palabras utilizadas estaban formadas por una raíz bisilábica seguida de un sufijo de una sílaba (/ti/ (INF), /-ki/ (CMPL), /-a/ (PP2) y /-dʒa/ (EV). Los sufijos /-ti/, /-ki/ y /-a/ sólo se pueden combinar con verbos. El sufijo /-a/ se usó con verbos cuyas raíces terminan en una consonante de modo que la segunda sílaba de la raíz quedaba abierta por la silabificación del sufijo. El sufijo /-dʒa/ se utilizó para crear una tercera sílaba en raíces de sustantivos bisilábicos.

El gráfico de barras muestra por separado las sílabas inacentuadas iniciales y segundas solo con el fin de respaldar la afirmación que tienen virtualmente la misma duración y de poder comparar este comportamiento similar entre vocales inacentuadas, cuya diferencia es la posición de la sílaba donde ocurren, con el comportamiento de las sílabas acentuadas, las cuales bajo las mismas circunstancias, manifiestan diferentes patrones duracionales.

Figura 2-9: Promedios de duración por el tipo de vocales (en segundos)



2.4 LA DURACIÓN DE LAS VOCALES LARGAS

Las vocales largas del shipibo solo pueden ocurrir en la sílaba inicial de las palabras. En todos estos casos, las vocales largas siempre llevan el acento primario. Este tipo de vocales están restringidas a raíces monosilábicas, con la excepción de unas pocas raíces bisilábicas, en las cuales las vocales largas también aparecen en sus

sílabas iniciales. Debido a esta restricción fonotáctica, el trabajo de investigación no pudo incluir mediciones de vocales largas más allá de la primera sílaba. No se puede utilizar prefijos para lograr que una sílaba inicial que tenga una vocal larga se vuelva una segunda sílaba. El shipibo tiene un conjunto bastante restringido de prefijos (los cuales corresponden a partes del cuerpo) y ninguno de ellos puede combinarse con raíces que tienen vocales largas. Otra propiedad de las vocales largas en raíces monosilábicas es que se mantienen fonológicamente largas inclusive cuando reciben sufijos.

Veamos cómo se llevaron a cabo la medición de las vocales largas. Fueron principalmente medidas dentro de palabras trisilábicas (en la mayoría de los casos formados por raíces monosilábicas seguidas de dos sufijos). Esto se hizo para que los resultados puedan ser directamente comparables con la duración de las vocales cortas (las cuales también fueron medidas en palabras trisilábicas). Como en el caso de las vocales cortas, todas las vocales largas pertenecieron a sílabas abiertas. Solo se incluyeron monosílabos con vocales largas que tenían una consonante final *si*, a través de sufijación, esta consonante aparecía como el arranque de la siguiente vocal.

Además de palabras trisilábicas, se decidió medir la duración de las vocales largas en palabras bisilábicas y monosilábicas. Las palabras bisilábicas mayormente incluyeron casos de raíces monosilábicas seguidas de un sufijo que añadía la segunda sílaba. La sílaba donde aparecía la vocal larga siempre fue abierta.

La Tabla 2-7 presenta la duración promedio (en segundos) de las vocales largas. La duración de las vocales largas en palabras trisilábicas tienen virtualmente la misma duración que las vocales largas de palabras bisilábicas. Tienden a durar en promedio 180 ms; es decir, son aproximadamente 80 ms más largas que las vocales cortas acentuadas de sílabas iniciales. Como en el caso de las vocales cortas, aquí también se encontró que la vocal [a] muestra la duración más larga comparada al resto de las vocales.

Las vocales largas presentan una duración particularmente larga en palabras monosilábicas; es decir, cuando una raíz monosilábica ocurre sin ningún sufijo. Véase la segunda columna de la Tabla 2-7. En ese contexto, estas vocales tienden a durar alrededor de 210 ms, lo cual quiere decir que son aproximadamente 30 ms más largas que las vocales largas de palabras bisilábicas y trisilábicas.

Es necesario indicar dos limitaciones que se encontraron al obtener las mediciones que presenta la Tabla 2-7 para las vocales largas. Primero, el tamaño de la muestra es mucho más pequeña que la que se usó para medir las vocales cortas. Combinando juntas las vocales largas de las palabras trisilábicas y bisilábicas, los resultados que se muestran en la tercera y cuarta columnas de la Tabla 2-7 corresponden a las mediciones de 1753 vocales (693 palabras trisilábicas y

**Tabla 2-7: Promedios de la duración de las vocales largas
(en segundos)**

	Vocales largas en palabras monosilábicas	Vocales largas en palabras bisilábicas	Vocales largas en palabras trisilábicas
[i:]	0,204 ^(0,045)	0,172 ^(0,034)	0,177 ^(0,030)
[ɪ:]	0,209 ^(0,044)	0,165 ^(0,033)	0,169 ^(0,032)
[u:]	0,205 ^(0,041)	0,174 ^(0,034)	0,180 ^(0,031)
[a:]	0,214 ^(0,047)	0,192 ^(0,041)	0,191 ^(0,035)

1060 bisilábicas). Los resultados presentados para las vocales largas de palabras monosilábicas corresponden a las mediciones de 201 vocales. Esta muestra más pequeña, comparada a las 8000 vocales medidas en el caso de las vocales cortas, se debe al pequeño número de raíces monosilábicas que tiene el shipibo.

Los datos en (3) presentan una lista de las raíces monosilábicas que tienen vocales largas. Esta lista se recopiló del diccionario de Loriot, Lauriault *et al.* 1993. Esta lista nos permite explicar la segunda limitación que encontró el estudio, la cual está relacionada al tipo de consonantes que rodean a la vocal larga. Aquellas vocales largas de monosílabos que empiezan con una semi-consonante, [j] o [w] (indicadas en las celdas grises oscuras en (3)), no se midieron por lo difícil que es determinar exactamente dónde termina la semi-consonante y dónde empieza la vocal. Otro conjunto problemático de arranques de sílabas fueron las consonantes [h] y [b^β] (indicadas en las celdas grises claras). La fase fricción de estos segmentos tiende a sobreponerse con el inicial de las vocales. Se tuvo que descartar aquellas instancias en las cuales no era posible establecer de un modo inequívoco las fronteras de las vocales largas.

(3) Raíces monosilábicas con vocales largas (I)

1. b ^β i:	‘zancudo’	17. mu:	(esp. de sapo)
2. b ^β i:n	‘pegajoso’	18. na:	‘nido’
3. b ^β u:	‘cabello, pelo’	19. nɪ:	(esp. de lagartija)
4. b ^β u: -ti	‘sanarse (una herida)’	20. nɪ: -ti	‘trepar’
5. b ^β u:ʂ	‘de cabello blanco’	21. ni:	‘monte, selva’
6. tʃa:	(esp. de pájaro)	22. pa:n	‘pan’ (del castellano [pan])
7. tʃi:	‘fuego’	23. dʒa:n -ti	‘enviar’
8. hɪ:	(esp. de hormiga)	24. dʒu:	(esp. de mono)
9. hu:	(esp. de sapo)	25. dʒu:n	‘ruido’
10. ki:n	‘deseo’	26. ʂu:	‘verde’
11. ki: -ti	‘esconderse detrás (de algo o alguien)’	27. ti:	‘trabajo’
12. ki:s -ti	‘mirar el muslo de alguien’		‘quitar, recoger, astillar,
13. ku:	‘pus’	28. tsa: -ti	quebrar’
14. ma:n -ti	‘caer (en masa)’	29. wa: -ti	‘cocinarse’
15. mi: -ti	‘tocar’	30. ja:n	‘garrapata’
16. mi:n -ti	‘enterrar’	31. ju: -ti	‘calentarse’

No solo ciertos arranques de sílaba presentan problemas para la segmentación de las vocales largas, sino también algunos sufijos. El shipibo tiene un conjunto de raíces verbales monosilábicas que siempre ocurren seguidas por los sufijos /-ʔiti/ (VZ2) y /-ʔati/ (VZ1). Véase los datos en (4). Ambos sufijos significan ‘hacer, causar’.³ Para el análisis, se excluyeron todas las vocales largas en (4). La razón para ello es que las realizaciones de los sufijos /-ʔiti/ y /-ʔati/ no permiten identificar con exactitud las fronteras de las vocales largas. Algunas veces estos sufijos se realizan como una oclusiva glotal ([-ʔiti] y [-ʔati]) y otras veces, se realizan sin la oclusiva glotal (es decir, como [-iti] y [-ati]).

Aún cuando las oclusivas glotales de esos sufijos se realicen, no es posible determinar dónde termina la vocal larga ya que la transición de la vocal larga hacia la primera vocal del sufijo ocurre sin ningún hito que permita separar con exactitud ambas vocales. La ocurrencia de la oclusiva glotal de los sufijos /-ʔiti/ y /-ʔati/ tampoco es de mucha ayuda. El problema es que la oclusiva glotal usualmente se manifiesta como voz glotalizada, la cual se extiende tanto sobre el final de la vocal larga como sobre el inicio de la siguiente vocal (véase la sección 2.5). Este fenómeno ocurre incluso cuando una consonante nasal sigue a la vocal larga ([si:n.ʔiti] ‘silbar’). En esos casos, la nasal y la oclusiva glotal tienden a coalescer tanto con la vocal que las precede como con la que las sigue haciendo aún más

³ El primero, /-ʔiti/ (VZ2), se usa cuando el verbo es intransitivo y el otro, /-ʔati/ (VZ1), cuando el verbo es transitivo.

difícil medir la duración de las vocales largas (para más información sobre este fenómeno de coalescencia, véase la sección 5.3).

(4) Raíces monosilábicas con vocales largas (II)

1. 'ma:n.ʔiti	'mugir (el ganado)'	7. 'su:.ʔiti	'hincharse'
2. 'sa:.ʔati	'soltar haciendo ruido'	8. 'su:n.ʔati	'hacer sonar (por ejemplo, una bocina)'
3. 'sa:.ʔiti	'soltarse haciendo ruido'	9. 'su:n.ʔiti	'silbar'
4. 'si:.ʔiti	'escurrirse, derramarse haciendo ruido'	10. 'ti:.ʔati	'tocar'
5. 'sin.ʔati	'pitar'	11. 'tu:.ʔati	'disparar'
6. 'sin.ʔiti	'silbar'	12. 'tsa:n.ʔati	'Hacer sonar (por ejemplo, una campana)'

Otro grupo de vocales largas, que tuvieron que ser excluidas de las mediciones debido a los segmentos que las rodean, está formado por una media docena de raíces bisilábicas cuya sílaba inicial tiene una vocal larga. Este tipo de raíz es extremadamente rara en shipibo. Solo sé de la existencia de las seis raíces que se muestran en (5). La mayoría de las palabras excluidas (véase las celdas grises) tienen arranques iniciales que hacen difícil la segmentación de la vocal que sigue. En el caso de la raíz /na:b^βa/, el análisis solo incluyó aquellas realizaciones en las cuales la vocal larga pudo ser claramente segmentada de la consonante africada que le sigue. Se descartó la palabra /dʒu:nb^βi/ (esp. de árbol) porque su sílaba inicial tiene una coda y por lo tanto, no se ajustaba al criterio de medir las vocales en sílabas abiertas.

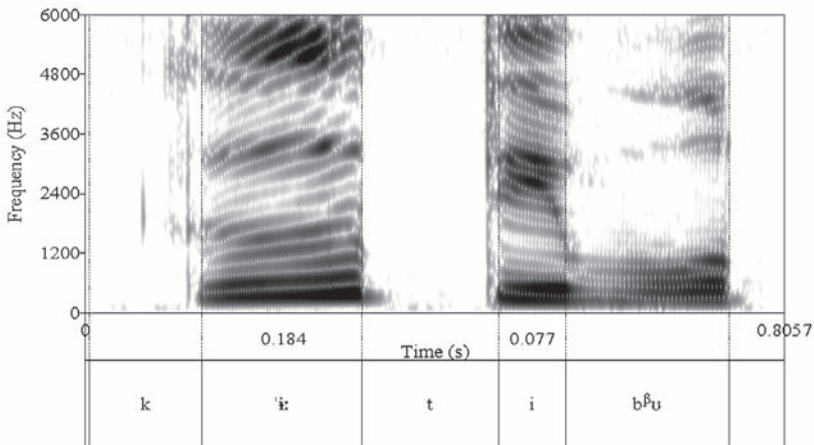
(5) Raíces polisilábicas con vocales largas

- a. na:b^βa -ti 'aumentar líquido' (INF)
- b. dʒa:ʂi -ti 'aclimatarse' (INF)
- c. dʒu:kun 'tos ferina'
- d. dʒu:nb^βi (esp. de árbol)
- e. wa:ta 'tía materna'
- f. wi:ki -ti 'entrar en la pierna' (INF)

2.4.1 Algunas Ilustraciones de las Vocales Largas

En esta sección voy a presentar algunos ejemplos de la duración de las vocales largas del shipibo. El espectrograma de la Figura 2-10 ilustra un ejemplo de la vocal larga [i:] dentro de la palabra [kʰi:ti.bʰu] ‘escudos’. Esta palabra trisilábica está formada por la raíz monosilábica /kʰi:/ ‘escondese detrás de (algo o alguien)’ y por los sufijos /-ti/ (NMLZ) y /-bʰu/ (PL). Compárese la duración de la vocal larga de la Figura 2-10 con la duración de la vocal inacentuada que le sigue. Hay una diferencia de casi 110 ms entre ellas.⁴

Figura 2-10: [kʰi:ti.bʰu] ‘escudos’



En la Figura 2-11, tenemos otro ejemplar de una palabra trisilábica. Sin embargo, esta vez la raíz que contiene la vocal larga es bisilábica: /na:i.bʰa/ ‘aumentar líquido’. Esta raíz aparece seguida por el sufijo /-ti/ (INF). Nótese que la duración de la vocal larga es similar a la de la Figura 2-10. En esta palabra, también podemos comparar las diferencias de duración entre una vocal [a:] larga acentuada y una vocal [a] corta inacentuada.

Mientras la Figura 2-10 y la Figura 2-11 nos mostraron vocales largas en palabras trisilábicas, la Figura 2-12 nos presenta una vocal larga en una palabra

⁴ La duración de esta vocal larga también puede compararse con la duración de la vocal corta acentuada [i] que ocurre en la palabra [sʰi.kʰi.ti] ‘quebrar’ (véase el capítulo 9, sección 9.2.3). Ambas vocales están acentuadas, pertenecen a una palabra trisilábica y aparecen en una sílaba inicial. La vocal larga [i:] dura aproximadamente 80 ms más que la vocal corta [i].

bisilábica. Una vez más, la duración de la vocal larga sigue siendo similar a los casos anteriores.

Figura 2-11: [¹na.bβa.ti] ‘aumentar líquido’

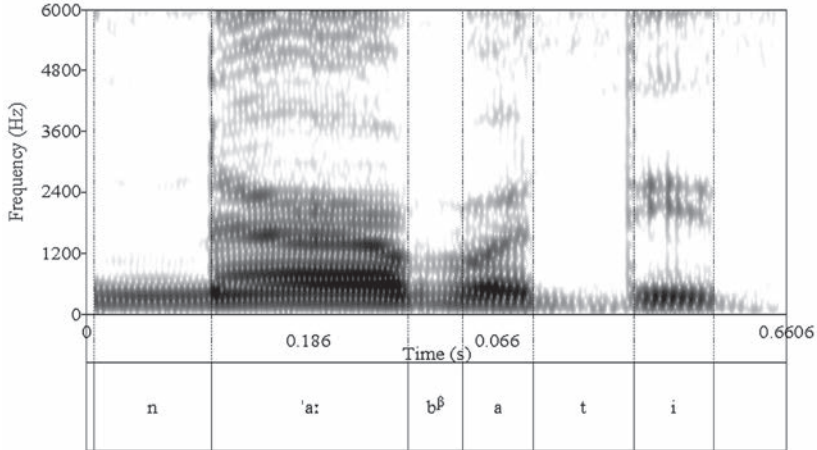
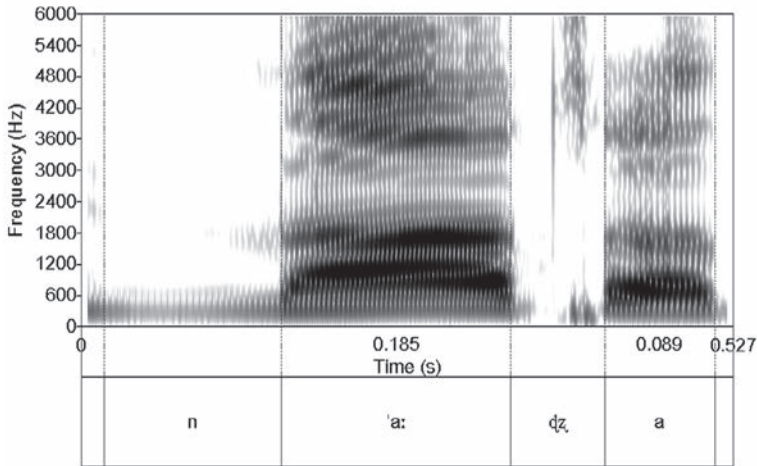
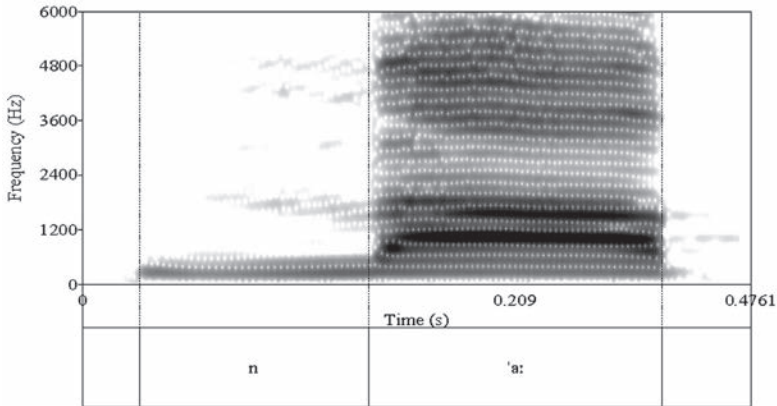


Figura 2-12: [¹na.ɟza] ‘nido’ (EV)



En el espectrograma de la Figura 2-13 tenemos una vocal larga en una palabra monosilábica. El espectrograma representa la palabra [¹na:] ‘nido’. En este contexto, si observamos que la vocal larga es aproximadamente 20 ms más larga que las vocales largas de la Figura 2-10 a la Figura 2-12. Ésta es un fenómeno que tiende a ocurrir en el shipibo y que es reflejado en los promedios de duración que se indican en la Tabla 2-7.

Figura 2-13: Vocal larga en palabra monosilábica - [ˈna:] ‘nido’



2.5 VOCALES EN CONTACTO CON OCLUSIVAS GLOTALES EPENTÉTICAS Y SUBYACENTES

Con la excepción de las palabras dentro de compuestos, las oclusivas glotales usualmente aparecen entre vocales adyacentes en fronteras de palabras, especialmente cuando no hay una pausa entre ellas. El espectrograma en la Figura 2-14 muestra un ejemplo de este tipo de oclusiva glotal (véase las estriaciones en el espectrograma entre las vocales adyacentes). Una oclusiva glotal, que se realiza como voz glotalizada, aparece insertada entre las vocales de las palabras adyacentes [dʒa.ˈbβi] ‘dos’ y [ˈa.ni] ‘grande’. El espectrograma de estas palabras se tomó de la oración [dʒa.ˈbβi ˈʔa.ni ˈbβa.wa.dʒa ˈju.bβi.kan ˈbβi.ki] ‘el brujo trajo dos loros grandes’ (literalmente: *dos grandes loros, brujo trajo*).

Al interior de morfemas, las oclusivas glotales nunca aparecen entre vocales adyacentes. El primer caso puede verse en el espectrograma de la Figura 2-15, el cual nos muestra la palabra [ˈti.a.mis] ‘adúltero’ (/tia/ ‘cometer adulterio’, /-mis/ AGTZ). Las líneas punteadas blancas en este espectrograma representan los formantes de las vocales. Los muestro a fin de hacer más fácil la identificación de las vocales adyacentes. Además, debido a la dificultad de segmentar vocales adyacentes, solo indico sus símbolos fonéticos aproximadamente en el medio de sus duraciones. Lo importante en este espectrograma es notar que en la sección correspondiente a la raíz /tia/ no ocurre ninguna oclusiva glotal entre las vocales.

Figura 2-14: [dʒa.ʼbβi ʼʔa.ni ...] ‘dos grandes...’

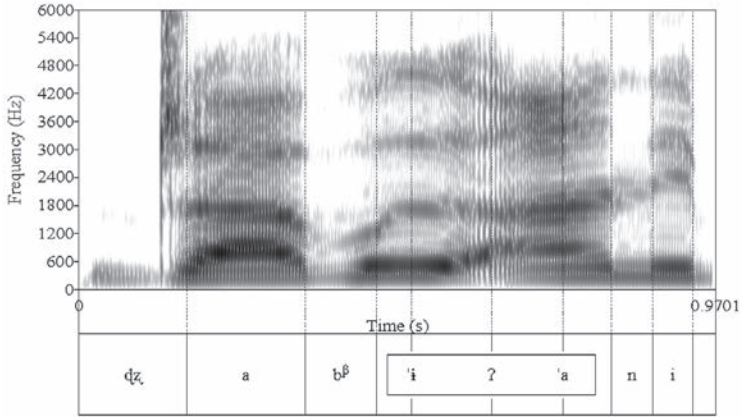
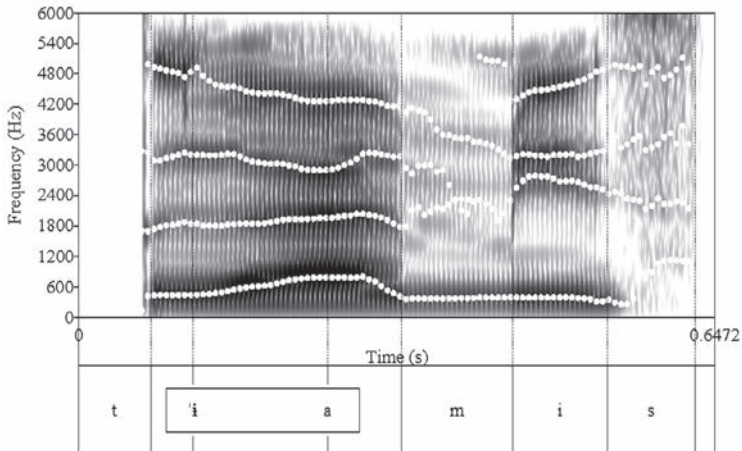
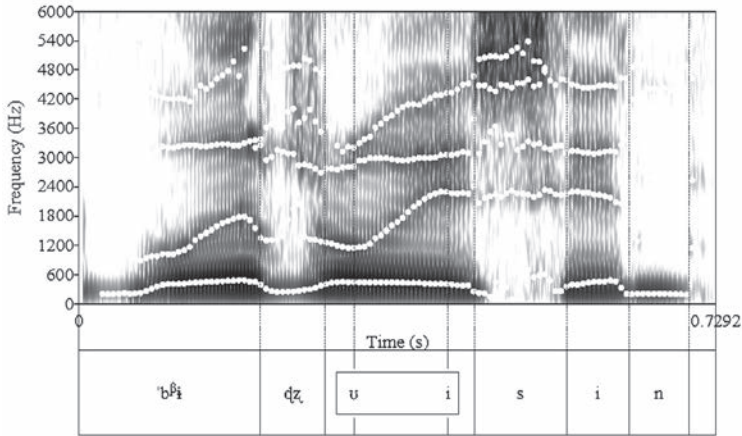


Figura 2-15: /tia -mis / → [ʼti.a.mis] ‘adultero’



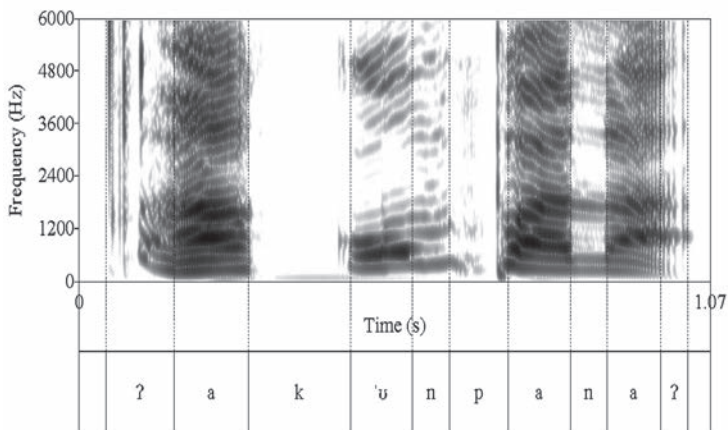
Las vocales adyacentes en las fronteras entre palabras dentro de un compuesto tampoco muestran oclusivas glotales. Es decir, con respecto a la oclusiva glotal, estas palabras se comportan como si fuesen parte de un solo morfema. El espectrograma en la Figura 2-16 presenta la palabra compuesta [ʼbβi.dʒʊ.i.sin] ‘conjuntivitis’, la cual está formada por las palabras /bβi.dʒʊ/ ‘ojo’ y /isin/ ‘enfermedad’. No aparece ninguna oclusiva glotal entre estas palabras, aunque la primera termina en una vocal y la siguiente empieza en una vocal.

Figura 2-16: /b^βi dʒu isin/ → [ˈb^βi.dʒu.i.sin] ‘conjuntivitis’



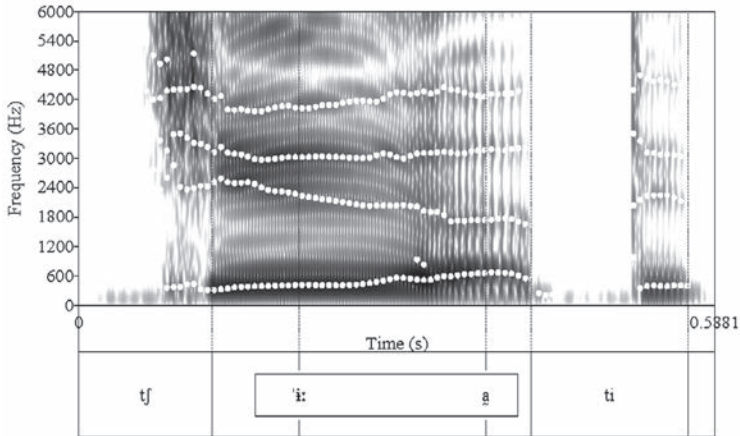
Además de las oclusivas glotales insertadas cuando dos vocales aparecen adyacentes en las fronteras de palabras, los hablantes de shipibo usualmente añaden una oclusiva glotal al inicio y final de frases si estas empiezan o terminan en una vocal. Lo mismo ocurre cuando las palabras se pronuncian en aislamiento (esto se debe a que palabras aisladas equivalen prosódicamente a frases). La Figura 2-17 nos presenta un ejemplo de este fenómeno: [ʔaˈkum.pa.naʔ] ‘cascabel’. Las oclusivas glotales aparecen en el espectrograma como estriaciones verticales.

Figura 2-17: /akunpana/ → [aˈkum.pa.na] ‘cascabel’
(realizado como [ʔaˈkum.pa.naʔ])



De este modo, como se puede apreciar, la mayoría de las oclusivas glotales son predecibles en shipibo. Esto contrasta con el estatus de la oclusiva glotal en otras lenguas pano como, por ejemplo, el capanahua (Elias-Ulloa 2009, Gonzalez (en preparación)). La única excepción a esta predictibilidad en el shipibo son algunos casos de sufijos en los cuales las oclusivas glotales son subyacentes (sobre la interacción entre las oclusivas glotales subyacentes con las nasales y semi-consonante, véase la sección 5.3). Este tipo de oclusiva glotal ocurre en sufijos como /-ʔati/ (VZ1), /-ʔiti/ (VZ2), /-ʔib^βat/ (PSD2) y /-ʔidza/ (INTENS). Los espectrogramas de la Figura 2-18 y la Figura 2-19 muestran ejemplos de los sufijos /-ʔati/ y /-ʔiti/, donde la oclusiva glotal tiende a realizarse como voz glotalizada.

Figura 2-18: /tʃi: -ʔati/ → [tʃi:.ʔa.ti] ‘volver anémico (a alguien)’
(realizado como [tʃi:.a.ti])



Algunos hablantes rara vez realizan las oclusivas glotales de sufijos como /-ʔati/ y /-ʔiti/. Un ejemplo de este tipo de hablante se da en el espectrograma de la Figura 2-20. Este espectrograma muestra la misma palabra de la Figura 2-19, [ʔa.ʔi.ʔi.ti] ‘eructar’, pero realizada como [ʔa.ʔi.i.ti], sin la oclusiva glotal entre las vocales [i] y [i].

Uno podría preguntarse si es que los sufijos /-ʔati/ (VZ1) y /-ʔiti/ (VZ2) realmente tienen una oclusiva glotal subyacente. Uno podría imaginarse un escenario alternativo donde se insertan las oclusivas glotales entre vocales adyacentes en las fronteras morfélicas. Sin embargo, hay evidencia fonológica que señala que esta alternativa es incorrecta. Primero, hay sufijos que realmente empiezan con una vocal (véase algunos ejemplos adicionales en (6)). Cuando esos sufijos se añaden a una raíz que termina en vocal, no aparece ninguna oclusiva glotal.

Figura 2-19: /a^h -ʔiti/ → [a.^hi.ʔi.ti] ‘eructar’
(realizado como [ʔa.^hi.i.ti])

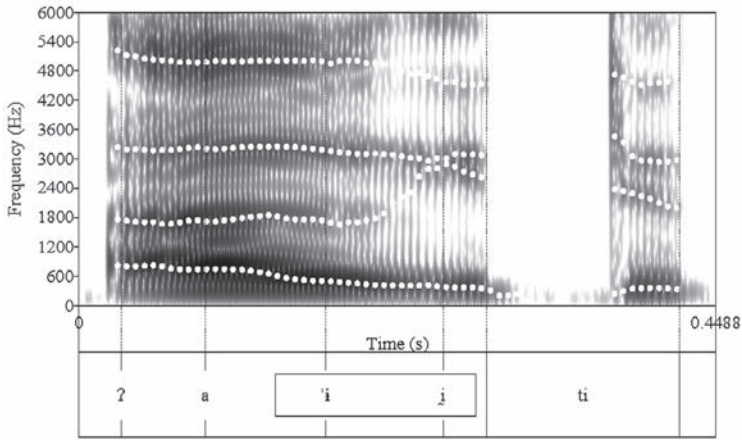
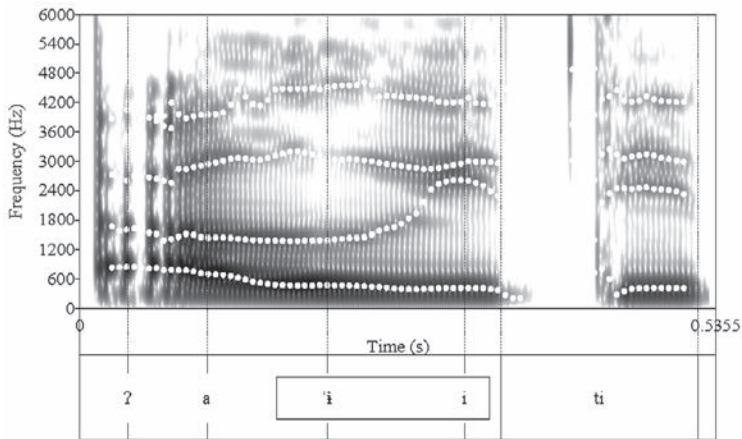


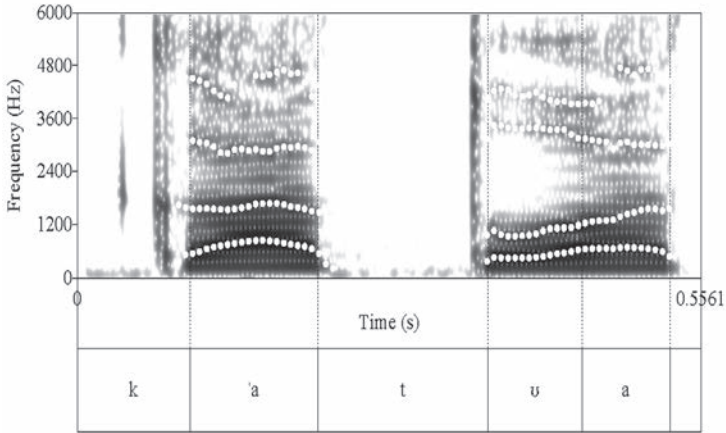
Figura 2-20: /a^h -ʔiti/ → [a.^hi.ʔi.ti] ‘eructar’ (realizado como [ʔa.^hi.i.ti])



Por ejemplo, tomemos el sufijo /-a/ (PP2) y agreguémoslo al verbo /katu/ ‘doblar’. La palabra resultante se muestra en la Figura 2-21. Obsérvese que no hay una oclusiva glotal entre las vocales hetero-morfémicas [u] y [a].

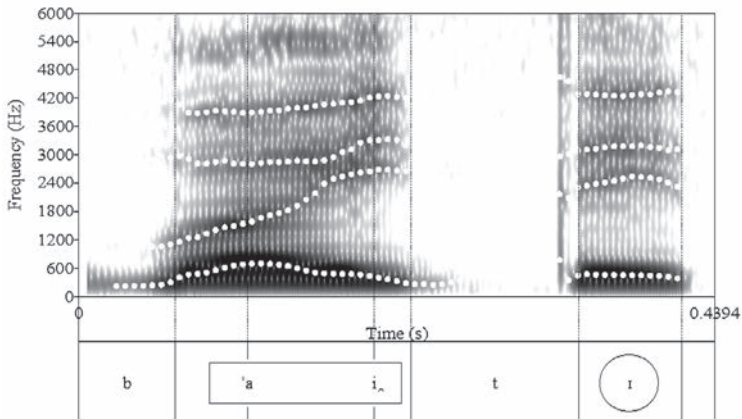
Puede encontrarse evidencia fonológica adicional en favor del estatus subyacente de las oclusivas glotales en sufijos como /-ʔati/ y /-ʔiti/ en la interacción entre la asignación del acento primario y la silabificación. En shipibo, las oclusivas glotales nunca se realizan cuando aparecen adyacentes a una fricativa. Si, por ejemplo, el sufijo /-ʔati/ se combinara con la raíz /b^βakuʃ/ ‘espuma’ la palabra

Figura 2-21: /katu -a/ → [k̟a.tu.a] ‘doblar’ (PP2)



resultante se realizaría como [b^βa.¹kuʂ.a.ti] ‘hacer espuma’. Esto se puede observar en el espectrograma de la Figura 2-22. Si la oclusiva glotal no fuese subyacente, entonces esperaríamos que la consonante sibilante ocurriese como el arranque de la siguiente vocal pero entonces, la segunda sílaba ya no sería cerrada y el acento aparecería sobre la primera sílaba, *[¹b^βa.ku.ʂa.ti], pero como indica el asterisco, esto no es aceptable en shipibo.

Figura 2-22: /bakuʂ -ʔati/ → [ba. kuʂ.ʔa.ti] ‘hacer espuma’
(realizado como [ba.¹kuʂ.a.ti])



En contraste, si reemplazamos el sufijo /-ʔati/ por un sufijo que no empieza con una oclusiva glotal, como el sufijo /-a/ PP2, y lo añadimos a la raíz /b^βakuʂ/ ‘espuma’, la fricativa sibilante esta vez sí se silabifica como el arranque de la

siguiente vocal y el acento cae sobre la primera sílaba porque la segunda es abierta: /b^βakuş -a/ → [b^βa.ku.ʂa] ‘formarse espuma’ (verbo intransitivo - PP2).

2.6 EL COMPORTAMIENTO DE LOS FORMANTES EN VOCALES ADYACENTES: ASIMILACIÓN Y COALESCENCIA

Las vocales que aparecen adyacentes influyen sus formantes mutuamente. En shipibo, este fenómeno es particularmente fuerte en vocales adyacentes dentro de sufijos. Es tan fuerte que las vocales muestran diferentes grados de asimilación y coalescencia. Los datos en (6) presentan una lista de sufijos con vocales adyacentes.

(6) Lista de sufijos que contienen vocales adyacentes

a.	/-ai/	(PP1)
b.	/-ʔain/	‘a, hacia’
c.	/-b ^β ain/	(AND2)
d.	/-b ^β ait/	‘todo el día’
e.	/-b ^β aun/	(FLPL)
f.	/-kain/	(AND1)
g.	/-kian/	(FRUSTR)
h.	/-mia/, /-kia/	‘de, desde’
i.	/-main/, /-miin/	(ESPC)
j.	/-nia/	‘de, desde’
k.	/-ua/	‘de, desde’
l.	/-painun/	‘ojalá que’
m.	/-pauni /	(REM2)
n.	/-pikaʉ/	‘después’
o.	/-tian/	‘cuando, en aquel tiempo cuando’
p.	/-tiu/	‘de tamaño similar a (un objeto)’
q.	/-jauma/	(PRIV)

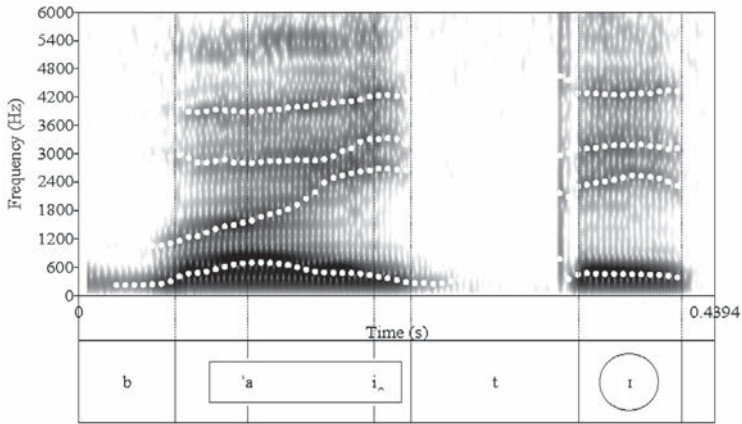
Tomemos el sufijo /-ai/ (PP1) para ejemplificar este fenómeno. Este sufijo tiene una serie de posibles realizaciones fonéticas. Puede aparecer como [-aĩ], [eĩ], [-e] o [-i]. Las últimas tres realizaciones ocurren más frecuentemente que la primera. Es importante resaltar que solo las vocales dentro de los sufijos sufren este proceso. La misma secuencia de vocales dentro de una raíz o en la frontera entre sufijos, por ejemplo, no van a comportarse así.

El espectrograma de la Figura 2-23 muestra un ejemplo claro de la asimetría entre vocales adyacentes en sufijos y vocales adyacentes en raíces. Se trata de la

forma verbal [¹ba.i.ta.i], la cual está formada por la raíz /bait/ ‘volverse torren-
so’ y el sufijo /-ai/ (PP1). Obsérvese que tanto la raíz como el sufijo contienen
la secuencia vocálica /ai/. Sin embargo, en el espectrograma, se puede observar
que mientras la secuencia /ai/ en la raíz verbal se realiza visiblemente como dos
segmentos, en el sufijo, la secuencia /ai/ se realiza como una sola vocal: [ɪ].

En la raíz verbal, se puede ver que el punto más alto del F1 el cual corres-
ponde a la vocal [a] y el punto más alto del F2 que corresponde a la vocal [i]. En
contraste, durante la duración del sufijo, el F2 se mantiene suficientemente alto
como para poder aseverar que se trata de una vocal frontal.

**Figura 2-23: /bait -ai/ → [¹ba.i.ta.i] ‘se vuelve torrenoso’ (PP1)
(realizado como: [¹baj.tɪ])**



Las vocales adyacentes en fronteras morfélicas se comportan como vocales
adyacentes dentro de raíces. Su grado de asimilación y coalescencia es mínimo
cuando se les compara con lo que se observa dentro de los sufijos. Los datos en
(7) a (9) muestran secuencias /ai/ que ocurren en diferentes ambientes. Ninguna
de ellas ocurre dentro de las fronteras de un sufijo. En (7), la secuencia /ai/ ocurre
en la frontera entre una raíz y un sufijo. En (8) y (9), ocurre en la frontera entre
sufijos. En todos los casos, tanto /a/ como /i/ se realizan como dos segmentos:
[a.i] ~ [aj].

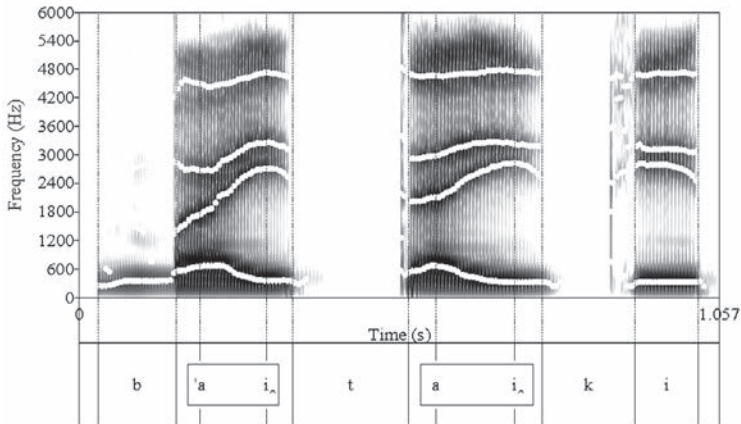
(7) /b^βana -i -wi/ → [¹b^βa.na.i.wi]
Sembrar -ir -IMP ‘¡Anda y siébralo!’

(8) / b^βana -ma -i -wi/ → [¹b^βa.na.ma.i.wi]
Sembrar -CAUS -ir -IMP ‘¡Anda y haz(lo) que (lo) siembre!’

(9) / b^βait -a -iki/ → ['b^βa.i.ta.i.ki]
 Volverse.torrentoso -PP2 -AUX 'Se volvió torrentoso' (pasado para narraciones)

El espectrograma de la Figura 2-24 muestra una instancia de la forma verbal de (9), realizada como: ['bại.tạ.ki]. Este ejemplo nos permite comparar la secuencia /ai/ dentro de la raíz verbal con la secuencia /ai/ formada por la concatenación de /-a/ (PP2) e /iki/ (AUX). En ambos casos, se puede ver claramente que sí se realiza cada una de las vocales de las secuencias /ai/ (obsérvese el comportamiento de los formantes F1 y F2). Lo que observamos aquí contrasta con el comportamiento de la secuencia /ai/ en el espectrograma de la Figura 2-23.

Figura 2-24: /bait -a -iki/ → ['ba.i.ta.i.ki] (realizado como: ['bại.tạ.ki])



El comportamiento observado en el sufijo /-ai/ (PP1) en el espectrograma de la Figura 2-23 también se encuentra en diferentes grados en otros sufijos de la lista en (6). Sin embargo, aquellos que tienen la secuencia /ai/ son los más conspicuos. En los sufijos que tienen la secuencia /au/, las vocales adyacentes tienden a realizarse como: [ạụ], [ə̣ụ] u [ọ] (la segunda realización es la que más comúnmente se escucha). Las secuencias /ia/, /ịụ/, /ụa/ y /ịa/ usualmente se realizan como, dentro de sufijos, como [ịạ], [ịụ], [ụạ] y [ịə̣], respectivamente.

2.7 RESUMEN

En este capítulo he examinado y determinado las propiedades acústicas de las vocales orales del shipibo (tanto de las vocales cortas [a, i, u] como de las largas [aː, iː, uː]). Las vocales se han caracterizado de acuerdo a su estructura de

formantes y su duración. Para todas las mediciones, se tomó en cuenta el género de los hablantes.

Con respecto a su estructura de formantes, cada vocal fue definida en términos de F1 (el formante que indica la altura vocálica) y F2 (el formante que revela el grado de posterioridad de la vocal). Como hemos visto, entre las vocales altas, la vocal frontal [i] representa una posición ligeramente más alta en el espacio vocálico que las vocales [ɨ] y [u]. Adicionalmente, el estudio encontró que la pertenencia a una sílaba acentuada o a una sílaba inicial es un factor importante que afecta la estructura de los formantes de las vocales. Las vocales en sílabas iniciales de palabra o en sílabas acentuadas tienden a ofrecer una mayor resistencia a la centralización. Las vocales inacentuadas que no pertenecen a una sílaba inicial muestran un mayor grado de centralización.

Como en el caso de los formantes, el acento es un factor que influye en la duración de las vocales cortas. Se mostró en los datos presentados que las vocales cortas tienden a durar un promedio de 100 ms mientras que las vocales cortas inacentuadas duran alrededor de 70 ms. Las vocales acentuadas en segundas sílabas son ligeramente más largas que aquellas que ocurren en sílabas iniciales. Las vocales largas, las cuales solo pueden ocurrir acentuadas y dentro de la sílaba inicial de las palabras, tienden a durar aproximadamente 180 ms cuando aparecen en palabras polisilábicas y 208 ms cuando aparecen en palabras monosilábicas.

También se examinó las realizaciones fonéticas de las vocales adyacentes dentro de sufijos. Estas vocales sufren diferentes grados de asimilación y coalescencia. Por otro lado, cuando las vocales adyacentes ocurren en la frontera entre palabras, se inserta una oclusiva glotal. En contraste, cuando las vocales en contacto se encuentran dentro de la raíz, éstas no coalescen ni tampoco se inserta una oclusiva glotal.

Al caracterizar el comportamiento de las vocales adyacentes, el capítulo abordó el tema de la realización fonética de las oclusivas glotales, las cuales usualmente se manifiestan por medio de voz glotalizada sobre el final de la vocal anterior y el inicio de la siguiente. La distribución restringida y la casi opcionalidad en la realización de las oclusivas glotales subyacentes del shipibo parecen indicar que son segmentos que están desapareciendo del nivel fonológico. Las oclusivas glotales subyacentes solo se han preservado en la sílaba inicial de ciertos sufijos. Su realización no es obligatoria incluso en casos en que son subyacentes y en general se puede observar que algunos hablantes tienden a realizar las oclusivas glotales subyacentes más que otros.

Habiendo descrito las propiedades acústicas de las vocales orales, ahora nos concentraremos en los segmentos que las rodean: las consonantes. El siguiente capítulo examina las consonantes oclusivas del shipibo.

3. LAS OCLUSIVAS

3.1 INTRODUCCIÓN

Como se muestra en la Figura 3-1, Shipibo tiene dieciséis consonantes. La oclusiva glotal, [ʔ], aparece entre paréntesis debido a la distribución tan restringida que tiene (véase la sección 2.5). Aunque tradicionalmente recibe el nombre de ‘oclusiva’, su realización fonética usual es la fonación glotalizada, la cual se extiende sobre la sección final de la vocal precedente y el comienzo de la siguiente vocal. El símbolo [ʃ] representa una sibilante álveo-palatal y [ʂ], una sibilante retrofleja sorda. El símbolo [b^β] refiere a una africada bilabial sonora; [tʃ], a una africada álveo-palatal sorda; y [dʒ], una africada retrofleja sonora.

Figura 3-1: Las consonantes del shipibo

p	t		k	(ʔ)
	s	ʃ	ʂ	h
b ^β	ts	tʃ	dʒ	
m	n			
w		j		

En la mayoría de los trabajos del shipibo, las consonantes [ʃ, ʂ, b^β, tʃ, dʒ] se representan ortográficamente por las letras ‘sh’, ‘sh’, ‘b’, ‘ch’ y ‘r’, respectivamente. Las semi-consonantes [j] y [w] usualmente se representan por las letras ‘y’ y ‘hu’ y las fricativa glotal sorda, [h], por la letra ‘j’. La oclusiva glotal solo se representa ortográficamente como ‘h’ cuando sigue a la letra ‘n’. En otros contextos, no recibe ninguna representación ortográfica (Faust 1990; Lorient, Lauriault *et al.* 1993).

Este capítulo discutirá las consonantes oclusivas, [p, t, k]. La oclusiva glotal fue presentada en el capítulo anterior (véase la sección 2.5 y la sección 5.3). El objetivo es la caracterización acústica de las oclusivas en términos de su (i) sonoridad, (ii) modo de articulación y (iii) punto de articulación.

3.2 LAS OCLUSIVAS

El shipibo tiene tres oclusivas sordas: [p, t, k]. Estas consonantes solo pueden aparecer como arranques de una sílaba. En los datos en (10) se ofrecen pares mínimos cuya única diferencia es la elección de la oclusiva sorda.

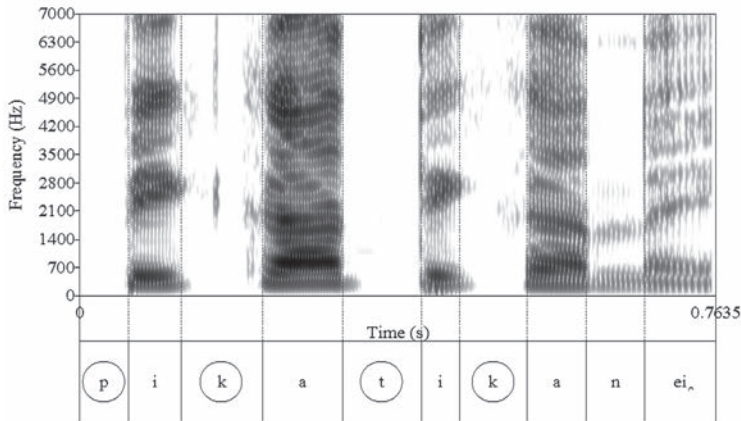
- (10) a. [pa.¹na] ‘cortina de estera’
 b. [ka.¹na] ‘relámpago’
 c. [¹ta.wa] ‘cañabrava (esp. de vegetación)’
 d. [¹ka.wa] (esp. de planta)
 e. [¹ta.ma] ‘maní’
 f. [¹pa.ma] (esp. de árbol)

La articulación de una consonante oclusiva involucra tres momentos o fases: (i) oclusión o cerrazón, (ii) explosión y (iii) tiempo de inicio de la voz (al cual se le conoce en el campo de la fonética acústica como VOT, debido a su sigla en inglés Voice Onset Time). Ya que en shipibo, las oclusivas son sordas, las tres fases involucradas en su articulación no están acompañadas por vibraciones de las cuerdas vocales. En la fase de oclusión, los articuladores entran en contacto en algún punto dentro de la cavidad oral e impiden la salida del aire por la boca. En los espectrogramas, la fase de oclusión de las oclusivas sordas se caracteriza por una carencia de energía acústica, lo cual visualmente se representa como un espacio blanco. Esta característica espectrográfica se puede observar en la Figura 3-2, la cual muestra la palabra shipiba [pi.¹ka.ti.ka.na.i] ‘comieron’ (realizada como [pi.¹ka.ti.ka.nej]).¹ La ausencia de la barra de voz indica la naturaleza sorda de las

¹ Para más información sobre la diptongación final que se observa en la realización de la palabra [pi.¹ka.ti.ka.nej], véase el capítulo 2, sección 2.6.

oclusivas del shipibo. Si las oclusivas del shipibo fuesen sonoras, mostrarían una barra de voz, es decir, aparecería energía acústica usualmente bastante debajo de los 500 Hz. Las oclusivas sordas [p, t, k] aparecen transcritas en la parte inferior del gráfico y encerradas en un círculo.

Figura 3-2: Consonantes oclusivas de la palabra [pi.'ka.ti.ka.na.i] ‘comieron’



3.2.1 Mediciones de duración de las oclusivas

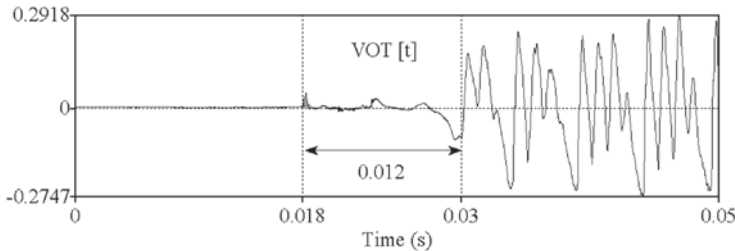
La duración de la fase de cerrazón de las consonantes oclusivas varía de acuerdo a si es velar o no. La oclusiva velar, [k], tiende a tener una fase de cerrazón aproximadamente 45 ms más corta que la fase de cerrazón que se observa en la oclusiva bilabial y en la coronal, [p] y [t]. La diferencia de duración en la fase de cerrazón entre las oclusivas velares y no-velares es consistente en todos los hablantes. Los resultados de la Tabla 3-1 ofrece la duración promedio en segundos de las fases de cerrazón de las oclusivas del shipibo. Estos resultados se basan en las mediciones de más de 3000 consonantes oclusivas. Todas ocurrieron en contextos intervocálicos y siempre como arranques de la segunda sílaba de la palabra. La razón por la cual solo se consideró oclusivas intervocálicas dentro de palabras al hacer las mediciones de las fases de cerrazón es que es extremadamente difícil determinar dónde empieza esta fase en una oclusiva que pertenece a la sílaba inicial de una palabra. La dificultad surge inclusive cuando la palabra fue elicitada dentro de una frase, como es el caso en este estudio, ya que los hablantes pueden insertar pausas, que aunque perceptualmente parecen pequeñas, tienen un efecto en las mediciones.

Tabla 3-1: Promedios de la duración (en segundos) de las fases de cerrazón de [p, t, k]

Oclusiva	Duración de la cerrazón
[p]	0,126 ^(0,021)
[t]	0,128 ^(0,021)
[k]	0,084 ^(0,020)

Luego de la fase de cerrazón sigue una explosión que indica el momento en que los articuladores se separan y dejan fluir el aire abruptamente a través de la boca (el análisis acústico de la explosión en las consonantes oclusivas se presentará en breve). Después de la explosión, las cuerdas vocales no empiezan a vibrar para la vocal que sigue hasta después que ha pasado un intervalo de tiempo. Este periodo de tiempo que pasa entre la explosión y el comienzo de las vibraciones de las cuerdas vocales se conoce como VOT (*Voice Onset Time*, el tiempo de inicio de la voz, Docherty 1992, Ladefoged 2003). El oscilograma de la Figura 3-3 presenta un ejemplo de VOT del shipibo. Este oscilograma pertenece a la consonante oclusiva [t] pronunciada dentro de la palabra [ʃa.ta.b^βu] (esp. de planta - plural). Se puede observar claramente la explosión al comienzo del VOT y el comienzo de las vibraciones de las cuerdas vocales al final del VOT. En este caso en particular, el VOT dura 12 ms.

Figura 3-3: VOT de la consonante oclusiva [t] en la palabra [ʃa.ta.b^βu]



Los promedios de la Tabla 3-2 corresponden a la duración del VOT de las oclusivas [p], [t] y [k]. Una vez más, como en el caso de las fases de cerrazón, encontramos una clara diferencia entre la oclusiva velar y las oclusivas no-velares. El VOT de la oclusiva velar [k], el cual dura alrededor de 35 ms, es aproximadamente tres veces más largo que el VOT de las oclusivas [p] y [t]. El VOT de estas últimas consonantes es virtualmente el mismo (aproximadamente 10 ms). Los resultados que se muestran en la Tabla 3-2 se basan en el mismo conjunto

de datos que se utilizaron para hacer las mediciones de la fase de cerrazón (véase la Tabla 3-1).

Tabla 3-2: Promedios de la duración (en segundos) del VOT en oclusiva

Oclusivas	Duración del VOT
[p]	0,010 ^(0,004)
[t]	0,011 ^(0,004)
[k]	0,034 ^(0,009)

Ahora que sabemos qué tan largas son las fases de cerrazón y de VOT en las consonantes oclusivas del shipibo, podemos averiguar cuál es la duración total de estos segmentos. Los promedios se ofrecen en la Tabla 3-3. Están basados en los mismos datos que las dos tablas anteriores. En general, la oclusiva velar tiende a ser más corta que [p] y [t], aunque recuérdese que su VOT es mucho más largo que el de [p] y [t].

Tabla 3-3: Promedio de la duración (en segundos) de las consonantes oclusivas

Oclusivas	Duración Total
[p]	0,136 ^(0,019)
[t]	0,139 ^(0,020)
[k]	0,118 ^(0,018)

Los oscilogramas y espectrogramas que se presentan de la Figura 3-4 a la Figura 3-6 nos permiten comparar las diferentes duraciones de las fases de cerrazón y de VOT de las oclusivas sordas del shipibo. Las tres figuras muestran la misma cantidad de tiempo: 180 ms. El inicio de la fase de oclusión en los tres casos empieza a los 20 ms en la escala de tiempo de todos los gráficos. Las oclusivas ocurren en el mismo contexto intervocálico. He incluido los espectrogramas para que el lector pueda identificar con mayor rapidez dónde se produce la explosión.

Es importante resaltar que las mediciones de duración de las oclusivas que se han presentado (particularmente aquellas de la duración de las fases de oclusión) fueron tomadas de palabras pronunciadas a una velocidad normal de habla y dentro de una frase. Bajo condiciones diferentes, estas mediciones podrían cambiar pero lo que se mantiene constante son las diferencias que se han encontrado entre oclusivas velares y no-velares.

Figura 3-4: Cerrazón y VOT de [p] en [sa.¹pa.ti] ‘aplanarse’

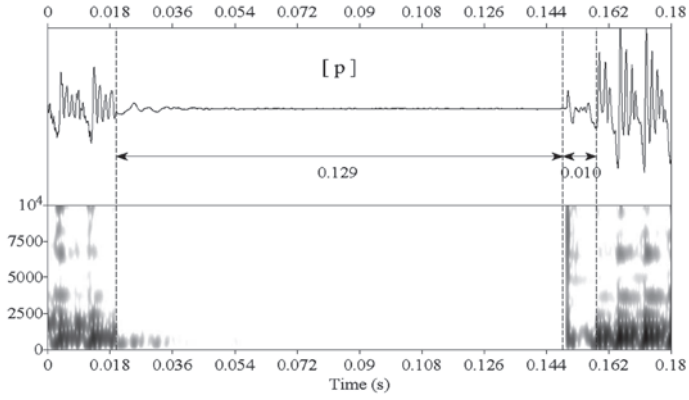


Figura 3-5: Cerrazón y VOT de [t] en [ma.¹ta.ti] ‘quitarse (algo) de la cabeza’

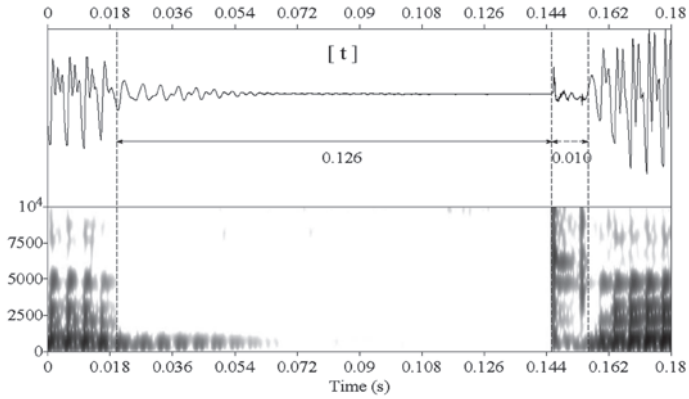
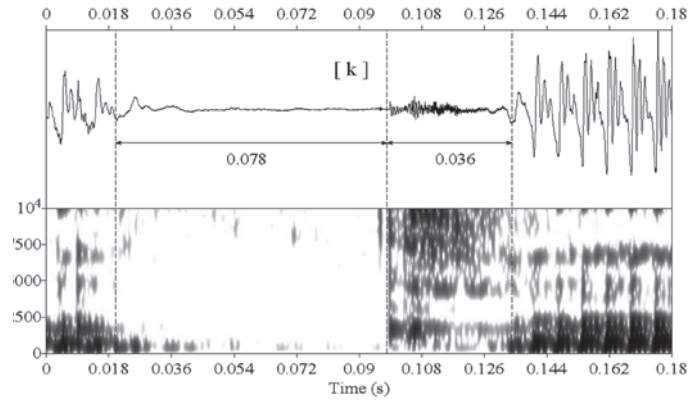


Figura 3-6: Cerrazón y VOT de [k] en [ja.¹ka.ti] ‘sentarse’



Los promedios de la Tabla 3-4 comparan la duración de las oclusivas (fase de cerrazón, VOT y duración total) en dos contextos diferentes: cuando las palabras que contienen las oclusivas ocurren dentro de frases y cuando se les pronuncia aisladamente. Los resultados que se han presentado hasta ahora en este capítulo son cuando las palabras que contienen las oclusivas ocurren dentro de frases.

Los resultados de cuando las palabras se pronuncian aisladamente vienen de un conjunto diferente de datos tomados de dos hablantes de shipibo durante la etapa del proyecto piloto para este estudio. En ambos casos, las oclusivas siempre ocurren en contextos intervocálicos. Como se observa, la duración del VOT permanece virtualmente sin cambios en ambas situaciones. Es la fase de cerrazón la que sufre ajustes. Aisladamente, la duración de la fase de cerrazón es más larga para los tres tipos de consonantes oclusivas. A pesar de estos ajustes, subsisten las diferencias entre oclusivas velares y no-velares. Las oclusivas velares mantienen un VOT más largo y una fase de oclusión más corta que las oclusivas [p] y [t].

**Tabla 3-4: Duración de las oclusivas (en segundos)
en dos contextos diferentes**

Palabras dentro de una Frase				Palabras en Aislamiento			
	Cerrazón	VOT	Total		Cerrazón	VOT	Total
[p]	0,126	0,010	0,136	[p]	0,160	0,007	0,167
[t]	0,128	0,011	0,139	[t]	0,161	0,009	0,170
[k]	0,084	0,034	0,118	[k]	0,131	0,027	0,158

3.2.2 Análisis espectral de la fase de explosión de las consonantes oclusivas

Ahora retornemos a la fase de explosión de las consonantes oclusivas. Esta fase contiene información acústica sobre el lugar dónde se articulan las oclusivas. Durante la explosión, la presión del aire aumenta detrás de la oclusión. Cuando los articuladores que detienen la salida del aire se separan, el aire escapa por la boca abruptamente. Esto crea un sonido transitorio, una explosión, en la cual la energía acústica se distribuye diferente a lo largo del rango de frecuencias dependiendo del lugar donde se produzca la liberación del aire. De este modo, los diferentes patrones de distribución de la energía acústica se correlacionan con el punto donde ocurrió la oclusión de la consonante.

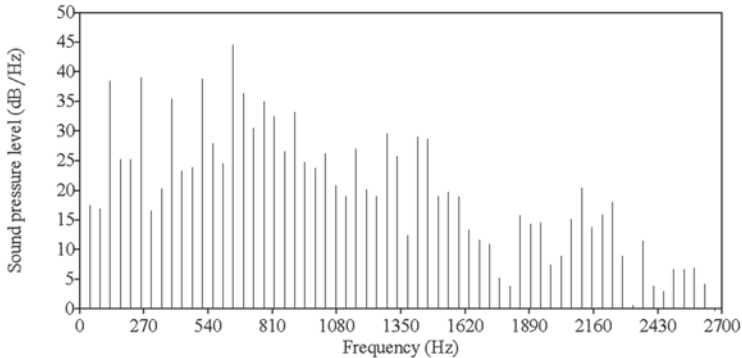
Recurriremos a análisis espectrales con el fin de estudiar cómo se distribuye la energía acústica a través de diferentes frecuencias durante la fase de explosión

de las consonantes oclusivas. Aunque los espectrogramas nos pueden brindar una idea intuitiva de esta distribución, su mayor utilidad, en realidad, yace en la representación detallada de información de las dimensiones de frecuencia (eje vertical) y tiempo (eje horizontal). En el caso de espectrogramas a blanco y negro, la información sobre la distribución de la energía acústica solo está vagamente codificada en los diferentes matices de los colores negro, gris y blanco. Contra más energía acústica se concentre en un rango de frecuencias, el espectrograma aparece mucho más oscuro e inversamente, contra menos energía haya en ciertas frecuencias, el espectrograma aparece mucho más claro.

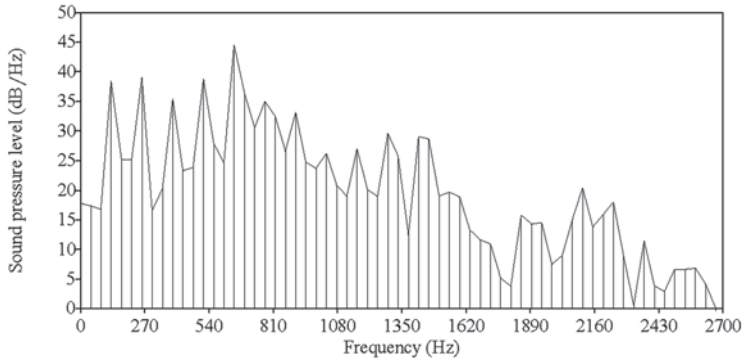
Sin embargo, usando espectros se puede obtener una imagen mucho más exacta de la distribución de la energía acústica a lo largo de diferentes frecuencias en un momento dado. Un espectro codifica la energía acústica en el eje vertical y las frecuencias en el eje horizontal dejando de lado la dimensión temporal (la cual esta vez solo va estar vagamente codificada en la porción del sonido que se usó para obtener el espectro).

La Figura 3-7 muestra el espectro de una porción de 10 ms tomados del medio de la vocal acentuada de la palabra [sa.'pa.ti] ‘aplanarse’. En este espectro, vemos los armónicos (las frecuencias componentes del sonido) de la vocal [a]. Estos aparecen como líneas verticales con diferentes alturas que indican su amplitud. El espectro que se muestra en la Figura 3-7 recibe el nombre de *Espectro de Líneas* (Hewlett y Beck 2006).

Figura 3-7: Espectro de líneas del medio de la vocal acentuada [a] en [sa.'pa.ti]



La estructura de los armónicos también se puede representar a través de un espectro conocido como *espectro FFT* (Fast Fourier Transform), en el cual cada polo del espectro de líneas se une a través de pequeñas líneas (véase la Figura 3-8). Esto hace que sea un poco más fácil reconocer los armónicos más prominentes.

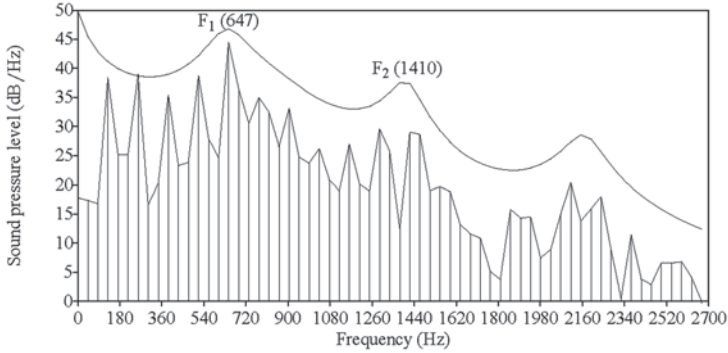
Figura 3-8: Espectro FFT del medio de la vocal acentuada [a] en [sa.¹pa.ti]

Aunque los espectros FFT ofrecen una imagen muy detallada de la relación entre la energía acústica y las frecuencias, los gráficos son difíciles de leer lo cual, a su vez, dificulta el poder determinar dónde exactamente están los picos más altos de energía (es decir, los formantes en el caso de sonidos periódicos como las vocales). Esta dificultad se puede superar usando un espectro conocido como *espectro LPC (Linear Predictive Coding)*. Este tipo de espectro muestra los picos más altos de energía a través de una línea alisada que corre sobre la dimensión de las frecuencias (Harrington y Cassidy 1999; Johnson 2003; Ladefoged 2003).

En la Figura 3-9, se ilustra un espectro LPC que aparece sobre el espectro FFT. El espectro LPC nos permite ver, por ejemplo, dónde se encuentran los formantes de la vocal [a]. De acuerdo al espectro LPC de la Figura 3-9, el formante F1 ocurre a 647 Hz y el F2, a 1410 Hz. El espectro LPC también muestra la presencia de un tercer formante (F3) que se ha ignorado ya que con la información de F1 y F2 es suficiente para poder describir correctamente las vocales del shipibo.² Si damos volvemos a revisar los valores de los formantes para la vocal [a] (véase la sección 2.1), veremos que son bastante cercanos a los valores que muestra la Figura 3-9. Además, los valores de F1 y F2 de este caso particular también indican correctamente que la vocal fue pronunciada por un hablante hombre.

² Para la mayoría de las vocales de las lenguas del mundo, F1 y F2 son suficientes para poder caracterizar su altura y grado de anterioridad/posterioridad en el espacio vocálico. Sin embargo, se sabe que la vocal [ə] del inglés, la cual resulta de la fusión de la vocal schwa [ə] con la aproximante retrofleja [ɻ], además necesita del formante F3 (Olive, J., A. Greenwood, *et al.* (1993)).

Figura 3-9: Espectro LPC sobre un FFT del medio de la vocal acentuada [a] en [sa.¹pa.ti]

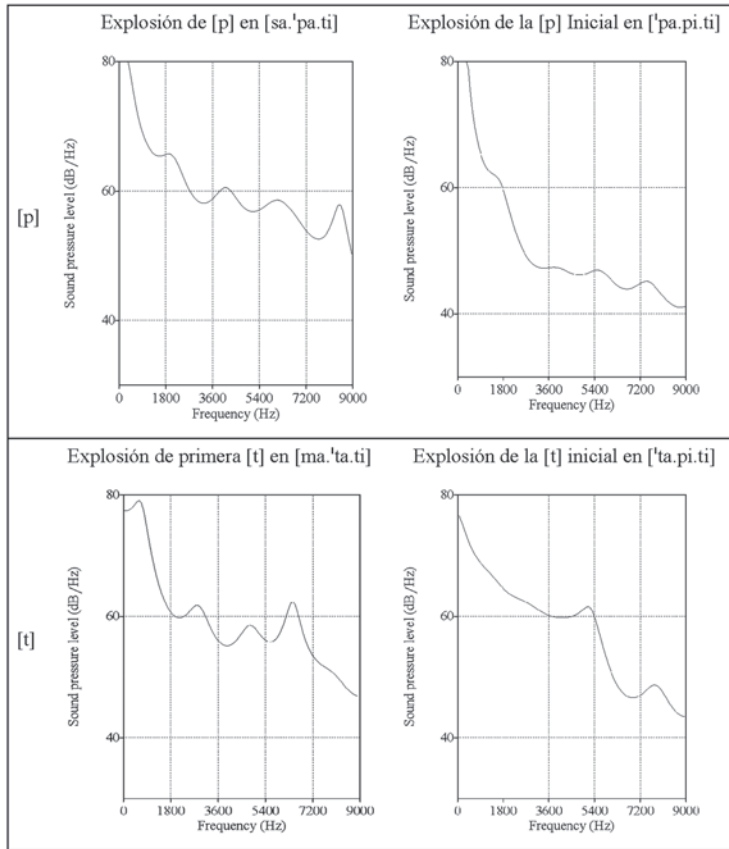


Como en el caso de las vocales, se puede usar un espectro LPC para analizar la fase de explosión de las consonantes oclusivas. Sin embargo, dado que estas fases no son sonidos periódicos, no encontraremos formantes en la distribución de la energía acústica. Esto se debe a que la fase de explosión de las oclusivas no tiene una estructura armónica ya que los armónicos solo ocurren a intervalos regulares, como se da en los sonidos vocálicos.

Cuando miramos a un espectro LPC de la fase de explosión de una oclusiva, estamos interesados en saber cómo la energía acústica se distribuye en diferentes frecuencias. El ruido (sonido aperiódico) que crea la fase de explosión es filtrado por la forma que toma la cavidad oral para articular una consonante oclusiva. Conforme esta forma varía de acuerdo a dónde se encuentra la oclusión de la consonante, la energía acústica se distribuye diferente a través del espectro de frecuencias.

La Figura 3-10 presenta espectros LPC de las fases de explosión de las consonantes [p] y [t]. En el caso de [p], los espectros vienen de la segunda sílaba de la palabra [sa.¹pa.ti] ‘aplanarse’, y de la sílaba inicial de la palabra [¹pa.pi.ti] ‘cargar en la espalda’. En el caso de [t], los espectros vienen de la segunda sílaba de [ma.¹ta.ti] ‘quitarse (algo) de la cabeza’ y la sílaba inicial de [¹ta.pi.ti] ‘escalera’. Aunque, en general, los espectros de [p] y [t] del shipibo vienen en una gran variedad de formas, se les puede caracterizar por tener una caída de la amplitud conforme las frecuencias suben. Sin embargo, una característica espectral que los distingue es la presencia de un pico notable en [t] en algún lugar entre 4500 Hz y 9000 Hz, el cual usualmente es seguido por una caída rápida de la amplitud.

Figura 3-10: Espectros LPC de las fases de explosión de [p] y [t]



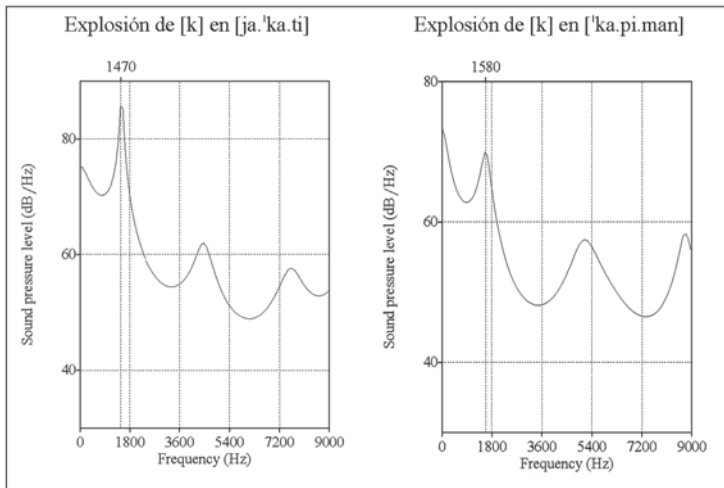
Las características espectrales de [t] en shipibo contrastan con aquellas usualmente reportadas en lenguas como el inglés. En particular, en inglés, el espectro de la fase de explosión de [t] se caracteriza por una subida de la amplitud conforme las frecuencias suben (Stevens and Blumstein 1978, 1981; Hewlett and Beck 2006). Esta diferencia se debe al hecho que en shipibo, para la mayoría de los hablantes, la consonante oclusiva coronal [t] es dental mientras, en inglés, es alveolar. Se han encontrado características espectrales similares a las del shipibo en oclusivas dentales en lenguas hindúes como Malayam (Lahiri, Gewirth *et al.* 1984) y Marathi (Veena Karjigi and Samudravijaya 2007).

La Figura 3-11 muestra dos espectros LPC de la fase de explosión de la consonante [k]. Los espectros corresponden a la oclusiva de la segunda sílaba de la palabra [ja.'ka.ti] 'sentarse' y de la sílaba inicial de ['ka.pi.man] (esp. de maleza). A diferencia de la variabilidad que se encuentra en las formas espectrales

de [p] y [t], el espectro de [k] es bastante regular. Siempre muestra dos picos prominentes. A veces se observa un tercer pico que ocurre más allá de 9000 Hz. El primer pico siempre se correlaciona con el grado de posterioridad de la siguiente vocal. En el ejemplo de la izquierda que se ofrece en la Figura 3-11, el primer pico de la fase de explosión de [k] se localiza a 1470 Hz. Este valor es bastante cercano al F2 de la siguiente vocal [a], 1410 Hz (hombres), la cual se examino en la Figura 3-9.

En el segundo ejemplo, el de la derecha, vemos otra [a] que sigue a la consonante [k]. Esta vez el primer pico en el espectro se encuentra a 1580 Hz. La frecuencia es mucha más alta que en el ejemplo anterior. Esto se debe a que la palabra fue pronunciada por una mujer (como se señaló en la sección 2.1, el promedio de F2 para la vocal [a] en hablantes mujeres es 1587 Hz). El segundo pico de la fase de oclusión de [k] tiene mucha menos amplitud y usualmente se localiza aproximadamente entre 4500 Hz y 6000 Hz.

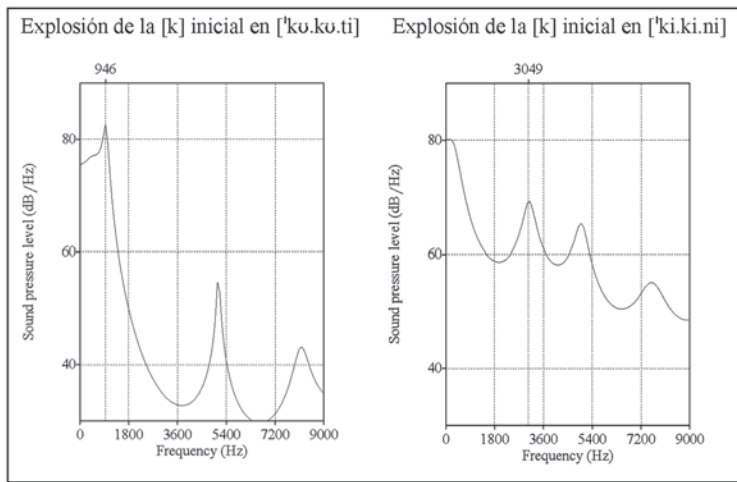
Figura 3-11: Espectros LPC de la fase de explosión de [k]



Mientras que en los espectros en la Figura 3-11, la consonante [k] aparecía seguida de la vocal central [a], para completar, la Figura 3-12 ilustra casos de espectros de la fase de explosión de [k] cuando es seguida por la vocal posterior [u] (espectro de la izquierda) y por la vocal frontal [i] (espectro de la derecha). El espectro de la explosión de [k] seguida por [u] corresponde a la oclusiva inicial de la palabra ['kũ.kũ.ti] 'comer fruta' y cuando es seguida por [i], corresponde a la oclusiva inicial de ['ki.ki.ni] 'verdaderamente, en realidad'.

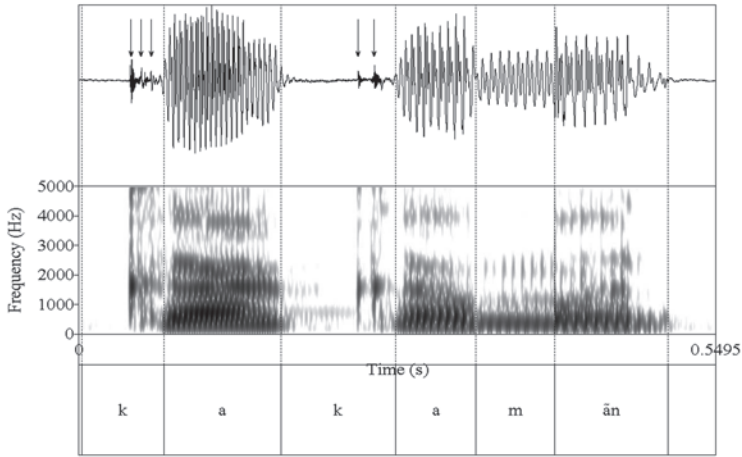
Cuando la vocal [u] sigue a [k], el primer pico del espectro se localiza a 946 Hz (el promedio de F2 para [u] es 964 Hz en mujeres). En el caso de [k] seguida por [i], la localización del primer pico de su fase de explosión siempre se encuentra mucho más alta que el F2 de la vocal [i]. En la Figura 3-12 (espectro de la derecha), vemos que el primer pico está a 3049 Hz. Recuérdese que el promedio de F2 para la vocal [i] es 2250 Hz en hombres y 2652 Hz en mujeres. En el ejemplo que se muestra, la palabra [ˈki.ki.ni] fue pronunciada por una hablante mujer. La vocal [i] que sigue a la consonante [k] tiene un F2 de 2567 Hz.

Figura 3-12: Espectro LPC de la fase de explosión de [k]



La oclusiva velar, además de tener un espectro de su fase de explosión bastante distintivo con respecto a [p] y [t], usualmente también manifiesta más de una explosión. Mientras esto es posible de encontrar en algunos casos en [p] y [t], su ocurrencia es bastante rara comparada con la frecuencia con que la consonante [k] muestra múltiples explosiones. La Figura 3-13 ilustra dos oclusivas velares con múltiples explosiones. El espectrograma y oscilograma de la ilustración pertenecen a la palabra [ˈka.ka.man] (la forma larga de [ka.¹kan] ‘joroba de sachavaca’).

Figura 3-13: Múltiples explosiones de [k]



3.2.3 Las oclusivas y las transiciones de los formantes de las vocales adyacentes

Además de las características espectrales de las fases de explosión de las oclusivas, el punto de articulación de las oclusivas puede estudiarse observando las *Transiciones de los Formantes* en las vocales adyacentes a estas consonantes. En particular, en la vecindad de otros segmentos, el F2 de las vocales se aleja máximamente del valor típico que idealmente obtiene al aproximarse al medio de la vocal. En la sección 2.1, se obtuvieron los valores promedio de F2 correspondientes al punto medio de cada vocal del shipibo. Usualmente se asume que el centro de una vocal es la zona donde los formantes llegan a estar lo más cercanamente posible a sus valores ideales. Entonces, volcaremos nuestra atención a los valores de F2 en los bordes de las vocales para poder identificar más claves fonéticas que se correlacionen con los puntos de articulación de las consonantes oclusivas.

En los espectrogramas de la Figura 3-14 a la Figura 3-16 tenemos a las tres oclusivas del shipibo ocurriendo en posiciones intervocálicas, rodeadas de dos vocales [i]. En la Figura 3-14, podemos ver el efecto que tiene [t] en los formantes de las vocales adyacentes. El F2 de las vocales [i] muestra una ligera caída al momento de la transición con [t]. Compárese esto con el espectrograma de la Figura 3-15. El F2 también baja en la oclusiva bilabial pero la caída es mucho más notable. Las consonantes labiales en general bajan el F2 en las transiciones con vocales de manera mucho más dramática que las consonantes coroneales. El espectrograma de la Figura 3-16 ejemplifica a la consonante velar en el mismo contexto intervocálico. A diferencia de [t] y [p], la consonante velar sube el F2

de las vocales en vez de bajarlo. Además de subir el F2, la consonante [k] tiende a bajar el tercer formante (F3). Esto causa muchas veces la impresión en el espectrograma que el F2 y el F3 se tocan (un *pellizco velar*).

Figura 3-14: El efecto de [t] en los formantes de las vocales en [ˈni.ti] ‘pararse’

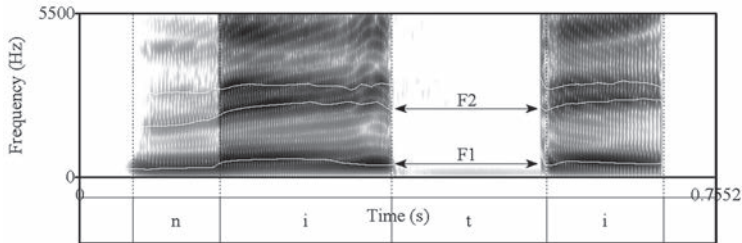


Figura 3-15: El efecto de [p] en los formantes de las vocales en [ˈʃi.pi] ‘mono’

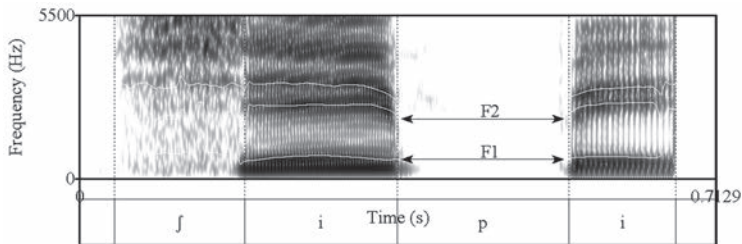
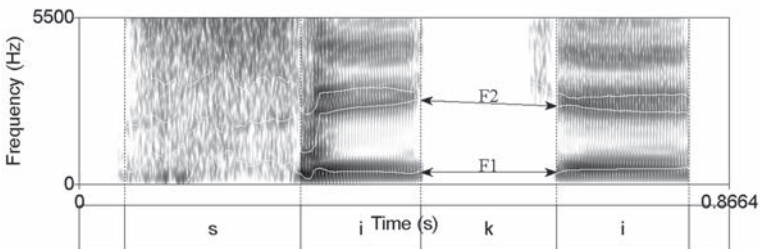


Figura 3-16: El efecto de [k] en los formantes de las vocales en [ˈsi.ki] (esp. de árbol)



Las Tabla 3-5 y Tabla 3-6 presentan los promedios de F2 tomados de las transiciones vocálicas con las consonantes [p], [t] y [k]. Los puntos de transición medidos comprenden tanto las transiciones de una oclusiva hacia una vocal (transiciones CV) como también las transiciones de una vocal hacia una oclusiva (transiciones VC). Los resultados se presentan separadamente para hombres y

mujeres y corresponden al mismo conjunto de datos utilizados para medir el F1 y F2 del medio de las vocales en el capítulo 2. La primera columna indica la vocal con la cual es adyacente la consonante oclusiva. Para facilitar la comparación de las transiciones de los formantes, he incluido en la primera columna el promedio de F2 que corresponde al centro de cada vocal (véase la sección 2.1).

Tabla 3-5: Promedios de F2 en las transiciones de formantes (mujeres)

	[p]	[t]	[k]
[i] - 2676	2245 Hz ⁽²²⁸⁾	2430 Hz ⁽¹⁴¹⁾	2678 Hz ⁽¹⁴⁹⁾
[a] - 1597	1372 Hz ⁽¹⁴⁴⁾	1748 Hz ⁽¹²⁹⁾	1630 Hz ⁽¹⁴¹⁾
[i̠] - 1704	1194 Hz ⁽¹⁵⁷⁾	1729 Hz ⁽¹⁶⁷⁾	1667 Hz ⁽¹⁴²⁾
[u] - 935	866 Hz ⁽¹⁰⁶⁾	1198 Hz ⁽¹⁷⁴⁾	910 Hz ⁽¹¹²⁾

Tabla 3-6: Promedios de F2 en las transiciones de formantes (hombres)

	[p]	[t]	[k]
[i] - 2244	1976 Hz ⁽¹⁷²⁾	2136 Hz ⁽¹⁴⁵⁾	2264 Hz ⁽¹⁴³⁾
[a] - 1418	1103 Hz ⁽⁹⁴⁾	1445 Hz ⁽¹²⁶⁾	1472 Hz ⁽¹³²⁾
[i̠] - 1539	1073 Hz ⁽¹¹²⁾	1559 Hz ⁽¹⁵⁰⁾	1517 Hz ⁽¹⁵²⁾
[u] - 896	817 Hz ⁽⁸⁸⁾	1082 Hz ⁽¹¹⁰⁾	879 Hz ⁽¹⁰⁰⁾

Las Figuras 3-17 y 3-18 representan gráficamente la información contenida en la Tabla 3-5 y en la Tabla 3-6, respectivamente. Ilustran los cambios que sufren los valores de F2 en la transición con consonantes oclusivas. Es decir, reportan los descensos, ascensos o falta de cambios que se dan en los valores de F2 en la transición con la oclusiva y en el medio de la vocal. Los diagramas muestran los valores de F2 en Hz. En este vertical, se puede apreciar la división tradicional de las vocales en frontales, centrales y posteriores. Como es de esperarse, los valores más altos de F2 corresponden a las vocales frontales y los valores más bajos a las vocales posteriores. Los intermedios van para las vocales centrales.

En el eje horizontal, se muestran las vocales divididas en dos grupos: las vocales altas [i, ɪ, u], y la vocal baja [a]. Dentro de cada grupo, la primera columna representa las vocales en contacto con [p], luego aquellas en contacto con [t], y finalmente, en la tercera columna, aquellas en contacto con [k].

Los círculos negros indican los valores de F2 en el medio de las vocales, los blancos, los valores de F2 en la transición con la oclusiva. Las flechas señalan la dirección del cambio desde el medio de la vocal. De este modo, una flecha hacia arriba muestra un incremento en el valor de F2 conforme se aproxima a la consonante oclusiva; una flecha hacia abajo indica una caída en el valor de F2. Cuando están muy cercanos o son iguales los valores de F2 tanto en el medio de la vocal como en la transición con una oclusiva, el diagrama utiliza un círculo que encierra ambos valores. He considerado que ambos valores de F2 se sobreponen si hay entre ellos una diferencia de menos de 100 Hz.

Figura 3-17: Transiciones de F2 para las consonantes oclusivas (hombres)

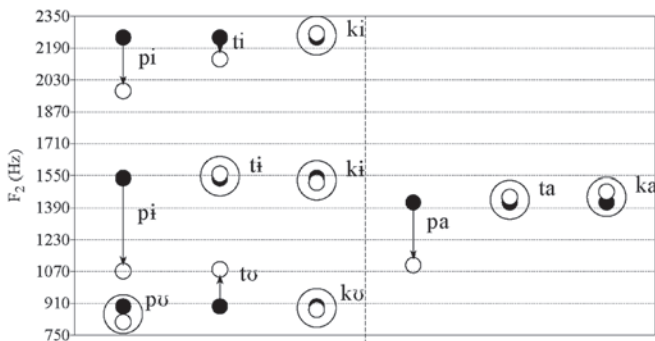
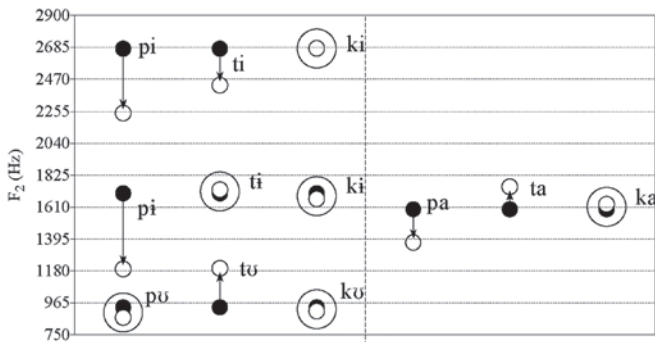


Figura 3-18: Transiciones de F2 para las consonantes oclusivas (mujeres)



De estos diagramas podemos obtener tres observaciones importantes. La primera es que la transición de F2 es una clave fonética particularmente fuerte para la oclusiva bilabial [p]. Tanto en hombres como en mujeres, F2 baja consistentemente conforme la vocal se aproxima a la consonante [p]. Esta es una tendencia general que también se ha observado en otras lenguas (Delattre, Liberman *et al.* 1955; Johnson 2003).

La segunda conclusión es que los valores promedio de F2 en la transición con la oclusiva velar [k] no ofrecen una clave fonética robusta para determinar de ellos el punto de articulación de la consonante. El problema no es que los valores de F2 no cambien en la transición con [k] (compárese los promedios de F2 en la columna en gris de la Tabla 3-5 y la Tabla 3-6 con aquellos de la primera columna). La falta de un incremento o caída del F2 podría ser la clave en sí misma. El problema es que este comportamiento coincide con el comportamiento de F2 en la transición con otras oclusivas (véase el comportamiento de F2 en las secuencias [kɪ] y [tɪ]; [ka] y [ta]; [kʊ] y [pʊ]).

La tercera conclusión es que para la oclusiva coronal [t], la transición de F2 sí brinda claves fonéticas más fuertes que para [k] pero no son tan robustas como las de [p]. En general, el F2 de una vocal frontal desciende en la transición con una oclusiva coronal pero sube si la vocal es posterior. Contra más frontal la vocal, mayor la caída de su F2 en la transición con [t]; y contra más posterior la vocal, mayor la subida de su F2 en dicha transición.

Examinemos un poco más el comportamiento de F2 a las transiciones con oclusivas pero esta vez desde el punto de vista de las vocales. El comportamiento del F2 de la vocal [i] indica el punto de articulación de la oclusiva que aparece adyacente. El F2 de [i] sufre un gran descenso para [p], uno pequeño para [t] y casi no sufre ningún cambio para [k]. En el caso de la vocal [ʊ], su F2 solo muestra un patrón diferente cuando está adyacente a la oclusiva [t]: el F2 sube conforme la vocal se acerca a la consonante oclusiva. Sin embargo, no se observa ningún cambio en el F2 de [ʊ] cuando está en contacto con [p] o [k]. El caso de las vocales centrales [ɪ] y [a] es un poco más complejo. En general, no se puede distinguir [t] y [k] desde el punto de vista de la transición del F2 de las vocales centrales. Lo cual quiere decir que cuando están adyacentes a vocales centrales, el contraste entre estas dos oclusivas debe depender de otras claves fonéticas. Lo mismo es cierto para los casos de [p] y [k] en contacto con [ʊ].

En las secciones previas, hemos visto que la mayoría de las claves fonéticas favorecen a [k]. Por ejemplo, tanto la duración de su fase de oclusión como la de su VOT la distinguen de las oclusivas no-velares. En el caso del VOT, es tres veces más largo en [k] que en [p] y [t]. El espectro de la fase de explosión de las oclusivas es otra clave importante. El espectro de [k] es absolutamente diferente

al de los espectros correspondientes a [p] y [t]. Son estas claves las que ayudan a mantener las distinciones entre [t] y [k] cuando el comportamiento del F2 no puede, es decir, cuando [t] y [k] están en contacto con [i] o [a], y entre [p] y [k], cuando aparecen adyacentes a [u].

3.2.4 Sobre la fonética de la labialización causada por [i]

Los resultados de la Tabla 3-5 y la Tabla 3-6 mostraron que el F2 de las vocales desciende en la transición con la oclusiva bilabial [p] (Delattre, Liberman *et al.* 1955; Johnson 2003). En particular, la caída del F2 es mucho más notable cuando [p] aparece adyacente a la vocal [i]. La caída es mucho menos dramática para la vocal central [a] y muy pequeña para la vocal posterior [u]. En este sentido, podría parecer sorprendente que la vocal central, [i], exhiba un mayor grado de descenso del F2 que la otra vocal central, [a], cuando ocurren adyacentes a [p]. Sin embargo, la diferencia en su comportamiento, se debe a un fenómeno tradicionalmente descrito como *Labialización* (Shell 1985; García-Rivera 1994). La generalización es que las consonantes bilabiales [p], [b^β] y [m] se labializan cuando aparecen seguidas de la vocal central *no-redondeada* [i]: [p^wi], [b^{βw}i] y [m^wi]. Esta sección ofrece una descripción acústica de este fenómeno.

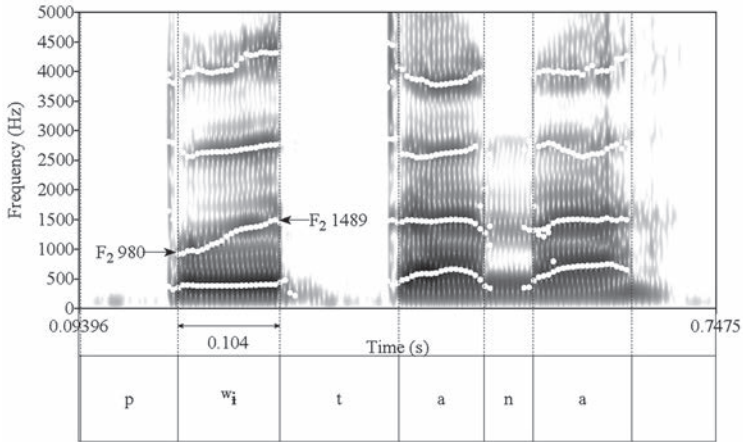
3.2.4.1 Una caracterización acústica de la labialización del shipibo

Los datos en (11) muestran que la labialización de las consonantes labiales causada por la vocal central [i] puede ocurrir tanto en sílabas iniciales como no-iniciales de una palabra así como también en sílabas acentuadas y no-acentuadas. Además, los datos muestran que las consonantes [t] y [k] no son afectadas por este fenómeno.

- (11) a. [p^wi.ta.na] ‘desviarse para evitar un obstáculo’ (PP2)
 b. [p^wi.'ka] ‘espalda’
 c. [b^{βw}i.ni] ‘macho, esposo’
 d. [b^{βw}i.p^wi.ki] ‘inclinarse la cabeza’ (CMPL)
 e. [ki.'m^wiʂ.ti] ‘arrancar los pelos del bigote o de la barba’ (INF)
 f. [m^wi.ti.ki] ‘muñeca de la mano’

El espectrograma de la Figura 3-19 presenta un ejemplo de este fenómeno. En este caso, la labialización ocurre en la sílaba inicial de la palabra [p^wi.ta.na] pronunciada por un hablante hombre. He indicado el valor de F2 al comienzo y al final de la vocal

Figura 3-19: Labialización - [p^wi.ta.na]
 ‘desviarse para evitar un obstáculo’ (PP2)



En la Figura 3-19, el F2 de [i] sube aproximadamente 500 Hz durante la duración de la vocal. La mayor parte de la subida ocurre durante la primera mitad de la vocal. Esto da la impresión auditiva de que fuese un diptongo [wi]. Sin embargo, existe evidencia de que lo que se observa no es ni un diptongo ni una secuencia de dos vocales hetero-silábicas. La primera pieza de evidencia proviene de la duración vocálica. La vocal de la secuencia [p^wi] es tan larga como el promedio de la vocal corta [i]. Por ejemplo, en el espectrograma de la Figura 3-19, [wⁱ] dura 104 ms (una duración bastante cercana a 98 ms, el promedio de una vocal [i] corta y acentuada en una sílaba inicial). En la Figura 3-20, la sílaba [p^wi] aparece acentuada y ocupa la segunda sílaba de la palabra (esta vez la palabra es pronunciada por una hablante mujer). En este caso, la secuencia [wⁱ] dura 119 ms (lo cual también es bastante cercano a 114 ms, el promedio de una vocal [i] corta y acentuada en la segunda sílaba de una palabra).

Los espectrogramas de la Figura 3-21 muestran casos de labialización de las otras dos consonantes bilabiales del shipibo: [b^{βw}i] y [m^wi]. Corresponden a las sílabas iniciales acentuadas de las palabras [b^{βw}i.pu.ti] ‘cubrirle los ojos a alguien’ y [m^wi.tu.ti] ‘dedo índice’.³

³ En la Figura 3-21, el F1 de la vocal [i] en [m^wi] muestra una frecuencia mucho más alta que el F1 de la vocal [i] en [b^{βw}i]. Esto se debe al efecto que tienen las consonantes nasales sobre las vocales adyacentes. Como se verá en el capítulo 5 (sección 5.2.2), la nasalización de vocales se correlaciona con un F1 más alto comparado a las vocales orales.

Figura 3-20: Labialización - [k^u.¹p^wɪ.kɪ] ‘perforar el labio inferior’ (CMPL)

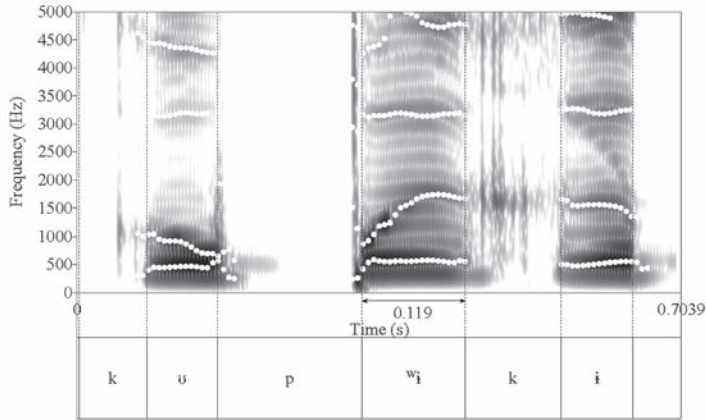
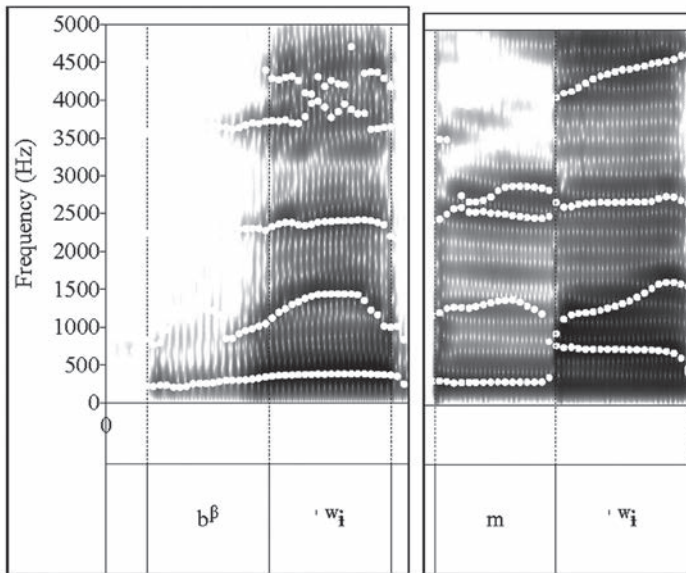


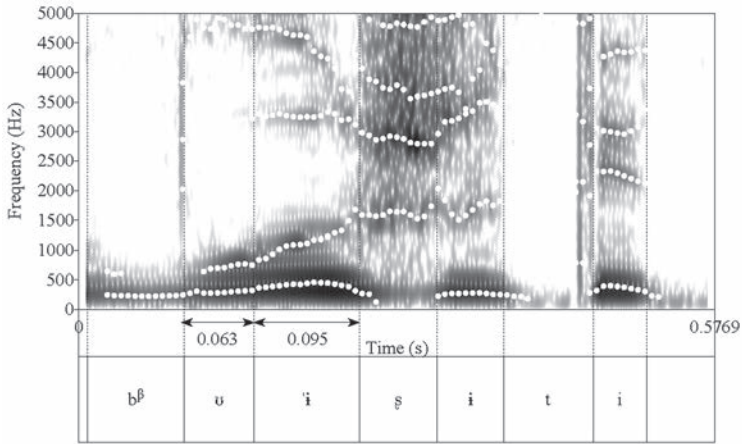
Figura 3-21: Labialización - [b^βwɪ] y [m^wɪ]



Ahora veamos por qué la labialización no es la secuencia de las vocales heterosilábicas [u.ɪ]. El espectrograma de la Figura 3-22 ejemplifica la palabra [b^βu.¹ɪ.ʂɪ.ti] ‘peinar’. Este ejemplo es importante porque nos permite contrastar las diferencias entre una secuencia de dos vocales heterosilábicas, la vocal redondeada [u] seguida de la vocal central [ɪ], y la secuencia formada por la labialización de la oclusiva bilabial seguida de la vocal [ɪ], por ejemplo, [wɪ] en la Figura 3-19. En ambos casos, la vocal central de [b^βu.¹ɪ] y de [p^wɪ], ocurren acentuadas. Si la vocal [u] en [b^βu.¹ɪ] fuese labialización, esperaríamos que su duración y contorno de F2 se

comporten igual que [w] en [p^wi]. Sin embargo, mientras en la labialización (véase Figura 3-19), F2 experimenta una subida continua e inmediatamente después de la consonante labial [p] hasta que alcanza una fase estable aproximadamente en el medio de la vocal [i], en la Figura 3-22, sabemos que no ocurre ninguna labialización justamente porque el F2 se muestra estable durante los primeros 63 ms, el tiempo que corresponde a la vocal inacentuada [u]. Después de la vocal [u], el F2 asciende para alcanzar el valor necesario para la vocal acentuada [i], la cual dura 95 ms.⁴ Las diferencias fonéticas observadas, tanto en términos de duración como del comportamiento de F2, inequívocamente apuntan a que la labialización en shipibo no equivale a la secuencia de dos vocales.

Figura 3-22: [b^βu. i.ʂi.ti] ‘peinar’



El espectrograma de la Figura 3-22 también nos da evidencia que la labialización causada por la vocal [i], no es una propiedad de las consonantes bilabiales del shipibo, es decir, las consonantes [p, b^β, m] no son subyacentemente labializadas. En el espectrograma, la consonante labial [b^β] aparece seguida de una vocal [u] pero no ocurre la labialización.

⁴ Obsérvese que el hablante que pronunció la palabra en la Figura 3-22 tiene una oclusiva bilabial sonora [b] en vez de la afrizada bilabial sonora [b^β] (el capítulo 4 discutirá esta variabilidad entre hablantes). Escogí presentar un ejemplar de este tipo de hablante en la Figura 3-22 porque es más fácil determinar dónde empieza la vocal [u], y de este modo, evitar las dificultades asociadas con la fase de fricción de la afrizada bilabial sonora. Transcribí [b^β] en lugar de [b] para mantener la consistencia entre transcripciones.

La Figura 3-23 representa la palabra [ˈpi.ti.ma] ‘cosa que no sirve para comer’ y la Figura 3-24, [ˈpa.pʷi.ta] ‘perforar la oreja’ (PP2). En las sílabas iniciales, la oclusiva bilabial [p] aparece seguida de una vocal alta frontal [i] y de una vocal baja central [a], respectivamente. En ambos casos, el F2 muestra un valor bajo en el punto de transición con la oclusiva pero luego asciende rápidamente hasta alcanzar el valor esperado para cada una de esas vocales. En ningún caso ocurre la labialización. Además, el espectrograma de la Figura 3-24 nos deja contrastar la secuencia [pa] de la sílaba inicial acentuada con la segunda sílaba inacentuada, [pʷi], la cual sí contiene las características típicas de la labialización: un incremento rápido y continuo del F2 hasta llegar al medio de la vocal.

Figura 3-23: [ˈpi.ti.ma] ‘cosa que no sirve para comer’

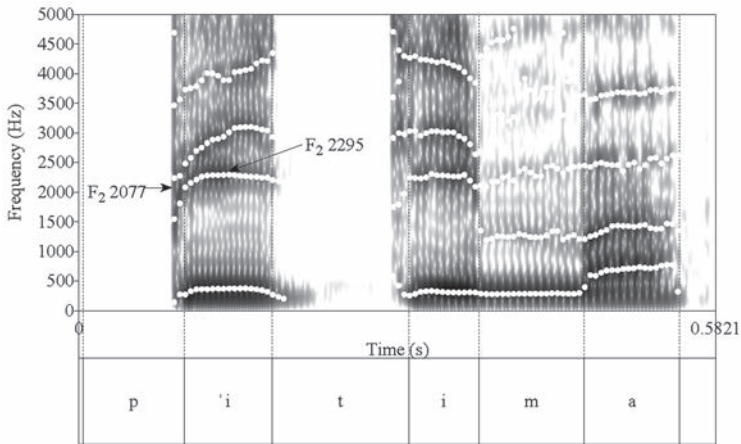
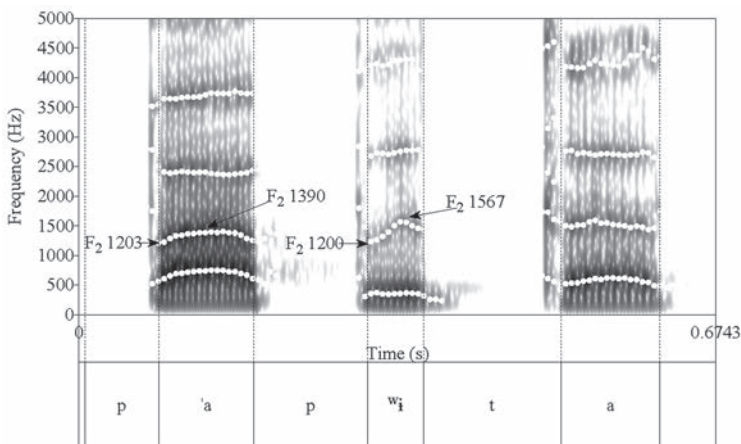
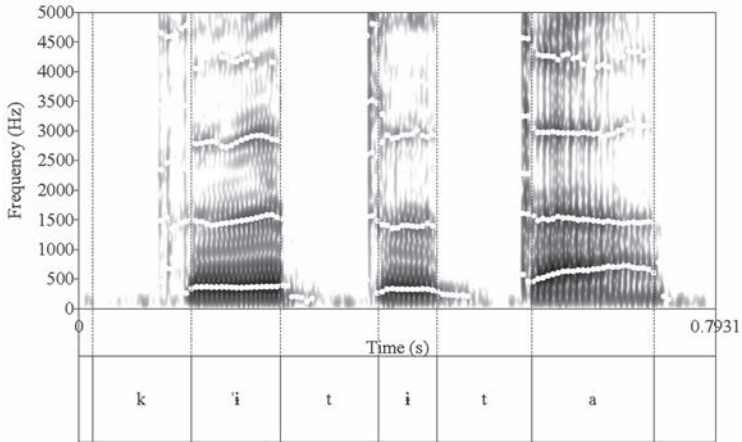


Figura 3-24: [ˈpa.pʷi.ta] ‘perforar la oreja’ (PP2)



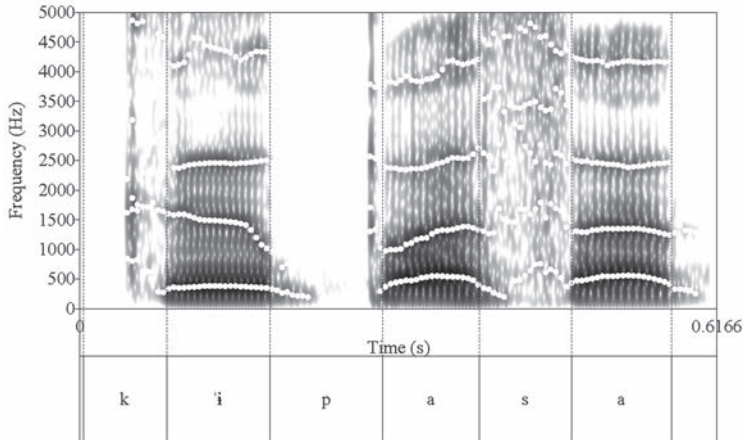
Hemos visto en los párrafos anteriores que la labialización no es una propiedad de las consonantes bilabiales del shipibo. Cuando éstas aparecen adyacentes a otras vocales, el fenómeno no ocurre. Una de las características más interesantes de este proceso es que la labialización tampoco es una propiedad que la vocal central [ɨ] provoque en general. Ésta vocal solo induce la labialización cuando sigue a consonantes bilabiales. Esto se puede observar en la Figura 3-25 donde las vocales [ɨ] siguen a [t] y [k], pero no ocurre labialización.

Figura 3-25: [ˈki.ti.ta] ‘incendiarse, arder’ (PP2)



Para ser exhaustivos, consideremos el espectrograma de la Figura 3-26. Éste muestra que para que la labialización se dé, la vocal central [ɨ] debe seguir a una consonante bilabial. La configuración inversa, es decir que una consonante bilabial siga a [ɨ], no causa labialización. El espectrograma corresponde a la palabra [ˈki.pa.sa] ‘dar una palmada en la boca de alguien’ (PP2) y contiene la secuencia hetero-silábica [ɨ.p]. El F2 de la vocal [ɨ] desciende al final de la primera sílaba conforme se aproxima a la consonante bilabial, pero obsérvese que a diferencia de los casos de labialización, el F2 se mantiene estable durante la mayor parte de la duración de la vocal.

Figura 3-26: ['ki.pa.sa] 'dar una palmada en la boca de alguien' (PP2)



3.3 RESUMEN

Este capítulo ha presentado una descripción acústica de las consonantes oclusivas del shipibo: [p, t, k]. Tomando en cuenta sus fases articulatorias, se ha caracterizado a las consonantes oclusivas en términos de la duración de sus fases de cerrazón y de su VOT. También se ha presentado un estudio detallado de las propiedades espectrales de la fase de explosión de estas consonantes. Con respecto a su duración, se encontró que las consonantes no-velares [p] y [t] tienden a ser más largas que [k]. Las consonantes [p] y [t] duran aproximadamente 138 ms mientras que [k] tiende a durar 118 ms. Esta asimetría también se manifiesta en la duración de sus fases de cerrazón, la cual dura en promedio alrededor de 127 ms para [p] y [t], pero 84 ms para [k]. En contraste, la duración del VOT de [k] (34 ms) tiende a ser tres veces más largo que el VOT de [p] y [t] (11 ms). Estos resultados indican que la duración juega un rol bastante importante como clave fonética para distinguir entre la oclusiva velar y las no-velares.

Las propiedades espectrales de la fase de explosión de las oclusivas son otro tipo de clave fonética que puede usarse para distinguir entre oclusivas velares y las no-velares. Los espectros de las explosiones de [k] son máximamente diferentes de los espectros para [p] y [t]. Los espectros de [k] se caracterizan por presentar dos o tres picos prominentes donde el primero se correlaciona con el grado de posterioridad de la siguiente vocal. En cambio, los espectros de las fases de explosión de [p] y [t] se definen en general por la caída rápida de la amplitud conforme las frecuencias suben. Además, los espectros de [t] típicamente muestran un pico encima de 4500 Hz seguido luego por un caída abrupta de la amplitud.

Los datos sobre las transiciones de los formantes nos brindan información adicional de qué claves fonéticas usan los hablantes de shipibo para distinguir los puntos de articulación de las consonantes oclusivas. El comportamiento de F2 es una clave especialmente robusta para la oclusiva labial [p]. En cambio, las diferencias de duración de VOT y de la fase de cerrazón de las oclusivas parecen ser las claves fonéticas más robustas encontradas para distinguir entre las consonantes no-velares, [p] y [t], y la consonante velar [k].

Además de la caracterización acústica de las oclusivas, se describió en términos espectrográficos el fenómeno de labialización del shipibo, un fenómeno que es causado solo por la vocal central no-redondeada [i] cuando sigue a una consonante bilabial.

El siguiente capítulo está dedicado al estudio acústico de otro subconjunto de consonantes del shipibo: las fricativas y africadas.

4. FRICATIVAS Y AFRICADAS

Shipibo tiene cuatro consonantes fricativas. Tres de ellas son sibilantes, [s, ʃ, ʂ] y una es laríngea, [h]. Además de las fricativas, hay cuatro consonantes africadas: [bʙ, ts, tʃ, dʒ] (no solo se distinguen por su punto de articulación pero también por su sonoridad). En este capítulo, se discute las propiedades acústicas de cada uno de estos segmentos. Las fricativas se abordan a continuación y las africadas a partir de la sección 4.3.

4.1 FRICATIVAS SIBILANTES SORDAS

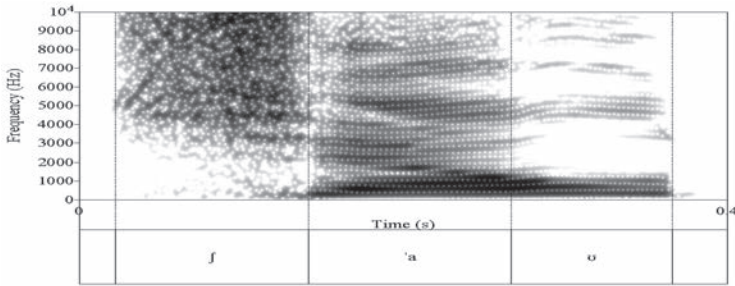
Las fricativas sibilantes sordas del shipibo son tres: una [s] dental, una [ʃ] alveo-palatal y una retrofleja [ʂ]. Los datos en (12) dan ejemplos de dos pares mínimos (y un par cuasi-mínimo) cuyos significados dependen de la opción de la sibilante.

- (12) a. [ʃi.pa.ti] ‘rozar’
 b. [ʃi.pa.ti] ‘desinflar’
 c. [ʂi.ki.ti] ‘quebrarse’
 d. [ʂi.ki] (esp. de iguana)
 e. [ʃa.ʊ] ‘anémico’
 f. [ʃa.ʊ] ‘hueso, espina’

La Figura 4-1 muestra el espectrograma de la palabra [ʃa.ʊ] ‘anémico’, la cual contiene la sibilante alveo-palatal [ʃ] como arranque de la primera sílaba. Acústicamente, las fricativas se caracterizan por presentar turbulencia (es decir, ruido aleatorio). En el caso de las sibilantes, la energía acústica se concentra en las frecuencias altas, que, en los espectrogramas, aparece como una concentración de

tonos oscuros sin patrones definidos en la parte superior. Conforme las frecuencias bajan, los tonos oscuros tienden a hacer mucho más claros. En shipibo, las sibilantes son sordas. Esto se indica en la Figura 4-1 por la ausencia de la barra de voz durante la sibilante.

Figura 4-1: La sibilante [ʃ] como arranque de la palabra [ʃa.ʊ] ‘anémico’



Las sibilantes son las únicas consonantes obstruyentes del shipibo que pueden aparecer en la posición de coda de las sílabas. Los espectrogramas de la Figura 4-2 a la Figura 4-4 muestran las tres sibilantes, [s, ʃ, ʂ], ocurriendo como codas al final de la palabra, después de la vocal [ʊ]. Las flechas serán mencionadas en la sección 4.1.1.

Figura 4-2: La sibilante [s] como coda de la palabra [b^βi.ʃu] ‘locura’

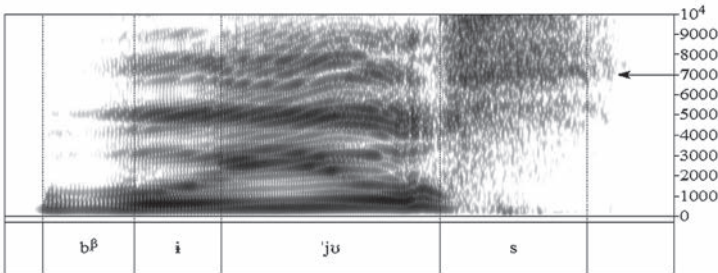


Figura 4-3: La sibilante [ʃ] como coda de la palabra [ki.ʃu] ‘de muslo rígido’

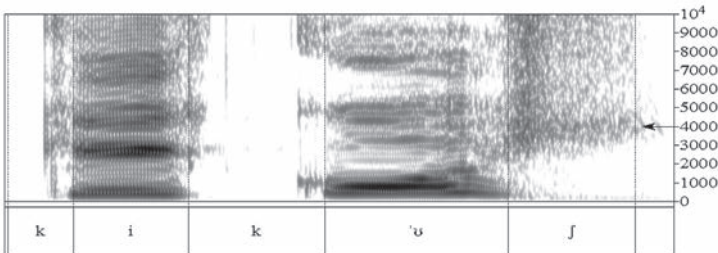
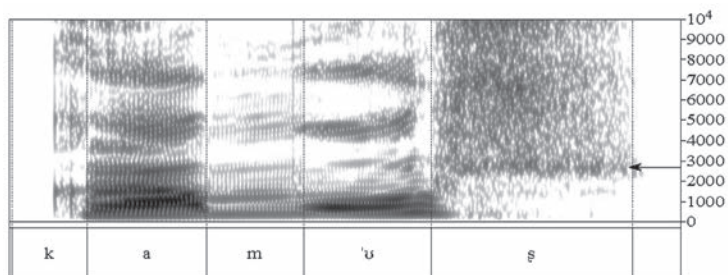


Figura 4-4: La sibilante [ʃ] como coda de la palabra [ka.ˈmuʃ] (esp. de serpiente)



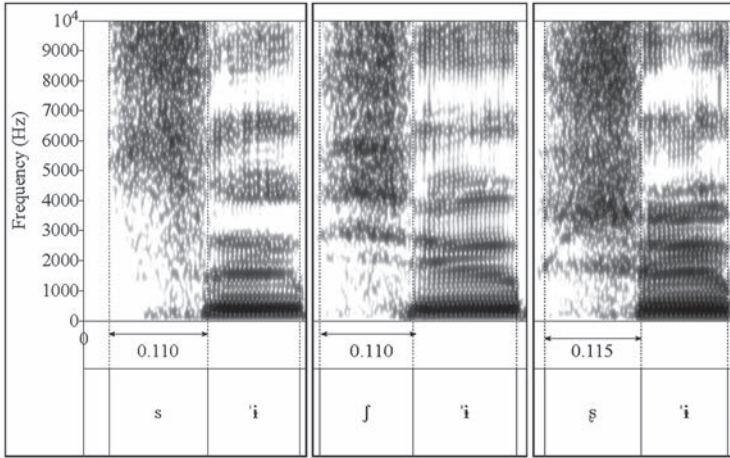
Las sibilantes en posición de coda tienden a ser más cortas que las que ocurren en posición de arranque de sílaba, excepto cuando ocurren en una sílaba final como es el caso en los espectrogramas de arriba. Discutiré la duración de las sibilantes en posición de coda en la sección 7.1.1. En este capítulo me concentraré en la duración de sibilantes en la posición de arranque. La Tabla 4-1 muestra la duración promedio en segundos de las sibilantes en posición de arranque. La duración de las tres consonantes es bastante similar, aproximadamente 110 ms. Los resultados se basan en las mediciones de 6000 sibilantes que ocurren ya sea en la primera o en la segunda sílaba de las palabras.

Tabla 4-1: Promedios de duración (en segundos) de las sibilantes en posición de arranque

[s]	0,112 ^(0,027)
[ʃ]	0,107 ^(0,024)
[ʒ]	0,108 ^(0,025)

Los espectrogramas representados en la Figura 4-5 presentan ejemplos de la duración de las tres sibilantes en posición de arranque. Corresponden a las sílabas iniciales de las palabras [ˈsi.pa.ti] ‘rozar’, [ˈʃi.pa.ti] ‘desinflar’, y [ˈʃi.pu.ti] ‘cerrar la puerta’.

Figura 4-5: Ejemplos de la duración de sibilantes en posición de arranque

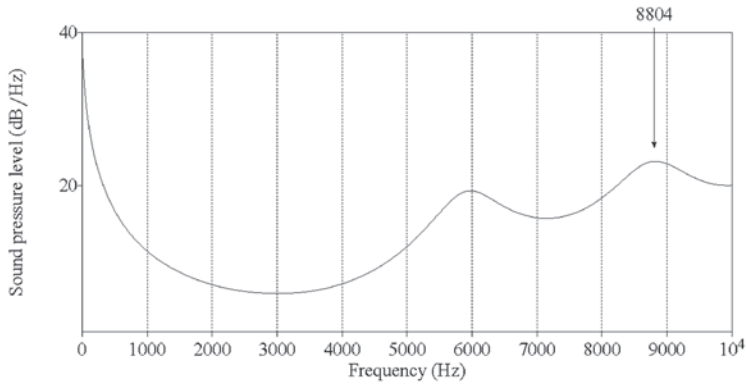


4.1.1 Análisis espectral de las sibilantes

Aunque la energía acústica de las consonantes fricativas tiene un patrón aleatorio (es decir, es ruido, turbulencia creado por fricción), la manera cómo se distribuye o dónde se concentra esa turbulencia no es aleatoria. Es decir, cada sibilante muestra un patrón específico de dónde se concentra la energía acústica a lo largo de las frecuencias. En general, conforme el punto de articulación se mueve del frente hacia atrás a lo largo de la cavidad oral, los picos más prominentes de energía acústica se mueve a frecuencias más bajas (Ladefoged y Maddieson 1996; Johnson 2003; Ladefoged 2003, 2005). Es posible observar esto rápidamente en los espectrogramas de arriba. Por ejemplo, en la Figura 4-2 a la Figura 4-4, el pico de [s] (indicado por una flecha localizada alrededor de los 7000 Hz) es más alto que el pico de [ʃ] y [ʂ], y el pico de [ʃ] (véase la flecha alrededor de 4000 Hz) es más alto que el pico de [ʂ] (véase la flecha alrededor de 3000 Hz).

Aunque los espectrogramas pueden brindar una idea intuitiva de cómo se distribuye la energía acústica en las consonantes fricativas por medio de las tonalidades de los colores negro, gris y blanco, los espectros FFT o LPC nos pueden brindar una imagen mucho más precisa (Harrington y Cassidy 1999; Johnson 2003; Ladefoged 2003). Como en el caso de las fases de explosión de las consonantes oclusivas, se presentará los espectros LPC para estudiar la distribución de la energía acústica en fricativas. La Figura 4-6 presenta el espectro LPC de la sibilante [s] de la palabra [ʔi.pa.ti] ‘rozar’. El espectro corresponde a una ventana de 46 ms localizada en el centro de la sibilante. La flecha indica dónde se sitúa el pico más prominente de la fricativa [s] dentro del rango de frecuencias 0-10000 Hz.

Figura 4-6: Espectro LPC de [s] en la palabra ['si.pa.ti] ‘rozar’



La posición promedio del pico más alto en el espectro de frecuencias en las sibilantes se muestra en la Tabla 4-2 (para hombres) y en la Tabla 4-3 (para mujeres). Los resultados se basan en el análisis espectral de sibilantes que ocurren como el arranque inicial de una palabra. Para realizar este análisis se tomó una ventana de 46 ms centrada en el medio de aproximadamente 3600 sibilantes. Los promedios mostrados toman en cuenta los efectos que la siguiente vocal pueda tener sobre las propiedades espectrales de estos segmentos.

La celda gris de cada tabla indica que no se encontró una palabra que contenga la secuencia *[ʃi]. Es posible encontrar algunas palabras donde la sibilante [ʃ] y la vocal [i] aparecen adyacentes. Sin embargo, en estos casos, la sibilante siempre precede a la vocal, por ejemplo, [ka.ʃi.i.ti] (from /kaʃ -ʔiti/) ‘carraspear’, [ʔʃi.ʃu] ‘nalga’ y [ʔʃi.ku] ‘demasiado corto’. Sin embargo, en general, inclusive la secuencia [iʃ] es rara en shipibo. Los promedios mostrados en las celdas grises se basan en los análisis espectrales de 54 sibilantes [ʃ] adyacentes a [i] de las tres palabras que acaban de mencionarse.

Las tablas muestran que la sibilante [s] tiene los picos más altos en el espectro de frecuencias. El pico para [s] se localiza aproximadamente a 7694Hz para hombres y a 8462 Hz, para mujeres. Las sibilantes [ʃ] y [ʃ] se encuentran más cercanas con respecto a la posición de sus picos más altos de energía acústica dentro del espectro de frecuencias. Sin embargo, la sibilante alveo-palatal [ʃ] claramente tiene un pico más alto de energía comparado con el de la sibilante retrofleja [ʃ]. El pico para [ʃ] está aproximadamente a 3927 Hz en hombres y a 4213 Hz en mujeres mientras que el pico para [ʃ] se localiza a 3134 Hz en hombres y a 3317 Hz en mujeres. Los picos de energía de las sibilantes son comparativamente más altos en mujeres que en hombres debido a las diferencias en el tamaño de las

cavidades orales, contra más pequeño el espacio donde el aire puede vibrar, más altas son las frecuencias.

Tabla 4-2: Promedio de la posición del pico más alto de energía en el espectro de las sibilantes - hombres

	[s]	[ʃ]	[ʂ]
[i]	8006 Hz ⁽¹⁴²³⁾	4110 Hz ⁽⁷¹³⁾	3302 Hz ⁽⁶⁸³⁾
[i̥]	7944 Hz ⁽¹⁵⁵²⁾	4067 Hz ⁽⁶⁴⁷⁾	3232 Hz ⁽⁴⁹⁰⁾
[a]	7788 Hz ⁽¹⁵⁹⁷⁾	4085 Hz ⁽⁶¹³⁾	3313 Hz ⁽⁴⁸⁷⁾
[u]	7037 Hz ⁽¹⁷⁶⁹⁾	3447 Hz ⁽⁶⁰⁸⁾	2856 Hz ⁽⁴⁸⁶⁾

Tabla 4-3: Promedio de la posición del pico más alto de energía en el espectro de las sibilantes - mujeres

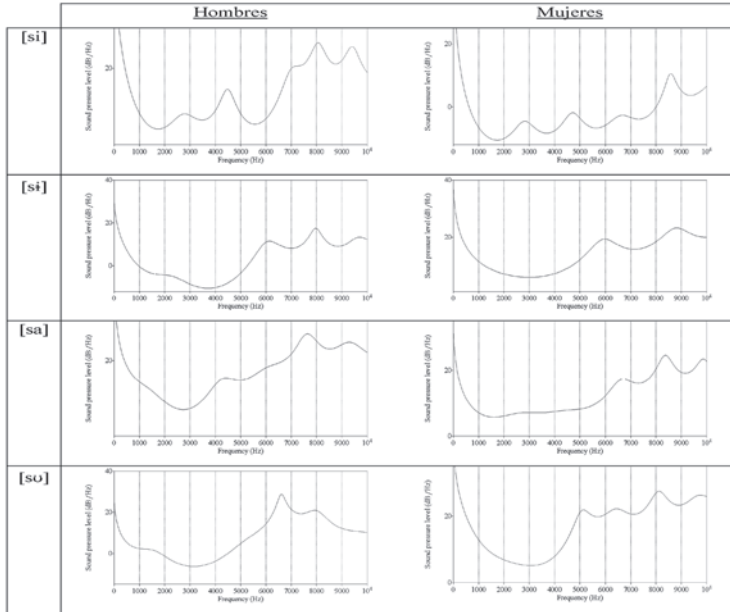
	[s]	[ʃ]	[ʂ]
[i]	8602 Hz ⁽¹²⁰³⁾	4514 Hz ⁽⁶⁶⁹⁾	3643 Hz ⁽⁴⁷⁹⁾
[i̥]	8579 Hz ⁽¹¹²⁶⁾	4325 Hz ⁽⁶³⁷⁾	3406 Hz ⁽²⁴³⁾
[a]	8526 Hz ⁽¹⁰⁹⁸⁾	4422 Hz ⁽⁶⁴³⁾	3558 Hz ⁽³⁴⁴⁾
[u]	8141 Hz ⁽¹⁴²⁶⁾	3591 Hz ⁽⁵²²⁾	2989 Hz ⁽²⁷²⁾

La Tabla 4-2 y la Tabla 4-3 también revelan que la vocal [u] baja el pico de energía de todas las sibilantes. Obsérvese que, por ejemplo, el pico de energía para la [s] seguida de las vocales [i], [i̥] o [a] en los hablantes hombres no muestra mayor diferencias (está localizado alrededor de 7913 Hz). En contraste, el pico de [s] seguida de la vocal [u] baja a 7037 Hz, una caída de 876 Hz. Lo mismo se repite para [ʃ] y [ʂ] tanto en hombres como en mujeres.

Los espectros LPC que se muestran a continuación corresponden a [s], [ʃ] y [ʂ]. Los espectros ilustran las diferencias indicadas en la Tabla 4-2 y la Tabla 4-3 entre hombres y mujeres con respecto a la posición del pico de energía de las sibilantes. Todos los espectros corresponden a las sibilantes en posición de arranque al inicio de las palabras. La primera columna indica la vocal que sigue a la sibilante. Además de la posición del pico de energía, obsérvese que la carac-

terística principal de los espectros de [s] (ver los espectros de la Figura 4-7) es que la energía aumenta conforme las frecuencias suben.

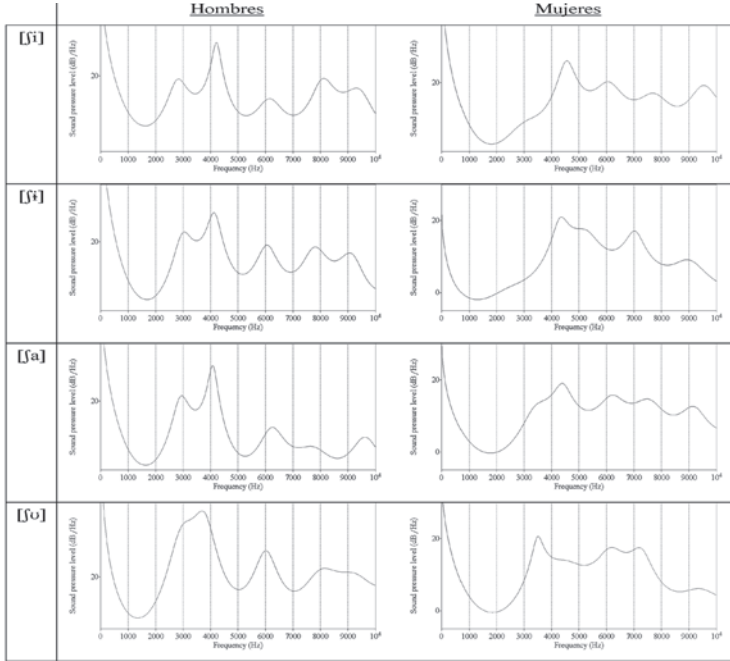
Figura 4-7: Espectros LPC de [s] en contextos vocálicos diferentes



Los espectros en la Figura 4-8 corresponden a la sibilante [ʃ], tomados bajo las mismas condiciones que los espectros de [s] en la Figura 4-7. La diferencia principal entre los espectros de [s] y [ʃ] es la posición del pico de energía. En el caso de [ʃ], se localiza en promedio entre 4000 Hz y 5000 Hz mientras que el pico para [s] generalmente ocurre por encima de 7000 Hz. La posición del pico más alto también afecta el comportamiento de la energía acústica en ambos tipos de sibilantes. Para [ʃ], la energía descende conforme pasa la frontera de los 5000 Hz mientras que para [s], la energía se mantiene incrementando después de los 5000 Hz.

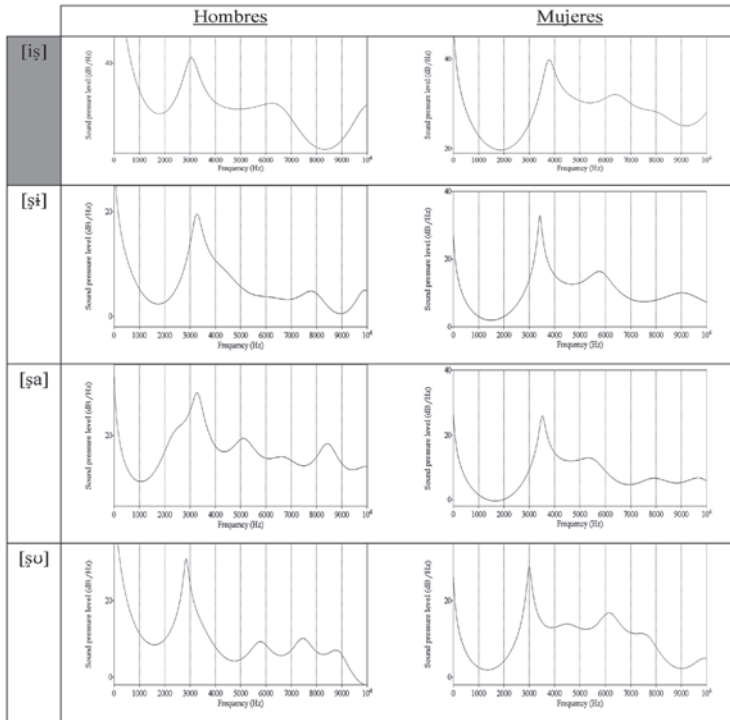
En la Figura 4-9 se muestran ejemplos de los espectros de [ʃ]. Las celdas en gris contienen los resultados de la secuencia [iʃ]. Los dos espectros de [ʃ] que aparecen seguidos de la vocal alta [i] corresponden a la palabra [ka.ʰi.ʃi.ti] (/kaʃ-ʔiti/) ‘carraspear’. La posición del pico de energía claramente distingue [ʃ] de [s]. En el caso de [ʃ], éste aparece entre 3000 Hz y 4000 Hz (contrástese esto con los 7000 Hz o más del pico para [s]). La posición promedio del pico de [ʃ] no se encuentra muy lejos del de [ʃ] (4000 Hz – 5000 Hz). Sin embargo, una

Figura 4-8: Espectros LPC de [ʃ] en contextos vocálicos diferentes



inspección meticulosa a los espectros de [ʃ] revela dos características únicas. El primer pico que muestra el espectro de [ʃ] corresponde al pico más alto. Esto es diferente a lo que encontramos en los espectros de [ʃ], los cuales pueden tener picos que anteceden o siguen al pico más alto. La segunda diferencia entre los espectros de [ʃ] y [ʃ] es que en el primero, la energía acústica del pico más alto se concentra en un rango de frecuencias pequeño y claramente domina en altura a todos los otros picos menores en el espectro. De este modo, aunque los picos de [ʃ] y [ʃ] son relativamente cercanos en el espacio de frecuencias, la forma de sus espectros claramente los distingue.

Figura 4-9: Espectros LPC de [ʃ] contextos vocálicos diferentes



4.1.2 Las sibilantes y las transiciones de formantes

Como en el caso de las consonantes oclusivas, las transiciones de los formantes nos brindan claves del punto de articulación de las sibilantes. La Tabla 4-4 y la Tabla 4-5 muestran los valores promedios de F2 para las vocales en el punto de transición con una consonante sibilante. Los resultados se basan en el análisis de 9000 vocales adyacentes tanto a sibilantes en posición de arranque como a sibilantes en posición de coda. Las vocales analizadas ocurren dentro de la primera o segunda sílaba de palabras trisilábicas. El promedio de F2 de la vocal [i] en la transición con la sibilante [ʃ] (véase las celdas grises en la tabla de abajo) solo se basa en el análisis de las tres palabras que contienen esa secuencia: [ka.ʰi.ʃ.i.ti] (/kaʃ -ʔiti/) ‘carraspear’, [ʰtʃi.ʃu] ‘nalga’ y [ʰtʃi.ʃ.ku] ‘demasiado corto’.

Tabla 4-4: Los promedios de F2 en las transiciones con sibilantes (hombres)

	[s]	[ʃ]	[ʂ]
[i] - 2244	1975 Hz ⁽²⁰⁴⁾	2122 Hz ⁽²⁰⁷⁾	1871 Hz ⁽²³⁴⁾
[a] - 1418	1437 Hz ⁽¹⁵³⁾	1827 Hz ⁽¹⁶⁶⁾	1529 Hz ⁽¹⁷⁰⁾
[ɨ] - 1539	1462 Hz ⁽¹⁶²⁾	1982 Hz ⁽²¹²⁾	1631 Hz ⁽¹⁷¹⁾
[u] - 896	1178 Hz ⁽¹³⁰⁾	1720 Hz ⁽¹⁸⁸⁾	1221 Hz ⁽¹⁵³⁾

Tabla 4-5: Los promedios de F2 en las transiciones con sibilantes (mujeres)

	[s]	[ʃ]	[ʂ]
[i] - 2676	2323 Hz ⁽¹⁹³⁾	2544 Hz ⁽¹⁴³⁾	1938 Hz ⁽²⁰⁴⁾
[a] - 1597	1620 Hz ⁽¹⁸²⁾	2185 Hz ⁽¹⁸²⁾	1752 Hz ⁽¹⁹³⁾
[ɨ] - 1704	1617 Hz ⁽¹⁹⁶⁾	2319 Hz ⁽²⁵⁰⁾	1710 Hz ⁽¹⁹⁹⁾
[u] - 935	1352 Hz ⁽¹⁵¹⁾	1966 Hz ⁽²⁰⁴⁾	1290 Hz ⁽¹⁵⁵⁾

Tanto la Figura 4-10 como la Figura 4-11 nos dejan ver gráficamente los cambios que sufre F2 en las transiciones con las sibilantes. Los cambios corresponden a los valores presentados en la Tabla 4-4 y Tabla 4-5, respectivamente. Los círculos negros indican el valor de F2 en el medio de la vocal y los círculos blancos indican el valor de F2 en la transición con la sibilante. Las flechas señalan la dirección del cambio desde el medio de la vocal. Cuando ambos valores de F2 son muy cercanos o iguales (menos de 100 Hz aparte), éstos aparecen encerrados en un círculo.

La Figura 4-10 y la Figura 4-11 muestran que el comportamiento de F2 en el punto de transición con la sibilante [ʃ] es bastante diferente al que se observa en la transición con [s] y [ʂ]. Cuando [ʃ] ocurre adyacente a [i], el F2 muestra un pequeño descenso mientras que para [s] y [ʂ], el descenso es mucho más pronunciado. Cuando [ʃ] está adyacente a las vocales centrales [ɨ] y [a], F2 muestra un importante incremento. En contraste, F2 casi se mantiene con el mismo valor que muestra en el medio de las vocales centrales cuando ocurre adyacente a las sibilantes [s] y [ʂ].

Figura 4-10: Transiciones de F2 para las sibilantes
(mujeres)

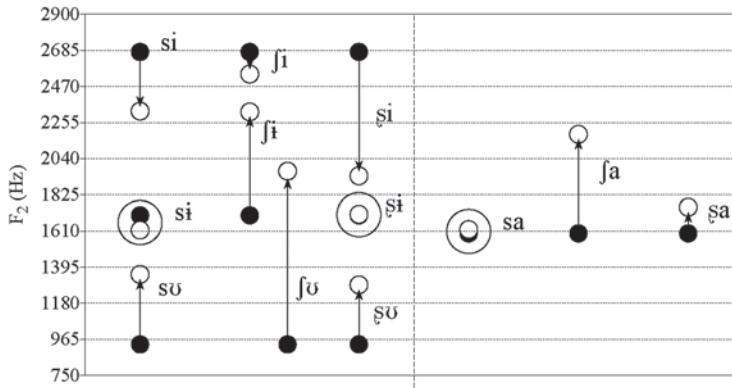
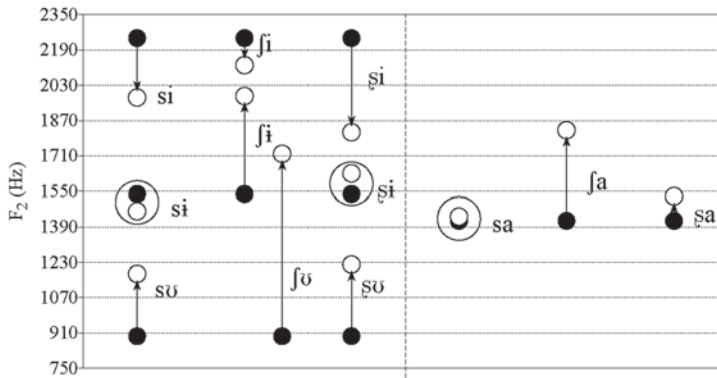


Figura 4-11: Transiciones de F2 para las sibilantes
(hombres)

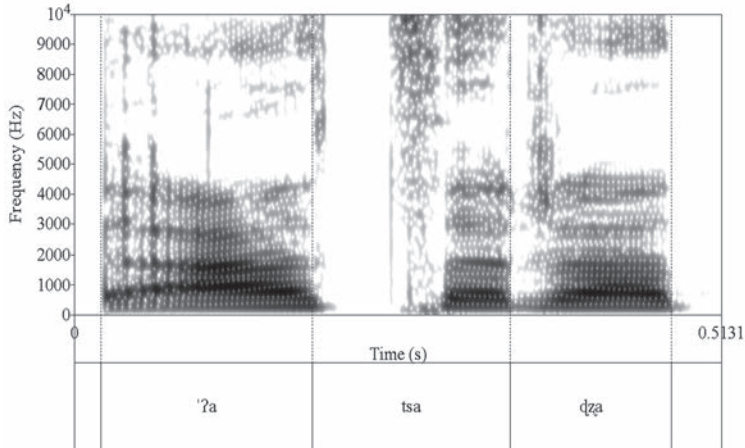


El valor de F2 se incrementa en la transición con cada una de las tres sibilantes cuando éstas ocurren adyacentes a la vocal [ʊ]. Sin embargo, mientras el incremento es bastante similar cuando [ʊ] están junto a [s] o [ʂ], el incremento de F2 es mucho más pronunciado cuando [ʃ] aparece adyacente a [ʊ]. De este modo, los cambios que sufre F2 en la transición con las consonantes sibilantes son una clave importante para determinar el punto de articulación de la sibilante [ʃ]. La distinción entre las sibilantes [s] y [ʂ] parece depender más de las propiedades espectrales establecidas en la sección 4.1.1.

4.2 LA FRICATIVA LARÍNGEA

Shipibo tiene una fricativa laríngea sorda, [h]. Hay palabras que solo se distinguen por la presencia o ausencia de la fricativa laríngea; como por ejemplo, [ˈa.tsa] ‘yuca’ y [ˈha.tsa] ‘fastidioso’. Este contraste se puede observar en los espectrogramas de la Figura 4-12¹ y la Figura 4-13. La ausencia de la barra de voz durante la producción de la consonante [h] en la Figura 4-13 indica su naturaleza sorda. Además, obsérvese que la energía acústica de la fricativa [h] se muestra mucho mejor distribuida a lo largo del rango entero de frecuencias que en las consonantes sibilantes. Esta es una de las características importantes de este segmento laríngeo que la distingue de otras fricativas.

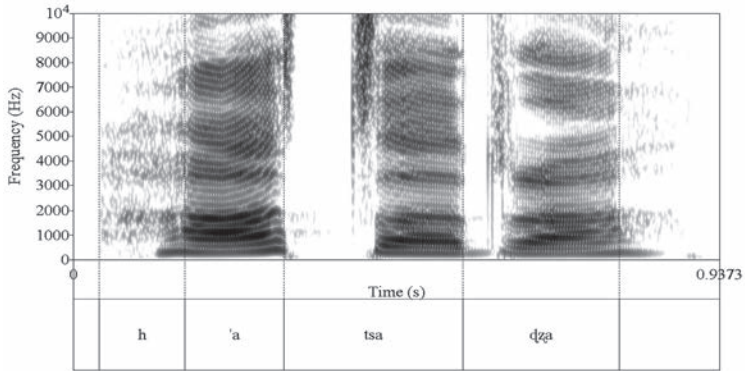
Figura 4-12: Espectrograma de la palabra /atsa -dʒa/ → [ʔa.tsa.dʒa] ‘yuca’ (EV)



En shipibo, la fricativa laríngea usualmente ocurre al inicio de palabras ([ˈha.ka] ‘pumagarza’ (esp. de pájaro), [hi.ˈti] ‘olor que produce estornudos’). Sin embargo, existen algunos pocos casos en los cuales este segmento ocurre en el interior en las palabras. Dentro del vocabulario nativo del shipibo, la fricativa laríngea puede encontrarse dentro de interjecciones [ˈhi.hi] (respuesta de los hombres a un saludo), [ˈhü.hü] (respuesta de las mujeres a un saludo) y en algunas pocas palabras onomatopéyicas ([hü.ˈhü.ʔa.ti] ‘ladrar’ (literalmente, hacer sonidos ‘hü.ˈhü’)).

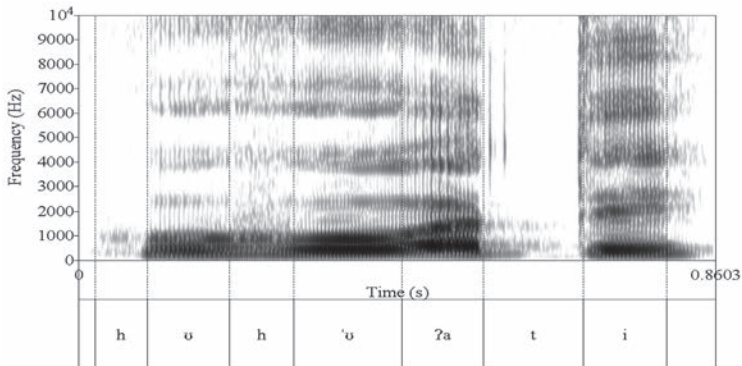
¹ La oclusiva glotal inicial que se observa en la Figura 4-12 se inserta porque la palabra está precedida por una palabra que acaba en vocal y luego una pequeña pausa (véase capítulo 2, sección 2.5).

Figura 4-13: Espectrograma de la palabra /hatsa -dʒa/ → [ˈha.tsa.dʒa] ‘fastidioso’ (EV)



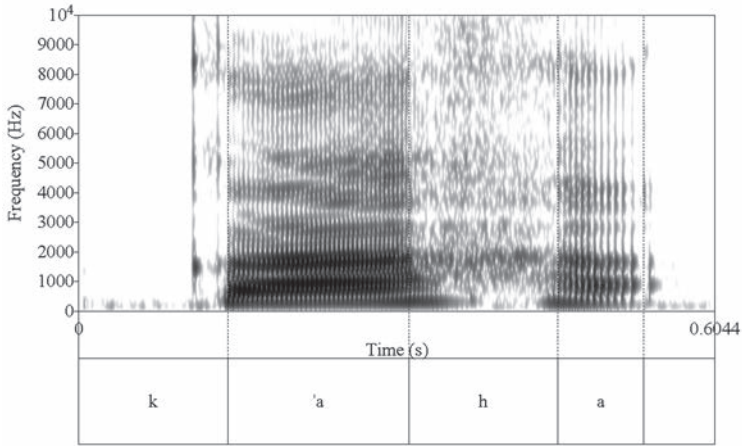
La mayoría de las ocurrencias de [h] al interior de palabras pertenece al vocabulario de préstamos obtenidos del castellano (por ejemplo, [ˈka.ha] ‘caja’ (del castellano [ˈka.xa])). En estos préstamos, la consonante [h] ocurre en vez de las fricativas del castellano [f] y [x] ([a.ˈsu.hi.dʒi] ‘azufre’ (del castellano [a.ˈsu.fre]), [ˈhʊ.kʊ] ‘foco’ (del castellano [ˈfo.ko])).² Los espectrogramas de la Figura 4-14 y la Figura 4-15 muestran la fricativa laríngea [h] al interior de una palabra. La Figura 4-14 ilustra la palabra [hʊ.ˈhʊ.ʔa.ti] ‘ladrar’ y la Figura 4-15, el préstamo [ˈka.ha] ‘caja’ (del castellano [ˈka.xa]).

Figura 4-14: La fricativa laríngea [h] en la palabra [hʊ.ˈhʊ.ʔa.ti] ‘ladrar’



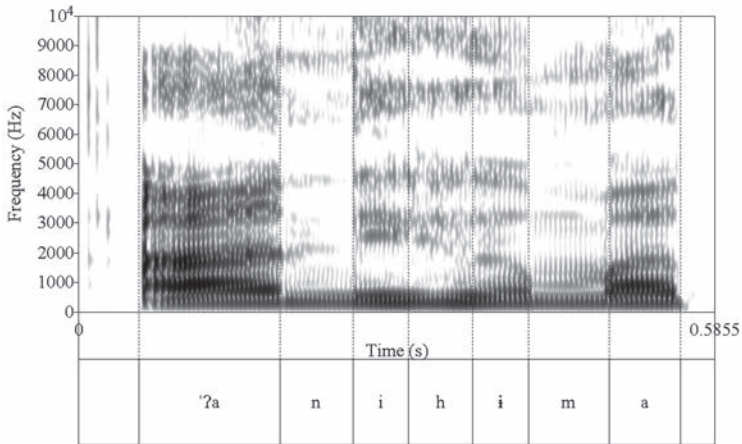
² Los préstamos del castellano muestran diferentes grados de asimilación a la fonología nativa del shipibo. Los hablantes más jóvenes tienden a pronunciar los préstamos tan cerca como les es posible a cómo esas palabras se pronuncian en castellano. De este modo, no es raro encontrar hablantes de shipibo que pronuncian, por ejemplo, la palabra ‘foco’ como [fo.ko] en lugar de [hʊ.kʊ].

Figura 4-15: La fricativa laríngea [h] en la palabra ['ka.ha] 'caja'



La fricativa laríngea puede también aparecer en la frontera entre palabras dentro de compuestos, como por ejemplo en la palabra ['a.ni.hi.ma] 'ciudad'. Este compuesto está formado por las palabras ['a.ni] 'grande' y ['hi.ma] 'pueblo, comunidad'.

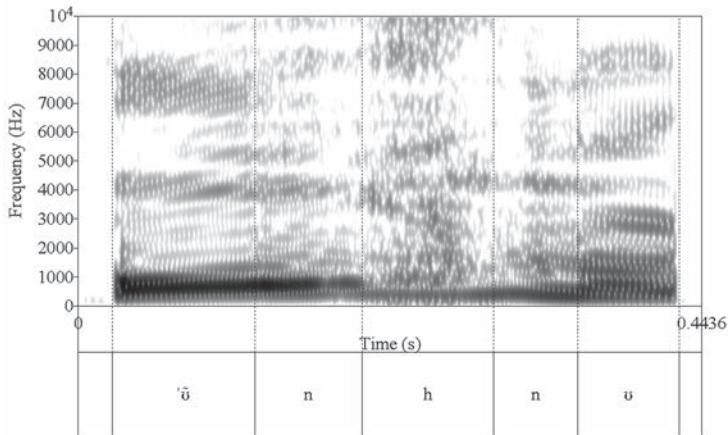
Figura 4-16: La fricativa laríngea [h] en la palabra ['a.ni.hi.ma] 'ciudad'



Shipibo no permite que la fricativa laríngea ocupe la posición de coda de las sílabas excepto por la palabra: ['unh.nu] 'allá muy lejos'. Ésta se muestra en el espectrograma de la Figura 4-17. La pronunciación de este adverbio alterna libremente con la forma ['un.nu] (véase la sección 7.2). Las codas complejas son

muy raras de encontrar en Shipibo pero cuando ocurren, la nasal [n] siempre es seguida por una sibilante. La forma [ʰũh.nũ] (usualmente realizada como [ũnh.nũ]) es el único caso que conozco de una coda compleja que tienen una fricativa laríngea, [h], en lugar de una sibilante.

Figura 4-17: La fricativa laríngea [h] en la palabra [ʰũh.nũ] ‘allá muy lejos’



La Tabla 4-6 presenta la duración promedio de la fricativa laríngea medida en segundos. Dada los pocos casos donde [h] ocurre al interior de una palabra y el hecho que algunas veces en ese contexto, la [h] coalesce completamente con las vocales que la rodean, lo cual hace difícil su medición, las medidas de duración que se presenta en la Tabla 4-6 solo corresponden a los casos donde la [h] aparece al inicio de la palabra. Los resultados de la Tabla 4-6 se basan en la medición de 600 instancias de la fricativa laríngea. En promedio, [h] dura 105 ms en sílabas acentuadas en posición inicial de palabra mientras tiende a ser aproximadamente 10 ms más corta en sílabas inacentuadas iniciales.

Tabla 4-6: Duración promedio de la fricativa laríngea

	Duración (s)
[h] – en sílaba acentuada inicial	0,105 ^(0,040)
[h] – en sílaba inacentuada inicial	0,094 ^(0,043)

4.2.1 Análisis espectral de [h]

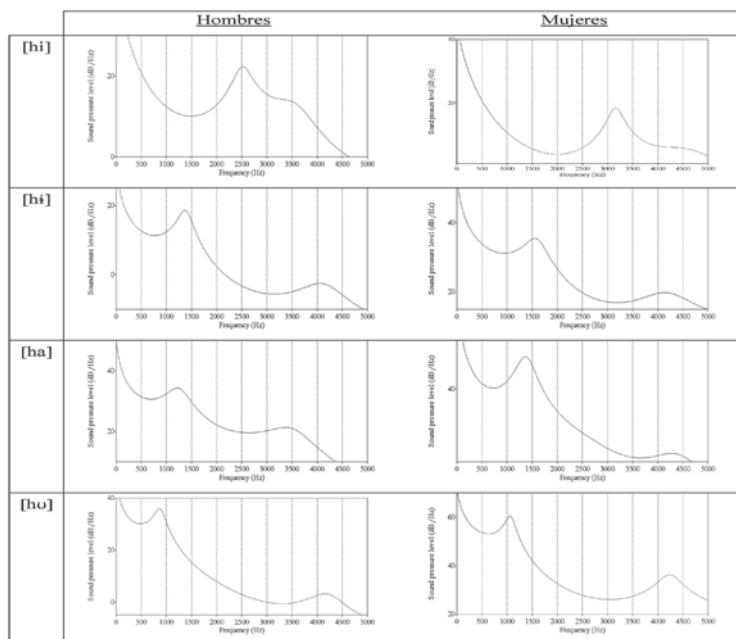
La Tabla 4-7 muestra dónde se localiza el pico más alto de energía en el espectro de una fricativa laríngea. El pico y la distribución de la energía acústica de [h] varían considerablemente dependiendo de los segmentos que la rodean. Por ejemplo, cuando [h] aparece entre dos vocales, su pico empieza muy cerca al valor de F2 de la vocal izquierda y conforme se aproxima a la siguiente vocal, su pico se mueve hacia el valor de F2 de esa vocal. Los resultados que se presentan en la Tabla 4-7 corresponden a las fricativas laríngeas en posición inicial de palabra y después de una pausa. Las mediciones fueron tomadas de la primera ocurrencia de la palabra objetivo en la frase que se usó en este estudio (véase la sección 1.3.1). Los análisis espectrales se realizaron dentro de una ventana de 46 ms centrada en el medio de la fricativa. Los resultados se basan en las mediciones de aproximadamente 600 instancias de [h].

Tabla 4-7: Promedio del pico más alto en el espectro de la fricativa laríngea – [h]

	[i]	[ɨ]	[a]	[u]
[h] – hombres	2508 Hz ⁽²¹⁰⁾	1516 Hz ⁽¹⁸²⁾	1493 Hz ⁽¹⁶⁴⁾	952 Hz ⁽¹⁰¹⁾
[h] – mujeres	3112 Hz ⁽²⁹⁷⁾	1551 Hz ⁽¹⁶⁶⁾	1530 Hz ⁽¹⁷¹⁾	1010 Hz ⁽¹¹⁴⁾

La Tabla 4-7 indica que el pico de la fricativa laríngea sigue bastante de cerca al F2 de la siguiente vocal (véase los promedios de F2 para las vocales en la sección 2.1). Cuando la [h] es seguida por la vocal [i], el pico se localiza en promedio entre 2500 Hz y 3000 Hz. Cuando la [h] es seguida por una vocal central, ya sea por [ɨ] o [a], el pico de [h] se localiza aproximadamente a 1500 Hz y cuando es seguida de la vocal posterior [u], el pico se encuentra localizado cerca de los 1000 Hz. La Figura 4-18 presenta una muestra de espectros LPC de la fricativa [h] tanto de hombres como de mujeres. La primera columna muestra qué tipo de vocal sigue a la fricativa laríngea. A diferencia de los espectros presentados hasta ahora, los cuales muestran un rango de frecuencia entre 0-10000 Hz, los espectros de la Figura 4-18 tienen un rango de frecuencia entre 0-5000 Hz. Este rango representa la mayor parte donde se concentra la energía acústica de [h].

Figura 4-18: Espectros LPC de [h] en diferentes contextos vocálicos



4.3 LAS AFRICADAS SORDAS

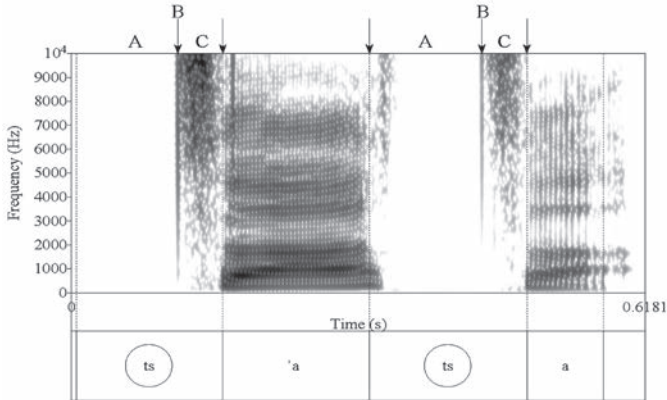
En esta sección, discutiré las dos africadas sordas del shipibo: [ts] y [tʃ]. En shipibo, las africadas solo pueden ocurrir como arranques de sílaba. Los datos en (13) presentan pares mínimos en los cuales el significado se distingue por el tipo de africada.

- (13) a. [ʰt͡sa.ka.ti] ‘pescar con flechas’
 b. [ʰt͡ʃa.ka.ti] ‘machucar’
 c. [ʰmʊ.t͡sa.ti] ‘triturar algo con o sin agua’
 d. [ʰmʊ.t͡ʃa.ti] ‘cantar’

En términos articulatorios, una africada consiste de tres momentos o fases: una fase de cerrazón, una fase de explosión y una fase de fricción. Mientras todas las africadas muestran una fase de cerrazón y de fricción, la fase de explosión no siempre está presente o no es suficientemente fuerte para ser detectada. En la Figura 4-19, se presenta dos casos de la africada alveolar [ts] en la palabra [ʰt͡sa.t͡sa] ‘pez’ (pronunciada en aislamiento): la primera africada [ts] ocurre en el arranque de la sílaba acentuada inicial y la otra, como el arranque de la sílaba inacentuada. El espectrograma muestra las tres fases de la africada. Las fases de

cerrazón se indican por la letra A, las de explosión por la letra B y las de fricción por la letra C. La ausencia de una barra de voz durante la producción de [ts] indica su naturaleza sorda.

Figura 4-19: La africada /ts/ en la palabra [t'sa.tsa] 'pez'



La Tabla 4-8 y la Tabla 4-9 presentan las duraciones promedio en segundos de las africadas sordas. La Tabla 4-8 muestra la duración de [ts] y [tʃ] cuando ocurren como arranques de sílabas acentuadas y la Tabla 4-9 muestra sus duraciones cuando ocurren en sílabas inacentuadas. Los resultados presentados en ambas tablas se basan en las mediciones de aproximadamente 2200 instancias de las africadas. Todas las africadas sordas que se midieron aparecen como arranques de la primera o segunda sílaba de palabras trisilábicas. Sin embargo, debido a la dificultad de medir la fase de cerrazón de las africadas cuando ocurren al inicio de palabras, las mediciones de esa fase solo se han considerado en aquellas africadas [ts] y [tʃ] que aparecen en contextos intervocálicos de modo que el inicio y el fin de la fase de cerrazón pudieran reconocerse con precisión.

De modo similar, los resultados para la duración total de las africadas [ts] y [tʃ] solo incluyen aquellas instancias en las cuales el comienzo y final de las africadas pudieron ser claramente establecidos. Esto solo se pudo lograr con certeza en contextos intervocálicos. En contraste, los resultados de la duración de la fase de fricción sí incluyen africadas sordas que ocurren tanto en posición inicial de palabra como aquellas que aparecen en posición intervocálica dentro de la palabra.

La Tabla 4-8 y la Tabla 4-9 indican que la duración de las africadas [ts] y [tʃ] son muy similares. Como arranques de sílabas acentuadas, tienden a durar aproximadamente 130 ms. En ese contexto, la fase de cerrazón dura cerca de 80 ms y la de fricción tiende a durar alrededor de 50 ms. En sílabas inacentuadas, la

duración de las africadas [ts] y [tʃ] tienden a ser 20 ms más corta que en sílabas acentuadas. Sin embargo, obsérvese que lo que disminuye en las sílabas inacentuadas es la fase de cerrazón de las africadas. La duración de la fase de fricción se mantiene virtualmente igual tanto en sílabas acentuadas como inacentuadas.

Tabla 4-8: Duración promedio de [ts] y [tʃ] en sílabas acentuadas (en segundos)

	Cerrazón	Fricción	Duración Total
[ts]	0,083 ^(0,019)	0,052 ^(0,010)	0,130 ^(0,019)
[tʃ]	0,081 ^(0,014)	0,054 ^(0,011)	0,131 ^(0,017)

Tabla 4-9: Duración promedio de [ts] y [tʃ] en sílabas inacentuadas (en segundos)

	Cerrazón	Fricción	Duración Total
[ts]	0,060 ^(0,013)	0,052 ^(0,011)	0,110 ^(0,013)
[tʃ]	0,058 ^(0,010)	0,051 ^(0,013)	0,107 ^(0,013)

Los espectrogramas de la Figura 4-20 y la Figura 4-21 nos permiten comparar el efecto duracional que tiene la presencia o ausencia del acento en dos instancias de la africada [tʃ], una en una sílaba acentuada y la otra, en una inacentuada. La duración total se muestra en la parte superior y la duración de cada una de las fases se muestra en la parte inferior.

Figura 4-20: [tʃ] en la sílaba acentuada de la palabra [mi.ˈtʃa.ma] ‘seco’

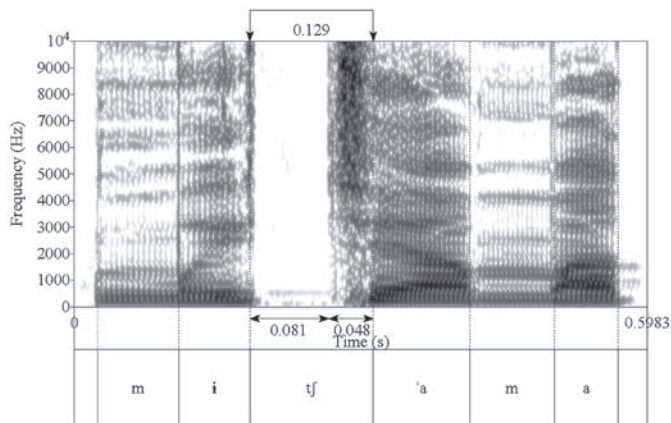
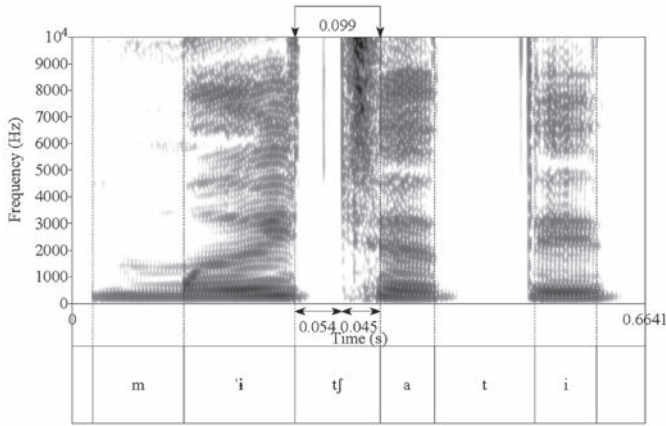


Figura 4-21: [tʃ] en la segunda sílaba inacentuada de [¹mi.tʃa.ti] ‘convertirse en un buen pescador’



4.3.1 Análisis espectral de la fase de fricción de las africadas sordas

Los resultados en la Tabla 4-10 y la Tabla 4-11 muestran la frecuencia del pico más alto de energía para las africadas sordas. Se obtuvieron de una ventana de 46 ms en el centro de la fase de fricción de [ts] y [tʃ]. En aquellos casos donde la fase de fricción fue más pequeña que la ventana de 46 ms, se utilizó la fase entera de fricción de la africada para calcular el pico más alto en el espectro. Ambas tablas toman en cuenta el efecto que tiene la vocal siguiente en las propiedades espectrales de la fase de fricción de la africada. Los análisis espectrales que se presentan en esta sección se basan en el mismo conjunto de datos usados para calcular la duración del periodo de fricción que se presentó en la sección anterior.

En general, la fase de fricción de las africadas [ts] y [tʃ] sigue la misma tendencia de las sibilantes [s] y [ʃ]. Las tablas muestran que, en promedio, el pico más alto para la africada alveolar [ts] es mucho más alto que el de la africada alveo-palatal [tʃ]. El pico para el primero se encuentra alrededor de 8000 Hz para los hablantes hombres y 8600 Hz para las hablantes mujeres mientras que para [tʃ], el pico se localiza aproximadamente a 4000 Hz para hombres y a 4500 Hz para mujeres. Como en el caso de [s] y [ʃ], cuando las africadas [ts] y [tʃ] son seguidas por la vocal [ʊ], el pico más alto en el espectro aparece en una frecuencia ligeramente más baja.

Tabla 4-10: El pico más alto en el espectro de las africadas sordas hombre

	[ts]	[tʃ]
[i]	8195 Hz ⁽¹³⁰⁴⁾	4099 Hz ⁽⁶³⁵⁾
[ɪ]	8105 Hz ⁽¹¹⁹⁸⁾	3924 Hz ⁽⁵³⁹⁾
[a]	8073 Hz ⁽¹³⁵³⁾	3977 Hz ⁽⁶⁷²⁾
[ʊ]	7600 Hz ⁽¹⁴⁵⁶⁾	3465 Hz ⁽⁶⁰⁹⁾

Tabla 4-11: El pico más alto en el espectro de las africadas sordas mujeres

	[ts]	[tʃ]
[i]	8729 Hz ⁽¹²⁵⁸⁾	4557 Hz ⁽⁵⁰⁶⁾
[ɪ]	8668 Hz ⁽¹⁰⁰²⁾	4582 Hz ⁽⁴⁵⁷⁾
[a]	8550 Hz ⁽¹⁰⁵⁵⁾	4550 Hz ⁽⁴⁴⁴⁾
[ʊ]	8270 Hz ⁽¹¹³¹⁾	3978 Hz ⁽⁴⁰⁹⁾

La Figura 4-22 y la Figura 4-23 muestran espectros LPC de [ts] y [tʃ], respectivamente. De modo similar al caso de [s], el espectro del periodo de fricción de [ts] se caracteriza por un incremento de energía conforme las frecuencias suben. La mayor parte de la energía acústica se concentra en las frecuencias más altas (de 5000 Hz para arriba).

Las características espectrales de la fase de fricción de [tʃ] son similares a las de la sibilante [ʃ]: el pico más alto aparece entre los 4000 Hz y 5000 Hz y podría aparecer acompañado de otros picos prominentes pero más pequeños.

Figura 4-22: Espectros LPC de la fase de fricción de [ts] en diferentes contextos vocálicos

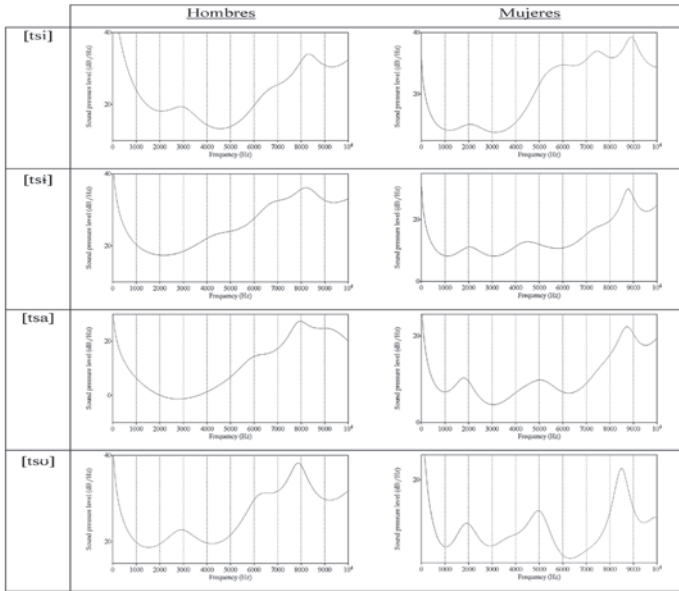
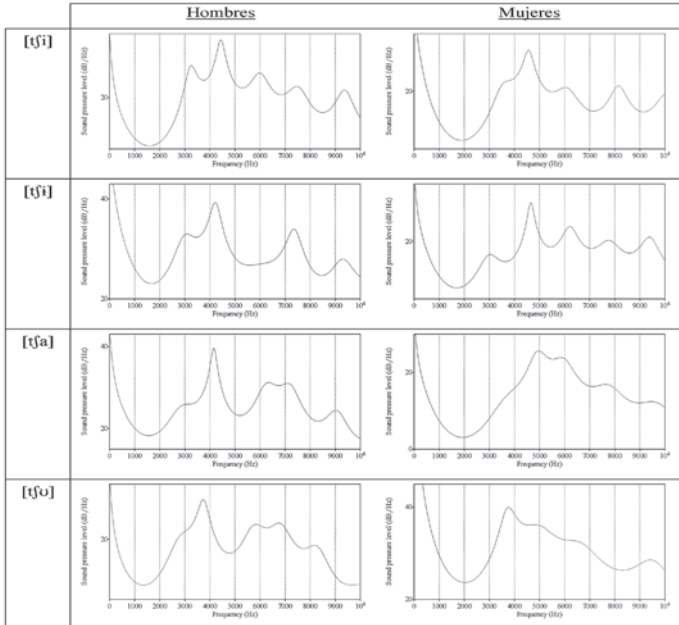


Figura 4-23: Espectros LPC de la fase de fricción de [tʃ] en diferentes contextos vocálicos



4.3.2 Las africadas sordas y la transición de formantes

La transición del F2 en las vocales adyacentes a las africadas sordas [ts] y [tʃ] es muy similar a la transición observada para [s] y [ʃ]. Con el fin de ser exhaustivo, muestro en la Tabla 4-12 y la Tabla 4-13 los resultados de los promedios de F2 al momento que ocurre la transición hacia las africadas sordas. Como es de costumbre, la primera columna indica la vocal adyacente a la africada y el valor ideal de F2 en el medio de la vocal de modo que pueda ser comparado con el valor de F2 en la transición con [ts] (segunda columna) o con [tʃ] (tercera columna). Los promedios de F2 en las transiciones incluyen vocales que ocurren antes y después de la africada. No he separado estos dos tipos de transiciones porque no se observa ninguna diferencia entre ellas con respecto a los valores de F2. No mostraré en un gráfico los cambios que sufre el F2 en las transiciones con las africadas sordas [ts] y [tʃ] ya que los resultados son similares a los que se ilustran en la Figura 4-10 y la Figura 4-11 para las sibilantes [s] y [ʃ].

Tabla 4-12: Valores de F2 en la transición de formantes de las africadas sordas (hombres)

	[ts]	[tʃ]
[i] - 2244	1952 Hz ⁽¹⁸⁶⁾	2116 Hz ⁽¹⁴²⁾
[a] - 1418	1396 Hz ⁽¹⁶³⁾	1800 Hz ⁽¹⁵³⁾
[ɪ] - 1539	147 Hz ⁽¹⁴⁰⁾	2013 Hz ⁽¹⁸⁸⁾
[ʊ] - 896	1121 Hz ⁽¹³⁵⁾	1605 Hz ⁽¹⁹⁹⁾

Tabla 4-13: Valores de F2 en la transición de formantes de las africadas sordas (mujeres)

	[ts]	[tʃ]
[i] - 2676	2253 Hz ⁽¹³⁵⁾	2494 Hz ⁽²²⁹⁾
[a] - 1597	1628 Hz ⁽¹⁸⁷⁾	2128 Hz ⁽¹⁶⁹⁾
[ɪ] - 1704	1607 Hz ⁽¹³⁶⁾	2358 Hz ⁽²⁴⁴⁾
[ʊ] - 935	1306 Hz ⁽¹⁶⁹⁾	2022 Hz ⁽²⁰⁶⁾

4.4 LA AFRICADA SONORA BILABIAL [b^β]

Shipibo tiene una consonante bilabial sonora que ha sido tradicionalmente descrita como una fricativa bilabial sonora (Loriot, Lauriault *et al.* 1993). Este segmento puede realizarse de formas muy variadas. Puede ocurrir con dos fases, una primera fase de cerrazón y una segunda fase de fricción. La segunda fase, aunque puede mostrar fricción, ésta no es obligatoria. Cuando no hay fricción la segunda fase se realiza como una aproximante, [β]. Además, cuando su fase de cerrazón se acorta, sobre todo en contextos intervocálicos, el segmento entero puede realizarse completamente como una aproximante. Sin embargo, la realización más común es que muestre algún grado de cerrazón antes de que se inicie su segunda fase. En este estudio voy a caracterizar a este segmento como una *africada* bilabial sonora, [b^β], debido a su habilidad de mostrar dos fases en su articulación y porque la segunda fase puede mostrar fricción. Dada la variabilidad de su segunda fase con respecto a si muestra fricción o no, he optado por representarla como un índice, [b^β], es decir, [b^β] en vez de [bβ].

Empecemos la discusión de esta consonante presentando algunos pares (casi-) mínimos que prueben su oposición al segmento [w], el segmento con más similaridad a [b^β] en relación con la sonoridad y la labialidad. Los datos en (14) muestran que el segmento [b^β] no es una variante de [w]. Se oponen para causar diferencias en el significado de las palabras.

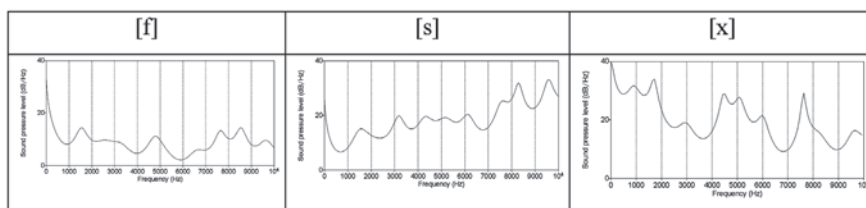
- (14) a. [^lwa.ka] ‘pez’ (o como préstamo del castellano ‘vaca’)
 b. [^lb^βa.ka] ‘suave’
 c. [wi.^lku.ti] ‘tambalearse’
 d. [b^βi.^lku] (esp. de árbol)

Ahora, la primera pregunta que hay que responder es por qué esta consonante no es una fricativa bilabial sonora como tradicionalmente ha sido descrita. Una consonante fricativa involucra la presencia de fricción (es decir, ruido, energía acústica en un patrón aleatorio) durante su producción. La fricción ocurre porque los articuladores involucrados en la producción de la fricativa se acercan lo suficiente como para crear un pasaje angosto. En el caso de consonantes sordas que tienen una fase de fricción (como las consonantes fricativas y africadas), tenemos una sola fuente de energía acústica; es decir, la fricción creada por el pase del aire a través de ese pasaje angosto en la cavidad oral.

La cantidad de turbulencia que crea este pasaje angosto y la distribución de la energía acústica de la turbulencia dependerán de qué tan estrecha es y dónde ocurre la obstrucción. La obstrucción por acanalamiento, como sucede por

ejemplo en la sibilante [s], concentra la energía acústica de la fricción del aire en las frecuencias más altas. La obstrucción velar, como en la fricativa [x], concentra la energía acústica en frecuencias más bajas (comparado a la obstrucción por acanalamiento) y la fricción laríngea, como en [h], tiende a concentrar la energía acústica en frecuencias aún más bajas. La razón por la cual las frecuencias bajan conforme la obstrucción se produce más atrás en la cavidad oral es debido a que el espacio en frente de la obstrucción se hace más grande y eso hace que el aire vibre a frecuencias más bajas. En el caso de la fricción labial, como en la fricativa [f], no hay una cavidad en frente de los labios, los articuladores que en este caso crean la obstrucción parcial. Esto causa que la energía acústica se distribuya más o menos uniformemente a lo largo de todas las frecuencias. Estos patrones de distribución de la energía acústica se pueden observar en la Figura 4-24 donde muestro los espectros de tres fricativas del castellano producidos por el autor.

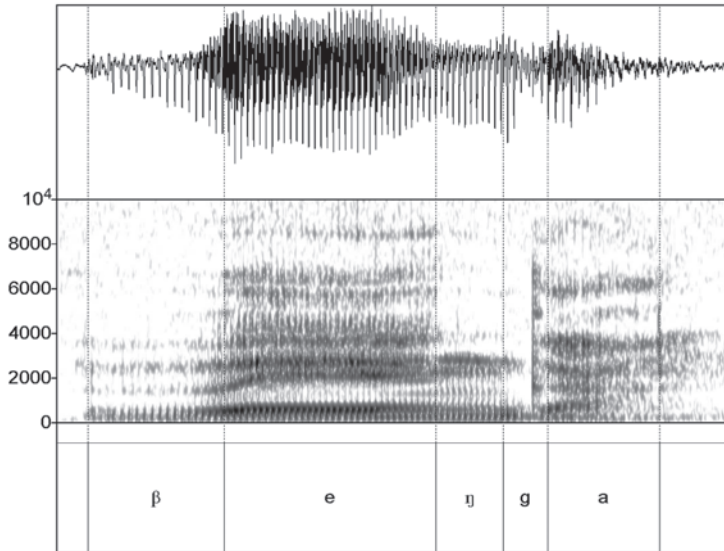
Figura 4-24: Espectros LPC de las fricativas del castellano



En el caso de las consonantes sonoras que tienen una fase de fricción (como las fricativas y africadas sonoras), las cosas se complican más. A diferencia de sus contrapartes sordas, éste tipo de segmentos poseen dos fuentes de energía acústica: las vibraciones de las cuerdas vocales y la fricción del aire al pasar por el pasaje estrecho creado por los articuladores. Las vibraciones periódicas de las cuerdas vocales durante la producción de la fase de fricción son responsables por la aparición de formantes. La posición de estos formantes durante la fase de fricción reflejan la forma de la cavidad oral y sobre todo al final de esta fase, los formantes reflejan los cambios que la forma de la cavidad va sufriendo conforme la lengua y los labios se acomodan para la vocal que sigue. En la Figura 4-25, muestro un ejemplo de una bilabial fricativa sonora [β] de la lengua africana venda (una lengua bantú). El oscilograma y espectrograma corresponden a la palabra venda [βɛŋga] ‘odiar’ (UCLA Phonetics Lab Archive).³ Nótese que es posible ver la fricción y la presencia de formantes durante la fricativa sono

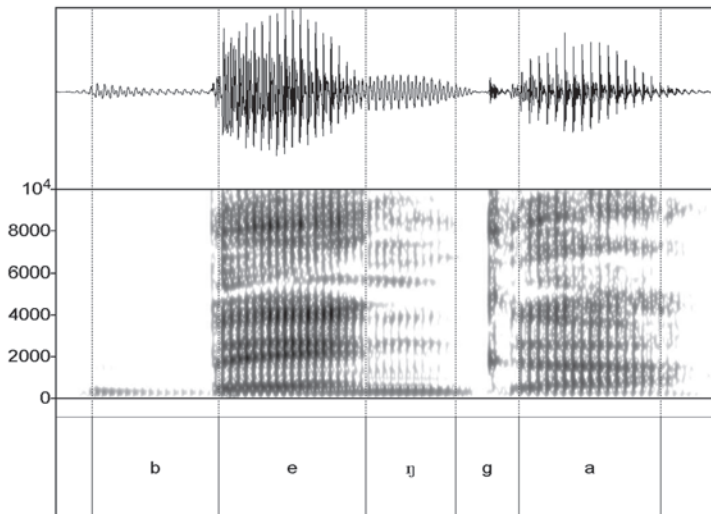
³ El archivo de audio usado para crear el oscilograma y espectrograma de la Figura 4-25 se obtuvo de *The UCLA Phonetics Lab Archive* (UCLA Department of Linguistics. <http://archive.phonetics.ucla.edu/>).

Figura 4-25: Fricativa bilabial sonora [β] de la lengua venda



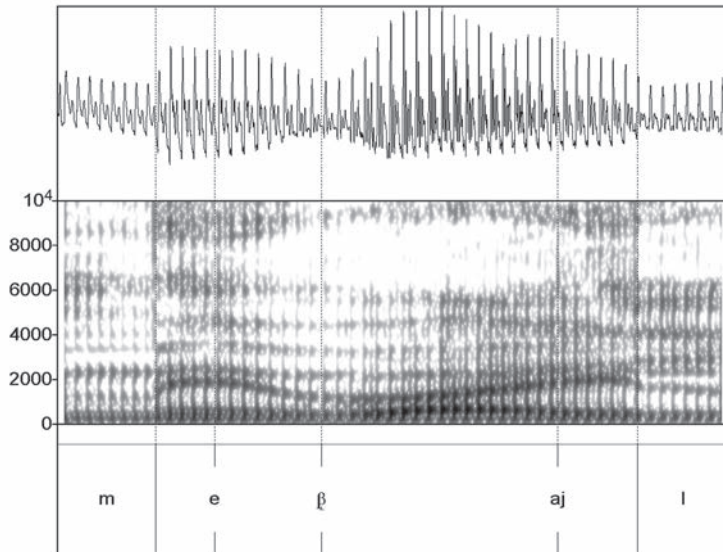
Ahora compárese la fricativa bilabial sonora, [β], de la Figura 4-25 con la oclusiva bilabial sonora, [b], de la palabra del castellano [benga] ‘venga’, pronunciada por el autor. En la oclusiva bilabial sonora, solo se observa la barra de voz en la parte inferior del espectrograma (véase la Figura 4-26).

Figura 4-26: Oclusiva bilabial sonora [b] del castellano



Antes de pasar a examinar el segmento [b^β] del shipibo, veamos otro tipo de segmento bilabial sonoro: la aproximante bilabial sonora, que se marca con el diacrítico [ɸ] debajo del símbolo [β] (es decir, [β^ɸ]). Éste es el segmento que usualmente ocurre en castellano en contextos intervocálicos (Hualde 2005). A diferencia de la consonante fricativa bilabial sonora [β] de la Figura 4-25, la aproximante bilabial sonora [β^ɸ] no muestra fricción (véase la Figura 4-27), y por supuesto, su estructura de formantes es mucho más fuerte que los formantes que se observan en su contraparte fricativa. De hecho, es tan fuerte que es difícil poder saber dónde empieza y dónde termina la consonante en relación con las vocales que le rodean. La [β^ɸ] que se muestra en la Figura 4-27 ocurrió dentro de la oración del castellano [xaj.me.βaj.la] ‘Jaime baila’.

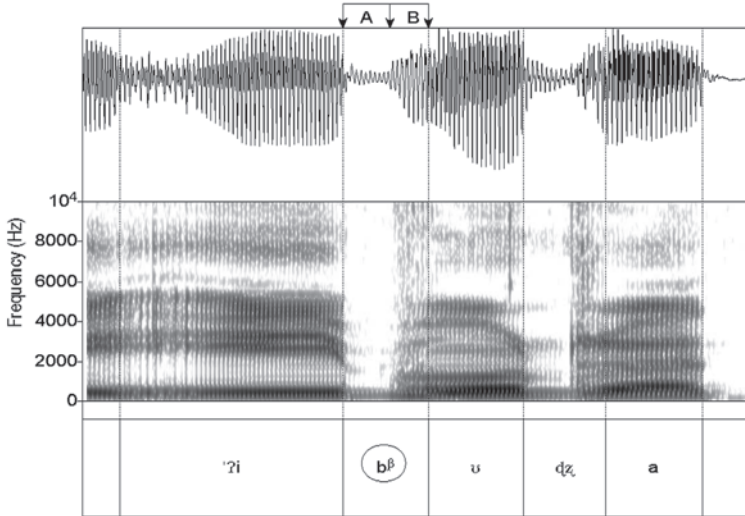
Figura 4-27: Aproximante bilabial sonora [β^ɸ] del castellano



Teniendo como referencia estos tres tipos de segmentos bilabiales sonoros, es decir, [b], [β] y [β^ɸ], veamos un primer ejemplo del segmento [b^β] del shipibo. El espectrograma en la Figura 4-28 muestra la ocurrencia de este segmento en la palabra [ʔi.b^βʊ.qza] ‘dueño’ (EV). Nótese la diferencia entre una fricativa bilabial sonora (Figura 4-25) y la africada bilabial sonora del shipibo. La consonante bilabial sonora del shipibo tiene un gesto de cerrazón (la sección señalada como “A”) que no posee una fricativa bilabial sonora. Tampoco es una consonante completamente oclusiva (como en la Figura 4-26) puesto que después de la fase de cerrazón sigue una segunda fase (la sección señalada como “B”). Esta segunda fase

puede mostrar fricción como en la Figura 4-28, donde se puede ver la fricción en el patrón aleatorio de la energía acústica encima de 4000 Hz. Véase qué diferente es el patrón de energía de la fase “B” con respecto al inicio de la siguiente vocal, en la cual la energía acústica no se distribuye aleatoriamente.

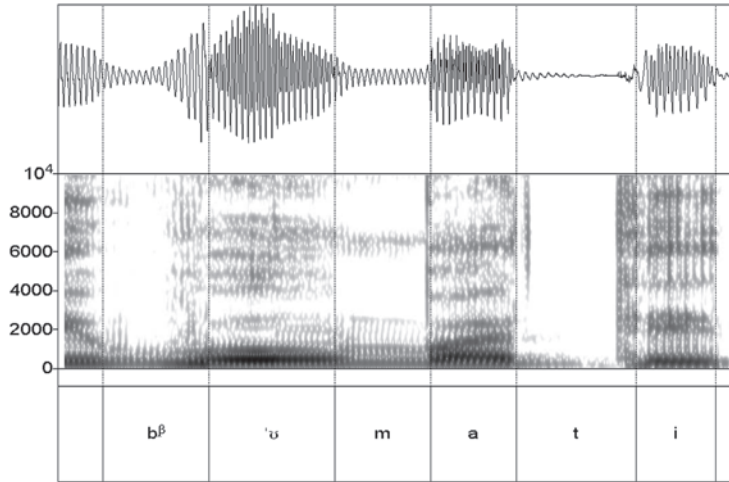
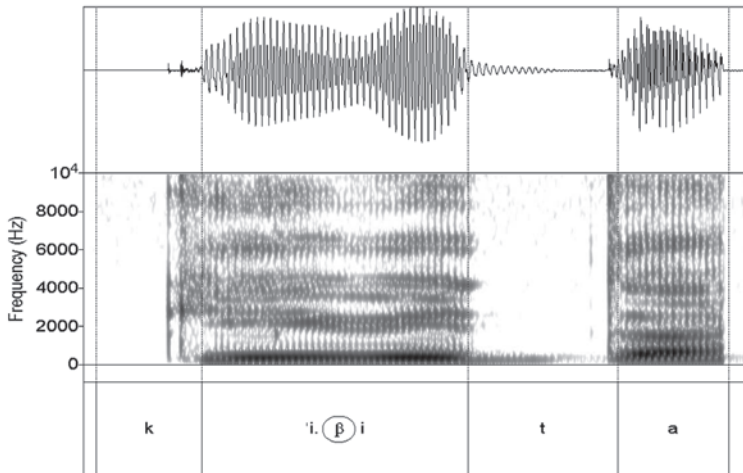
Figura 4-28: La africada [b^β] del shipibo en /ib^βu -dza / → [ʔi.b^βu.dza] ‘dueño’ (EV)



A diferencia de la Figura 4-28, en la cual [b^β] ocurre al interior de una palabra, la Figura 4-29 la muestra al inicio de la palabra [ʔb^βu.ma.ti] ‘encomienda’. Obsérvese que en el espectrograma de la Figura 4-29, la africada pertenece a la sílaba acentuada mientras que en el espectrograma de la Figura 4-28, es parte de la sílaba inacentuada. Esto muestra que la fase de fricción de esta obstruyente no se debe a un proceso de fortalecimiento causado por el acento.

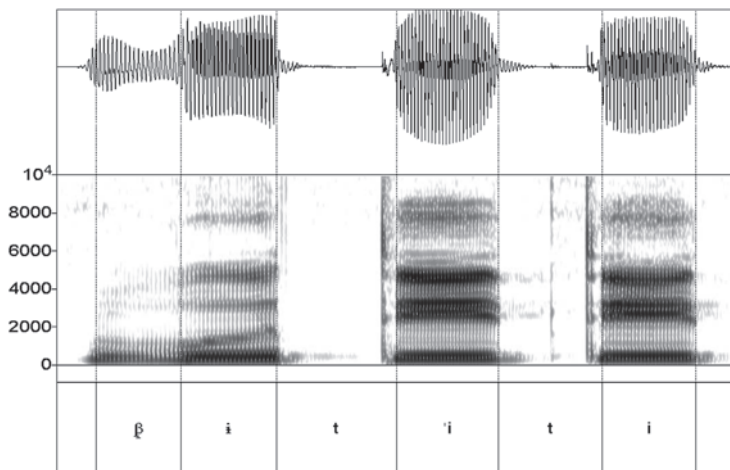
En contextos intervocálicos, la fase de cerrazón de [b^β] tiende a acortarse, aunque no obligatoriamente. Cuando esto pasa y dependiendo de que tanto la fase de cerrazón se acorte, el segmento podría realizarse como una aproximante bilabial, [β]. Esto se muestra en la Figura 4-30 (compáresele con la Figura 4-27). El espectrograma representa la palabra [ʔki.b^βi.ta] (realizado como [ʔki.βi.ta]) ‘quitarse uno (por ejemplo, una venda) del muslo’ (PP2).

La consonante [b^β] también puede realizarse como una aproximante al inicio de una palabra. Esto se muestra en la Figura 4-31 a través de la palabra [βi.ʔti.ti] (realizada como [βi.ʔti.ti]) ‘cerrarse los ojos (de uno mismo)’ (INTVZ, INF).

Figura 4-29: La africada [b^β] in ['b^βu.ma.ti] 'encomienda'Figura 4-30: La africada /b^β/ realizada como [β] en un contexto intervocálico

Aunque la realización de la africada bilabial [b^β] como aproximante es posible, no es la más común. Los hablantes de shipibo tienden a preservar el gesto de cerrazón aunque su implementación fonética no siempre llega a ser una cerrazón completa de los labios, especialmente en contextos intervocálicos. La Figura 4-32 y la Figura 4-33 muestran las mismas palabras que la Figura 4-30 y la Figura 4-31 pero con la forma más común en que se realiza la consonante [b^β]. A diferencia de los espectrogramas de la Figura 4-30 y la Figura 4-31, se puede ver esta vez

Figura 4-31: La africada /b^β/ realizada como [β]
en un contexto inicial de palabra



que [b^β] tiene dos fases. En ambos casos, la fase de cerrazón no es total, se puede observar la presencia de energía acústica en las frecuencias bajas por encima de 500 Hz. Sin embargo, hay una clara disminución de la intensidad (el espacio blanco) en comparación con la vocal anterior que tomo como evidencia de la implementación fonética de esta fase.⁴ Nótese que en ambos casos el segmento [b^β] ocurre entre dos vocales. Si la consonante fuese una fricativa (como en la Figura 4-25), es difícil poder explicar por qué la primera mitad de su articulación carece de fricción. Si solo fuese una aproximante (como en la Figura 4-27), entonces sería un misterio porque hay una caída consistente de la intensidad en la primera parte de su producción.

¿En lugar de que este sonido se le considere africado podría ser una oclusiva bilabial sonora? Si la consonante que en este estudio represento con el símbolo [b^β] fuese una oclusiva [b], sería una oclusiva bilabial diferente a la [b] del castellano o del inglés. Primero, nótese que en las lenguas que tienen el segmento oclusivo [b], el inicio de la siguiente vocal es *casi* inmediato, abrupto (véase la oclusiva bilabial sonora del castellano en la Figura 4-26). No hay una etapa de transición gradual de alrededor de 45 ms en que los labios se van abriendo poco a poco, como se da en shipibo. Es decir, la oclusiva [b] no presenta la segunda fase que posee el segmento [b^β] del shipibo. Pero podría ser que el segmento bilabial

⁴ En la Figura 4-31, el segmento [b^β] tiene una vocal que le precede porque la palabra se produjo dentro de la frase del estudio y el hablante no hizo ninguna pausa antes de la palabra objetivo.

Figura 4-32: La africada /b^β/ realizada como [b^β] en un contexto intervocálico

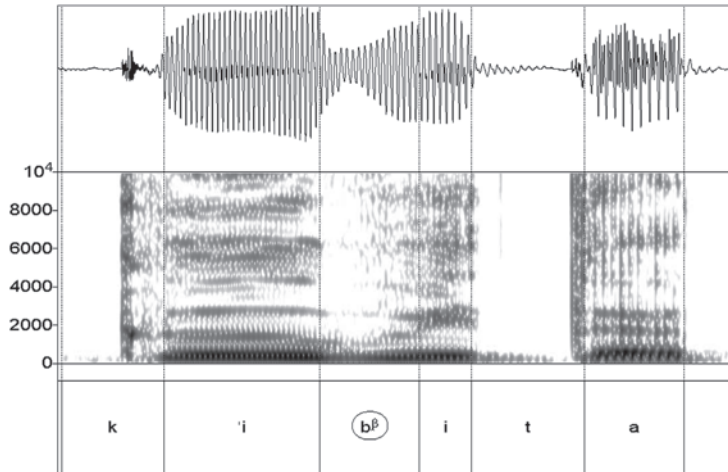
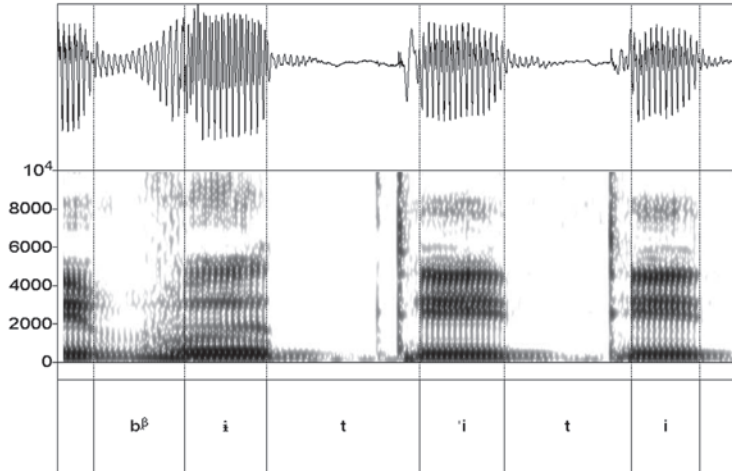


Figura 4-33: La africada /b^β/ realizada como [b^β] en un contexto inicial de palabra (I)

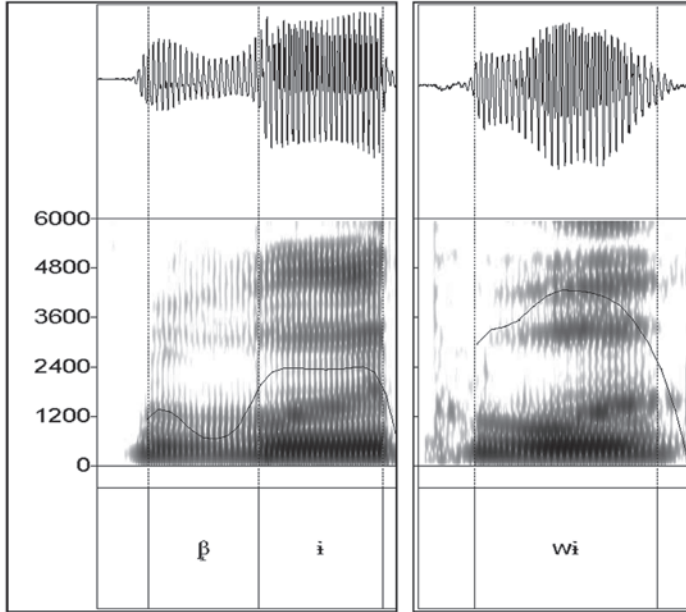
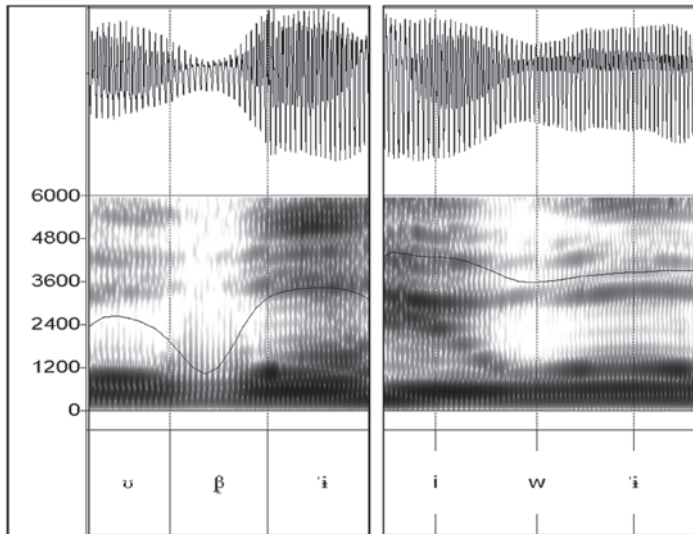


sonoro del shipibo sea una oclusiva diferente, cuya realización además incluye la sobreposición del gesto bilabial sobre el inicio de la siguiente vocal. El problema con este análisis es que esperaríamos que las vocales que sigan al segmento [b^β] posean algún grado de labialización, después de todo, la coarticulación consistiría en la sobreposición del gesto labial sobre el inicio de la siguiente vocal.

Sin embargo, esto no ocurre en shipibo (con la excepción de la vocal [i], la cual no solo se labializa en contacto son [b^β] sino también en contacto con [p] y [m]). La fase indicada como “B”, por ejemplo, en la Figura 4-28, solo pertenece al segmento consonántico [b^β], no a la vocal que le sigue.

Voy a terminar esta sección comparando [b^β] con el segmento [w]. Esta comparación es importante porque si el segmento [b^β] puede mostrar características de una aproximante cuando su fase de cerrazón se acorta y su segunda fase no muestra fricción (véase la Figura 4-30 y la Figura 4-31), entonces surge la pregunta de cómo diferenciarlo de la aproximante [w] con la cual comparte las propiedades de ser bilabial y sonoro. Además del gesto dorsal que pertenece exclusivamente a [w], una de las claves fonéticas más consistentes para diferenciar entre estos segmentos es la presencia de una caída notable de la intensidad seguida de una subida. Estos dos eventos producen un contorno característico de la intensidad en [b^β] en forma de “U”. El contorno de intensidad en “U” se da inclusive cuando el segmento [b^β] se realiza con todas las características de un segmento aproximante. Este patrón no se observa en el contorno de la intensidad de [w]. Para [w], la intensidad puede subir hasta alcanzar el nivel de intensidad de la siguiente vocal o puede mostrar un nivel de intensidad tan alto como el de las vocales circundantes. La Figura 4-34 compara [b^β], el cual se ha realizado como la aproximante [β] con el segmento [w]. En el primer caso, [b^β] ocurre dentro de la palabra [b^βi.¹ti.ti] ‘cerrarse los ojos (de uno mismo)’ (el espectrograma completo se mostró en la Figura 4-31). El segmento [w] ocurrió en la palabra [wi.¹ku.ti] ‘tambalearse’. El contorno de la intensidad aparece como una línea oscura sobrepuesta al espectrograma.

La Figura 4-35 compara nuevamente [b^β] con [w]. Como en el caso anterior, el segmento [b^β] se ha realizado como una aproximante, [β]. Sin embargo, esta vez ocurre en posición intervocálica al interior de la palabra. La consonante [b^β] corresponde al tercer segmento de la palabra [b^βu.¹b^βis.ti] ‘trepar hasta la cima (de algo)’. La semi-consonante [w] ocurre como el tercer segmento de la palabra [ʃi.¹wi.wi] (esp. de garza). Como se puede apreciar, el contorno de intensidad de ambas consonantes es diferente. El segmento [b^β] muestra el contorno de intensidad característico en forma de “U”. Lo que causa la caída de la intensidad de [b^β] en la Figura 4-34 y la Figura 4-35 es el acercamiento de los labios tratando de implementar la fase de cerrazón de [b^β]. Sin embargo, en estos dos casos, los labios no logran acercarse lo suficiente como para bloquear momentáneamente la salida del aire dando como resultado un segmento aproximante, [β].

Figura 4-34: Contorno de intensidad de [b^β] y [w] en inicio de palabraFigura 4-35: Contorno de intensidad de [b^β] y [w] en posición intervocálica

4.4.1 La duración de [b^β]

La duración de la segunda fase de la consonante [b^β] es bastante variable y presenta diferentes grados de fuerza y de fricción, lo cual hace difícil medirla con exactitud. Esto a su turno, hace que las mediciones de su fase de cerrazón también sean difíciles de llevar a cabo. En algunas ocurrencias, debido a la sobreposición parcial de los gestos articulatorios dentro de este segmento, no es sencillo determinar con exactitud dónde acaba la fase de cerrazón y dónde empieza la fase de fricción o dónde termina la fase de fricción y dónde se inicia la vocal que sigue.

Con esos problemas en mente, la Tabla 4-14 muestra la duración promedio de la africada sonora bilabial. Los resultados se basan en la medición de alrededor de 1500 1050 ejemplares. Todas las [b^β] que se midieron ocurren como arranques ya sea de la primera o de la segunda sílaba de palabras trisilábicas. Los promedios de duración de las fases de cerrazón y fricción solo incluyen aquellos casos en los cuales no hubo ambigüedad al determinar las fronteras de dichas fases. Hubo casos en que no fue claro dónde exactamente se localizaban la frontera entre la fase de cerrazón y la de fricción pero era posible determinar dónde empezaba y terminaba la africada, como segmento consonántico. En esos casos, se incluyó en los análisis la duración de la africada pero no la duración individual de sus fases. Los resultados de la Tabla 4-14 indican que en promedio la africada bilabial sonora tiende a durar aproximadamente 100 ms en sílabas acentuadas pero solo 80 ms en sílabas inacentuadas. La tabla también indica que la fase de cerrazón de la africada bilabial tiende a ser igual o ligeramente más larga que la fase de fricción.

**Tabla 4-14: Duración promedio (en segundos)
de la africada bilabial sonora [b^β]**

	Fase de cerrazón	Fase de fricción	Duración total
[b ^β] – en sílaba acentuada	0,053	0,045	0,095 ^(0,016)
[b ^β] – en sílaba inacentuada	0,044	0,037	0,078 ^(0,015)

4.4.2 Análisis espectral de la fase de fricción de la africada bilabial [b^β]

La Tabla 4-15 muestra los promedios de la frecuencia del pico más alto en los espectros de la fase de fricción de la africada bilabial. Dada la variabilidad observada en la duración de la fase de fricción de [b^β], los espectros analizados corresponden a un punto tomado en el medio de dicha fase. Aunque la Tabla 4-15 presenta los promedios tomando en cuenta el efecto que podría tener la siguiente vocal

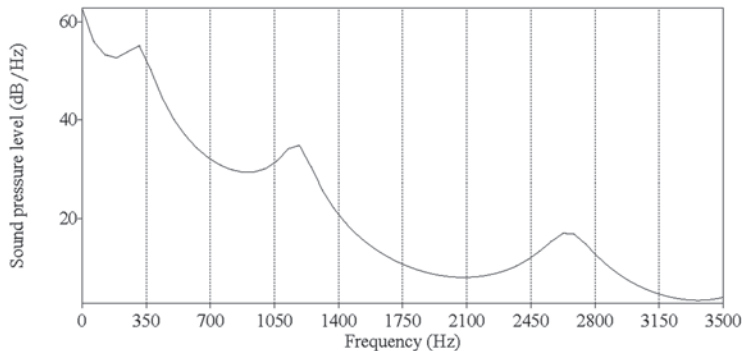
sobre la africada, se observa que el pico se localiza aproximadamente a 250 Hz sin importar el timbre de la vocal que seguía a la africada. Este es otra pieza de evidencia que indica que la segunda fase del segmento [b^β] no es una fase de coarticulación con la siguiente vocal.

Tabla 4-15: Promedio del pico más alto en el espectro de a fase de fricción de [b^β]

	[i]	[ī]	[a]	[u]
[b ^β] – hombres	247 Hz ⁽⁴⁸⁾	239 Hz ⁽⁴⁵⁾	250 Hz ⁽⁴⁵⁾	253 Hz ⁽⁴⁹⁾
[b ^β] – mujeres	261 Hz ⁽³⁸⁾	248 Hz ⁽³²⁾	262 Hz ⁽⁴⁸⁾	247 Hz ⁽³⁵⁾

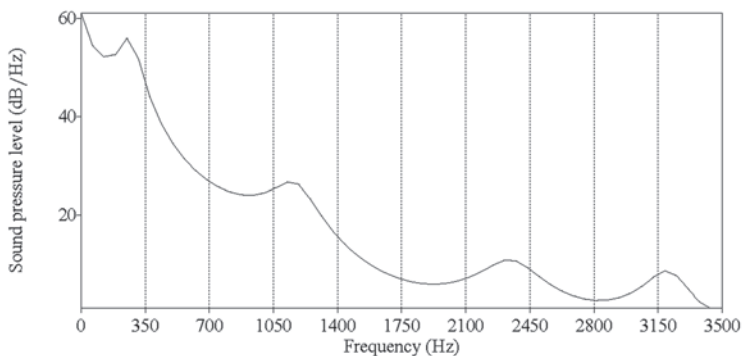
La Figura 4-36 y la Figura 4-37 muestran dos espectros LPC de la fase de fricción de [b^β] en la palabra [ˈʎi.b^βu.ɔza] ‘dueño’ (EV). Estos espectros corresponden a la pronunciación de una hablante mujer y de un hablante hombre, respectivamente. Además de la posición del pico más alto, el espectro de la fase de fricción de la africada bilabial se caracteriza por tener la energía acústica concentrada en las frecuencias más bajas, usualmente debajo de 1500 Hz. La energía acústica rápidamente desciende conforme las frecuencias suben.

Figura 4-36: Espectro LPC de la fase de fricción de [b^β] en la palabra [ˈʎi.b^βu.ɔza] (mujer)



El lector puede preguntarse a este punto porque el segmento [b^β] no muestra un espectro llano como lo hace la consonante fricativa labiodental del castellano (véase la Figura 4-24 y la discusión que antecede a esa figura). La razón es que el espectro que obtenemos de las fricativas y africadas labiales sonoras es bastante

Figura 4-37: Espectro LPC de la fase de fricción de [b^β] en [ʔi.b^βu.dza] (hombre)



diferente del espectro de sus contrapartes sordas. En el caso de las fricativas y africadas labiales sonoras, además de la distribución uniforme de la energía acústica que crea la fricción labial, tenemos la energía que viene de las vibraciones periódicas de las cuerdas vocales filtradas por la forma de la cavidad oral. Es la forma de esta cavidad la que concentra la energía acústica en las frecuencias bajas debajo de los 3500 Hz.

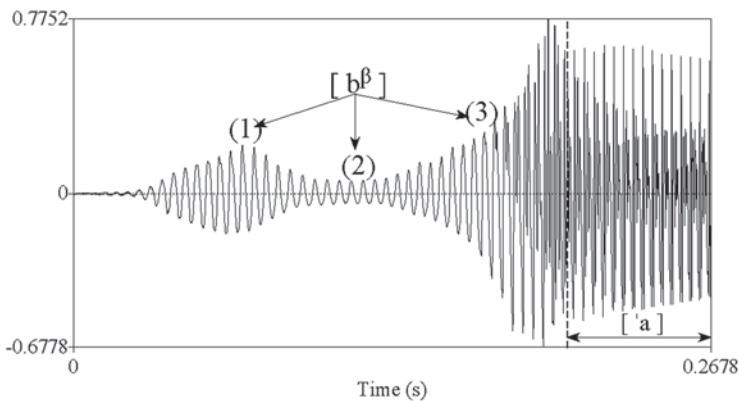
4.4.3 La africada bilabial sonora en contextos después de pausa

La africada bilabial sonora muestra una realización inusual cuando ocurre después de una pausa, un contexto que puede también describirse como de frase inicial. En este contexto, el segmento [b^β] tiende a mostrar un movimiento labial complejo que puede describirse que consiste en tres fases: (i) primero, los labios empiezan ligeramente abiertos y el aire fluye a través de la boca; (ii) luego, los labios se juntan y crean la fase de cerrazón de la africada, y (iii) finalmente, los labios se separan para empezar la segunda fase, la cual puede contener fricción. Aunque este alófono de la africada bilabial no es obligatorio, se le encuentra regularmente en algunos hablantes en sílabas acentuadas. Las tres fases de este alófono pueden describirse por medio de dos gestos articulatorios, uno ocurre dentro del otro: un gesto de ‘continuo’ (el aire sale por la boca) y un gesto de ‘cerrazón’ (el aire deja de fluir por la boca). El gesto de ‘cerrazón’ empieza ligeramente después del gesto de ‘continuo’ y termina antes de que termine el gesto de ‘continuo’. De este modo, el gesto de ‘cerrazón’ se puede interpretar como una breve interrupción del flujo de aire durante la articulación de la africada bilabial.

Los cambios en la salida del aire por la boca descritos arriba se pueden observar fácilmente en el oscilograma de la Figura 4-38. El oscilograma corresponde a la

parte inicial de la palabra [l^βa.wa] ‘loro’, pronunciada en aislamiento. La Figura 4-38 solo muestra el oscilograma de la africada [b^β] y una porción de la siguiente vocal acentuada [a]. Las tres fases del segmento africado aparecen numeradas e indicadas por flechas. Obsérvese que el sonido empieza con cierta cantidad de aire saliendo por la boca, y luego, conforme los labios se cierran en la segunda fase, la amplitud de la onda de sonido disminuye. Finalmente, durante la tercera fase, conforme los labios se separan nuevamente, la amplitud se incrementa notablemente.

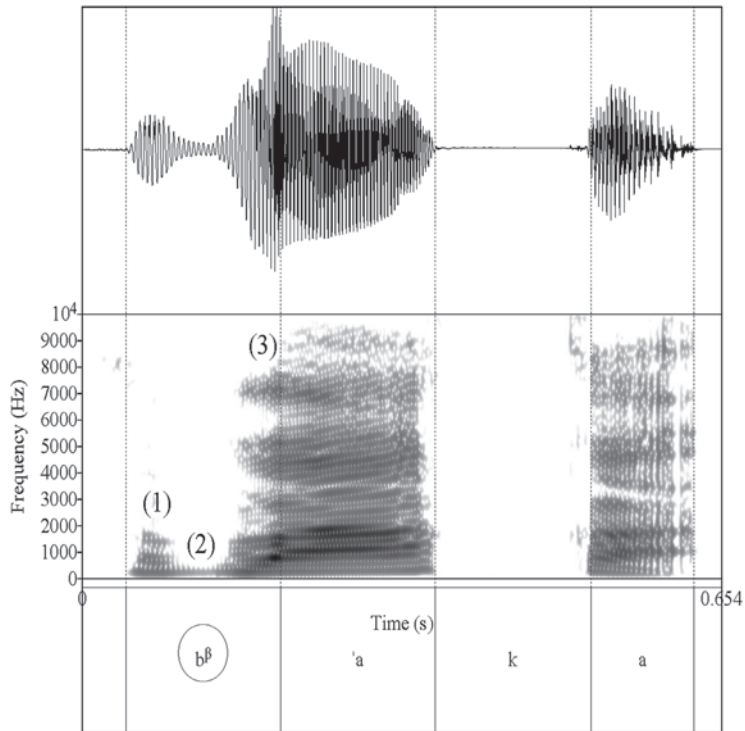
Figura 4-38: Oscilograma de la africada [b^β] después de pausa



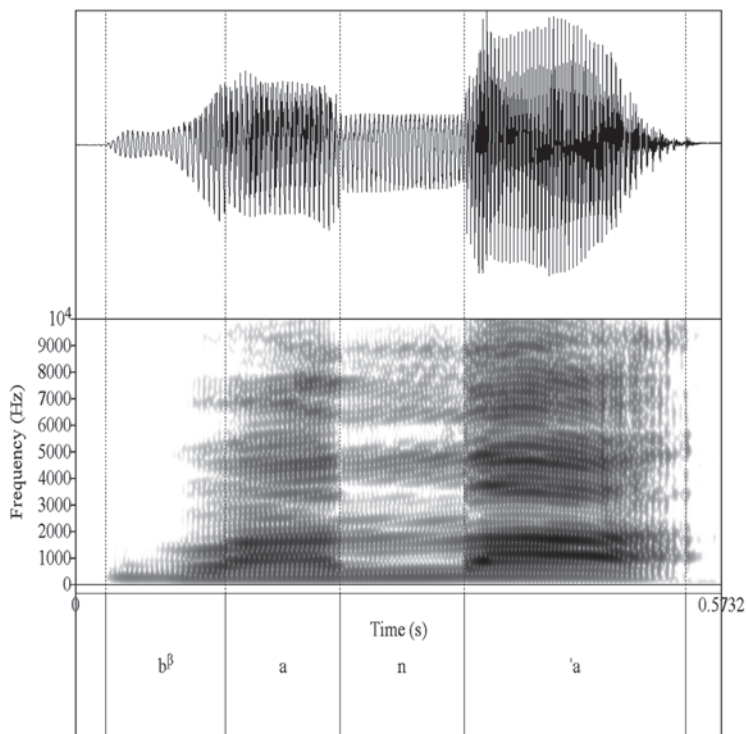
El espectrograma de la Figura 4-39 muestra el oscilograma y espectrograma de la palabra [l^βa.ka] ‘suave’ pronunciada en un contexto inicial de frase (después de una pausa). El espectrograma también nos deja ver las tres fases de la consonante [b^β] en este contexto desde la perspectiva de cómo se distribuye la energía acústica a lo largo de las frecuencias. Las fases aparecen numeradas como en la Figura 4-38. La mayor parte de la energía acústica de la primera fase de la africada se concentra usualmente debajo de 2000 Hz.

Para mis oídos, como hablante nativo de castellano, esta fase inicial es casi imperceptible. Cuando el audio de esta fase inicial se escucha repetidamente en forma aislada, uno obtiene la impresión auditiva de un sonido parecido a una vocal bastante reducida y casi tan breve como una vocal schwa en inglés. En promedio, esta fase inicial no dura más de 30 ms. Sin embargo, tan minúsculo como parece su duración, está presente en casi todos los hablantes y su aparición, aunque no es obligatoria, está condicionada por la presencia de una pausa al inicio de una frase.

Figura 4-39: Oscilograma y espectrograma de [b^βa.ka] ‘suave’



En sílabas inacentuadas al inicio de una frase, la africada bilabial tiende a ocurrir sin la fase inicial que aparece en la Figura 4-39. El oscilograma y espectrograma de la Figura 4-40 ilustran esto a través de la palabra [b^βa.'na] ‘planta’ pronunciada después de una pausa. En esta palabra, la africada bilabial pertenece a una sílaba inacentuada.

Figura 4-40: Oscilograma y espectrograma de [b^βa.¹na] ‘planta’

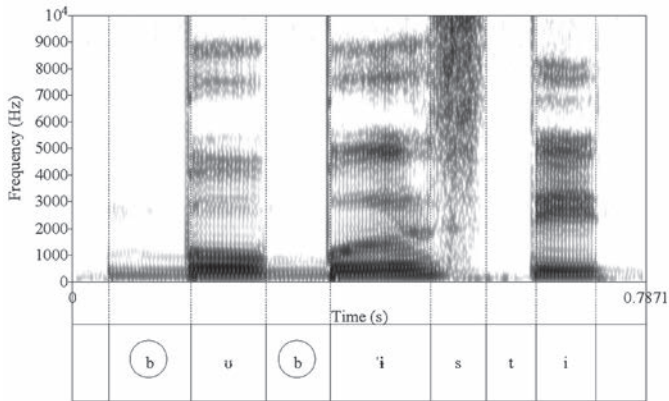
4.4.4 [b^β] y [b]: Variabilidad entre hablantes

Algunos hablantes de shipibo no tienen la africada bilabial sonora, [b^β]. En lugar de esta consonante, tienen una *oclusiva* bilabial sonora: [b]. Esto no parece deberse a una diferencia dialectal. Inclusive dentro de una misma familia, uno puede encontrar hablantes que tienen la oclusiva bilabial [b] en lugar de la africada bilabial [b^β]. De un total de 30 hablantes de shipibo, escogidos al azar, todos mayores de 35 años, y entrevistados entre el año 2000 y el 2008, 8 de ellos tienen consistentemente una oclusiva bilabial sonora en todos los ambientes donde se espera la africada bilabial sonora. De este modo, hay dos tipos de hablantes de shipibo: aquellos que tienen la africada bilabial [b^β] descrita arriba, y para quienes, este segmento puede alternar libremente con la aproximante [β], sobre todo en contextos intervocálicos, y aquellos que tienen la oclusiva bilabial [b] en vez de una consonante africada.

La Figura 4-41 muestra el espectrograma de la palabra [bu.¹bis.ti] ‘trepar a la cima de algo (por ejemplo, un árbol)’, pronunciada por un hablante de shipibo que tiene la oclusiva bilabial sonora en lugar de la africada bilabial. Obsérvese

que tanto la consonante bilabial que ocurre al inicio de la palabra como la que ocurre en posición intervocálica no muestran ninguna fase de fricción. Ambas consonantes son sin lugar a dudas oclusivas sonoras, como lo evidencia la presencia de la barra de voz en la parte inferior de las secciones correspondientes del espectrograma.

Figura 4-41: Espectrograma de [bu.'bis.ti] ‘trepar a la cima de algo’



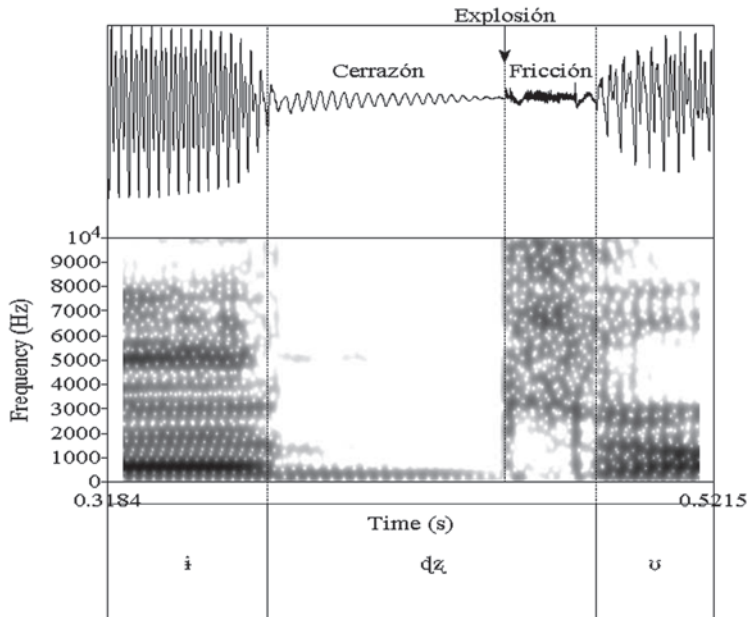
Para aquellos hablantes que tienen la oclusiva bilabial sonora, como aquel que produjo la palabra de la Figura 4-41, la africada ha perdido su fase de fricción. La han reinterpretado como una simple oclusiva. Esto provee evidencia adicional para analizar fonológicamente la bilabial sonora como una africada para el resto de los hablantes de shipibo y no como una bilabial fricativa, como se ha asumido tradicionalmente. Puesto en otras palabras, el sonido es evidentemente percibido por los hablantes nativos de shipibo como que posee un gesto articulatorio de cerrazón, el cual bloquea la salida del aire. Para algunos, éste es el único gesto que importa de modo que han filtrado la fase de fricción original de la africada.

Este fenómeno no parece estar relacionado al grado de bilingüismo shipibo-castellano que la persona pueda tener. En primer lugar, los hablantes pueden tener la oclusiva bilabial sin importar su dominio del castellano. Segundo, en castellano peruano, la oclusiva bilabial sonora se comporta diferente. Siempre se realiza como una aproximante bilabial en contextos intervocálicos (por ejemplo, [ˈlo.βo] ‘lobo’). En contraste, los hablantes de shipibo que poseen la oclusiva bilabial sonora también la tienen en ambientes intervocálicos. Para algunos hablantes de castellano peruano, inclusive la oclusiva bilabial sonora libremente alterna con la aproximante bilabial en posición inicial de palabra (por ejemplo, [ˈbo.ka] ~ [ˈβo.ka] ‘boca’). Éste no es el caso de la oclusiva bilabial sonora del shipibo. Los hablantes la usan consistentemente en todos los contextos.

4.5 LA AFRICADA RETROFLEJA SONORA

Shipibo tiene otra africada sonora, la africada retrofleja: [dʒ]. Como otras africadas en shipibo, [dʒ] solo puede ocurrir en arranque de sílaba. Este segmento ha sido uno de los más difíciles de identificar. Tradicionalmente se le ha descrito como un tipo de vibrante simple, lo cual explica por qué se le representa ortográficamente por la letra 'r'. García 1994 fue el primero en identificar su naturaleza retrofleja pero la clasificó como una fricativa. Elias-Ulloa 2000a lo clasificó como una africada retrofleja y Valenzuela, Márquez-Pinedo *et al.* 2001 presentaron un primer acercamiento a las complejidades de su realización fonética. En este estudio, la africada retrofleja sonora se representa por el símbolo fonético [dʒ]. La Figura 4-42 muestra el oscilograma y espectrograma de esta africada ocurriendo dentro de la palabra [ˈbʲi.dʒu] 'ojo' pronunciada en aislamiento. Tanto el oscilograma como el espectrograma solo muestran la parte final de la vocal [i], la africada y el inicio de la vocal [u]. En la sección de la africada, la Figura 4-42 indica los tres hitos de este sonido consonántico: (i) la fase de cerrazón, que es sonora (véase las fluctuaciones en la presión del aire en el oscilograma así como también la barra de voz en el espectrograma indicada por la presencia de energía acústica dentro del rango de 0-500 Hz), (ii) la fase de explosión, y (iii) la fase de fricción, que muestra la energía acústica aproximadamente a 3000 Hz para arriba.

Figura 4-42: Oscilograma y espectrograma de la africada [dʒ]



La duración promedio de la africada retrofleja sonora aparece en la Tabla 4-16. Los resultados se basan en las mediciones de 900 ejemplares de esta africada. Todas ocurren como el arranque de la primera o segunda sílaba de palabras trisilábicas. Como es usual, todas las palabras se pronunciaron dentro de la frase usada en este estudio (véase la sección 1.3.1). La africada retrofleja sonora tiende a durar aproximadamente 80 ms en sílabas acentuadas y 60 ms en sílabas inacentuadas. Como en el caso de las africadas sordas, la posición prosódica de la sílaba que contiene la africada retrofleja no tiene ningún efecto notable en la duración de su fase de fricción. Ésta tiende a durar alrededor de 30 ms. En contraste, la fase de cerrazón de [dʒ] sí muestra efectos de duración. En sílabas acentuadas, dura aproximadamente 50 ms pero como arranque de sílabas inacentuadas, dura alrededor de 30 ms.

**Tabla 4-16: Promedio de duración (en segundos)
de la africada retrofleja sonora [dʒ]**

	Fase de cerrazón	Fase de fricción	Duración total
[dʒ] – en sílaba acentuada	0,047 ^(0,013)	0,032 ^(0,011)	0,079 ^(0,016)
[dʒ] – en sílaba inacentuada	0,032 ^(0,009)	0,031 ^(0,010)	0,063 ^(0,018)

El espectrograma de la Figura 4-43 nos permite apreciar un ejemplo de cómo la duración de la fase de cerrazón de [dʒ] se ve afectada por la presencia o ausencia del acento. El espectrograma muestra la palabra [ˈdʒi.dʒa.ti] ‘cortar a hachazos’. Esta palabra contiene dos africadas retroflejas sonoras. Una ocurre al inicio de la palabra como el arranque de la sílaba acentuada y la otra ocurre en un ambiente intervocálico como el arranque de una sílaba inacentuada. La duración de las fases de cerrazón así como también la duración de sus fases de fricción se indican en la parte superior del espectrograma de cada africada.

La Figura 4-44 muestra que la fase de cerrazón de la africada [dʒ] tiene una duración más larga en sílabas acentuadas inclusive si esa sílaba no aparece al inicio de la palabra. El espectrograma de la Figura 4-44 ilustra la palabra [ki.ˈdʒi.ti] ‘cortar a hachazos (el borde de un objeto)’ (esta palabra está formada por el prefijo /ki-/ ‘en el labio, en el borde’, el sufijo /-ti/ (INF), y la raíz verbal, mostrada en la Figura 4-43, /dʒi.dʒa/ ‘cortar a hachazos’). La africada retrofleja sonora [dʒ] ocurre en un contexto intervocálico como el arranque de la segunda sílaba. Obsérvese que mientras la duración de su fase de fricción es alrededor de 30 ms, su fase de cerrazón es 28 ms más larga que la fase de cerrazón en la [dʒ]

intervocálica de la Figura 4-43. La diferencia entre ambas [dʒ] intervocálicas es que la primera pertenece a una sílaba inacentuada mientras que la segunda, pertenece a una acentuada.

Figura 4-43: Espectrograma de la palabra ['dʒi.dʒa.ti]
'cortar a hachazos'

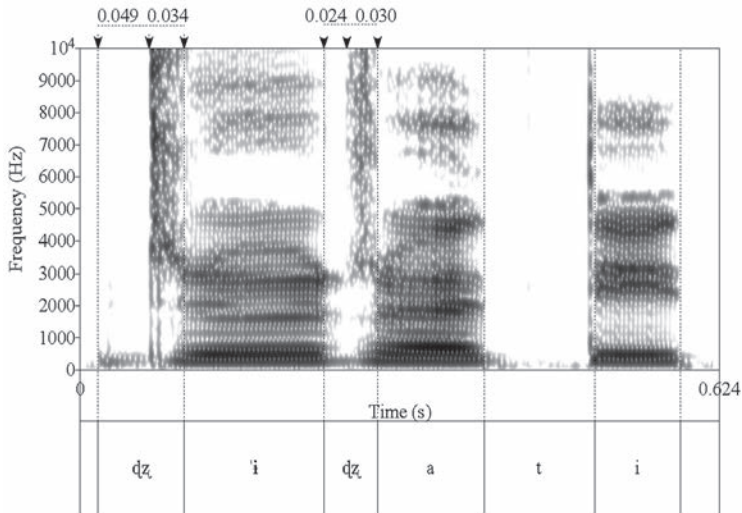
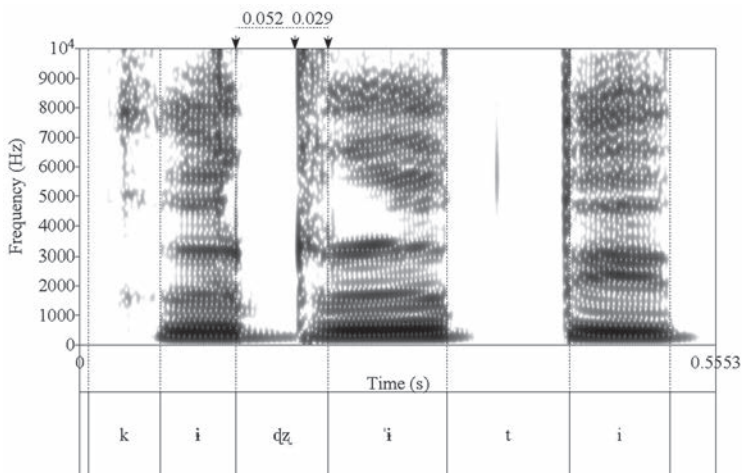


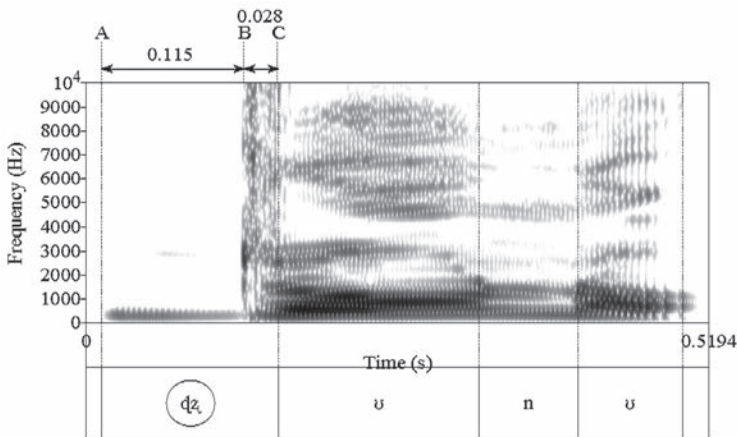
Figura 4-44: Espectrograma de [ki.'dʒi.ti] 'cortar a hachazos
(el borde de un objeto)'



La duración de la fase de cerrazón sonora de la africada [dʒ] también se ve afectada por otros factores, lo cual explica su considerable variación. Además de la posición del acento, la duración de la fase de cerrazón de [dʒ] puede verse afectada por si el segmento pertenece a una palabra pronunciada en aislamiento o dentro de una frase, o por el número de sílabas de la palabra. Sin embargo, a pesar de toda esa variación que muestra su fase de cerrazón, la fase de fricción de la africada [dʒ] se queda constante a 30 ms en promedio.

La Figura 4-45 y la Figura 4-46 muestran a la africada [dʒ] ocurriendo tanto al inicio de la palabra [ˈdʒʊ.nʊ] ‘serpiente’ y entre vocales en la palabra [ˈbʲi.dʒʊ] ‘ojo’, respectivamente. Los espectrogramas muestran la duración de las fases de cerrazón y fricción de [dʒ] cuando pertenece a palabras bisilábicas pronunciadas en aislamiento. En el primer caso, la africada aparece como el arranque de una sílaba acentuada. En el segundo caso, aparece como el arranque de una sílaba inacentuada. Cuando la africada retrofleja sonora ocurre en una palabra pronunciada en aislamiento, la duración de su fase de cerrazón varía entre 65 ms y 155 ms. En promedio, mide 90 ms en una sílaba acentuada y 70 ms en una sílaba inacentuada.

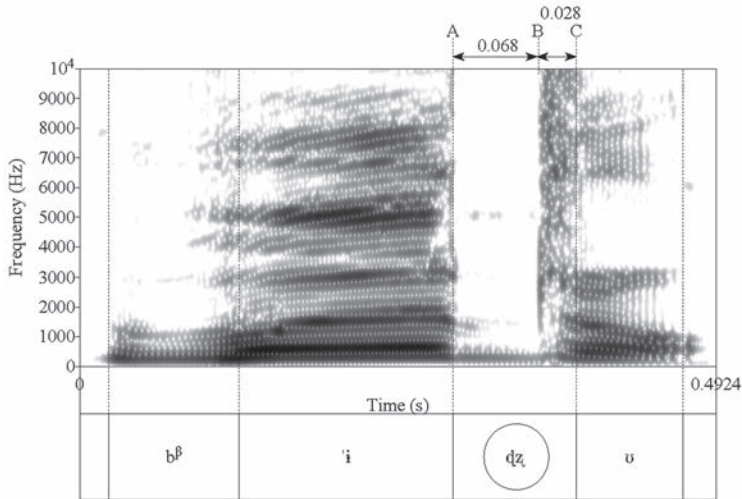
Figura 4-45: Espectrograma de [ˈdʒʊ.nʊ] ‘serpiente’,
(pronunciada en aislamiento)



La africada [dʒ] tiene un alófono [z], una sibilante retrofleja sonora. Tiende a ocurrir en el habla rápida en vez de [dʒ], y particularmente en contextos intervocálicos después de una sílaba acentuada. Además, en posiciones intervocálicas lejos del acento, la africada retrofleja sonora tiende a debilitarse aún más hasta el punto de obtener casi un estatus de aproximante. Esto es paralelo a lo que se

observó para la africada bilabial sonora (véase la Figura 4-30, sección 4.4). El espectrograma de la Figura 4-47 muestra ejemplos de la realización de los alófonos de /dʒ/. Tanto el alófono de la sibilante retrofleja como de la aproximante se representan por el símbolo [z] en la Figura 4-47.

Figura 4-46: Espectrograma de ['b^βi.dʒu] 'ojo', (pronunciada en aislamiento)

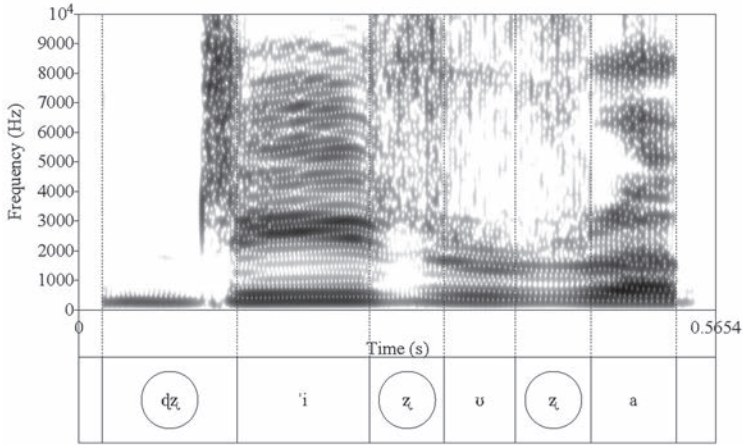


El espectrograma representa una muestra de la realización del morfema /dʒi.dʒu/ (esp. de mono) acompañado por el sufijo /-dʒa/ (EV). La palabra se pronunció al inicio de una frase como parte de una oración. Obsérvese que la [dʒ] del inicio de la palabra pertenece a una sílaba acentuada mientras que las otras dos ocurrencias de la africada retrofleja ocupan sílabas débiles. En esta posición, la africada tiene claramente todas sus fases características. Sin embargo, la africada retrofleja que aparece inmediatamente después de la sílaba acentuada se realiza como una sibilante retrofleja sin una fase de cerrazón evidente, y la que ocurre en la última sílaba muestra señales de haberse debilitado aún más, es casi un sonido aproximante retroflejo.

4.5.1 Análisis espectral de la fase de fricción de la africada [dʒ]

La Tabla 4-17 presenta la posición promedio del pico más alto de la africada [dʒ] durante su fase de fricción. Debido a la corta duración de esta fase, los análisis espectrales se tomaron de un punto centrado en el medio de la fase de fricción. Ya que la africada retrofleja sonora solo puede ocurrir como arranque de sílaba,

Figura 4-47: /dʒidʒu -dʒa/ → [ˈdʒi.zu.za] (esp. de mono - EV)



la primera columna de la tabla muestra la vocal que sigue al periodo de fricción. Los resultados se muestran tanto para hablantes hombres (segunda columna) como para mujeres (tercera columna) y se basan en el mismo conjunto de datos usados para determinar la duración de [dʒ] en la sección anterior.

Las posiciones promedio del pico mostradas en la Tabla 4-17 son muy similares a aquellos que se encontró para la sibilante retrofleja [ʃ], en particular, aquellos que corresponden a las hablantes mujeres (véase la Tabla 4-3, sección 4.1.1). La excepción en este caso es la posición del pico en la fase de fricción de [dʒ] cuando ocurre seguida de la vocal [i]. El pico se localiza aproximadamente a 3302 Hz para las mujeres cuando la [i] sigue a la sibilante [ʃ], pero tiende a centrarse en 3952 Hz cuando la [i] sigue a la africada [dʒ]. La diferencia, sin embargo, parece deberse a los pocos ejemplares disponibles para analizar el primer caso. Como se indicó en la sección 4.1.1, hay poco casos de la ocurrencia de [ʃ] en contacto con [i] en shipibo. Sin embargo, la secuencia [dʒi] no es para nada rara. Hay muchas palabras que poseen esa secuencia inclusive al inicio de las palabras.

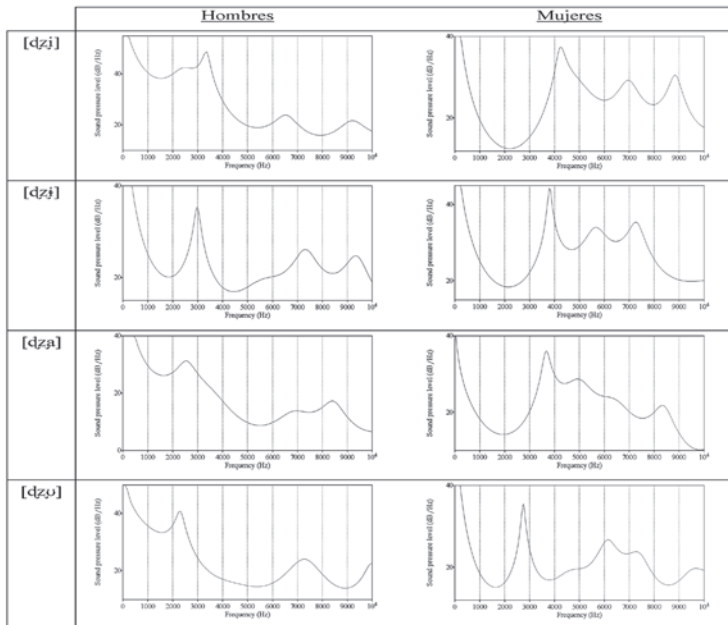
En el caso de los hablantes hombres, los promedios presentados en la Tabla 4-17 son ligeramente más bajos que los que se encontraron para la sibilante retrofleja [ʃ] (véase la Tabla 4-2). Sin embargo, estos promedios aún son manifiestamente mucho más bajos que las frecuencias que muestra la fase de fricción de las africadas [ts] y [tʃ] (véase la sección 4.3.1, Tabla 4-10 y Tabla 4-11), lo cual indica, como se espera para una retrofleja, que el punto de articulación de [dʒ] es mucho más posterior dentro de la cavidad oral que para [tʃ].

Tabla 4-17: El pico más alto en el espectro de la fase de fricción de [dz̥]

	[dz̥] - hombres	[dz̥] - mujeres
[i]	3477 Hz ⁽⁴⁷⁴⁾	3952 Hz ⁽⁴⁰⁴⁾
[í]	3005 Hz ⁽⁴⁵⁷⁾	3471 Hz ⁽³¹⁰⁾
[a]	2904 Hz ⁽⁴⁶³⁾	3533 Hz ⁽³⁰⁶⁾
[u]	2523 Hz ⁽³⁸⁷⁾	2954 Hz ⁽²⁹⁶⁾

La Figura 4-48 ilustra ejemplos de los espectros de la fase de fricción de la africada [dz̥]. Como en el caso de la sibilante retrofleja [ʃ] (véase los espectros en la Figura 4-9), el primer pico de los espectros es el más alto y agudo. Notablemente domina en altura a todos los otros picos que aparecen después de él. Contrástese esto con la forma del espectro de la fase de fricción de [tʃ] (véase la Figura 4-23), donde el pico más alto no es necesariamente el primero y además se puede observar la presencia de otros picos menores prominentes.

Figura 4-48: Espectros LPC de la fase de fricción de [dz̥] en diferentes contextos vocálicos



4.5.2 Transición de formantes y la africada retrofleja sonora [dʒ̠]

Tanto la Tabla 4-18 como la Tabla 4-19 presentan los promedios de F2 de las vocales cuando aparecen adyacentes a una africada retrofleja [dʒ̠]. La primera los presenta para los hablantes hombres y la segunda, para las hablantes mujeres. La primera columna de cada tabla da el valor ideal de F2 en el medio de las vocales. Los resultados incluyen tanto las vocales que siguen como las que preceden a la africada y se basan en las africadas retroflejas sonoras que ocurren como arranques de sílabas iniciales de palabra o como arranques de segundas sílabas. Los resultados obtenidos son bastante similares a aquellos presentados para la sibilante retrofleja [ʂ̠] (véase la Tabla 4-4 y la Tabla 4-5 así como también la Figura 4-10 y la Figura 4-11 en la sección 4.1.2).

Tabla 4-18: Los promedios del formante F2 en las transiciones con [dʒ̠] (hombres)

	[dʒ̠]
[i] - 2244	2005 Hz ⁽¹⁴¹⁾
[a] - 1418	1593 Hz ⁽¹⁵⁴⁾
[i̠] - 1539	1627 Hz ⁽¹⁶⁹⁾
[u̠] - 896	1273 Hz ⁽¹⁴⁴⁾

Tabla 4-19: Los promedios del formante F2 en las transiciones con [dʒ̠] (mujeres)

	[dʒ̠]
[i] - 2676	2304 Hz ⁽²³¹⁾
[a] - 1597	1719 Hz ⁽¹⁶²⁾
[i̠] - 1704	1733 Hz ⁽¹⁹⁹⁾
[u̠] - 935	1206 Hz ⁽¹¹⁴⁾

4.6 RESUMEN

En este capítulo, he presentado las características acústicas de las consonantes fricativas y africadas del shipibo. Se han discutido estos sonidos en términos de su sonoridad, duración y de sus propiedades espectrales. Con respecto a las sibilantes [s, ʃ, ʂ̠], se encontró que como arranques de sílaba, duran aproximadamente 110 ms. Compárese esto con la duración más corta encontrada para estos mismos segmentos (véase el capítulo 7) cuando ocurren como codas, aproximadamente 70 ms. La posición del acento no parece afectar la duración de las sibilantes.

El análisis espectral de las sibilantes revela cómo la energía acústica se distribuye de manera diferente para cada una de ellas. Contra más hacia adelante en la cavidad oral se encuentre el punto de articulación de la sibilante, más altas son las frecuencias donde se concentra la energía acústica ya que la cavidad delante de la obstrucción es más pequeña. De este modo, la energía acústica para la sibilante [s] tiende a concentrarse en promedio en 7000 Hz para arriba mientras que para [ʃ], se concentra entre 4000 Hz y 4500 Hz. La energía acústica de la sibilante [ʒ] se concentra entre 3000 Hz y 3500 Hz.

Además de las características espectrales de las sibilantes, otra clave fonética que puede servir para distinguir sus puntos de articulación es la transición del formante F2. Sin embargo, la transición de F2 en vocales adyacentes a sibilantes no parece ser una clave tan fuerte como son las propiedades espectrales de estos segmentos. La transición de F2 es bastante buena para distinguir entre [ʃ] y las otras dos sibilantes pero no es muy eficiente cuando se quiere distinguir entre [s] y [ʒ].

En este capítulo también se discutió la duración promedio de la fricativa laríngea [h], la cual es bastante similar a la de las sibilantes en posición de arranque de sílaba: 100 ms (la fricativa laríngea en sílabas iniciales inacentuadas tiende a ser aproximadamente 10 ms más corta). Sus espectros muestran la presencia de un pico de energía localizado cerca del F2 de la siguiente vocal (debido a problemas con su segmentación, solo se incluyeron en los análisis espectrales mediciones de la [h] en posición inicial de palabra).

Además, se determinó la duración de las africadas sordas [ts] y [tʃ]. Como arranques de sílabas acentuadas, duran en promedio 130 ms. En sílabas inacentuadas, su duración se acorta a aproximadamente 110 ms. Sin embargo, la duración de sus fases articulatorias no se acortan proporcionalmente en estos casos. Mientras que la fase de fricción tiende a no acortarse sin importar si la africada está en una sílaba con acento o no (50 ms), la fase de cerrazón sí se acorta en sílabas inacentuadas. Dura alrededor de 80 ms en sílabas con acento pero solo 60 ms en sílabas sin acento. La estabilidad de la duración de la fase de fricción en las africadas con respecto a la posición del acento refleja el mismo patrón encontrado en la duración de las sibilantes. Dura virtualmente lo mismo en sílabas acentuadas e inacentuadas. Los espectros de la fase de fricción de las africadas y la transición del formante F2 juntas son las mejores claves acústicas que pueden distinguir entre ellas. Estos resultados son similares a los que se encontraron para las sibilantes [s] y [ʃ].

Tradicionalmente, se pensaba que el shipibo, como otras lenguas pano, tenía una consonante bilabial fricativa y una vibrante simple, representadas ortográficamente por las letras 'b' y 'r', respectivamente. Este capítulo ha demostrado que

estos segmentos realmente son las africadas [b^β] y [dʒ] y que poseen un rango complejo de realizaciones fonéticas descritas en detalle e ilustradas con la ayuda de espectrogramas.

En este capítulo también se halló que la duración promedio de la africada [b^β] es 100 ms cuando ocurre en sílabas acentuadas pero aproximadamente 80 ms en sílabas inacentuadas. Con respecto a sus características espectrales, su pico más alto de energía ocurre a baja frecuencias: aproximadamente a 250 Hz.

Además de sus manifestaciones como una africada bilabial sonora, que puede ocurrir al inicio de la palabra o en ambientes intervocálicos, la africada bilabial sonora [b^β] también puede realizarse como una fricativa o una aproximante (usualmente en ambientes intervocálicos). Tal vez el alófono más interesante que muestra ocurre en contextos al inicio de frases, después de pausa. En este ambiente, los gestos de ‘continuo’ y de ‘cerrazón’ de la africada se sobreponen de tal modo que el segmento se inicia con el aire fluyendo por la boca, luego sigue una cerrazón labial y finalmente, reaparece el flujo de aire a través de la boca. Otro hallazgo interesante con respecto a la africada bilabial [b^β] es que algunos hablantes la están reanalizando como una oclusiva bilabial sonora, [b]. Este segmento aparece en todos los contextos donde otros hablantes muestran la africada [b^β].

He presentado datos que muestran que la duración de la africada retrofleja sonora [dʒ] era aproximadamente 80 ms en sílabas acentuadas y 60 ms en sílabas inacentuadas. Esto la hace la africada más corta del shipibo. Las africadas sordas son las más largas. La africada [dʒ] muestra el mismo comportamiento que las africadas sordas con respecto a la fase de fricción. Se mantiene estable sin importar que la sílaba que la contenga sea acentuada o no: 30 ms. El acento tiende a afectar la duración de la fase de cerrazón. En promedio, dura 50 ms en sílabas acentuadas pero solo 30 ms en las inacentuadas. Las propiedades espectrales de la fase de fricción de la africada [dʒ] y la transición de F2 son similares a aquellas encontradas para la sibilante retrofleja [ʃ]. Esta africada también muestra una variación compleja de realizaciones fonéticas. Puede realizarse como [dʒ] en todos los contextos. Uso este hecho para argumentar a favor de su naturaleza subyacente africada. Además, en el habla rápida y particularmente en contextos intervocálicos, tiende a realizarse como una fricativa retrofleja sonora, [ʒ], o inclusive como una consonante casi aproximante, donde desaparece la mayor parte de su fase de fricción.

Hasta ahora en este estudio se ha caracterizado en términos acústicos las oclusivas, fricativas y africadas del shipibo. El siguiente capítulo se concentrará en las consonantes nasales.

5. NASALES

En los arranques de sílaba, el shipibo presenta dos posibles consonantes nasales: [m] y [n]. Además, shipibo también muestra vocales nasalizadas, las cuales presentan varios grados diferentes de coalescencia con la coda consonante nasal que les sigue. La coda nasal puede variar desde fusionarse completamente con la vocal precedente hasta solo fusionarse parcialmente con ella. En el último caso, se puede observar una coda nasal después de la vocal. En este capítulo, primero discuto las nasales en posición de arranque de la sílaba y luego, en la sección 5.2, las vocales nasalizadas.

5.1 LAS CONSONANTES NASALES COMO ARRANQUES DE SÍLABA

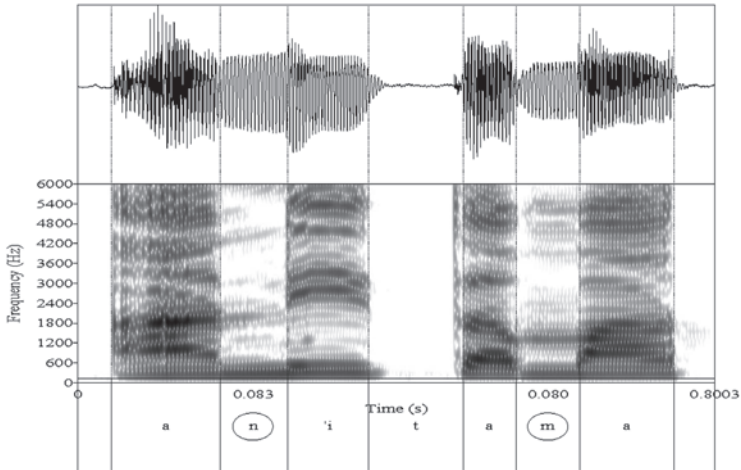
En shipibo, ambas consonantes nasales, [m] y [n], posee fonación modal (no son ni veladas ni glotalizadas). Articulatoriamente, las consonantes nasales se hacen creando una cerrazón completa en algún punto en la cavidad oral, de modo que el aire no puede pasar a través de la boca, y al mismo tiempo, al bajar el velo, se deja escapar el aire por las fosas nasales.

En los espectrogramas, el *murmullo nasal* (el efecto creado por una cerrazón oral sostenida mientras el aire sale por la nariz) aparece con menos energía acústica comparada con la cantidad de energía acústica que se observa en las vocales. Esto ocurre debido a la interacción entre las resonancias nasales y las anti-resonancias orales, que lleva a la energía acústica a cancelarse mutuamente (Huffman 1990; Huffman y Krakow 1993; Harrington y Cassidy 1999; Johnson 2003; Silverman 2006).

Un ejemplo de murmullo nasal de las consonantes [n] y [m] se ilustra en el espectrograma de la palabra [a.¹ni.ta.ma] ‘muy pequeño’ en la Figura 5-1. Además

de mostrar en general menos energía acústica que las vocales, la mayor parte de la energía que muestran las nasales tiende a concentrarse en las frecuencias bajas. Esto se debe a dos propiedades de las paredes de la cavidad nasal: son suaves y tienen un área de superficie extensa llena de pasajes intrincados. De este modo, la estructura de la cavidad nasal absorbe la mayor parte de la energía de ciertas frecuencias, particularmente la energía acústica de las frecuencias más altas. Esta pérdida de energía en las frecuencias más altas puede verse en el espectrograma de las consonantes nasales por la presencia de varios manchas blancas desde los 1000 Hz para arriba. La pérdida general de energía también puede observarse en el oscilograma al comprar la amplitud de la onda de sonido en las consonantes nasales con la de las vocales vecinas.

Figura 5-1: Espectrograma de la palabra [a.¹ni.ta.ma] ‘muy pequeño’



La Tabla 5-1 muestra la duración promedio de las consonantes nasales en posición de arranque de sílaba. Los promedios que se muestran corresponden a las nasales [m] y [n] como arranques ya sea de la primera o segunda sílaba de palabras trisilábicas. Como parte del estudio, se midió aproximadamente 2600 ejemplares de las consonantes nasales. La tabla no presenta la duración por separado de estas nasales en sílabas acentuadas e inacentuadas ya que su duración no se vio afectada por la posición del acento.

Tabla 5-1: Duración promedio de las consonantes nasales en posición de arranque

	Duración (en segundos)
[m]	0,087 ^(0,014)
[n]	0,081 ^(0,017)

5.1.1 Análisis espectral de las consonantes nasales

Las características espectrales de las nasales revelan una interacción intrincada de formantes nasales, anti-formantes orales y formantes orales. Empecemos por la característica constante que define a las nasales. Ésta es la presencia de un primer formante nasal ($F_n 1$) en frecuencias muy bajas. Cuando se articula una consonante nasal, el velo del paladar baja. Esta acción crea lo que podemos asumir como un tubo largo que empieza en la laringe y termina en las fosas nasales. Este tubo de resonancia es principalmente responsable por la frecuencia de $F_n 1$. Ya que este tubo es largo, la frecuencia de $F_n 1$ es muy baja. Además, dado que no podemos crear ninguna obstrucción dentro de la cavidad nasal, la frecuencia de $F_n 1$ permanece virtualmente invariable para todas las consonantes nasales. En los datos que se recolectaron del shipibo, el $F_n 1$ se localiza alrededor de 286 Hz

Tabla 5-2: Posición promedio de $F_n 1$ en consonantes nasales

	[m]	[n]
$F_n 1$ – hombres	273 Hz ⁽²⁴⁾	287 Hz ⁽²⁵⁾
$F_n 1$ – mujeres	278 Hz ⁽⁴⁴⁾	301 Hz ⁽⁶¹⁾

La Figura 5-2 y la Figura 5-3 ilustran ejemplos del espectro de las nasales [m] y [n]. El espectro LPC se muestra encima del espectro FFT. En cada caso se indica la posición del $F_n 1$. Los espectros muestran claramente que la mayor parte de la energía acústica de las nasales se concentra en las frecuencias bajas y esta rápidamente decae conforme suben las frecuencias.

Obsérvese que ambos espectros en la Figura 5-2 y la Figura 5-3 tienen anti-formantes. Además de la cavidad principal (el tubo laríngeo-nasal), las consonantes nasales tienen una cavidad secundaria: la boca. Esta segunda cavidad también actúa como resonador, lo cual significa que tiene su propia estructura de formantes. Esta cavidad forma un tubo que va desde la faringe hasta el punto donde se produce la obstrucción dentro de la cavidad oral (los labios en el caso de [m], los

Figura 5-2: Espectros FFT y LPC de la nasal [m]

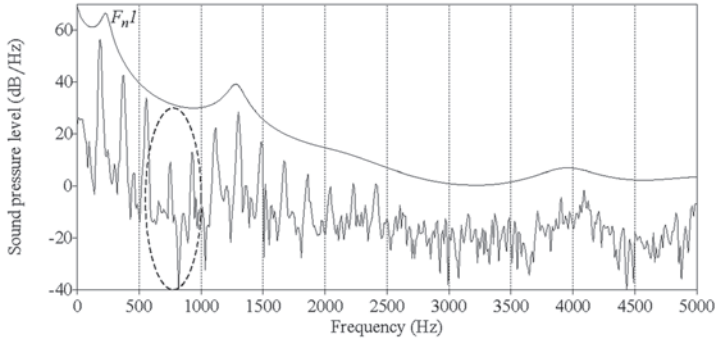
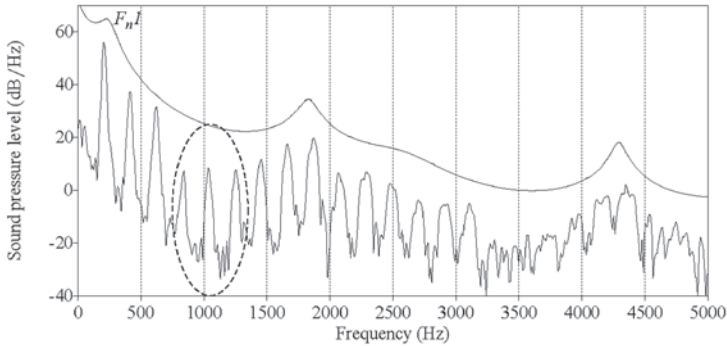


Figura 5-3: Espectros FFT y LPC de la nasal [n]



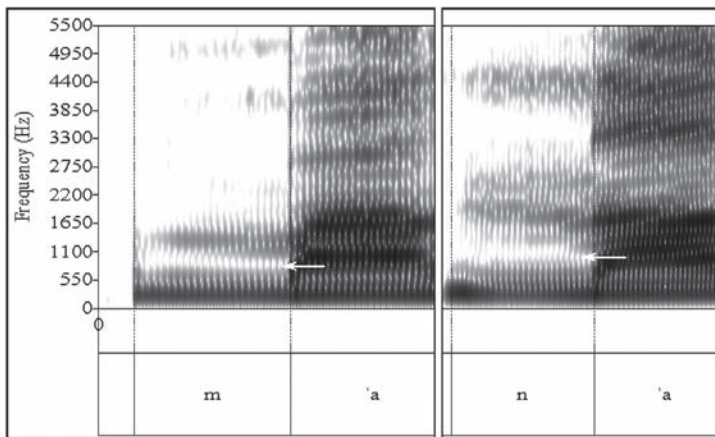
alvéolos en el caso de [n], etcétera). La estructura de formantes aquí dependerá de qué tan largo es este tubo resonador secundario. Contra más largo sea, más bajo es el formante que añade; e inversamente, contra más pequeño, más alta la frecuencia del formante. Sin embargo, ya que la boca está cerrada durante la articulación de una consonante nasal, la energía acústica dentro de la boca está atrapada, y por lo tanto es reflejada en la cavidad principal (el tubo laríngeo-nasal), donde se cancela con resonancias de frecuencias similares. De este modo, la interacción de las resonancias nasales y de las anti-resonancias orales crean regiones de frecuencias en las cuales la amplitud del sonido disminuye lo cual crea anti-formantes. Estos aparecen como valles de energía notables en los espectros al contrario de los formantes, lo cuales aparecen como picos de energía.

Como se indica mediante la elipse punteada en la Figura 5-2 y la Figura 5-3, el primer anti-formante en el espectro de [m] aparece en el rango de frecuencias entre 500 Hz y 1000 Hz. El punto más bajo del anti-formante se localiza aproximadamente a 800 Hz. El valle creado por el primer anti-formante puede verse en la disparidad entre el espectro FFT y el espectro LPC. Lo mismo ocurre en los

espectros de [n] en la Figura 5-3. El anti-formante se localiza en el rango de frecuencias entre 900 Hz y 1250 Hz (véase el área encerrada por la elipse punteada). El punto más bajo está aproximadamente en 1100 Hz. Estas frecuencias, 800 Hz y 1100Hz, son las posiciones típicas del primer anti-formante en las consonantes nasales [m] y [n], respectivamente. Recuérdese que la razón por la cual la posición del primer anti-formante tiende a ser más alta para [n] en comparación a [m] se relaciona con la longitud de la cavidad oral. La longitud de la cavidad oral para [m] es mayor a la longitud de la misma cavidad para [n] (Marple 1987; Qi 1989; Johnson 2003; Silverman 2006).

En los espectrogramas, los anti-formantes aparecen como franjas blancas horizontales cuando no hay ningún formante en una frecuencia cercana. De otro modo, tiende a debilitar al formante haciéndolo aparecer mucho más claro, es decir, con menos energía en el espectrograma (Johnson 2003). La franja blanca dejada por el primer anti-formante puede observarse en los espectrogramas de la Figura 5-4, el cual corresponde a las sílabas acentuadas iniciales de las palabras ['ma.ki.dʒa] 'piraña' (EV) y ['na.ta.ti] 'vaciar'. Las posiciones de los anti-formantes se indican con una flecha blanca. Los espectros presentados en la Figura 5-2 y la Figura 5-3 se basan en estas dos nasales en particular.

Figura 5-4: Espectrogramas de [m] y [n]



5.1.2 Las transiciones de formantes y las nasales

La Tabla 5-3 y la Tabla 5-4 presentan las frecuencias promedio de F2 en la transición de vocales con consonantes nasales adyacentes [m] y [n]. La primera tabla presenta los promedios para hablantes hombres y la segunda para hablantes

mujeres. Se utilizó el mismo conjunto de datos usado para obtener las mediciones de duración presentadas en la Tabla 5-1. Ya que los resultados muestran el mismo comportamiento de F2 observado en las oclusivas orales [p] y [t], no presentaré un gráfico.

El F2 de las vocales presenta un descenso claro cuando éstas están adyacentes a la nasal bilabial [m]. Esto se nota particularmente cuando [m] está adyacente a las vocales [i], [i̥] y [a]. Cuando [m] aparece adyacente a [u], el F2 tiende a permanecer estable durante toda la duración de la vocal. En el caso de la nasal [n], hay un incremento del F2 para [i] y [i̥] conforme se acercan al punto de transición con la consonante nasal. Sin embargo, la disminución del F2 para [n] no es tan dramática como la que se observa cuando las vocales aparecen adyacentes a [m]. El F2 de la vocal baja [a] tiende a permanecer sin cambios cuando ocurre vecina a la consonante [n] o muestra un pequeño incremento. El F2 de la vocal [u] se incrementa conforme se acerca a la transición con la consonante [n]. Estas diferencias marcadas que se observan en el comportamiento de F2 en las vocales en los puntos de transición con las nasales sirven como claves fonéticas fuertes para distinguir los puntos de articulación de las nasales.

Tabla 5-3: Promedios de las transiciones de F2 para las consonantes nasales (hombres)

	[m]	[n]
[i] - 2244	1806 Hz ⁽²⁰¹⁾	2010 Hz ⁽²⁰³⁾
[a] - 1418	1019 Hz ⁽¹⁰⁴⁾	1455 Hz ⁽¹⁴⁴⁾
[i̥] - 1539	1058 Hz ⁽¹⁰⁵⁾	1473 Hz ⁽¹⁵⁹⁾
[u] - 896	889 Hz ⁽¹¹⁷⁾	1218 Hz ⁽¹²²⁾

Tabla 5-4: Promedios de las transiciones de F2 para las consonantes nasales (mujeres)

	[m]	[n]
[i] - 2676	2196 Hz ⁽¹⁹⁶⁾	2426 Hz ⁽¹⁷⁷⁾
[a] - 1597	1305 Hz ⁽¹⁴⁰⁾	1640 Hz ⁽¹⁵⁴⁾
[i̥] - 1704	1094 Hz ⁽¹²⁶⁾	1580 Hz ⁽¹⁸⁵⁾
[u] - 935	939 Hz ⁽¹³⁵⁾	1302 Hz ⁽¹⁶²⁾

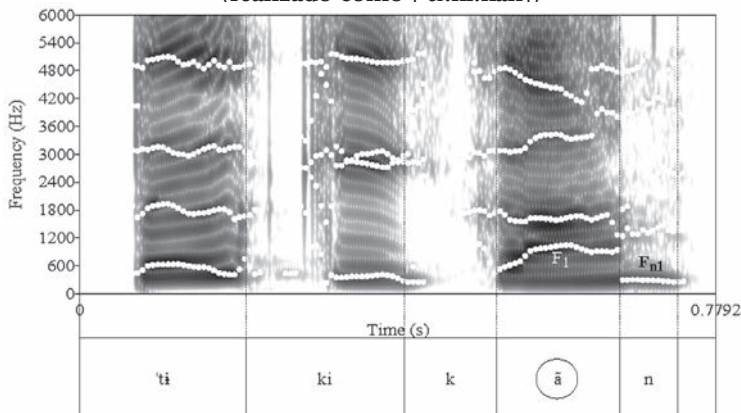
5.2 VOCALES NASALIZADAS Y LAS CODAS NASALES

Además de las vocales orales, shipibo tienen posesiones vocales con un fuerte grado de nasalización: [ã, ã̃, ã̄, ã̅]. En este estudio, tomo la posición que representacionalmente, son la secuencia de una vocal oral seguida de una coda nasal. Esta posición también es compartida por otros investigadores de la lengua (García-Rivera 1994; Valenzuela 2003b). Como se describirá en breve, en el nivel fonético, esta secuencia se realiza con varios grados de coalescencia.

Comencemos por presentar un ejemplo de una vocal completamente nasalizada. El espectrograma de la Figura 5-5 ilustra la vocal [ã] en la sílaba final de la palabra [ˈt̪i.ki.kan] ‘coyuntura’ (FL). La nasalización de la vocal tiende a incrementar el valor de formante F1. Esto quiere decir que las vocales nasales tienden a ser más bajas en altura que las vocales orales. Contra más fuerte el grado de nasalización, el valor de F1 se vuelve más alto. La palabra en la Figura 5-5 fue pronunciada por una hablante mujer. En promedio, una vocal oral [a] para hablantes mujeres tiene un F1 de 724 Hz. Sin embargo, en la vocal nasalizada [ã] que se muestra en el espectrograma de la Figura 5-5, el F1 es aproximadamente 300 Hz más alto. Es decir, el F1 de esta vocal tiene un valor de 1040 Hz.

Las vocales nasalizadas del shipibo son seguidas de una consonante nasal en posición de coda (se debe recordar que la coda nasal puede mostrar diferentes grados de coalescencia con la vocal precedente - sección 5.3). Obsérvese que la coda nasal aparece inclusive en posición final de palabra, sin ningún material segmental que la siga. Como se indicó en la Figura 5-5, en esta posición, la coda nasal ocurre como un segmento coronal: [n]. Durante el murmullo nasal de [n] en la Figura 5-5, se puede ver el formante nasal F_n 1 apareciendo alrededor de 280 Hz.

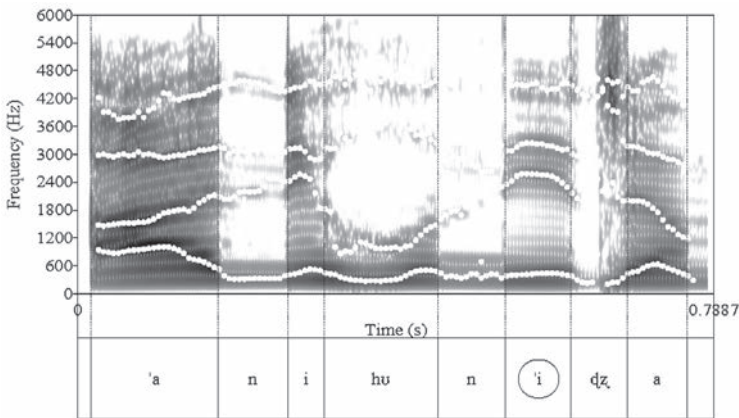
**Figura 5-5: Espectrograma de [ˈt̪i.ki.kan] ‘coyuntura (FL)’
(realizado como [ˈt̪i.ki.kã̃n])**



Algunas veces los hablantes no realizan la coda nasal. Cuando esto ocurre la vocal precedente usualmente tampoco obtiene nasalización. Esto tiende a suceder especialmente en el habla rápida. El espectrograma de la Figura 5-6 se muestra la frase [ʼa.ni hu.ʼnin.dʒa] ‘hombre grande’ (EV). Esta frase¹ está formada por la palabras /ani/ ‘grande’, /huni/ ‘hombre’, el sufijo /-n/² (FL) y el evidencial /-dʒa/. La palabra para ‘hombre’ sin el sufijo /-n/ (FL) se pronuncia [ʼhu.ni], el acento cae en la sílaba inicial ya que la segunda sílaba es abierta. Cuando se añade el sufijo /-n/, la segunda sílaba se vuelve cerrada y entonces atrae el acento: [hu.ʼnin].

En el ejemplo de la frase [ʼa.ni hu.ʼnin.dʒa] ‘hombre grande’ (EV) ilustrado en la Figura 5-6, la palabra [hu.ʼnin.dʒa] ‘hombre’ (FL, EV) se realiza como [hu.ʼni.dʒa], sin la coda nasal o nasalización de la vocal precedente en la segunda sílaba. Sin embargo, la posición del acento provee una clave importante a los hablantes de shipibo indicando que la segunda sílaba es fonológicamente cerrada y en este caso, que la frase nominal está en caso ergativo.

Fiura 5-6: Coda nasal dejada sin realizar: [hu.ʼnin] ‘hombre (FL)’ ([hu.ʼni])



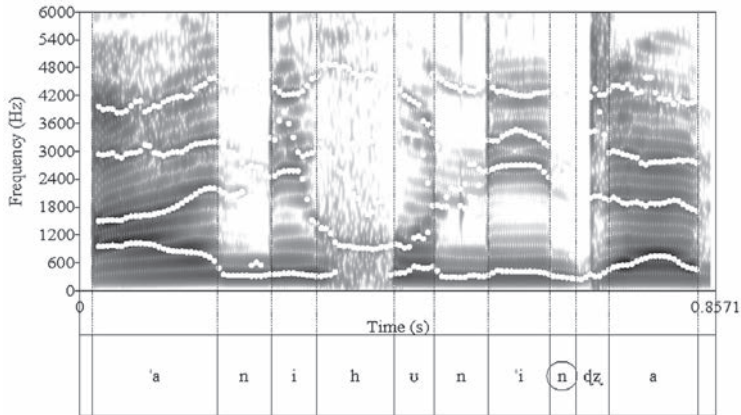
El mismo hablante cuando se le pidió producir la oración [ʼa.ni hu.ʼni.dʒa ma.ʼmi ʼb^βa.na.ki] ‘el/un hombre grande sembró mamey’ en aislamiento sí realizó la coda nasal en [hu.ʼnin]. Esta producción se muestra en la Figura 5-7.

¹ La frase se tomó de la oración [ʼa.ni hu.ʼnin.dʒa ma.ʼmi ʼb^βa.na.ki] ‘el/un hombre grande sembró mamey’.

² En shipibo, sustantivos y adjetivos muestran una forma de alternancia usualmente llamada ‘formas largas’ (FL). Las formas largas se marcan con los sufijos /-(a)n/ y /-nin/. Estos sufijos señalan varias funciones sintácticas. La más común es la de servir como marcador de caso ergativo, como ocurre en la oración [ʼa.ni hu.ʼnin.dʒa ma.ʼmi ʼb^βa.na.ki] ‘el/un hombre grande sembró mamey’. Sin embargo, otras funciones incluyen el instrumental, posesivo y locativo.

Esto proporciona mayor evidencia que la representación fonética de la palabra para ‘hombre’ en caso ergativo es siempre [hũ.'nin]. Durante el habla cuidada, los hablantes tienden a realizar la coda nasal mientras en el habla rápida podrían no hacerlo.

Figura 5-7: Coda nasal realizada en la palabra [hũ.'nin] ‘hombre (FL)’



5.2.1 La duración de las vocales nasalizadas y de las codas nasales

La Tabla 5-5 muestra la duración promedio (en segundos) de las vocales nasalizadas en tres posiciones: vocales en sílabas iniciales acentuadas, en segundas sílabas acentuadas y en sílabas inacentuadas. Los resultados se basan en el análisis de 2400 vocales nasalizadas. Todas ellas pertenecen a palabras trisilábicas pronunciadas dentro de la frase que uso este estudio. Los resultados de la Tabla 5-5 no incluyen la duración de la coda nasal que sigue a las vocales nasalizadas (la duración de las codas nasales se presenta en la Tabla 5-6).

Los resultados de la Tabla 5-5 muestran que las vocales nasalizadas son fonéticamente más largas que las vocales orales. En sílabas inacentuadas, las vocales orales tienden a durar alrededor de 75 ms. Las vocales nasalizadas en el mismo ambiente prosódico duran aproximadamente 15 ms más. En sílabas acentuadas, las vocales orales tienden a ser ligeramente más largas en segundas sílabas (116 ms en promedio) que en sílabas iniciales (102 ms en promedio). En contraste, las vocales nasalizadas acentuadas, ya sea en sílabas iniciales o en segundas sílabas, tienden a durar el mismo tiempo: aproximadamente 120 ms.

Tabla 5-5: Duración promedio de las vocales nasalizadas (en segundos)

	Sílaba Inicial Acentuada	Sílaba Acentuada Segunda	Sílaba Inacentuada
[ĩ]	0,122 ^(0,018)	0,121 ^(0,019)	0,090 ^(0,024)
[ĩ̃]	0,121 ^(0,012)	0,123 ^(0,018)	0,090 ^(0,014)
[ũ]	0,119 ^(0,012)	0,122 ^(0,020)	0,087 ^(0,013)
[ũ̃]	0,129 ^(0,020)	0,126 ^(0,020)	0,091 ^(0,014)

La Tabla 5-6 muestra la duración promedio (en segundos) de las codas nasales. La primera columna indica el timbre de la vocal precedente. En promedio, las codas nasales duran cerca de 60 ms. Los resultados indican que los diferentes contextos prosódicos no parecen afectar la duración de las codas nasales. Si combinamos las vocales nasalizadas con las codas nasales, la duración de las rimas que forman es aproximadamente 180 ms en sílabas acentuadas y 150 ms en sílabas inacentuadas.

Tabla 5-6: Duración promedio de las codas nasales (en segundos)

	Sílaba Inicial Acentuada	Sílaba Acentuada Segunda	Sílaba Inacentuada
i_.	0,062 ^(0,013)	0,062 ^(0,015)	0,059 ^(0,018)
ĩ_.	0,056 ^(0,013)	0,056 ^(0,011)	0,061 ^(0,015)
ũ_.	0,063 ^(0,020)	0,059 ^(0,011)	0,062 ^(0,015)
ũ̃_.	0,056 ^(0,013)	0,062 ^(0,014)	0,060 ^(0,017)

Los espectrogramas de la Figura 5-8 y la Figura 5-9 nos permiten comparar la duración de una vocal nasalizada acentuada cuando aparece seguida de una coda nasal con la duración de una vocal oral acentuada. Ambas palabras contienen la raíz verbal /kʊʃki/ ‘colocar a través’. La valencia de este verbo debe indicarse por un sufijo. El sufijo /-n/ (TRVZ) se usa para indicar que el verbo es transitivo (por ejemplo, /kʊʃki -n -ti/ → [kʊʃ.ˈkin.ti] ‘colocar a través’ (TRVZ, INF)) y el sufijo /-t/ (INTVZ)³ se usa para marcar el verbo como intransitivo (por ejemplo, /kʊʃki -t -ti / → [kʊʃ.ˈki.ti] ‘atravesarse’ (INTVZ, INF)). La vocal nasalizada en la segunda sílaba acentuada del verbo en la Figura 5-8 dura 128 ms, 10 ms más que su contraparte oral en la Figura 5-9.

³ El sufijo /-t/ (INTVZ) sólo aparece cuando el siguiente sufijo posee una vocal inicial (/kʊʃki -t -a/ → [kʊʃ.ki.ta] ‘atravesarse’ (INTVZ, PP2)).

Figura 5-8: Espectrograma de [kʊ].¹kin.ti ‘colocar a través’

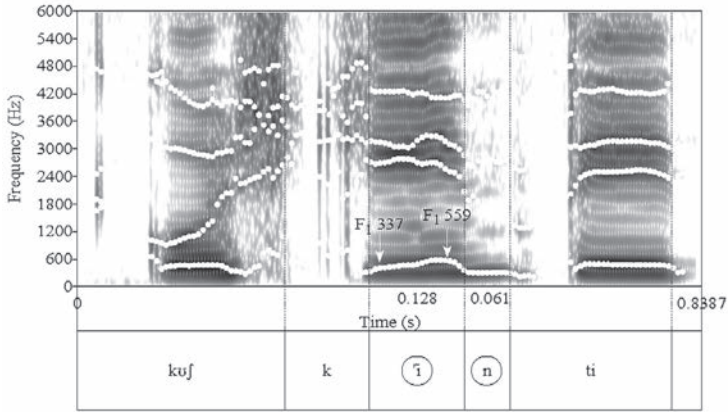
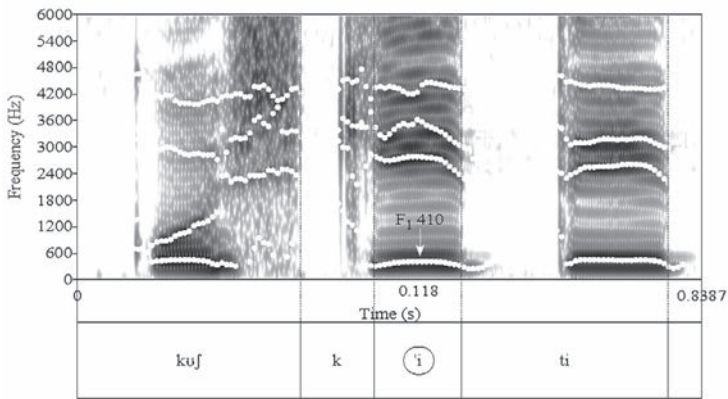


Figura 5-9: Espectrograma de [kʊ].¹ki.ti ‘atrasarse’



5.2.2 La estructura de formantes de las vocales nasalizadas

La subida de F1 en una vocal adyacente a una coda nasal brida una clave importante para poder determinar si la vocal está nasalizada. Obsérvese los comportamientos contrastantes de F1 en las sílabas acentuadas de los verbos en los espectrogramas de la Figura 5-8 y la Figura 5-9. En el espectrograma de la Figura 5-8, la vocal acentuada [i] aparece seguida de una coda nasal y su F1 se incrementa (de 337 Hz a 559 Hz), particularmente durante la segunda mitad de la vocal. El incremento de F1 se correlaciona con la percepción de que la vocal acentuada es nasalizada en la Figura 5-8. En contraste, en la Figura 5-9, la vocal oral acentuada [i] tiene un F1 estable (410 Hz).

En las vocales seguidas por una coda nasal, F1 sube con el tiempo conforme la vocal se aproxima a la coda. Su valor más bajo usualmente se encuentra durante la primera mitad de la vocal nasalizada y su valor más alto durante la segunda mitad (véase la Figura 5-8). La Tabla 5-7 y la Tabla 5-8 presentan los valores promedio de F1 para cada una de las vocales nasalizadas del shipibo. La Tabla 5-7 muestra los promedios para los hablantes hombres y la Tabla 5-8, los promedios para las hablantes mujeres. Cada tabla considera las vocales acentuadas e inacentuadas por separado. Los resultados presentados se basan en el análisis del mismo conjunto de datos usados para obtener la duración promedio de las vocales nasalizadas en la Tabla 5-6.

Los resultados de la Tabla 5-7 y la Tabla 5-8 se obtuvieron dividiendo cada vocal seguida de una coda nasal en tres partes: dos regiones de transición de formantes, cada una localizada a los bordes de la vocal y una región intermedia. Cada zona de transición de formantes fue igual a 10% de la duración de la vocal. Las regiones de transición de formantes fueron excluidas de las mediciones de F1 ya que principalmente reflejan información sobre el punto de articulación de los segmentos vecinos, y no necesariamente información sobre nasalidad.

La región intermedia se dividió en dos partes. Dentro de la primera mitad, el análisis obtuvo el valor de F1 más bajo. El promedio del valor más bajo de F1 se muestra en la tercera columna de las tablas bajo el nombre ‘Min F1’. Dentro de la segunda mitad, el análisis calculó el valor más alto de F1 (los promedios aparecen en la cuarta columna bajo ‘Max F1’). El promedio de cuánto se incrementa F1 aparece en la quinta columna de cada tabla bajo el nombre ‘Incremento de F1’.

Tabla 5-7: Comportamiento de F1 en vocales nasalizadas (hombres)

		Min. F1	Max. F1	Incremento de F1
[ĩ]	Acentuado	324 Hz ⁽⁴²⁾	436 Hz ⁽⁵²⁾	112 Hz ⁽⁴⁴⁾
	Inacentuado	288 Hz ⁽³⁵⁾	388 Hz ⁽⁴⁷⁾	100 Hz ⁽³⁷⁾
[ĩ̃]	Acentuado	375 Hz ⁽⁴⁵⁾	504 Hz ⁽⁵⁴⁾	129 Hz ⁽⁵¹⁾
	Inacentuado	346 Hz ⁽⁵³⁾	452 Hz ⁽⁶⁵⁾	106 Hz ⁽⁵¹⁾
[ũ]	Acentuado	376 Hz ⁽⁶⁴⁾	527 Hz ⁽⁶⁹⁾	151 Hz ⁽⁴⁶⁾
	Inacentuado	338 Hz ⁽⁵³⁾	475 Hz ⁽⁶⁴⁾	137 Hz ⁽⁵¹⁾
[ã]	Acentuado	593 Hz ⁽⁷¹⁾	694 Hz ⁽⁶⁸⁾	101 Hz ⁽⁴⁸⁾
	Inacentuado	489 Hz ⁽⁶⁴⁾	598 Hz ⁽⁷¹⁾	109 Hz ⁽⁴⁹⁾

Tabla 5-8: Comportamiento de F1 en vocales nasalizadas (mujeres)

		Min. F1	Max. F1	Incremento de F1
[ĩ]	Acentuado	376 Hz ⁽⁵⁹⁾	497 Hz ⁽⁶⁴⁾	121 Hz ⁽³⁷⁾
	Inacentuado	347 Hz ⁽⁴⁴⁾	456 Hz ⁽⁴²⁾	109 Hz ⁽³²⁾
[ɨ]	Acentuado	475 Hz ⁽⁷⁵⁾	642 Hz ⁽¹²⁶⁾	167 Hz ⁽⁷¹⁾
	Inacentuado	440 Hz ⁽⁷¹⁾	606 Hz ⁽¹²⁷⁾	166 Hz ⁽⁷⁵⁾
[ũ]	Acentuado	476 Hz ⁽⁶⁹⁾	657 Hz ⁽¹³¹⁾	181 Hz ⁽⁹⁷⁾
	Inacentuado	440 Hz ⁽⁶²⁾	649 Hz ⁽¹²⁹⁾	209 Hz ⁽⁹⁷⁾
[ã]	Acentuado	698 Hz ⁽⁹⁷⁾	859 Hz ⁽¹²³⁾	161 Hz ⁽⁸⁷⁾
	Inacentuado	653 Hz ⁽⁹⁵⁾	770 Hz ⁽¹³³⁾	117 Hz ⁽⁶⁵⁾

La Tabla 5-9 presenta los promedios de F2 en vocales nasalizadas tanto para hablantes hombres (tercera columna) como mujeres (cuarta columna). La tabla también distingue el comportamiento de F2 de acuerdo a si la vocal nasalizada está acentuada o no. Las mediciones de F2 de cada vocal se obtuvieron promediando sus valores de F2 dentro de una ventana equivalente a 30% de la duración de la vocal y centrada en el medio del segmento.

Tabla 5-9: Valores promedio de F2 en vocales nasalizadas

		F2 (Hombres)	F2 (Mujeres)
[ĩ]	Acentuado	2200 Hz ⁽¹⁷²⁾	2572 Hz ⁽¹³⁹⁾
	Inacentuado	2149 Hz ⁽²⁰⁵⁾	2468 Hz ⁽¹²⁸⁾
[ɨ]	Acentuado	1414 Hz ⁽¹⁰⁰⁾	1538 Hz ⁽¹⁹⁹⁾
	Inacentuado	1339 Hz ⁽¹⁸⁸⁾	1427 Hz ⁽¹⁶¹⁾
[ũ]	Acentuado	873 Hz ⁽¹²¹⁾	1042 Hz ⁽¹⁶⁶⁾
	Inacentuado	922 Hz ⁽¹⁴⁷⁾	1059 Hz ⁽¹²⁶⁾
[ã]	Acentuado	1392 Hz ⁽¹⁰⁴⁾	1596 Hz ⁽²⁴⁰⁾
	Inacentuado	1458 Hz ⁽¹³⁶⁾	1585 Hz ⁽¹⁹⁹⁾

Los resultados presentados en la Tabla 5-7 a la Tabla 5-9 se pueden apreciar mejor en los gráficos de la Figura 5-10 a la Figura 5-13. Los gráficos representan el espacio vocálico de las vocales nasalizadas formado por la dimensión de la altura vocálica (F1) y el grado de frontalidad/posterioridad de la vocal (F2). El gráfico

de la Figura 5-10 muestra los resultados de las vocales nasalizadas acentuadas para las hablantes mujeres y el gráfico de la Figura 5-11 para los hablantes hombres. Para poder hacer comparaciones, he incluido la posición de las vocales acentuadas orales, las cuales se indican por medio de círculos grises. Los círculos negros y blancos representan las vocales nasalizadas. Los círculos negros indican el valor más bajo de F1 durante la primera mitad de la vocal nasalizada mientras que los círculos blancos representan el valor más alto de F1 alcanzado durante la segunda mitad de la vocal. Las flechas indican la subida de F1 (la cual, en términos de la altura vocálica, significa que las vocales nasalizadas se hacen más bajas).

Nótese que los movimientos de F1 que muestran las vocales nasalizadas ocurren dentro del espacio delimitado por sus contrapartes orales. De este modo, tomando como referencia la posición promedio de las vocales altas orales en las sílabas acentuadas, las vocales altas nasalizadas ([ĩ], [ĩ̃] y [ũ̃]) constantemente bajan sus alturas conforme se aproximan a la coda nasal. Sin embargo, aunque la vocal nasalizada [ã] en sílabas acentuadas también disminuye su altura con el tiempo, ésta empieza como una vocal central y baja como si estuviese tratando de alcanzar el F1 ideal de su contraparte oral [a].

Con respecto a F2, las vocales acentuadas nasalizadas muestran casi el mismo grado de frontalidad/posterioridad que las vocales orales. La única excepción tanto para hablantes hombres como para mujeres es la vocal central acentuada [ĩ̃], la cual ocupa una posición ligeramente más posterior en el espacio vocálico que su contraparte oral [ĩ]. En las hablantes mujeres, la vocal nasalizada acentuada [ũ̃] tiende a ser ligeramente más centralizada que la vocal oral acentuada [u]. Ésta es una tendencia que no se observa en los hablantes hombres.

Figura 5-10: Vocales acentuadas nasalizadas y sus contrapartes orales (mujeres)

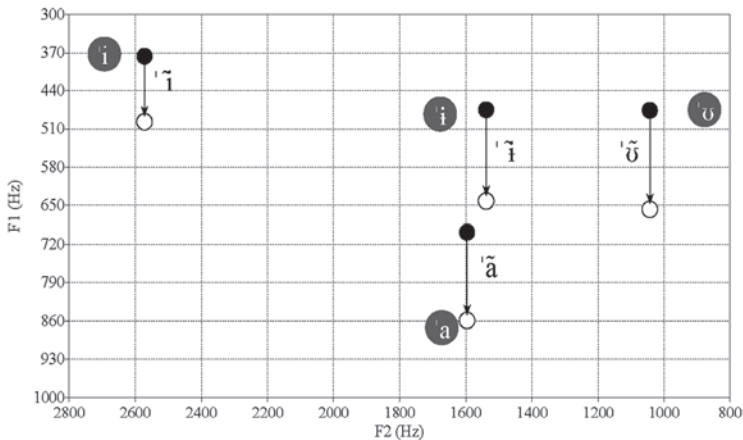
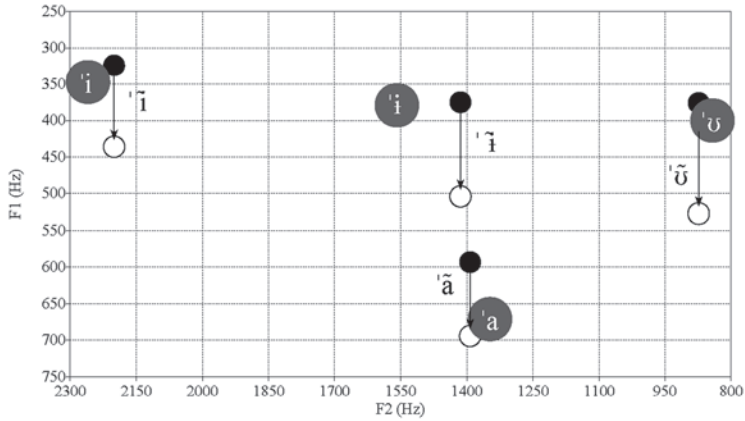


Figura 5-11: Vocales acentuadas nasalizadas y sus contrapartes orales



Los gráficos de la Figura 5-12 y la Figura 5-13 ilustran la distribución de las vocales inacentuadas nasalizadas dentro del espacio vocálico tanto para los hablantes hombres como para las hablantes mujeres. Las vocales orales inacentuadas aparecen en círculos grises. Aunque las mismas observaciones hechas para las vocales acentuadas nasalizadas también pueden hacerse con respecto a las inacentuadas, uno además puede ver que el espacio de las vocales nasalizadas y orales es mucho menos definido que en el caso de las vocales acentuadas. Tienen a sobreponerse mucho más. Además, debido a la centralización que sufren las vocales en sílabas inacentuadas, el espacio vocálico ocupado por las vocales nasalizadas inacentuadas es mucho más pequeño que el de las vocales nasalizadas acentuadas.

Figura 5-12: Vocales inacentuadas nasalizadas y sus contrapartes orales (mujeres)

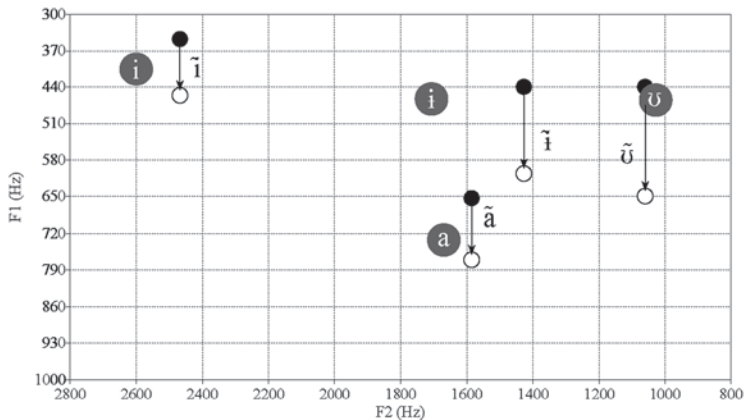
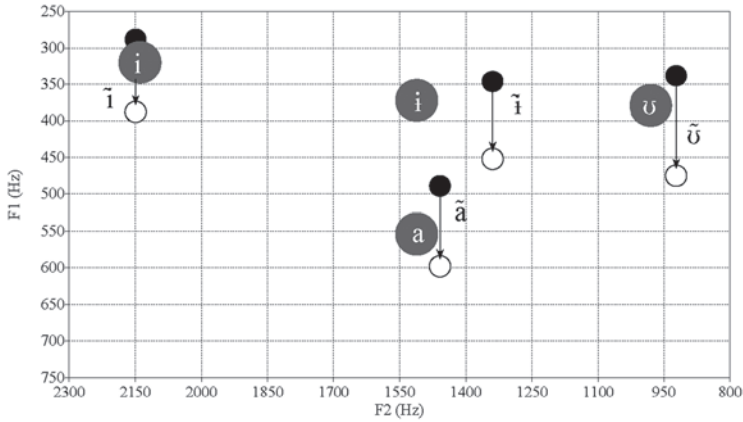


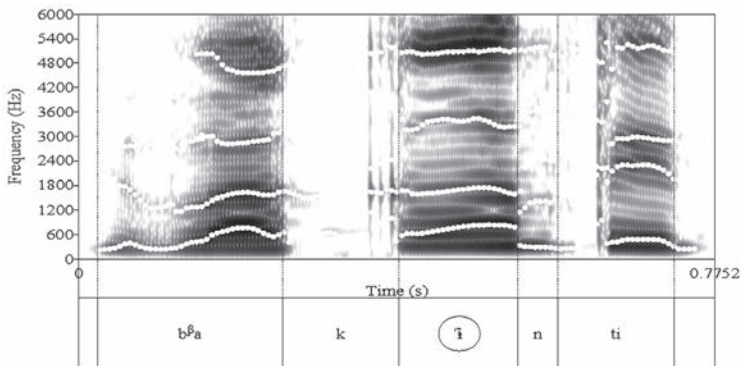
Figura 5-13: Vocales inacentuadas nasalizadas y sus contrapartes orales (hombres)



5.2.2.1 *El comportamiento de F1 y las asimetrías entre las vocales acentuadas e inacentuadas*

En vocales fuertemente nasalizadas, F1 tiende a subir notablemente conforme la vocal se aproxima a la coda nasal. Mientras el espectrograma de la Figura 5-5 muestra este fenómeno para una vocal inacentuada, el espectrograma de la Figura 5-14, [b^βa.¹kĩn.ti] ‘dar a luz’, lo ilustra para una vocal acentuada. En el último caso, [b^βa.¹kĩn.ti], el F1 de la vocal es en promedio 520 Hz durante la primera mitad de la vocal acentuada [i], y luego sube a 700 Hz durante la segunda mitad. Compárese esto con el promedio de F1 de una vocal oral [i] para una hablante mujer: 465 Hz.

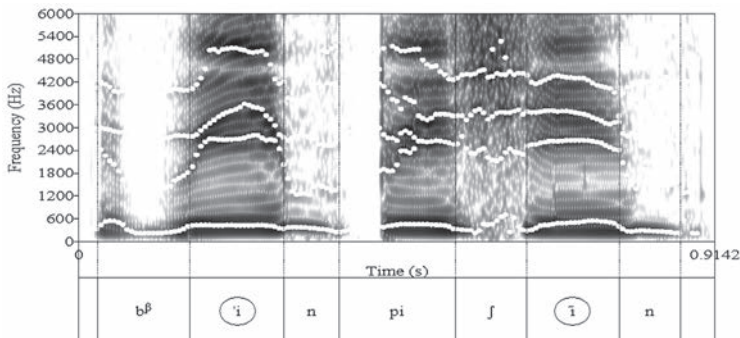
Figura 5-14: Espectrograma de [b^βa.¹kĩn.ti] ‘dar a luz’ (realizado como [b^βa.¹kĩn.ti])



Aunque las vocales acentuadas seguidas de una coda nasal pueden aparecer completamente nasalizadas, como en el ejemplo que acaba de discutirse, hay una tendencia en shipibo a que las vocales que tienen acento primario muestren poca nasalización o mucho menos nasalización que las vocales que tienen acentos secundarios. El espectrograma de la palabra [ˈbβin.pi.ɰin] (esp. de planta - FL) en la Figura 5-15, pronunciada por una hablante mujer, nos permite observar cómo difieren las vocales que llevan acento principal de aquellas que llevan acento secundario con respecto a su tratamiento de la nasalización, inclusive dentro de la misma palabra. La primera y la última vocal [i] de la palabra ocurren seguidas de una coda nasal. La primera vocal [i] lleva acento primario mientras la segunda lleva acento secundario. La vocal con el acento primario tiene un timbre claramente oral durante casi toda su duración. Si no fuese por la presencia de la coda nasal, no sería posible saber que la sílaba tiene un segmento nasal. Esta vocal muestra un F1 estable. De hecho, el F1 de la vocal [i] que lleva el acento principal en [ˈbβin.pi.ɰin], aunque está seguida de una coda nasal, se comporta como el F1 de la segunda vocal [i], la cual ocurre en un contexto oral. Inclusive sus valores de F1 son similares: 409 Hz para la vocal [i] que lleva el acento primario y 429 Hz para la vocal [i] inacentuada.

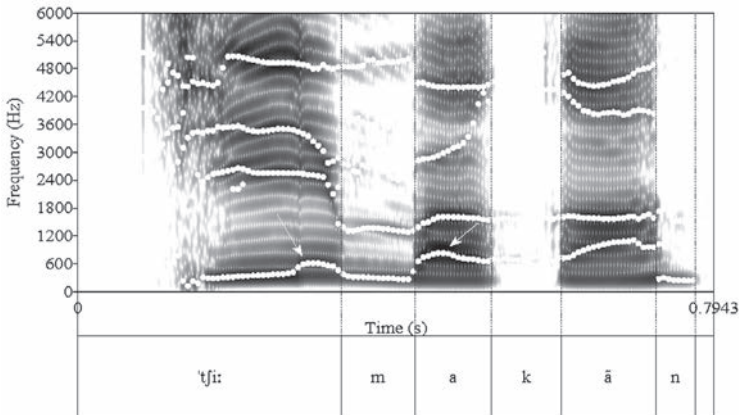
El comportamiento de la nasalidad en la vocal [i] que llevan el acento secundario en la palabra [ˈbβin.pi.ɰin], es bastante diferente a la vocal [i] con el acento primario. Esta vocal tiene un timbre nasal fuerte. Incluso si se sacase la coda nasal, se podría percibir su estatus nasalizado. A diferencia de la vocal [i] con el acento primario o de la vocal oral [i] inacentuada, el F1 de la tercera vocal [i] se incrementa constantemente. La primera mitad de la vocal posee en promedio un F1 de 370 Hz mientras que en la segunda mitad conforme la vocal se acerca a la coda nasal, su F1 salta a 506 Hz (compárese estos valores al promedio de F1 para la vocal oral [i] en hablantes mujeres: 391 Hz).

Figura 5-15: Espectrograma de [ˈbβin.pi.ɰin] (esp. de planta)
(realizado como [ˈbβim.pi.ɰin])



Las consonantes nasales en posición de arranque tienen un efecto similar en las vocales vecinas. Suben el F1 de las vocales de modo similar a las codas nasales. Sin embargo, a diferencia de las codas nasales, los arranques nasales solo tienden a afectar el F1 de una sección bastante localizada de las vocales adyacentes. En contraste a la fuerte nasalización observada en vocales seguidas de codas nasales, las vocales en contacto con consonantes nasales en posición de arranque muestran mucho menos nasalización. Esto se ilustra en el espectrograma de la Figura 5-16, el cual muestra la palabra [tʃi:ma,kan] ‘piedra de chispa’. Cada una de las vocales de esta palabra está en contacto con una consonante nasal y en cada caso el F1 de las vocales sube. En el caso de la vocal [i:], el F1 sube de pronto conforme se aproxima al arranque de la siguiente sílaba: [m]. El F1 de la vocal [a] adyacente a [m] sube inmediatamente después del arranque nasal y luego rápidamente desciende conforme la vocal se aleja del arranque. En contraste, el F1 de la vocal [a] adyacente a la coda nasal [n] se mantiene subiendo desde el inicio del segmento y esto se correlaciona con la fuerte nasalización que presenta la vocal.

Figura 5-16: Espectrograma de [tʃi:ma,kan] ‘piedra de chispa’
([tʃi:ma,kān])



5.3 SOBRE EL PUNTO DE ARTICULACIÓN DE LAS CODAS NASALES

Al interior de las palabras, las codas nasales se asimilan al punto de articulación de la siguiente consonante (por ejemplo, [p̃im.p̃in] ‘mariposa’, [dʒ̃an.tuŋ.ku] ‘rótula’). Esto puede observarse en los datos en (15). Al final de las palabras, las codas nasales tienden a realizarse como un segmento coronal: [n]. El punto de articulación de las codas nasales en esos contextos se puede identificar mirando el comportamiento del formante F2 en la transición entre una vocal y la coda nasal adyacente (véase la sección 5.1.2).

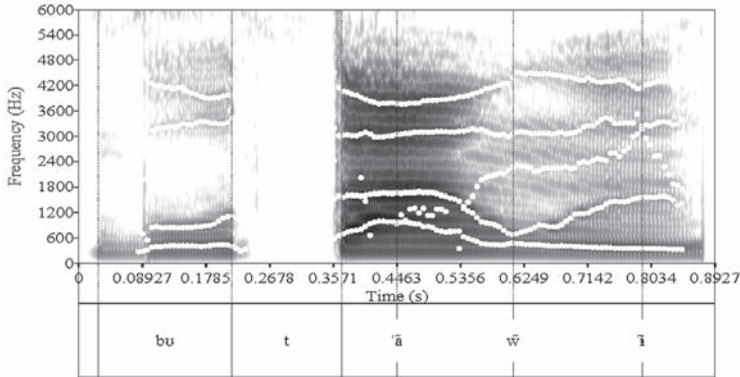
- (15) a. /tʊʃmʊn -ti/ → [tʊʃ.ˈmʊn.ti] ‘formarse una ampolla’
(INF)
- b. /isin -kan -ki/ → [i.ˈsiŋ.kan.ki] ‘se enfermaron’ (/isin/
‘enfermedad’, /-kan/ 3ra
persona plural, /-ki/
CMPL)
- c. /jʊnan -paʊni -ki/ → [jʊ.ˈnam.pa.ʊ.ni.ki] ‘ahumar’ (/jʊnan/
‘ahumar’, /-paʊni/ acción
remota incompleta, /-ki/
CMPL)

Sin embargo, las codas nasales no aparecen como segmentos independientes cuando son seguidas por una semi-consonante o una oclusiva glotal. En ese contexto, coalescen completamente con la vocal precedente. Cuando esto ocurre, la vocal nasalizada resultante es tan larga como una secuencia de una vocal seguida de una coda nasal. Este fenómeno no solo ocurre en la frontera de morfemas sino que también ocurre al interior de morfemas (esto excluye a las oclusivas glotales ya que solo pueden ocurrir como el primer segmento de un morfema).

Observemos un ejemplo de una semi-consonante que sigue a una coda nasal. El espectrograma de la Figura 5-17 muestra la palabra [bʊ.ˈtan.wi] ‘janda y toma(lo)’ (realizada como [bʊ.ˈtã.wĩ]). La palabra está formada por la raíz verbal /bʊ/ ‘tomar’, el sufijo /-tan/ (ve y regresa) y el sufijo /-wi/ (IMP). La concatenación de los sufijos /-tan/ y /-wi/ crea una secuencia de una coda nasal seguida de una semi-consonante a través de una frontera morfémica. Como se puede ver, la coda nasal no aparece como un segmento independiente. En vez de ello, coalesce con la vocal precedente, [a], la nasaliza y la nasalización se propaga sobre la semi-consonante siguiente, [w], así como también sobre la vocal que sigue a la semi-consonante, [i]. Obsérvese que en la Figura 5-17 no ha quedado ninguna huella de la coda nasal.

Shipibo tiene algunos sufijos como /-ʔati/ (VZ1), /-ʔiti/ (VZ2), /-ʔib^βat/ (PSD1), /-ʔiɟza/ (INTENS) que empiezan con una oclusiva glotal subyacente. Cuando se agregan a una raíz o base que termina en una coda nasal, la nasal se comporta exactamente como cuando sigue a una semi-consonante. Coalesce completamente con las vocales precedentes y la nasalidad se propaga sobre todos los segmentos sonantes adyacentes. Esto se puede observar en los datos de (16) a (19). Los hablantes de shipibo no siempre realizan la oclusiva glotal de estos sufijos (véase la sección 2.5).

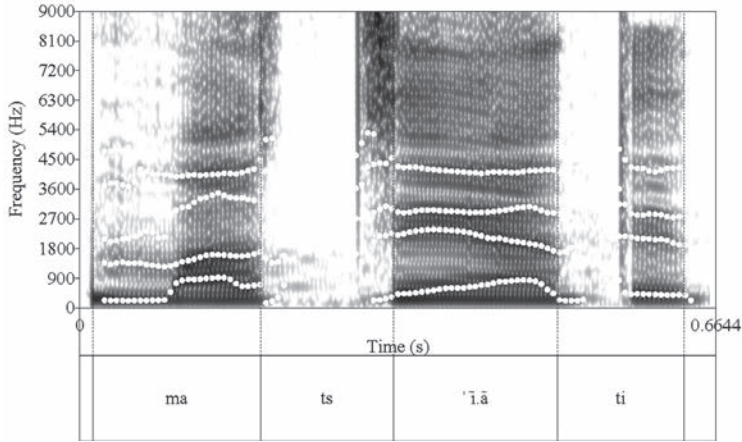
Figura 5-17: /bU -tan -wi/ → [bU.¹tan.wi] ‘janda y toma(lo)!’
(realizado como [bU.¹tã.wĩ])



- (16) /b^βikun -ʔati/ → [b^βi.kun.ʔa.ti] (realizado como: [b^βi.kũ.ʔã.ti]
~ [b^βi.kũ.ã.ti])
‘empañar, anublar’ (/b^βikun/ ‘con mala vista, miope’, /-ʔati/ VZ1)
- (17) /b^βidzɔn -ʔiti/ → [b^βi.dzɔn.ʔi.ti] (realizado como: [b^βi.dzũ.ʔĩ.ti]
~ [b^βi.dzũ.ĩ.ti])
‘tener dolor de ojos’ (/b^βidzɔ/ ‘ojo’, /-n/ FL, /-ʔiti/ VZ2)
- (18) /isin -ʔib^βat -ki/ → [i.sin.ʔib^βa.ki] (realizado como: [i.¹sĩ.ʔĩ.b^βa.ki]
~ [i.¹sĩ.ĩ.b^βa.ki])
‘se enfermó el otro día’ (/isin/ ‘enfermedad’, /-ʔib^βat/ (PSD1), /-ki/ CMPL)
- (19) /ikun -ʔidzɔ/ → [i.kun.ʔi.dzɔ] (realizado como: [i.¹kũ.ʔĩ.dzɔ] ~
[i.¹kũ.ĩ.dzɔ])
‘muy cierto’ (/ikun/ ‘verdad, verdadero’, /-ʔidzɔ/ INTENS)

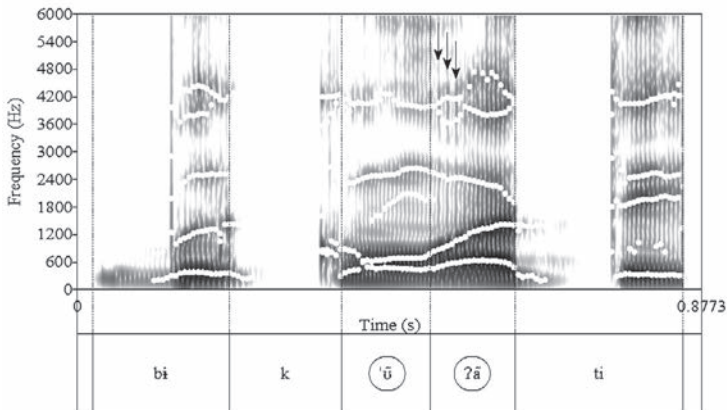
El espectrograma de la Figura 5-18 muestra el verbo [ma.¹tsin.ʔa.ti] ‘sentir frío’ (la base es [ma.¹tsin], la cual es la forma larga de [¹ma.tsi] ‘frío’). En este ejemplo, el verbo se ha realizado como [ma.¹tsĩ.ã.ti], es decir, la nasal final de la base se ha fusionado con la vocal acentuada que le precede y no se ha realizado la oclusiva glotal del verbalizador /-ʔati/ (VZ1). La nasalización se ha propagado sobre los segmentos sonantes vecinos (éste es el mismo patrón que se observó en casos como [b^βU.¹tã.wĩ] de la Figura 5-17).

Figura 5-18: /matsin-ʔati/ → [ma.¹tsin.ʔa.ti] ‘sentir frío’
(realizado como [ma.¹tsĩ.ã.ti])



El espectrograma de la Figura 5-19 ilustra el verbo [bi.¹kũn.ʔa.ti] ‘empañar, anublar’ (la raíz es /bikũn/ ‘miope’). En este ejemplo, sí se ha realizado la oclusiva glotal subyacente del sufijo /-ʔati/: [bi.¹kũ.ʔã.ti]. Las flechas indican dónde está la oclusiva glotal (véase las estriaciones verticales). Como en el caso previo, la coda nasal coalesce con la vocal que le precede y la nasalización se esparce sobre los segmentos vecinos.

Figura 5-19: /bikũn-ʔati/ → [bi.¹kũn.ʔa.ti] ‘empañar, anublar’ ([bi.¹kũ.ʔã.ti])



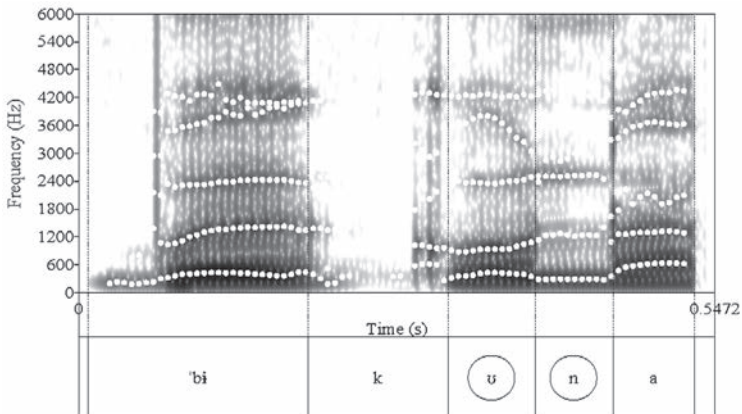
Las consonantes nasales en posición final de morfema ocurren como un arranque coronal [n] cuando aparecen seguidas de un sufijo que empieza en vocal, excepto cuando la nasal pertenece a una raíz nominal o adjetival (véase la

sección 9.2.3.4). En ese caso, la nasal final del morfema debe aparecer como un arranque nasal bilabial, [m]. Los datos en (20) ilustran casos en que la nasal final de morfema ocurre como [n].

- (20) a. /b^βu -tan -a/ → ['b^βu.ta.na] 'ir y tomar' (/tan/ 've y regresa', /-a/ PP2)
 b. /tuʃmũn -a/ → ['tuʃ.mũ.na] 'formarse una ampolla' (/a/ PP2)
 c. /isin -ai/ → ['i.si.na.i] 'enfermarse' (/ai/ PP1)
 d. /pi -kan -ai/ → ['pi.ka.na.i] 'comen' (/kan/ CON.3PL, /ai/ PP1)
 e. /hakũn -i/ → ['ha.kũ.ni] 'verdaderamente' (de [ha.^hkũn] 'verdad', /-i/ CR)

El espectrograma de la Figura 5-20 ilustra la realización de una nasal como un arranque [n] en posición final de morfema. Esto se muestra a través de la palabra ['bi.kũ.na] 'nublarse la vista' (PP2). Esta palabra contiene dos morfemas, la raíz /bikũn/ y el sufijo /-a/ (PP2). Debido a la silabificación, la nasal final de /bikũn/ ocurre como el arranque de la siguiente vocal, el sufijo /-a/. Debido a que la segunda sílaba es ahora abierta, el acento principal aparece sobre la sílaba inicial (compárese esto con /bikũn -ti/ → [bi.^hkũn.ti] 'nublarse la vista' (INF), en el cual la segunda sílaba es cerrada por una nasal y en consecuencia, atrae el acento).

Figura 5-20: ['bi.kũ.na] 'nublarse la vista' (PP2)

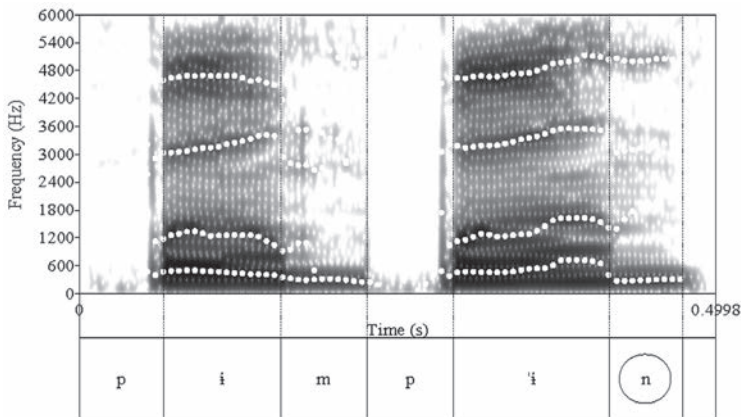


Ahora observemos los datos en (21) que muestran casos de raíces nominales y adjetivales cuyas consonantes nasales finales ocurren como un arranque nasal bilabial [m] cuando se añade el sufijo /-an/ (FL).

- (21) a. /p̥in̥p̥in̥ -an/ → [p̥in̥.p̥i.man] ‘mariposa’ (forma larga de [p̥in̥p̥in̥])
 b. /tuʃm̥un̥ -an/ → [tuʃ.m̥u.man] ‘ampolla’ (forma larga de [tuʃ.ʔm̥un̥])
 c. /isin̥ -an/ → [i.si.man] ‘enfermedad’ (forma larga de [i.ʔsin̥])
 d. /j̥unan̥ -an/ → [j̥u.na.man] ‘ahumado’ (forma larga de [j̥u.ʔnan̥])
 e. /n̥un̥un̥ -an/ → [n̥u.n̥u.man] ‘pato’ (forma larga de [n̥u.ʔn̥un̥])
 f. /b̥i̥k̥un̥ -an/ → [b̥i̥.k̥u.man] ‘miope’ (forma larga de [b̥i̥.ʔk̥un̥])
 g. /hak̥un̥ -an/ → [ha.k̥u.man] ‘bueno’ (forma larga de [ha.ʔk̥un̥])

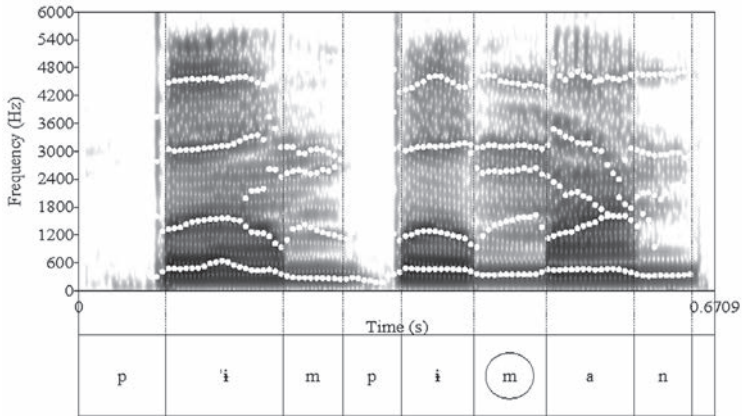
Los espectrogramas de la Figura 5-21 y la Figura 5-22 ilustran la forma corta y larga de la palabra para ‘mariposa’: [p̥in̥.ʔp̥in̥] y [p̥in̥.p̥i.man]. En la Figura 5-21, la nasal final de la raíz de /p̥in̥p̥in̥/ se realiza como [n̥] mientras que en la Figura 5-22, se realiza como [m]. El punto de articulación bilabial de la nasal final del morfema en la Figura 5-22 se puede identificar visualmente mirando el comportamiento de F2 en las vocales adyacentes (véase la Tabla 5-3 y la Tabla 5-4).

Figura 5-21: /p̥in̥p̥in̥/ → [p̥in̥.ʔp̥in̥] ‘mariposa’ (realizado como [p̥im̥.ʔp̥in̥])



Antes de terminar esta sección, hay un caso adicional que presentaré relacionado al punto de articulación de las nasales finales de morfema cuando ocurren al final de una frase. Recuérdese que para este estudio, las palabras que se recolectaron se grabaron dentro de la frase: “___ riqui westiora Shipibo joi. Enra ja joi ___ Tomás axeacai” (“___ es una palabra del shipibo. Le enseñó la palabra ___ a Tomás” - véase la sección 1.3.1). Las palabras que se analizaron fueron aquellas que ocurren en el segundo espacio en blanco de la frase, entre las palabras “joi” y “Tomás”. En consecuencia, la nasal final de una palabra no está en posición final de frase ya que siempre es seguida por la consonante oclusiva [t] de [tu.mas].

Figura 5-22: /pinpín -an/ → [ˈpin.pi.man] ‘mariposa’ (FL),
realizado como [ˈpim.pi.man]



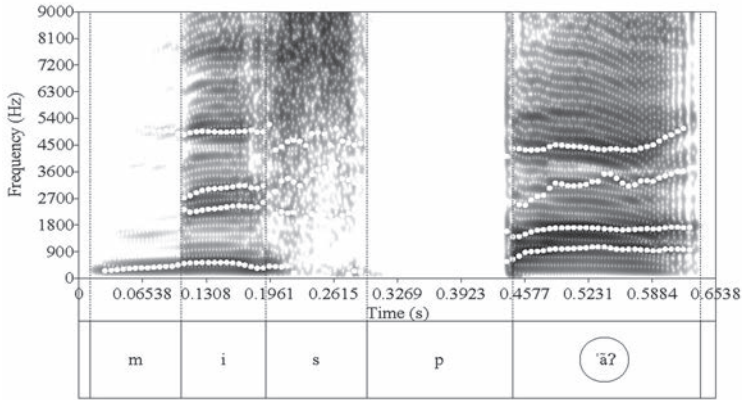
En posición final absoluta de frase, las codas nasales tienden a coalescer completamente con la vocal precedente. Esta coalescencia usualmente no deja ninguna huella del segmento nasal. Además, frecuentemente, se inserta una oclusiva glotal al final de la vocal nasalizada.⁴ El espectrograma de la Figura 5-23 ilustra un ejemplo de este caso a través de la palabra [mis.ˈpan] ‘tamal’ pronunciada en aislamiento. Las estriaciones verticales al final de la vocal [ã] indican la presencia de una oclusiva glotal. La coda nasal aparece completamente fusionada con la vocal [a] lo que da como resultado una vocal nasalizada: [ã]. La glotalización de la vocal nasal solo ocurre hacia el final. El resto de la vocal no se glotaliza. Compárese esta realización de una coda nasal al final de frase con la realización al final de una palabra pero dentro de una oración o frase (véase los espectrogramas de la Figura 5-5, la Figura 5-15, la Figura 5-16, la Figura 5-21 y la Figura 5-22).

5.4 RESUMEN

Este capítulo ha presentado una caracterización acústica de los segmentos nasales del shipibo. Se determinó que la duración de las consonantes nasales en posición de arranque de sílaba es aproximadamente 85 ms tanto en sílabas acentuadas como inacentuadas. Como codas, sin embargo, las consonantes nasales tienden a ser más cortas. Sin importar que la sílaba que las contiene sea acentuada o no, las codas nasales tienden a durar 60 ms.

⁴ Sin embargo, la aparición de la oclusiva glotal no es única al contexto de vocal nasalizada de final de frase. La oclusiva glotal también puede aparecer en posición final de frase después de una vocal oral.

Figura 5-23: /mispan/ → [mis.¹pan] ‘tamal’ (realizado como [mis.¹pāʔ])



Todos los espectros de las nasales indican la presencia de su F_{n1} a aproximadamente 286 Hz. Además de F_{n1} y la transición de F_2 , la consonante nasal bilabial [m] puede distinguirse de la nasal alveolar/dental [n] por medio de la posición del primer anti-formante, el cual usualmente se localiza a 800 Hz para la [m] y a 1100 Hz para la [n].

Como hemos visto, las vocales seguidas de codas nasales muestran diferentes grados de nasalización. Las vocales nasalizadas resultantes tienden a ser más largas que sus contrapartes orales. En las sílabas acentuadas, las vocales nasalizadas tienden a durar aproximadamente 120 ms mientras que en las vocales orales en sílabas iniciales acentuadas duran en promedio solo 100 ms. En sílabas inacentuadas, las vocales nasalizadas duran cerca de 90 ms mientras que las vocales orales inacentuadas duran aproximadamente 75 ms. Las vocales nasalizadas también se caracterizan por un F_1 más alto que el F_1 de las vocales orales. Además, este estudio mostró que el F_1 de las vocales nasalizadas presenta un movimiento ascendente que se desarrolla conforme la vocal se aproxima a la coda nasal. En general, en términos de la altura vocálica, las vocales nasalizadas ocupan una posición más centralizada que las vocales orales.

Además de la caracterización acústica de los segmentos nasales, este capítulo ha presentado información sobre el punto de articulación que muestran las codas nasales en diferentes ambientes. Ahora estamos listos para proceder a la discusión del último conjunto de segmentos del shipibo: las semi-consonantes.

6. GLIDES

6.1 LOS FORMANTES F1 Y F2 DE LAS GLIDES

La estructura de formantes de las glides es bastante similar a la de las vocales altas. Esto se muestra en la Tabla 6-1 y la Tabla 6-2. Ellas presentan los valores promedios de F1 y F2 para las semi-consonantes [j] y [w], respectivamente. La tercera columna en cada tabla da los resultados para los hablantes hombres y la cuarta columna, para las hablantes mujeres. Ambas tablas nos permiten comparar los resultados obtenidos para las semi-consonantes con los valores promedios de F1 y F2 en las vocales altas [i] y [u]. Los valores de F1 y F2 de [j] y [w] en la Tabla 6-1 y la Tabla 6-2 se basan en el análisis de 850 ejemplares de cada semi-consonante. En el caso de [j], se obtuvieron los valores de sus formantes de un punto dentro de la semi-consonante donde el F2 obtenía su valor más alto. Para [w], el punto donde se midió el valor de sus formantes fue el lugar donde F2 obtuvo su valor más bajo.

La Tabla 6-1 muestra que los valores de F1 y F2 están muy cercanos, casi idénticos, tanto para [j] como para [i]. Los resultados presentados en la Tabla 6-2 muestran una diferencia importante entre [w] y [u]. El valor promedio de F1 para [w] es más bajo que el F1 de [u], lo cual indica, como se esperaba para una semi-consonante, que [w] tiende a realizarse como un segmento mucho más alto que la vocal [u]. Obsérvese también que el F1 de [w] es virtualmente el mismo que el F1 de [j]. De este modo, a diferencia de [i] y [u], que muestran una diferencia en altura (véase la sección 2.1), las semi-consonantes [j] y [w] tienen la misma altura.

Tabla 6-1: Formantes de la semi-consonante [j] y la vocal [i]

		Hombres	Mujeres
[j]	F1	337 Hz ⁽⁵²⁾	416 Hz ⁽⁹³⁾
	F2	2142 Hz ⁽¹⁶³⁾	2575 Hz ⁽¹⁷²⁾
[i]	F1	326 Hz	387 Hz
	F2	2244 Hz	2676 Hz

Tabla 6-2: Formantes de la semi-consonante [w] y la vocal [u]

		Hombres	Mujeres
[w]	F1	323 Hz ⁽⁵⁷⁾	403 Hz ⁽⁶¹⁾
	F2	1035 Hz ⁽²²⁹⁾	1081 Hz ⁽²⁵⁹⁾
[u]	F1	397 Hz	470 Hz
	F2	896 Hz	935 Hz

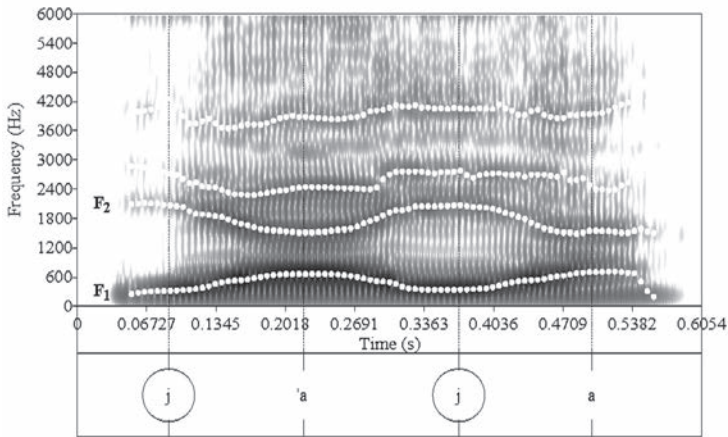
Las semi-consonantes son conocidas por la dificultad en determinar sus fronteras con exactitud. Son segmentos transitorios de o hacia una vocal adyacente (Olive, Greenwood *et al.* 1993; Raphael, Borden *et al.* 2007). Como segmentos transitorios, usualmente no poseen un periodo donde sus formantes aparezcan estables. En shipibo, es también evidente este comportamiento de las glides, aunque la semi-consonante [j] puede alternativamente mostrar un periodo estable para sus formantes si se realiza como [j̥], una fricativa palatal sonora que posee una ligera cantidad de fricción (algunos hablantes pueden realizarla como una fricativa sibilante sonora [ʒ]). La realización de la semi-consonante /j/ como una fricativa [j̥] sucede frecuentemente.

Otra diferencia entre las semiconsonantes y las vocales altas es que las primeras muestran menos energía acústica, es decir, aparecen ligeramente más claras en los espectrogramas que las vocales. Esto se debe a que las semi-consonantes tienden a tener una constricción más angosta en la cavidad oral que las vocales, especialmente cuando se realizan como una fricativa. Veamos algunos ejemplos de las semi-consonantes [j] y [w] en espectrogramas.

El espectrograma de la Figura 6-1 ilustra la palabra [ˈja.ja] ‘tía materna de una mujer’, la cual contiene dos glides palatales [j]: una al inicio de la palabra y otra en posición intervocálica. En términos de formantes, la [j] es muy similar a la vocal [i]. Tiene un F1 muy bajo (lo cual indica que es un segmento alto) y

un F2 muy alto (lo que indica que es un segmento articulado en la parte frontal de la cavidad oral). En la Figura 6-1, los movimientos de los formantes F1 y F2, conforme las glides [j] van hacia las vocales [a], son bastante conspicuos. No he hecho ningún intento de determinar las fronteras de las glides. En lugar de eso, solo he indicado la posición de cada glide y vocal colocando su símbolo fonético en el medio del segmento. En este ejemplo en particular, ambas semi-consonantes /j/ han sido realizadas fonéticamente como una fricativa palatal sonora, [j], lo cual les da una clara calidad consonántica. Los formantes aparecen en la Figura como una línea horizontal de pequeños círculos blancos.

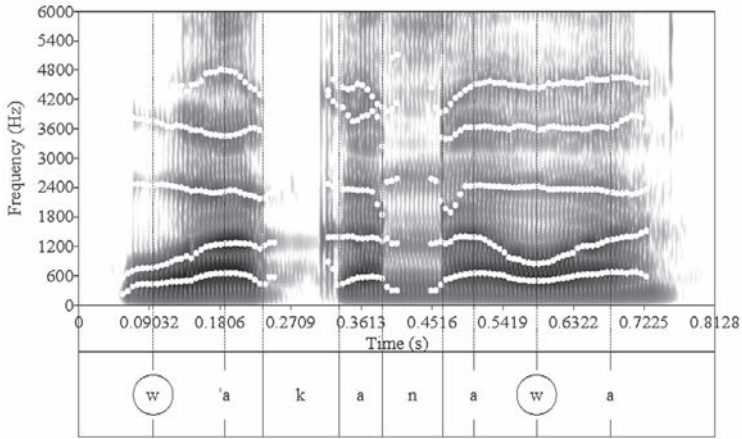
Figura 6-1: Espectrograma de ['ja.ja] 'tía materna de una mujer' (realizado como [j̥a.ja])



El espectrograma de la Figura 6-2 ilustra la palabra ['wa.kɑ.nɑ.wɑ] (aglomeración de peces migrando de una zona del río a otra). Este compuesto está formando por las palabras ['wa.kɑ] 'pez' y ['nɑ.wɑ] 'extraño, foráneo' y posee dos glides labio-dorsales: [w]. Como en el caso de [j] en la Figura 6-1, una ocurre al inicio de la palabra y la segunda, en un ambiente intervocálico. La glide [w] se caracteriza por tener un F1 y un F2 muy bajos (lo cual indica que el segmento es alto y posterior). Es similar a la vocal alta redondeada [u] pero con una constricción más angosta. Como en la Figura 6-1, la glide [w] aparece rodeada por la vocal [ɑ].

6.2 SOBRE LA FONÉTICA DE LAS VOCALES ALTAS Y LAS GLIDES

En el nivel subyacente, shipibo posee tanto las vocales /i/ y /u/ como las semi-consonantes /j/ y /w/. Fonéticamente, sin embargo, el panorama es más complejo. Las vocales /i/ y /u/ además de realizarse como segmentos vocálicos, también

Figura 6-2: Espectrograma de la palabra [^hwa.kɑ.nɑ.wɑ]

pueden ocurrir como las semi-vocales [ɨ] y [ʉ] o las semi-consonantes [j] y [w]. En contraste, las semi-consonantes subyacente /j/ y /w/ solo pueden ocurrir como [j]/[j̥] y [w], respectivamente.

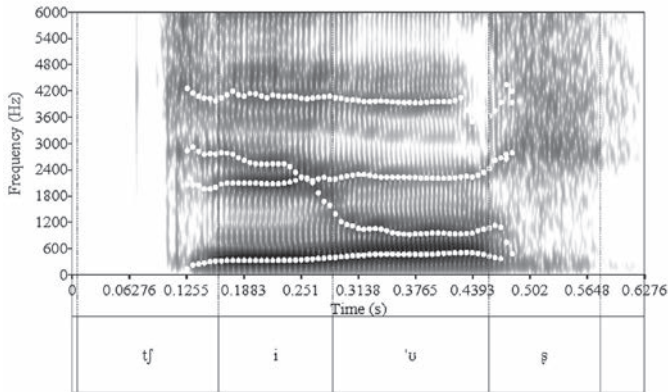
Distingo entre las semi-vocales ([ɨ, ʉ]) y las semi-consonantes ([j, w]) en que las semi-vocales son glides que mantienen una calidad muy similar a las de las vocales mientras que las semi-consonantes son glides con una constricción mucho más estrecha en la cavidad oral y en consecuencia, tienden a tener una calidad más consonántica. A diferencia de las semi-vocales, las semi-consonantes del shipibo pueden mostrar una ligera cantidad de fricción, particularmente cuando /j/ se realiza como la fricativa palatal sonora [j]. Uso el símbolo [j] para representar tanto la realización como una semi-consonante, [j̥] y la realización como una fricativa, [j̥].

6.2.1 Fuera de la fonología: las vocales /i/ y /u/ como semi-vocales

Empecemos por presentar la realización de las vocales /i/ y /u/ como las semi-vocales [ɨ] y [ʉ]. Este es un fenómeno fonético que tiende a ocurrir cuando las vocales inacentuadas [i] o [u] aparecen adyacentes al núcleo de otra sílaba. Por ejemplo, esto ocurre en la palabra [tʃi.ʉʂ] ‘de nalga blanquizca’. La vocal inacentuada [i] opcionalmente puede ocurrir como la semi-vocal [ɨ]: [tʃɨʉʂ]. Sin embargo, usualmente cuando a los hablantes se les pide pronunciar la palabra en aislamiento, la vocal [i] mantienen su estatus silábico y la palabra se realiza como [tʃi.ʉʂ]. Los espectrogramas de la Figura 6-3 y la Figura 6-4 ilustran estas dos posibilidades.

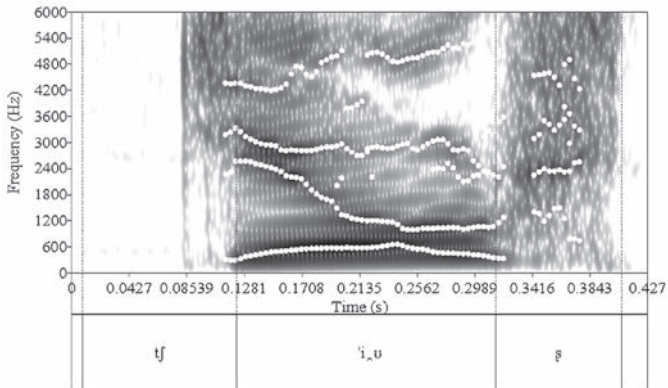
En el espectrograma de la Figura 6-3, los formantes de la vocal [i] son bastante estables durante los primeros 95 ms de la vocal. En este caso, la vocal [i] ha evitado convertirse en una semi-vocal. Ha logrado mantenerse como el núcleo de su propia sílaba. Es solo después del periodo estable de [i] que F2 empieza a bajar hacia la vocal acentuada [u].

**Figura 6-3: Espectrograma de la palabra / tʃiʊʂ / → [tʃi.ʊʂ]
‘de nalga blanquizca’**



En contraste, en el espectrograma de la Figura 6-4, la vocal [i] de [tʃi.ʊʂ] ocurre como una semi-vocal, [i̯]. La palabra se ha reducido en el nivel fonético a una sola sílaba: [tʃiʊʂ]. El formante F2 de [i̯] baja desde el inicio sin mostrar un periodo estable. Es solo cuando F2 alcanza el valor para la vocal acentuada [u] que se observa un periodo estable en los formantes.

**Figura 6-4: Espectrograma de la palabra / tʃiʊʂ / → [tʃiʊʂ]
‘de nalga blanquizca’**



6.2.2 Las vocales /i/ y /u/ como glides en posición inicial de palabra

Shipibo evita acentuar la vocal inicial de palabras que empiezan con dos vocales adyacentes. En el caso de las vocales /i/ y /u/ en posición inicial de palabra, este fenómeno tiende a volverlas glides. Los datos en (22) listan las realizaciones de todas las palabras que empiezan con dos vocales reportadas por Lorient, Lauriault *et al.* 1993. He incluido la transcripción ortográfica del diccionario en letra cursiva junto a la glosa. En la ortografía, la vocal acentuada se indica por una tilde aguda.

Es tentador generalizar este fenómeno en términos de palabras que empiezan en las vocales /i/ y /u/. La lista en (22) revela que la mayoría de las palabras que empiezan con dos vocales tienen ya sea /i/ o /u/ como su segmento inicial. Una excepción es la existencia de dos palabras de este tipo cuyo segmento inicial no es ni /i/ ni /u/. Esas palabras son: [a.¹u] (especie de árbol) y [a.¹i.ʔi.ti] ‘eructar’. No es claro si el patrón acentual de esas dos palabras (es decir, el acento principal sobre la segunda sílaba) es solo una coincidencia o si refleja un patrón más general de evitar acentuar la vocal inicial. No he encontrado ninguna palabra que empiece con dos vocales y que el acento aparezca sobre la primera vocal.

(22) Palabras del shipibo que empiezan con dos vocales

a.	[a. ¹ u]	<i>áo</i> (esp. de árbol)
b.	[a. ¹ i.ʔi.ti]	<i>áéiti</i> ‘eructar’ (VZ2)
c.	[¹ ja] ~ [‘ja] ~ [i. ¹ a]	<i>ía</i> ‘piojo’
d.	[¹ ja.maʃ] ~ [‘ja.maʃ] ~ [i. ¹ a.maʃ]	<i>iamāsh</i> ‘es decir, así’
e.	[¹ ja.ti] ~ [‘ja.ti] ~ [i. ¹ a.ti]	<i>íati</i> ‘cargar algo en los hombros’ (INF)
f.	[¹ jín] ~ [‘jín] ~ [i. ¹ ín]	<i>ién</i> ‘de sien calva’
g.	[¹ iú.ti] ~ [‘jü.ti] ~ [i. ¹ ü.ti]	<i>íoti</i> ‘tomar, dirigir’ (INF)
h.	[¹ ua] ~ [‘wa] ¹	<i>óa</i> ‘ese, ése, aquel, aquél’
i.	[¹ ui:] ~ [‘wi:] ~ [ü. ¹ i]	<i>óe</i> ‘inusualmente grande’
j.	[¹ úi:] ~ [‘wi:] ~ [ü. ¹ i]	<i>ói</i> ‘lluvia’
k.	[¹ ui.na] ~ [‘wi.na] ~ [ü. ¹ i.na]	<i>óina</i> ‘ver’ (PP2)
l.	[¹ úi.pua] ~ [‘wi.pua] ~ [ü. ¹ i.pua]	<i>óipoa</i> (esp. de caracol)
m.	[¹ úi.fü.dzu] ~ [‘wi.fü.dzu] ~ [ü. ¹ i.fü.dzu]	<i>óishoro</i> (esp. de pájaro)
n.	[¹ úi.ti] ~ [‘wi.ti] ~ [ü. ¹ i.ti]	<i>óiti</i> ‘quemar (cerámica cruda)’ (INF)

¹ No he encontrado ninguna instancia de este elemento funcional realizado como [ü.¹a].

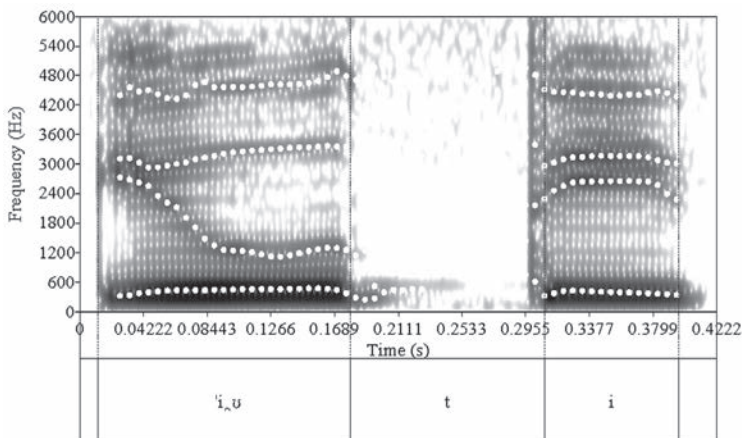
Es importante notar que con la excepción de (22.a-b) y (22.e-f), reportadas por Lorient, Lauriault *et al.* 1993 con el acento sobre la segunda vocal, las otras palabras parecen haber experimentado un cambio en la posición del acento de la sílaba inicial a la segunda sílaba. Por ejemplo, ‘óipoa’ (esp. de caracol) y ‘óti’ ‘tomar, dirigir’ han sido reportadas por Lorient, Lauriault *et al.* 1993 con el acento sobre la vocal inicial, [u] y [i], respectivamente. Sin embargo, los colaboradores shipibos entrevistados para este estudio acentúan la vocal [i] en /uipua/ → [ˈʷi.ɸu̘a] ~ [wi.ɸu̘a], y la vocal [u] en /iu -ti/ → [ˈi̘u.ti] ~ [j̘u.ti].

El acento cae sobre la segunda vocal inclusive en los casos donde el segmento inicial se realiza como un segmento vocálico: /uipua/ → [u.ˈi.ɸu̘a] y /iu -ti/ → [i.ˈu.ti]. Se debe recordar que aunque el diccionario de Lorient, Lauriault *et al.* fue publicado en 1993, los datos fueron originalmente recolectados entre 1940 y 1960, lo cual parece indicar que el evitar acentuar la vocal inicial de las palabras que empiezan con dos vocales podría ser un cambio reciente en la lengua.

Como se muestra en los datos en (22), la mayoría de veces, las vocales iniciales de palabra /i/ y /u/ se realizan como semi-vocales cuando ocurren seguidas por otra vocal (por ejemplo, /iu -ti/ → [ˈj̘u.ti] ‘tomar, dirigir’ (INF)). Sin embargo, también se pueden realizar como semi-consonantes (por ejemplo, /iu -ti/ → [ˈj̘u.ti]), e inclusive como segmentos vocálicos inacentuados (por ejemplo, /iu -ti/ → [i.ˈu.ti]). Los espectrogramas de la Figura 6-5 a la Figura 6-7 ilustran cada uno de estos casos, los cuales ocurren en variación libre.

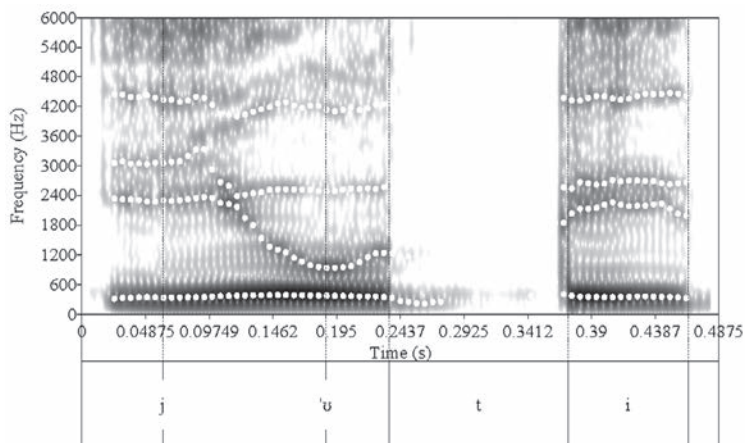
En el espectrograma de la Figura 6-5, se puede ver claramente la vocal inicial de /iu -ti/ realizada como una semi-vocal: un segmento vocálico sin fricción, [i̘], cuyo F2 esta en constante movimiento hacia el F2 de la vocal acentuada [u]. Ese movimiento le toma alrededor de 60 ms.

Figura 6-5: Espectrograma de /iu -ti/ realizado como [ˈi̘u.ti] ‘tomar, dirigir’



El espectrograma de la Figura 6-6 muestra la misma palabra que el espectrograma anterior. Sin embargo, esta vez, la vocal inicial /i/ de /iʊ-ti/ posee una calidad consonántica. Aunque la transcribo como una semi-consonante palatal [j] en el espectrogramas, se puede observar la presencia de una ligera cantidad de fricción desde los 3000 Hz para arriba, lo cual indica que el segmento se ha realizado como una fricativa palatal sonora [j].

**Figura 6-6: Espectrograma de /iʊ-ti/ → [ˈjʊ.ti] ‘tomar, dirigir’
(realizado como [ˈjʊ.ti])**

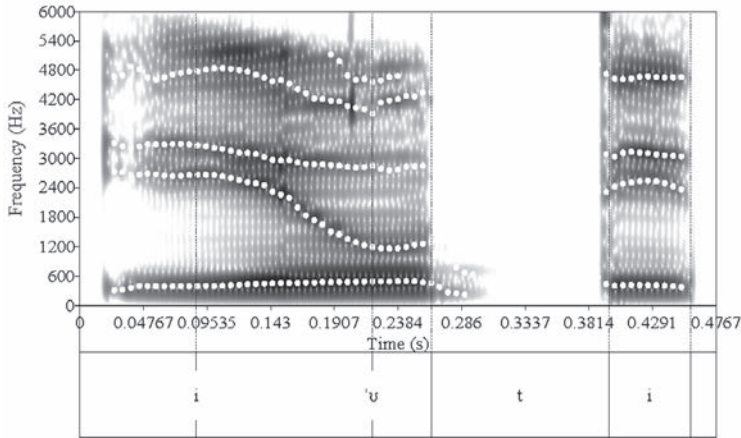


Obsérvese que a diferencia de la Figura 6-5 y la Figura 6-6, esta vez en la Figura 6-7, la vocal inicial /i/ de /iʊ-ti/ ocurre como un segmento completamente vocálico. Aunque la vocal inicial [i] de la Figura 6-7 casi tiene la misma duración y tiene formantes estables como [j] en la Figura 6-6, no muestra ningún tipo de fricción y claramente posee un timbre vocálico.

Antes de continuar, es importante mencionar que el evitar acentuar la sílaba inicial de palabras que empiezan con dos vocales no afecta a palabras que empiezan con una vocal. De este modo, el shipibo posee palabras con acento inicial como [ˈi.βʊ] ‘dueño’, [ˈʊ.ta] ‘sombra’, [ˈa.ni] ‘grande’ y [ˈi.pa] ‘tío paterno’. Tampoco afecta a palabras que tienen dos vocales en contacto siguiendo a un arranque inicial de palabra. En una palabra como [ˈpi.a] ‘flecha’, el acento siempre ocurre sobre la vocal [i]. De hecho, shipibo tiene otra palabra cuya única diferencia observable es que el acento cae sobre la segunda vocal: [pi.ˈa] ‘sobrino, sobrina’.²

² La palabra [ˈpi.a] ‘flecha’ subyacentemente es representada como /pia/ pero la palabra [pi.ˈa] ‘sobrino, sobrina’ viene de /piak/.

Figura 6-7: Espectrograma de /iʊ -ti/ → [i.ˈu.ti] ‘tomar, dirigir’



La existencia de palabras como [ˈti.ta] ‘madre’ o [na.βi] ‘pared interior’ aporta evidencia adicional de que la asignación de acento en shipibo no es sensible a la sonoridad de las vocales. La ocurrencia del acento principal sobre la segunda vocal en palabras como [i.ˈu.ti] ‘tomar, dirigir’ o [i.ˈa.ti] ‘llevar algo sobre los hombros’ parece solo evitar que el acento principal aparezca adyacente al margen izquierdo de las palabras, un problema que se evita cuando la palabra posee un arranque inicial ([ˈpi.a] ‘flecha’). Sin embargo, si la palabra no tiene un arranque inicial y la segunda sílaba tiene un arranque pero no una coda que pueda atraer el acento, entonces el acento primario es forzado a ocurrir al inicio de la palabra ([ˈu.ta] ‘sombra’).

6.3.2 Las vocales /i/ y /u/ como semi-consonantes en contextos intervocálicos

Las vocales /i/ y /u/ se realizan obligatoriamente como las semi-consonantes, [j] y [w], cuando ocurren en un ambiente intervocálico. Un ejemplo de este fenómeno se puede observar en la alternancia que presentan las raíces verbales /pui/ ‘defecar’ y /iʊ/ ‘tomar, dirigir’. Cuando estas raíces están seguida por el sufijo /-ti/ (INF), usualmente se realizan como [ˈpui.ti] y [ˈju.ti] ~ [ˈji.ti] ~ [i.ˈu.ti]. El espectrograma de la Figura 6-8 muestra la realización de /pui -ti/ y los espectrogramas de la Figura 6-5 a la Figura 6-7 muestran las posibles realizaciones de /iʊ -ti/.

Sin embargo, cuando el sufijo de infinitivo /-ti/ se reemplaza por el sufijo /-a/ (PP2), la /i/ de la raíz /pui/ obligatoriamente ocurre como una semi-consonante palatal [j] debido a que ahora la vocal /i/ aparece en un contexto intervocálico: /pui -a/. Esto se muestra en el espectrograma de la Figura 6-9.

Figura 6-8: Espectrograma de /pui -ti/ → [p̥u.i.ti] ‘defecar’
(realizado como: [p̥u̇i.ti])

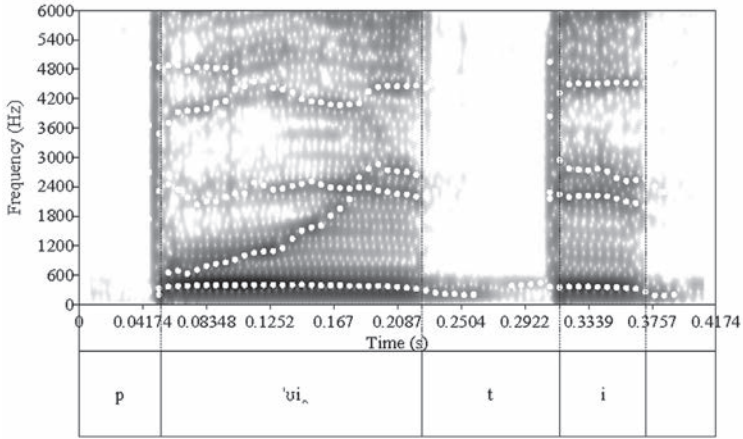
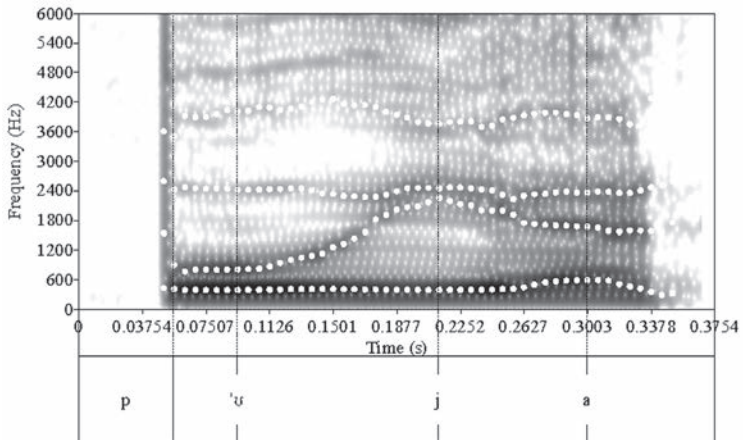
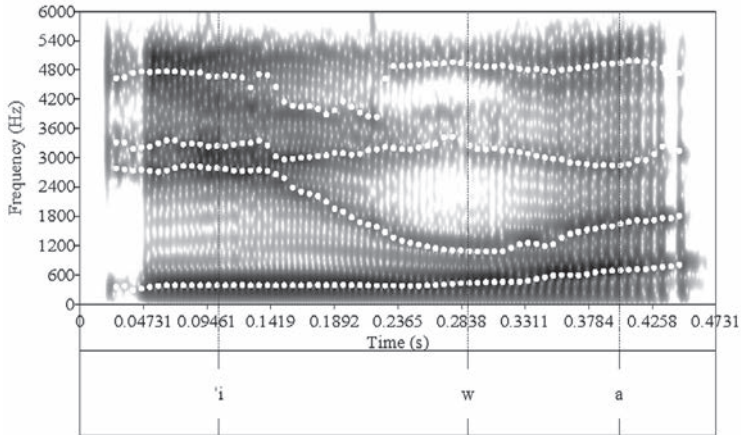


Figura 6-9: Espectrograma de /pui -a/ → [p̥u.ja] ‘defecar’ (PP2)



Lo mismo ocurre con la vocal /u/ de la raíz /iu -a/. Ésta se realiza obligatoriamente como la semi-consonante [w]: [i.wa]. Esto se muestra en el espectrograma de la Figura 6-10 (para más información sobre este fenómeno, ver la sección 9.3.1.1)

Figura 6-10: Espectrograma de /iʊ -a/ → [i.wa] ‘tomar, dirigir’ (PP2)



6.3 RESUMEN

En este capítulo, he presentado un estudio de los segmentos glides del shipibo. Como parte de este estudio, se ha determinado los valores promedio de los formantes de F1 y F2 de esos segmentos. En el caso de la glide palatal [j], los valores de F1 y F2 son muy similares a aquellos que se encontraron para la vocal [i]. El F1 y F2 de la glide labio-dorsal [w] corresponden a una vocal alta posterior [u]. Las glides se distinguen de las vocales altas por su naturaleza transitoria, la cual hace difícil poder determinar con exactitud sus fronteras y su duración.

He presentado evidencia espectrográfica que indica que las glides, especialmente la glide palatal, pueden realizarse como segmentos consonánticos fricativos (/j/ → [j]). Cuando esto sucede, usualmente ocurren acompañados por algún grado de fricción y una caída notable de la intensidad. Además, hemos mostrado que las glides y las vocales afectan mutuamente su estructura de formantes. Sin embargo, estos efectos son menos fuertes cuando los segmentos pertenecen a sílabas iniciales y acentuadas.

También se presentó evidencia fonética que muestra que el shipibo tiene dos tipos de glides: aquellas que provienen de las vocal subyacentes /i/ y /ʊ/ y otras que provienen de las semi-consonantes subyacente /j/ y /w/. Se brindó abundante evidencia de esto en los espectrogramas sobre las realizaciones de las glides en diferentes ambientes. Los contextos que se examinaron fueron los siguientes: glides que provienen de vocales altas subyacentes que aparecen adyacentes a otra vocal; y glides intervocálicas e iniciales de palabra que provienen de vocales altas subyacentes. También se ha presentado evidencia espectrográfica de las realizaciones de glides subyacentes.

De este modo, hasta ahora, nos hemos concentrado en las propiedades acústicas de las vocales y consonantes del shipibo. El siguiente capítulo examina las propiedades fonéticas de los constituyentes silábicos que forman los segmentos cuando se agrupan. En particular, me concentraré en determinar las propiedades duracionales de las rimas y codas de sílabas (in)acentuadas y (no)inicial.

7. LAS SÍLABAS Y LAS PROPIEDADES DURACIONALES DE LAS RIMAS

Considero la sílaba un concepto principalmente de naturaleza fonológica; es decir, el agrupamiento abstracto de segmentos en un constituyente fonológico. Sin embargo, hay importantes propiedades acústicas que se correlacionan con la posición que ocupan los segmentos dentro de la sílaba. En los capítulos anteriores, se ha estudiado los segmentos del shipibo en relación con ciertas posiciones dentro de la sílaba. De este modo, por ejemplo, el capítulo sobre vocales estudia los núcleos vocálicos posibles del shipibo y sus características acústicas (es decir, estructura de formantes y duración). Los capítulos sobre consonantes presentan la caracterización acústica de las consonantes en posición de arranque (es decir, sus propiedades espectrográficas, espectrales y duracionales). Esos capítulos también han aportado información sobre las propiedades acústicas presentes en la transición de vocales a consonantes y viceversa.

Sin embargo, hay preguntas que han sido dejadas sin respuesta hasta ahora, particularmente con respecto a los patrones duracionales de las codas y de las rimas de las sílabas. Este capítulo presenta y discute esos patrones: la duración de los segmentos codas, la duración de las vocales cuando ocurren en sílabas cerradas y la duración de las rimas de las sílabas.

7.1 ESTRUCTURA DE LA SÍLABA Y SUS PROPIEDADES DURACIONALES

7.1.1 Duración de las codas

Solo las sibilantes [s, ʃ, ʂ] y la nasal [n] pueden ocurrir como codas en shipibo. Los arranques muestran una duración más larga que la de las codas. De este modo, los arranques nasales tienden a durar en promedio entre 80 ms y 90 ms mientras que

las codas nasales usualmente duran entre 20 ms y 30 ms menos (véase la sección 5.1 y la sección 5.2.1). La sibilantes tienden a durar aproximadamente 110 ms cuando ocurren como arranques (véase la sección 4.1). Como codas, la duración de las sibilantes es similar a la que se encuentra en las nasales en posición de coda. Duran aproximadamente 65 ms. Esto se muestra en la Tabla 7-1.

Los resultados en la Tabla 7-1 se basan en la medición de aproximadamente 2000 sibilantes en posición de coda en la primera o segunda sílaba de palabras trisilábicas pronunciadas dentro de la frase usada en el estudio. Todas las codas sibilantes ocurrieron seguidas de una consonante oclusiva. Las mediciones incluyen codas de sílabas iniciales acentuadas e inacentuadas y codas de segundas sílabas acentuadas.¹

Ya que las tres codas sibilantes (es decir, [s, ʃ, ʂ]) presentan duraciones similares, la Tabla 7-1 muestran sus promedio agrupados. La vocal precedente se indica en la primera columna. La Tabla 7-1 también separa los promedios de acuerdo a si la coda sibilante pertenece a la primera-segunda sílaba acentuada o a una sílaba inicial inacentuada.

Tabla 7-1: Duración promedio de las sibilantes en posición de coda (en segundos)

	Sílaba Inicial Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[i]__.	0,070 (0,017)	0,073 (0,013)	0,076 (0,017)
[ị]__.	0,063 (0,019)	0,065 (0,012)	0,070 (0,018)
[u]__.	0,060 (0,018)	0,067 (0,013)	0,072 (0,015)
[a]__.	0,066 (0,017)	0,065 (0,012)	0,075 (0,017)

Como punto de referencia, repito en la Tabla 5-6 las mediciones de duración de las codas nasales presentadas en la sección 5.2.1.

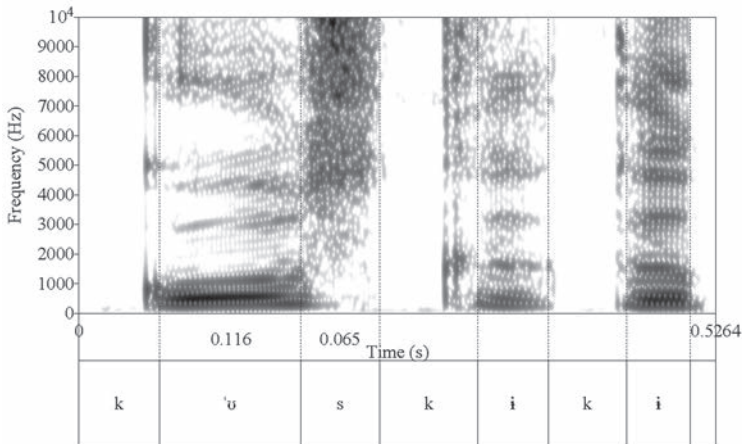
¹ En general shipibo no tiene segundas sílabas cerradas inacentuadas. Las segundas sílabas cerradas usualmente ocurren acentuadas debido a los principios de asignación del acento. La única excepción es cuando la sílaba inicial tiene una vocal larga como en el caso de una raíz monosilábica con una vocal larga que es seguida por sufijos que forman una sílaba cerrada. En ese caso, encontramos la segunda sílaba cerrada inacentuada (por ejemplo, /tʃiː -dʒɪs/ → [tʃiː.dʒɪs] ‘sólo el fuego’ (/tʃiː/ ‘fuego’ y /-dʒɪs/ ‘sólo’)). Sin embargo, hay tan pocos de estos casos que no los he incluido en las mediciones de la duración de las sibilantes en posición de coda.

Tabla 7-2: Duración promedio de la nasal [n] en posición de coda (en segundos)

	Sílaba Inicial Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
i_.	0,062 ^(0,013)	0,062 ^(0,015)	0,059 ^(0,018)
í_.	0,056 ^(0,013)	0,056 ^(0,011)	0,061 ^(0,015)
ü_.	0,063 ^(0,020)	0,059 ^(0,011)	0,062 ^(0,015)
a_.	0,056 ^(0,013)	0,062 ^(0,014)	0,060 ^(0,017)

Los espectrogramas de la Figura 7-1 a la Figura 7-3 ilustran ocurrencias de las sibilantes en los tres ambientes que se tomaron en cuenta al medir la duración de las codas. El espectrograma de la Figura 7-1 muestra la coda sibilante en una sílaba acentuada inicial conforme ocurre en la palabra [ˈkus.ki.ki] (/kuskí/ ‘romper la mandíbula de alguien’ /-ki/ CMPL). He indicado en la Figura 7-1 la duración de la coda sibilante así como también la duración de la vocal de la sílaba inicial.

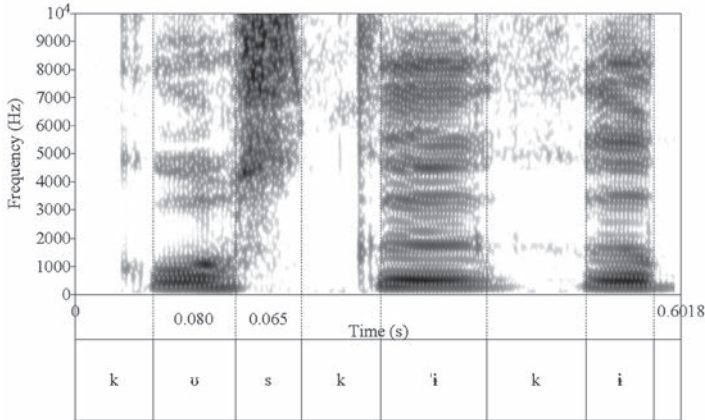
Figura 7-1: Espectrograma de [ˈkus.ki.ki] ‘romper la mandíbula de alguien’ (CMPL)



El espectrograma de la Figura 7-2 muestra una coda sibilante en la sílaba inacentuada inicial de la palabra [kus.ˈki.ki] ‘romperse la mandíbula’ (CMPL). Esta palabra contiene la raíz /kuskí/ ‘romper la mandíbula’ y el sufijo /-ki/ como en la palabra [ˈkus.ki.ki] (la cual se muestra en la Figura 7-1). También contiene el sufijo /-t/ (INTVZ) (/kuskí -t -ki/ → [kus.ˈkit.ki] → [kus.ˈki.ki]). Obsérvese que la coda sibilante mantiene la misma duración si ocurre en una sílaba acentuada

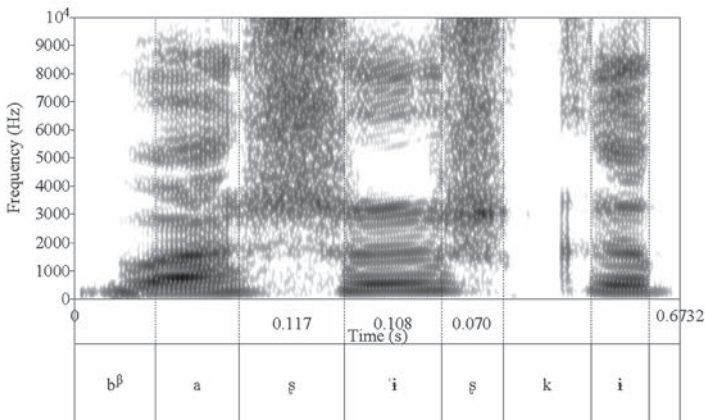
o inacentuada. La vocal muestra diferencias duracionales que se correlacionan con la presencia o ausencia del acento. Como se espera, las vocales son más largas en sílabas acentuadas que en las inacentuadas.

Figura 7-2: Espectrograma de la palabra [küs.'ki.ki] ‘romperse la mandíbula’ (CMPL)



El espectrograma de la Figura 7-3 ilustra la ocurrencia de una sibilante coda en una segunda sílaba inacentuada: [b^βa.ʃiʃ.ki] (/b^βaʃiʃ/ ‘hablar en voz baja’ y /-ki/ CMPL). Como en los casos anteriores, la duración de la coda sibilante fluctúa alrededor de 65 ms. En esta ocurrencia en particular, la coda dura 70 ms. La Figura 7-3 también nos permite comparar la duración de una sibilante en posición de arranque con la duración de una sibilante en posición de coda. Las sibilantes en arranque son más largas que sus contrapartes en codas.

Figura 7-3: Espectrograma de la palabra [b^βa.ʃiʃ.ki] ‘hablar en voz baja’ (CMPL)



7.1.2 duración de las vocales cortas en sílabas cerradas

La sección 2.3 presentó la duración de las vocales cortas en sílabas abiertas y la sección 5.2, la duración de las vocales cortas en sílabas cerradas por una coda nasal. Los promedios de duración de estas vocales los repito en la Tabla 7-3 y la Tabla 5-5, respectivamente, para que puedan servir como punto de comparación con la Tabla 7-5.

Tabla 7-3: Duración de vocales en sílabas abiertas (en segundos)

	Sílaba Inicial Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[i]	0,094 ^(0,021)	0,109 ^(0,027)	0,066 ^(0,022)
[ɪ]	0,098 ^(0,020)	0,112 ^(0,031)	0,068 ^(0,026)
[ū]	0,101 ^(0,023)	0,113 ^(0,031)	0,068 ^(0,022)
[a]	0,111 ^(0,024)	0,127 ^(0,033)	0,080 ^(0,023)

Tabla 7-4: Duración de vocales en sílabas cerradas por una nasal (en segundos)

	Sílaba Inicial Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[ĩ]	0,122 ^(0,018)	0,121 ^(0,019)	0,090 ^(0,024)
[ɨ]	0,121 ^(0,012)	0,123 ^(0,018)	0,090 ^(0,014)
[ũ]	0,119 ^(0,012)	0,122 ^(0,020)	0,087 ^(0,013)
[ã]	0,129 ^(0,020)	0,126 ^(0,020)	0,091 ^(0,014)

La Tabla 7-5 ofrece la duración promedio de las vocales cortas (orales) en sílabas cerradas por una coda sibilante. Los promedios se obtuvieron del mismo conjunto de datos usados para obtener los resultados mostrados en la Tabla 7-1.

Primero, como es de esperarse, las vocales inacentuadas muestran una duración más corta que las vocales acentuadas. La segunda observación es que las vocales nasalizadas (Tabla 5-5) tienden a ser en general ligeramente más largas que las vocales orales. Ésta es una tendencia que se ha observado en otras lenguas (Hajek 1997). Obsérvese también que las vocales de las sílabas acentuadas cerradas (Tabla 5-5 y Tabla 7-5) duran virtualmente lo mismo ya sea en sílabas iniciales o

en segunda sílabas. Esta observación contrasta con los resultados de la duración de las vocales acentuadas en sílabas abiertas (Tabla 7-3). En sílabas abiertas, las vocales cortas inacentuadas en sílabas abiertas, los promedios de sus duraciones bordean los 70 ms. En el mismo tipo de sílabas, las vocales cortas acentuadas en segundas sílabas tienden a ser ligera pero consistentemente más largas que las vocales cortas acentuadas en sílabas iniciales (una diferencia de aproximadamente 15 ms). En el caso de las vocales cortas acentuadas de sílabas iniciales, sus promedios de duración se acercan a 98 ms. Pero para las vocales cortas acentuadas de sílabas segundas, los promedios de duración se acercan a 112 ms (excepto para la vocal [a] que muestra una duración aún más larga).

Tabla 7-5: Duración de vocales en sílabas cerradas por sibilante (en segundos)

	Sílaba Inicial Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[i]	0,099 ^(0,017)	0,104 ^(0,016)	0,074 ^(0,015)
[i̥]	0,107 ^(0,017)	0,109 ^(0,014)	0,072 ^(0,014)
[u]	0,111 ^(0,019)	0,115 ^(0,015)	0,079 ^(0,012)
[a]	0,119 ^(0,016)	0,120 ^(0,014)	0,088 ^(0,011)

La palabra ilustrada en la Figura 7-2 nos da un buen ejemplo de un caso de una coda latente en la segunda sílaba: /kuski -t -ki/ → [kus.ˈkɪt.ki] → [kus.ˈkɪ.ki] ‘romperse la mandíbula’ (INTVZ, CMPL).² Aunque, en la superficie, un segmento latente puede hacer que una sílaba parezca abierta, es claro que no lo es. No solo la fonología de la lengua la trata como cerrada pero la misma sílaba muestra comportamientos fonéticos que corresponden a las sílabas cerradas.

7.1.3 Duración de las rimas de las sílabas

Hasta ahora, me he concentrado en la duración de las vocales y las codas por separado. Veamos ahora la duración de las rimas (es decir, el constituyente formado

² Aunque los segmentos latentes no pueden aparecer en la superficie como codas porque no son sibilantes o nasales, ellos sí tienen influencia sobre la asignación del acento. Hacen que la segunda sílaba sea pesada. El acento es atraído a esta sílaba que en la superficie luce abierta. Una vez que se cambia el sufijo que sigue a la consonante latente por un sufijo que empieza en vocal, el segmento latente ocurre como arranque y el acento cae sobre la sílaba inicial ya que la segunda sílaba es ahora abierta: /kuski -t -a/ → [ˈkus.ki.ta] ‘romperse la mandíbula’ (INTVZ, PP2).

por la unión de un núcleo vocálico y una coda, si hay alguna). Para las vocales de sílabas abiertas, la duración de la rima es igual a la duración de sus vocales (véase la primera y tercera columna de la Tabla 7-3, la columna en gris en esa tabla se refiere a sílabas con una coda latente).

La Tabla 7-6 y la Tabla 7-7 muestran la duración promedio de las rimas en sílabas cerradas. La Tabla 7-6 indica la duración promedio de las rimas de sílabas cerradas por una coda sibilante (rimas VS). Los promedios de la Tabla 7-6 se obtuvieron sumando los resultados de la Tabla 7-1 y la Tabla 7-5. Los promedios de la Tabla 7-7 representan la duración promedio de las rimas de las sílabas cerradas por una coda nasal (rimas VN). Esos promedios se obtuvieron sumando la duración promedio de las vocales nasalizadas (Tabla 5-5) y la duración de las codas nasales (Tabla 5-6).

Tabla 7-6: Duración de las rimas de sílabas cerradas por una sibilante (en segundos)

	Sílaba Primera Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[i]	0,169	0,177	0,150
[í]	0,170	0,174	0,142
[u]	0,171	0,182	0,151
[a]	0,185	0,185	0,163

Tabla 7-7: Duración de las rimas de sílabas cerradas por una nasal (en segundos)

	Sílaba Primera Acentuada	Sílaba Segunda Acentuada	Sílaba Inacentuada
[ĩ]	0,184	0,183	0,149
[ĩ]	0,177	0,179	0,151
[ũ]	0,182	0,181	0,149
[ã]	0,185	0,188	0,151

Al combinar los resultados de la Tabla 7-6 y la Tabla 7-7, se revela una tendencia a que las rimas de las sílabas cerradas acentuadas tiendan a durar

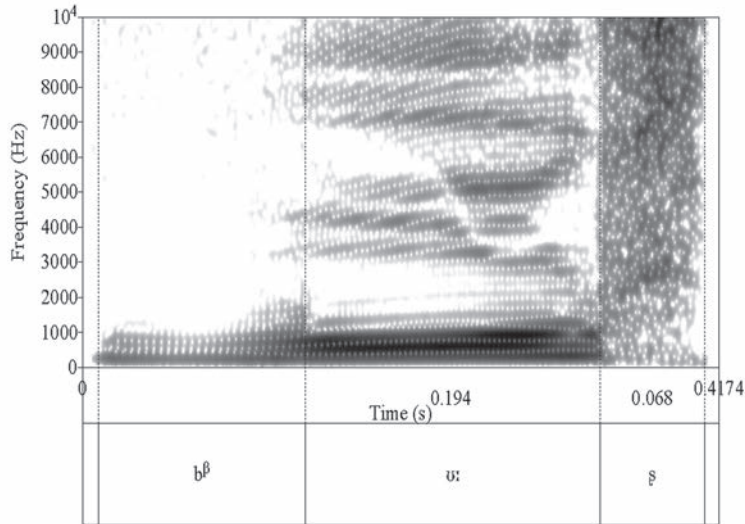
en promedio 180 ms mientras que las rimas de sílabas cerradas inacentuadas muestran una duración de 30 ms menos (es decir, duran aproximadamente 150 ms). El acento se perfila como el factor principal para explicar estos 30 ms de diferencia.

Shipibo tiene pocos casos de sílabas cerradas con vocales largas. Sé solo de quince raíces en total que contienen este tipo de sílaba. Once de ellas son raíces monosilábicas cerradas por una nasal (véase (23.a-k)), una es una raíz bisilábica cuya sílaba inicial está cerrada por una nasal (véase (23.l)) y tres son raíces monosilábicas cerradas por una sibilante (véase (23.m-o)).

(23) a.	[^l b ^β i:n]	‘pegajoso’
b.	[^l ki:n]	‘deseo’
c.	/ma:n -ti/ → [^l ma:n.ti]	‘caer en cantidad’ (INF)
d.	/mi:n -ti/ → [^l mi:n.ti]	‘enterrar’ (INF)
e.	[^l pa:n]	‘pan’ (del castellano [pan])
f.	/si:n -ʔiti/ → [^l si:n.ʔi.ti]	‘silbar’ (VZ2)
g.	/ʂu:n -ʔati/ → [^l ʂu:n.ʔa.ti]	‘hacer sonar (una bocina)’ (VZ1)
h.	/tsa:n -ʔati/ → [^l tsa:n.ʔa.ti]	‘hacer sonar una campana’ (VZ1)
i.	/dʒa:n -ti / → [^l dʒa:n.ti]	‘enviar’ (INF)
j.	[^l dʒu:n]	‘ruido’
k.	[^l ja:n]	‘garrapata’
l.	[^l dʒu:n.b ^β i]	(especie de árbol)
m.	[^l b ^β u:ʂ]	‘de cabello blanco’
n.	/ki:s -ti/ → [^l ki:s.ti]	‘mirar los muslos de alguien’ (INF)
o.	/nu:ʂ -ʔiti/ → [^l nu:ʂ.i.ti]	‘sorber ruidosamente’ (VZ2)

Una mirada a la duración de las codas de sílabas cerradas que contienen vocales largas revela que muestran la misma duración que las codas de sílabas cerradas con vocales cortas. Esto se puede observar en la Figura 7-4, la cual presenta el espectrograma de la palabra [^lb^βu:ʂ] ‘de cabello blanco’. La vocal larga dura 194 ms, lo cual es cercano a la duración promedio de una vocal larga [u:] en una palabra monosilábica: 205 ms (véase la sección 2.4). La misma observación puede hacerse con respecto a la coda, que dura 68 ms (la Tabla 7-1 indica que la duración promedio de una coda sibilante que sigue a una vocal [u] es 60 ms).

Figura 7-4: Una vocal larga en una sílaba cerrada por una coda sibilante



7.2 CODAS COMPLEJAS

Como en el caso de los arranques, las codas no son obligatorias. Aunque el shipibo abrumadoramente prefiere tener una sola consonante como coda de una sílaba, hay casos de sílabas con codas complejas.³ En estos casos, la coda compleja tiene dos consonantes: la primera siempre es una nasal y la segunda, una sibilante. Sé de una excepción a esta generalización: la palabra [ʰnh.nu] ‘allá muy lejos’ (véase el espectrograma en la sección 4.2). La sílaba inicial tiene una fricativa laríngea en vez de una sibilante como la segunda consonante de su coda compleja.

Las sílabas con codas complejas son raras en shipibo. La lista en (24) contiene todas las palabras que he encontrado con este tipo de sílabas en el diccionario de Lorient, Lauriault *et al.* (1993). Obsérvese que en todos los ejemplos, las codas complejas solo aparecen como parte de la sílaba inicial de la palabra y, aunque puede ser un accidente, no se encuentran codas complejas que tengan la sibilante palatal [ʃ]. Todas muestran ya sea [s] o [ʃ] (o incluso [h]) como la segunda consonante de la coda compleja.

³ En general, shipibo evita tener arranques complejos. la única excepción a esta generalización ocurre usualmente con hablantes jóvenes quienes tienen un alto grado de bilingüismo Shipibo-Castellano que otras generaciones. Ellos tratan de preservar los arranques complejos de las palabras que prestan del castellano (por ejemplo, [ʰpri.su] del castellano [ʰpre.so] *preso* - los hablantes más viejos tienden a pronunciar esta palabra como [ʰpi.dzi.su], sin el arranque complejo).

El origen de las palabras en (24) viene de un proceso morfológico antiguo que no parece ser productivo en la lengua a nivel sincrónico. Este mecanismo prefijaba a una raíz un afixo que se refiere a una parte del cuerpo. Las palabras en (24) contienen prefijos que terminan en [n]: /in-/ ‘en la sien, en la mejilla, en el lado’, /han-/ ‘en la lengua, en la boca’, /hin-/ ‘en la cola, en el pene’, /pün-/ ‘en el brazo’ (una lista exhaustiva de prefijos puede encontrarse en el capítulo 2). Como consecuencia de este proceso, la raíz que recibió el prefijo perdió una sílaba. En los casos en (24) (con la excepción de (24.g)), todas las raíces empiezan con una sibilante: [ʃa.b^βa.ti] ‘bostezar’, [ʃa.ti.ti] ‘cortar’, [ʃi.ki.ti] ‘romper algo’. Durante la operación morfológica, perdieron la primera vocal (por ejemplo, /hin- ʃati -ti/ → [hinʃ.ti.ti] ‘cortar en la cola’. A nivel sincrónico, ya no es posible identificar cuales fueron las raíces de (24.a-b). La palabra en (24.g), [ʰunh.nu] ‘allá muy lejos’ (véase la sección 4.2),⁴ la cual alterna con la forma [ʰun.nu], parece seguir un patrón diferente. Perteneció a un paradigma de adverbios ([ʰni.nu] ‘aquí’, [ʰu.nu] ‘allí’ y [ʰun.nu] ‘allá’).

- | | | |
|---------|----------------------------|---------------------------|
| (24) a. | [ʰins.pi] | ‘lado de la cabeza’ |
| b. | [ʰinʃ.pi] | ‘con un hueco en el lado’ |
| c. | [hanʃ.b ^β a.ti] | ‘abrir la boca’ (INF) |
| d. | [ʰhinʃ.ti.ti] | ‘cortar en la cola’ (INF) |
| e. | [ʰpünʃ.ti.ti] | ‘cortar el brazo’ (INF) |
| f. | [püns.ki.ti] | ‘romperse un brazo’ (INF) |
| g. | [ʰunh.nu] | ‘allá muy lejos’ |

El espectrograma de la Figura 7-5 ilustra una palabra con una coda compleja: [püns.ki.ti] ‘romperse un brazo’ (/pün- ʃiki -t -ti/ en.el.brazo- cortar - INTVZ -INF). La palabra fue pronunciada por una hablante mujer. La vocal de la sílaba inicial es bastante nasalizada, [ũ], esto se puede ver por lo alto que es su F1, alrededor de 639 Hz. Se puede ver la coda nasal que sigue a la vocal nasalizada por la presencia del formante nasal (F_n), típico de las consonantes nasales (véase la sección 5.1.1). En este caso en particular, el F_n de la coda nasal tiene una frecuencia de 279 Hz.

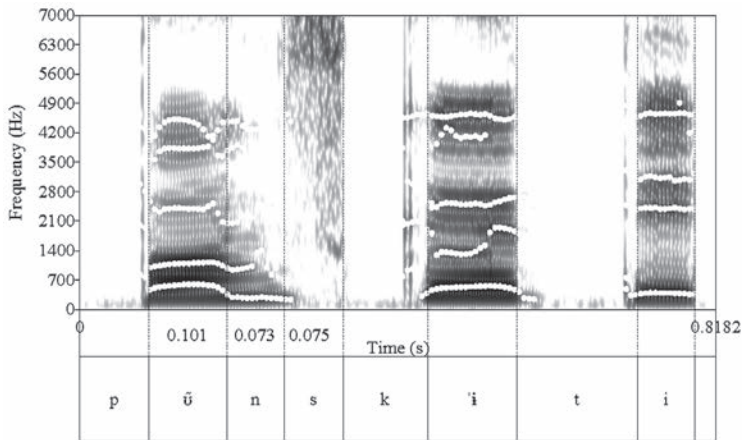
La realización fonética de la secuencia /Vns/ varía considerablemente (V se refiere a una vocal). Los hablantes muestran diferentes grados de coalescencia entre la vocal y la coda nasal. De este modo, por ejemplo, la sílaba con la coda

⁴ Shell (1975:41) reporta que el shipibo inserta una [h] en posición de coda para dar énfasis. Aunque no he encontrado casos de este fenómeno, es posible que la [h] que se observa en la palabra [ʰunh.nu] sea una muestra de lo que se refiere Shell, entendiendo “énfasis” en este caso como un tipo de intensificador de la distancia del adverbio: [ʰun.nu] ‘allá’ a [ʰunh.nu] ‘allá más lejos’.

compleja de una palabra como [pũns.¹ki.ti] puede realizarse fonéticamente entre [pũns] y [pũs]. Dependiendo de qué tanto la coda nasal coalesce con la vocal que le precede, la duración de la coda nasal puede verse afectada considerablemente. En la Figura 7-5, la coda nasal dura 73 ms (medida desde el comienzo de su F_n hasta el inicio de la sibilante que le sigue). Sin embargo, en otras ocurrencias, cuando la coda nasal se fusiona parcialmente con la vocal precedente, la coda nasal puede durar apenas 30 ms.

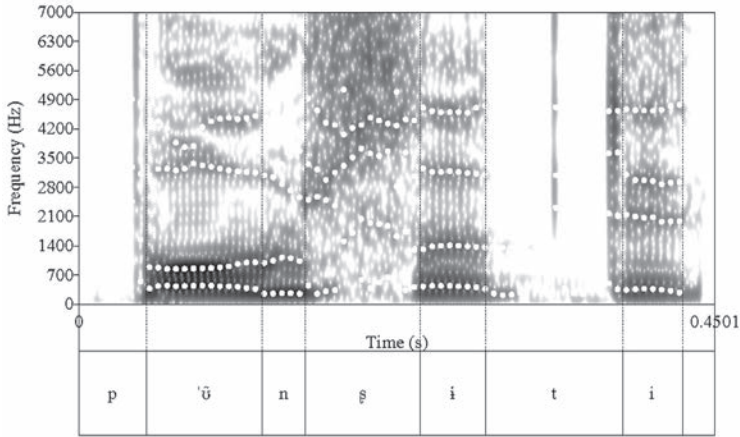
Es importante observar que la duración de las codas sibilantes es bastante estable inclusive dentro de codas complejas donde uno esperaría que se redujesen ya que están compartiendo la posición de coda con otra consonante. En la Tabla 7-1, el promedio de duración de una coda sibilante que sigue a la vocal [u] en una sílaba inacentuada es 72 ms. Como se puede observar en la Figura 7-5, la duración de la coda sibilante (75 ms) es bastante cercana a ese promedio.

Figura 7-5: [pũns.¹ki.ti] ‘romperse el brazo’ (realizado como [pũns.¹ki.ti])



Cuando el arranque que sigue a una coda compleja es la oclusiva coronal, [t], algunos hablantes tienden a disolver la coda compleja elidiendo la consonante [t]. Por ejemplo, la palabra [¹pũnʃ.ti.ti] ‘cortar el brazo’ algunas veces se realiza como [¹pũn.ʃi.ti] o [¹pũ.ʃi.ti]. Este fenómeno se ilustra en el espectrograma de la Figura 7-6. Relacionado a este proceso de simplificación de codas complejas, es importante mencionar que hay hablantes quienes han optado por abandonar completamente las consonantes nasales en codas complejas. Para ellos, por ejemplo, una palabra como [¹ins.pi] ‘lado de la cabeza’ se puede realizar como [¹is.pi]. En general, estos patrones indican la tendencia a evitar márgenes silábicos complejos.

Figura 7-6: [^hpũŋ.ʂi.ti] ‘cortar el brazo’ (realizado como [^hpũn.ʂi.ti])



7.3 RESUMEN

En este capítulo se ha determinado la duración promedio de las codas, vocales en sílabas cerradas y rimas de las sílabas del shipibo. Aunque me he concentrado mayormente en esos constituyentes de las sílabas cerradas, los resultados también se compararon a aquellos encontrados en los capítulos anteriores para las sílabas abiertas.

Con respecto a las codas, se encontró que en general las codas sibilantes y nasales tienden a durar entre 60 ms y 65 ms. En el capítulo 5, se estableció el promedio de duración de las codas nasales en 60 ms y en este capítulo, que las codas sibilantes duran aproximadamente 65 ms. En contraste, los mismos segmentos cuando ocurren como arranques muestran una duración más larga. Duran alrededor de 100 ms. La presencia o ausencia de acento no es un factor que influya la duración de las codas.

A diferencia de las codas, las vocales cortas sí muestran diferencias de acuerdo a si pertenecen a una sílaba acentuada o no. Las vocales cortas en sílabas acentuadas cerradas duran más que aquellas que ocurren en sílabas cerradas inacentuadas y las vocales en sílabas cerradas tienden a ser ligeramente más largas que las vocales en sílabas abiertas. En sílabas acentuadas abiertas, las vocales orales duran alrededor de 100 ms mientras que en sílabas cerradas por una sibilante, duran aproximadamente 110 ms. Cuando una coda latente cierra la sílaba, las vocales duran cerca de 115 ms y en sílabas cerradas por una coda nasal, las vocales (nasalizadas) duran en promedio 120 ms. El mismo patrón duracional se repite

en las sílabas inacentuadas. En sílabas abiertas sin acento, las vocales duran 70 ms. Cuando la sílaba inacentuada es cerrada por una sibilante, las vocales duran aproximadamente 80 ms y cuando aparecen cerradas por una coda nasal, las vocales (nasalizadas) duran cerca de 90 ms.

Las rimas reflejan las diferencias en la duración encontradas entre vocales acentuadas e inacentuadas. Cuando el acento principal cae sobre la primera o segunda sílaba cerrada, la rima de la sílaba tiende a durar 180 ms. En sílabas cerradas inacentuadas, la rima solo dura alrededor de 150 ms. En sílabas abiertas, la duración de las rimas es igual a la duración de la vocal de la sílaba. Estas rimas son en promedio 70 ms más cortas (110 ms en sílabas acentuadas, y 80 ms en sílabas inacentuadas).

Se discutió la existencia de codas complejas en Shipibo y se ilustró sus realizaciones con espectrogramas. Todas las codas complejas se caracterizan por ocurrir como parte de la sílaba inicial de la palabra. Las codas complejas siempre están formadas por una consonante nasal seguida de una fricativa. Sin embargo, la realización de las codas complejas aún revela la tendencia general del shipibo a evitarlas. En particular, cuando los arranques que siguen a la coda compleja es la oclusiva sorda [t], éste segmento tiende a elidirse y la fricativa de la coda compleja se resilabifica como el nuevo arranque de la siguiente sílaba.

El siguiente capítulo se aleja del tema de los segmentos con el fin de discutir el comportamiento de la frecuencia fundamental y su correlato, el tono perceptual (pitch), dentro de las palabras del shipibo.

8. LA FRECUENCIA FUNDAMENTAL, EL TONO PERCEPTUAL Y EL ACENTO

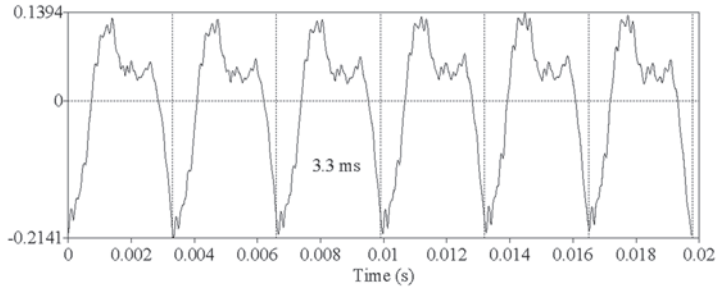
Este capítulo se focaliza en el comportamiento de la *Frecuencia Fundamental* (F_0) y su correlato, el tono perceptual (pitch), dentro del dominio de la palabra. En shipibo, el F_0 sirve como una de las claves fonéticas más importantes que indica el estatus de acentuado de una sílaba. Como se discutió a lo largo de los capítulos 2 y 7, otras claves fonéticas también incluyen la duración y la centralización de las vocales.

F_0 se refiere al número de veces que una onda periódica de sonido se repite en un segundo. Las repeticiones se relacionan directamente a qué tan rápido vibran las cuerdas vocales.¹ Contra más rápido vibran, más alto es el número de repeticiones por segundo y en consecuencia, más alto es el F_0 . Así por ejemplo, la Figura 8-1 muestra una sección de 20 ms del oscilograma de la vocal acentuada [i] cuando ocurre en la palabra ['ki.ki.ti] 'tener pesadillas' pronunciada por una hablante mujer.

En la onda de sonido de la Figura 8-1, podemos observar que hay seis repeticiones (o ciclos) del mismo patrón (se indican sus fronteras por medio de líneas punteadas verticales). Cada ciclo dura 3 ms, redondeado al milisegundo más cercano. En un segundo (1000 ms), hay 333 ciclos (es decir, $1000/3$). Ya que el F_0 se mide en Hercios, decimos que el F_0 de la vocal acentuada [i] en este ejemplo es 333 Hz.

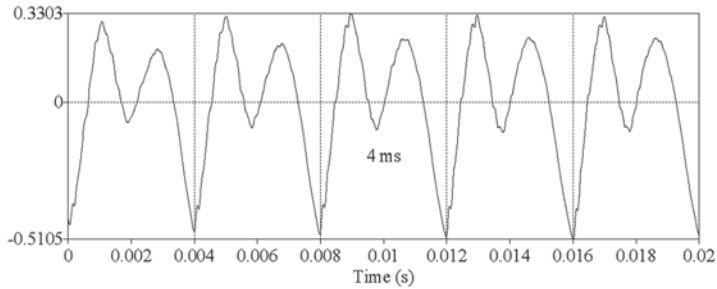
¹ Otros factores, como el tamaño y grosor de las cuerdas vocales y el tamaño de la cavidad laríngea también pueden afectar la velocidad de vibración de las cuerdas vocales y en consecuencia tienen un impacto sobre el F_0 . De este modo, las mujeres, quienes tienen una cavidad laríngea y unas cuerdas vocales más pequeñas, se caracterizan por tener un F_0 más alto que los hombres (Johnson 2003, Ladefoged 2003).

Figura 8-1: Ciclos por segundo en la vocal [i] de la palabra [kɪ.kɪ.ti] ‘tener pesadillas’



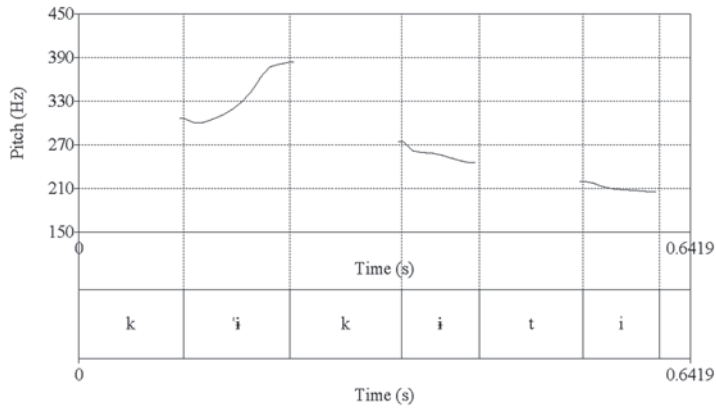
Calculemos ahora el F0 de la vocal inacentuada [i] en la misma palabra [kɪ.kɪ.ti] ‘tener pesadillas’. Se muestra una sección de 20 ms de la onda de sonido de esta vocal en la Figura 8-2. Esta vez observamos solo cinco ciclos de 4 ms cada uno. De este modo, el F0 de la vocal inacentuada [i] es más bajo que el de la vocal de la Figura 8-1. Su F0 es 250 Hz (es decir, $1000/4$).

Figura 8-2: Ciclos por segundo en la vocal [i] de la palabra [kɪ.kɪ.ti] ‘tener pesadillas’



La mejor manera de visualizar los cambios de F0 conforme pasa el tiempo es a través de un contorno como el que se muestra en la Figura 8-3 para la misma palabra [kɪ.kɪ.ti] ‘tener pesadillas’. Las frecuencias se muestran en el eje vertical y el tiempo en el eje horizontal. Aunque lo que ilustra la Figura 8-3 es el contorno de F0, de ahora en adelante me referiré a él indistintamente como contorno de F0 o contorno tonal.

Figura 8-3: El contorno de F0 de la palabra ['ki.ki.ti] 'tener pesadillas'



El contorno tonal es la impresión perceptual que obtenemos cuando nuestra audición procesa cambios en la frecuencia fundamental (F0). No todos los cambios del F0 en un tiempo dado pueden ser percibidos como diferencias de tono perceptual. Se han propuesto varias escalas perceptuales para determinar cuando los seres humanos podemos percibir un cambio en el F0. En este estudio, usaré la *Escala de Semitonos*, la cual se usa en música occidental. Un semitono (ST) es la distancia perceptual más pequeña que separa dos notas musicales adyacentes en la escala de tonos (Johnson 2003; Fant 2004; Hewlett y Beck 2006). Si dos valores de F0 son perceptualmente diferentes, deben por lo menos estar separados por un 1 ST. Se da en (25) la fórmula para comparar dos valores de F0 en Hercios y determinar su distancia en semitonos. F0^{alto} se refiere al F0 con el valor más alto y F0^{bajo}, al F0 con el valor más bajo.

$$(25) \text{ Distancia en ST} = 39,863 * \log (F0^{\text{alto}}/F0^{\text{bajo}})$$

De este modo, si uno quiere averiguar cuál es la distancia en ST entre el F0 de la vocal acentuada ['i] mostrada en la Figura 8-1 y el F0 de la vocal inacentuada [i], mostrada en la Figura 8-2, la fórmula sería: $39,863 * \log (333/250)$. Esto quiere decir que entre las frecuencias 333 Hz y 250 Hz hay 5 ST (es decir, $39,863 * \log (1,332) = 39,863 * 0,125$). La primera frecuencia será percibida como más alta tonalmente que la segunda.

8.1 EL ACENTO PRIMARIO EN SHIPIBO

Antes de continuar la discusión del comportamiento del F0 en el shipibo, es útil presentar una caracterización rápida de sus patrones de acento primario y dar

ejemplos representativos de cómo se comporta el contorno tonal alrededor de las sílabas con acento principal dentro de las palabras. Esto ayudará a contextualizar la caracterización acústica de los patrones del contorno tonal que se discuten más adelante en este capítulo. En general, las vocales con acento primario en shipibo tienen una duración más larga, evitan ser centralizadas y siempre se asocian con un tono alto mientras que las vocales inacentuadas son cortas, sí sufren varios grados de centralización (particularmente en sílabas no iniciales) y se asocian con tonos bajos.

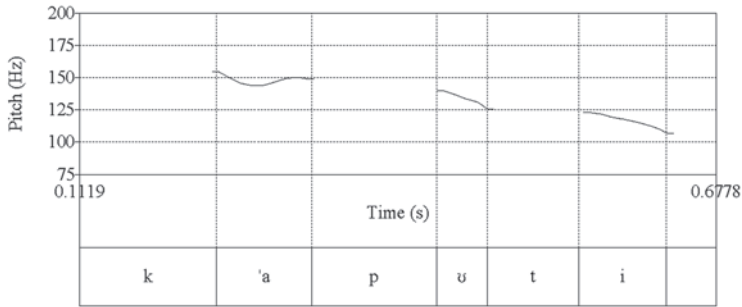
El acento primario es mayormente en predecible en shipibo. Con la excepción de algunas palabras que se discutirá en breve, el acento principal debe aparecer dentro de las dos primeras sílabas de una palabra (ésta es la ventana de acentuación del shipibo). El acento cae sobre la segunda sílaba si es cerrada, de otro modo, cae sobre la sílaba inicial.² Por defecto, el shipibo prefiere acentuar la sílaba inicial de las palabras. Los datos en (26) muestran palabras cuyo acento primario cae sobre la sílaba inicial. Todas ellas tienen la segunda sílaba abierta. La sílaba inicial puede o no tener una coda.

- | | | |
|---------|----------------------------|---------------------------|
| (26) a. | [¹ i.sa] | ‘erizo’ |
| b. | [¹ is.ku] | ‘páucar (esp. de pájaro)’ |
| c. | [¹ ma.tsi] | ‘frío’ |
| d. | [¹ a.ta.pa] | ‘gallina’ |
| e. | [¹ ʃun.ta.ku] | ‘mujer joven’ |
| f. | [¹ pi.ti] | ‘comer’ |
| g. | [¹ tʃi.ku.ti] | ‘derramar’ |
| h. | [¹ pa.ki.ti] | ‘dejar caer’ |
| i. | [¹ dʒiʃ.ki.ti] | ‘chicotear’ |

La Figura 8-4 presenta el comportamiento del F0 en una palabra shipiba que tiene el acento principal en la sílaba inicial. Un tono alto, indicado por el valor más alto del F0, se asocia con esa sílaba. Conforme la palabra se desarrolla y pasa a través de las sílabas inacentuadas, el F0 disminuye (es decir, el tono perceptual se vuelve más bajo).

² Este patrón acentual es compartido por otras lenguas pano, como el capanahua, por ejemplo (Elias-Ulloa 2004, Gonzalez 2003, Loos 1969).

Figura 8-4: El contorno tonal de la palabra [ˈka.pu.ti] ‘doblar’

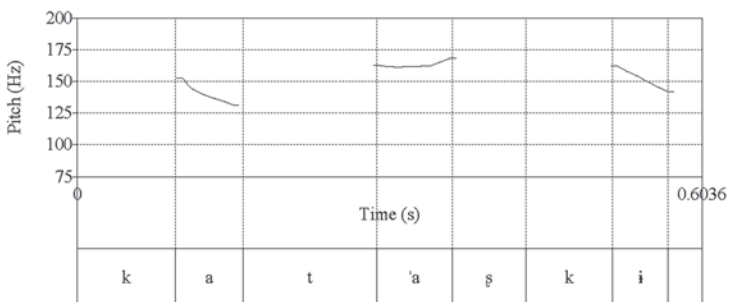


Los datos en (27) muestran ocurrencias del acento principal sobre la segunda sílaba, cuando esta sílaba es cerrada por una coda sibilante o una coda nasal.

- (27) a. [wi.ˈtaʃ] ‘pierna’
 b. [ki.ˈbβis] ‘salpullido en el interior de los labios’
 c. [ʃi.ˈnan.ja] ‘inteligente’
 d. [tʃi.ˈtun.ti] (tipo de falda usada por mujeres)
 e. [ha.ˈkũn.ti] ‘sentirse mejor’
 f. [ma.ˈpas.ti] ‘poner la mano en la cabeza de alguien’
 g. [ki.ˈmũs.ti] ‘apretar los labios a alguien’

La Figura 8-5 representa un ejemplo del comportamiento del F0 en una palabra shipiba que tiene el acento principal sobre la segunda sílaba. Como en el caso anterior, encontramos un tono alto alineado con la sílaba acentuada. En ambas sílabas inacentuadas, el F0 disminuye continuamente.

Figura 8-5: El contorno tonal de la palabra [ka.ˈtaʃ.ki] ‘poner lado a lado’ (CMPL)

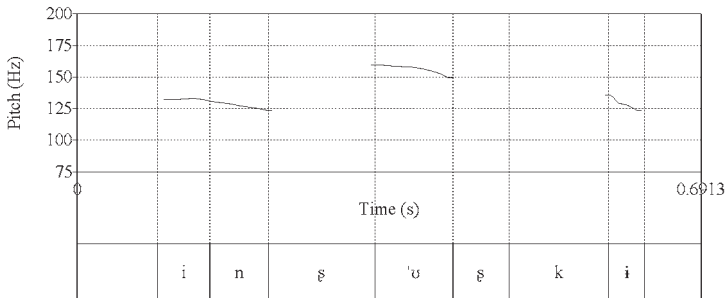


Vemos en los datos en (26) y (27) que dentro de la ventana de acentuación del shipibo, cuando dos sílabas abiertas compiten por el acento primario, la sílaba de más a la izquierda lo obtiene. Cuando una sílaba abierta compite con una cerrada, la sílaba cerrada obtiene el acento principal. Siguiendo este patrón, uno esperaría que cuando dos sílabas cerradas compiten por el acento primario, la sílaba de más a la izquierda lo recibe dado que ambas tienen la misma estructura silábica. Sin embargo, una de las características más interesantes del shipibo es que en este caso, el acento principal siempre se asigna a la segunda sílaba de la palabra. Es decir, las sílabas iniciales cerradas se comportan como si fuesen abiertas al tener que ceder el acento a la segunda sílaba (Elias-Ulloa 2005, 2006). Los datos en (28) ilustran este patrón.

- (28) a. [waʂ.ˈmɪn] ‘algodón’
 b. [niʂ.ˈniʂ] (esp. de pájaro)
 c. [kim.ˈpiʂ] ‘mal aliento’
 d. [pʊn.ˈtʊn.kʊ] ‘codo’
 e. [mis.ˈpan.ti] ‘preparar tamales’
 f. [pan.ˈʃin.ti] ‘amarillarse’
 g. [pʊn.ˈwiʂ.ti] ‘raspar el brazo de alguien’

La Figura 8-6 muestra un ejemplo de un contorno tonal en una palabra cuya primera y segunda sílaba es cerrada. El tono más alto que indica dónde se localiza el acento primario aparece asociado con la segunda sílaba. El contorno tonal que se muestra es el mismo que ocurre en palabras con una sílaba abierta inicial seguida por una cerrada (véase la Figura 8-5).

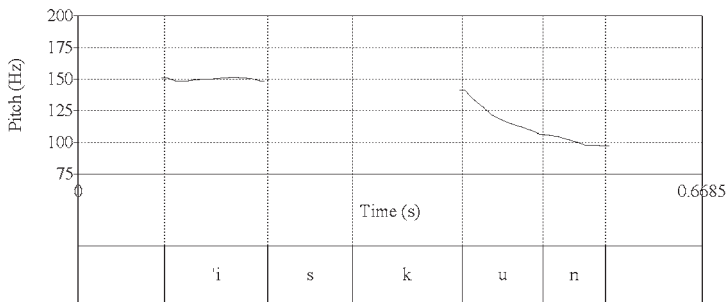
Figura 8-6: El contorno tonal de [in.ˈʂuʂ.ki] ‘crecer junto a algo’ (CMPL)



8.1.1 Casos de acento léxico en raíces

Obsérvese que los ejemplos presentados en (27), la segunda sílaba es cerrada y, como consecuencia, el acento principal cae sobre esa sílaba. Excepto por la existencia de dos palabras, [ˈkan.tʃis] ‘siete’ y [ˈis.kun] ‘nueve’, shipibo no tiene palabras en las cuales la segunda sílaba es cerrada y el acento aparece sobre la sílaba inicial. El patrón acentual extraño de estas palabras se debe a su origen. Fueron incorporadas en el shipibo del quechua donde se pronunciaban con el acento primario sobre la sílaba inicial: [ˈqan.tʃis] y [ˈis.qun]. La Figura 8-7 muestra el F0 de la palabra [ˈis.kun] ‘nueve’. El patrón de F0 es el mismo que se encuentra en palabras cuya segunda sílaba es abierta (véase la Figura 8-4).

Figura 8-7: El contorno tonal de la palabra [ˈis.kun] ‘nueve’



La mayoría de los casos de acento léxico que se encuentran en el shipibo son casos de préstamos de otras lenguas. Además de [ˈkan.tʃis] ‘siete’ y [ˈis.kun] ‘nueve’ que vienen del quechua, hay palabras con acento léxico que vienen de lenguas Arawak y del castellano. Por ejemplo, la palabra [u.ˈtʃi.ti] ‘perro’ viene del ashéninka-asháninka, donde se pronuncia como [o.ˈtsi.tsi] ~ [o.ˈtsi.ti] (Lenaerts 2006). Como en los casos del quechua, el acento también se preserva en su posición original a pesar de que contraviene los principios generales de asignación de acento del shipibo (es decir, ya que la segunda sílaba de la palabra es abierta, uno esperaría que el acento ocurra sobre la sílaba inicial). Las palabras prestadas del castellano también siguen este patrón: el acento se preserva en la segunda sílaba a pesar de que la sílaba es abierta (por ejemplo, [kas.ˈti.ku] del castellano [kas.ˈti.ɣo] ‘castigo’). La Figura 8-8 y la Figura 8-9 muestran el F0 de las palabras [u.ˈtʃi.ti] ‘perro’ y [kas.ˈti.ku] ‘castigo’.

Figura 8-8: El contorno tonal de la palabra [u.¹tʃi.ti] ‘perro’

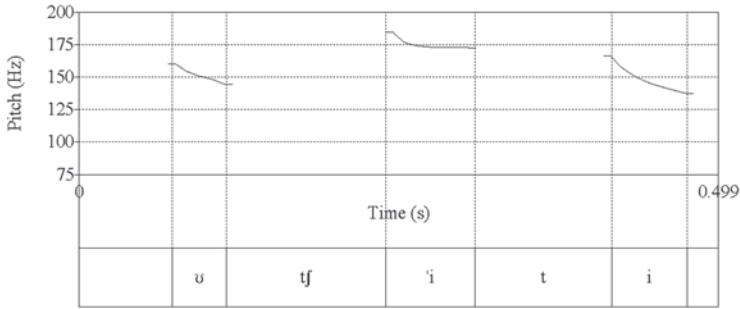
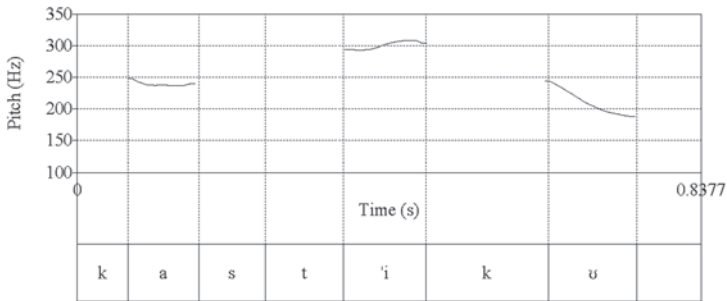


Figura 8-9: El contorno tonal de la palabra [kas.¹ti.ku] ‘castigo’



También hay algunas palabras del vocabulario nativo del shipibo que tienen acento léxico pero son muy pocas. Una revisión a las primeras doscientas páginas del diccionario de Loriot, Lauriault *et al.* 1993 arrojan solo 17 palabras nativas con acento léxico (por ejemplo, [a.¹wa.pa] ‘gato montés’, [ka.¹b^βu.dʒi] (esp. de tortuga), [wa.¹ka.pu] (tipo de madera)). Dado que cada página tiene alrededor de dieciséis entradas, diecisiete palabras con acento léxico representan aproximadamente 0,5% de las palabras examinadas.

8.1.2 El acento primario y los segmentos latentes

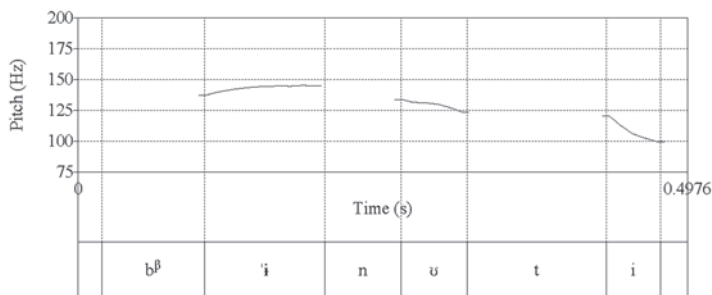
Es importante distinguir entre casos de palabras con acento léxico y palabras con segmentos latentes. A nivel fonológico, los segmentos en posición de coda no solo son las fricativas sibilantes (es decir, [s, ʃ, ʒ]) y la nasal [n]. Shipibo además tiene codas latentes. Estos son segmentos oclusivos que, aunque pueden ocupar la posición de coda de la sílaba y afectar la posición del acento como lo hace

cualquier otra coda, no se realizan a nivel fonético (véase la sección 9.2.3). Los segmentos latentes también muestran efectos fonéticos. Bajo las mismas condiciones de acento, las vocales que preceden a los segmentos latentes tienden a ser ligeramente más largas que aquellas que no preceden a un segmento latente (véase la sección 2.3 y la sección 7.1.2). Además, tan pronto como un segmento latente logra resilabificarse como el arranque de una vocal, aparece como una consonante oclusiva. Esto se ejemplifica en los datos en (29): /tʃiku/ ‘derramar’, /-t/ (INTVZ), /-ki/ (CMPL), /-a/ (PP2). Algunos ejemplos adicionales a este patrón se pueden encontrar en el capítulo 2.

- (29) a. / tʃiku -ki/ → [tʃi.ku.ki] ‘derramar’ (CMPL)
 b. / tʃiku -a/ → [tʃi.ku.a] ‘derramar’ (PP2)
 c. / tʃiku -t -ki/ → [tʃi.ku.ki] ‘derramarse’ (INTVZ, CMPL)
 d. / tʃiku -t -a/ → [tʃi.ku.ta] ‘derramarse’ (INTVZ, PP2)

Los contornos tonales mostrados en la Figura 8-10 a la Figura 8-12 muestran los cambios del acento que crean los segmentos latentes del shipibo. Los ejemplos usan la raíz verbal /b^βinʉ/ ‘perder’ y los sufijos /-t/ (INTVZ), /-ti/ (INF) y /-a/ (PP2). En la Figura 8-10, no hay un segmento latente en la segunda sílaba. Es una sílaba abierta. En este caso, como se espera, el acento principal cae sobre la sílaba inicial. Los valores más altos del F0 indican el estado acentuado de la sílaba inicial.³

Figura 8-10: / b^βinʉ -ti / → [b^βi.nu.ti] ‘perder’ (INF)

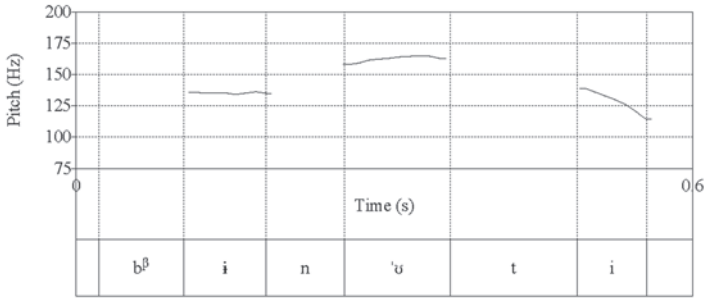


Compárese el comportamiento del F0 en la Figura 8-10 con lo que sucede en la Figura 8-11. A nivel fonológico, hay un segmento latente como la coda

³ Los contornos de F0 presentados en este capítulo sólo muestran el F0 asociado con las vocales y las codas nasales.

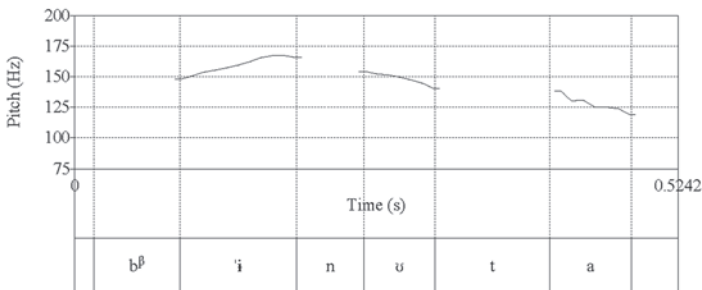
de la segunda sílaba de la palabra. Como consecuencia, la fonología del shipibo considera esa sílaba cerrada, y el acento cae sobre ella tal como lo indica el tono alto que lleva la segunda sílaba.

Figura 8-11: /b^βi.nʊ -t -ti/ → [b^βi.ˈnʊt.ti] → [b^βi.ˈnʊ.ti]
 ‘perderse’ (INTVZ, INF)



El F0 que se muestra en la Figura 8-12 corresponde a una palabra que tiene la misma raíz, /b^βi.nʊ/ ‘perder’, y el sufijo /-t/ (INTVZ) como en la Figura 8-11. Esta vez, sin embargo, se realiza el sufijo /-t/, el cual es seguido por una vocal (el sufijo /-a/ PP2) y, por lo tanto, la consonante latente logra silabificarse como una arranque de sílaba. Esto significa que a diferencia de lo que se presentó en la Figura 8-11, la segunda sílaba de la palabra en la Figura 8-12 es realmente abierta. El tono alto es llevado por la vocal [i], la vocal de la sílaba que tiene el acento principal. Nótese que el contorno de F0 que se observa es exactamente el mismo que muestra la Figura 8-10, la cual tampoco tenía un segmento latente en su segunda sílaba.

Figura 8-12: /b^βi.nʊ -t -a/ → [ˈb^βi.nʊ.ta] ‘perderse’ (INTVZ, PP2)



8.2 EL F0 DE LOS TONOS ALTOS EN SÍLABAS CON ACENTO PRIMARIO

Ahora retornemos a la caracterización acústica del F0 teniendo en cuenta el comportamiento del acento como ha sido descrito en la sección previa. Los contornos de F0 que muestran las palabras del shipibo se pueden describir mediante dos niveles tonales básicos: alto y bajo. Las sílabas con acento primario siempre atraen un tono alto mientras que las sílabas inacentuadas usualmente reciben un tono bajo (pero véase la sección 8.4). El tono alto asociado con las sílabas que tienen acento primario muestra tres realizaciones fonéticas básicas, las cuales ocurren en variación libre: nivelado, ascendente y descendente. Me referiré a estas realizaciones como tono alto-nivelado, tono alto-ascendente y tono alto-descendente, respectivamente. Las siguientes subsecciones discuten cada una de esas realizaciones. La frecuencia de ocurrencia de cada realización se da en la Tabla 8-1. Estos resultados se basan en el análisis de aproximadamente 5000 sílabas acentuadas (incluye tanto primeras como segundas sílabas) en palabras trisilábicas pronunciadas por hablantes hombres y mujeres.

Tanto las sílabas iniciales acentuadas como las segundas sílabas acentuadas prefieren un tono alto-nivelado (74% de los casos examinados). Sin embargo, esto no se da para el tono alto-ascendente o para el tono alto-descendente. Solo 24% de las sílabas acentuadas mostraron un tono alto-ascendente. 80% de esos casos ocurrieron en sílabas iniciales acentuadas. La realización de un tono alto-descendente en una sílaba acentuada es en general bastante raro (solo 2% de los casos examinados). Podría parecer sorprendente a primera vista que un tono alto se pueda realizar con una inclinación descendente. Este tipo de realización se refiere a un tono que cae durante la sílaba acentuada y en la cual el valor de F0 más alto es más alto que los valores más altos de F0 en las otras vocales de la palabra (véase la sección 8.2.3). 66% de ese 2% de casos de tonos alto- descendentes en sílabas acentuadas ocurrió cuando la segunda sílaba de la palabra llevaba el acento primario.

Tabla 8-1: Preferencia de las realizaciones fonéticas del tono alto en sílabas acentuadas

	Sílabas acentuadas	Primera sílaba acentuada	Segunda sílaba acentuada
Tono alto-nivelado	74%	47%	53%
Tono alto-ascendente	24%	80%	20%
Tono alto-descendente	2%	10%	90%

8.2.1 Tono alto-nivelado

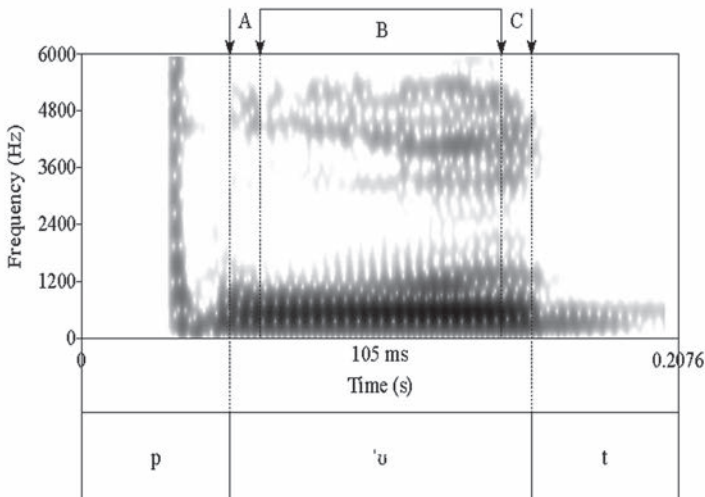
Veamos cuál es el rango de F0 para cada una de las realizaciones fonéticas del tono alto en sílabas acentuadas. La Tabla 8-2 muestra el valor promedio de F0 en un tono alto-nivelado de hablante hombres y mujeres en frases elicitadas usando un habla normal.

Tabla 8-2: Promedio de F0 en tono alto-nivelado de sílabas acentuadas

Hablantes Hombres	138 Hz ⁽²³⁾
Hablantes Mujeres	230 Hz ⁽²⁰⁾

Se usó el siguiente procedimiento para obtener los resultados de la Tabla 8-2, así como también aquellos de las tablas que se muestran más adelante en este capítulo. Se excluyeron las vocales acentuadas que presentaban fonación no-modal (por ejemplo, fonación glotalizada). Usando un script de Praat, las vocales se dividieron en tres secciones como se muestra en la Figura 8-13, la cual presenta la vocal acentuada de [ˈpu.tu.kɨ] ‘volverse polvoriento’ (CMPL). A fin de evitar cualquier efecto sobre el F0 que puedan tener las consonantes adyacentes a las vocales acentuadas (Lee 2008), se excluyó del análisis un área equivalente a 10% de la duración de la vocal a cada lado de la misma. Éstas son las secciones etiquetadas como A y C en la Figura 8-13. Dado que en este ejemplo, la vocal dura 105 ms, cada sección excluida es de 10,5 ms.

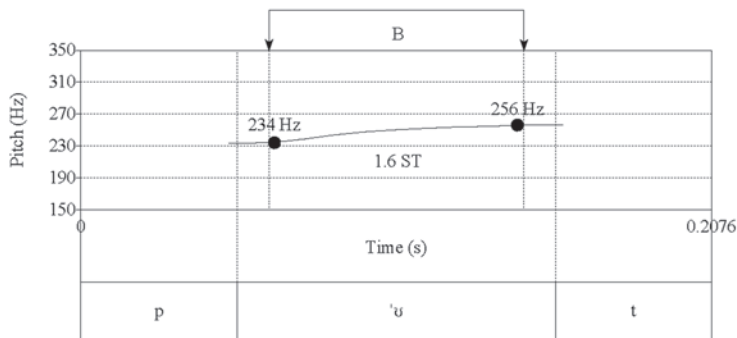
Figura 8-13: División de las vocales en tres áreas para el análisis de F0



Una vez que se identificó el área etiquetada como B en la Figura 8-13, el programa Praat obtuvo el punto más alto y más bajo del F0 dentro de la vocal. La distancia perceptual entre estos puntos se calculó en *Semitonos (ST)*. Si la distancia entre el F0 más alto y el F0 más bajo fue menos de un semitono, se consideró que el contorno del F0 era *nivelado*. Si la distancia fue más de un semitono, entonces se consideró el contorno del F0 ya sea como *ascendente* o como *descendente*, de acuerdo a la relación temporal entre el punto más alto y el más bajo del F0.

La Figura 8-14 indica el valor más alto y el más bajo del F0 de la sección B mostrada en la Figura 8-13. La distancia perceptual entre 234 Hz y 256 Hz es 1.6 ST. Dado que la diferencia es mayor a 1 ST y el valor más alto del F0 sigue al valor más bajo, el tono de la vocal con el acento primario se clasificó como un tono alto-ascendente.

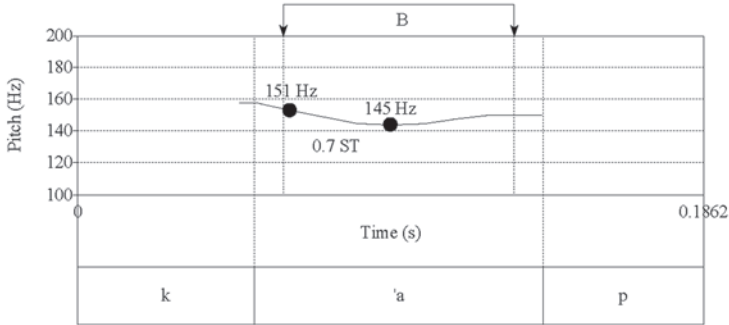
Figura 8-14: Tono alto-ascendente en la vocal acentuada de ['pʊ.tʊ.ki] (CMPL)



La Figura 8-15 muestra un ejemplo de un tono alto-nivelado asociado con una vocal con acento primario. El contorno de F0 corresponde a la vocal con el acento primario de la palabra ['ka.pʊ.ti] 'doblar' (véase el contorno tonal completo de esta palabra en la Figura 8-4). Se puede observar que aunque el F0 disminuye durante la primera parte de la sección B en el gráfico, el tono alto se considera *nivelado* porque la distancia perceptual entre el valor más alto y el más bajo de F0 es menos de 1 ST.

El valor en Hercios que se le asigna a un tono alto-nivelado es el promedio entre el punto más alto y el más bajo del F0. El tono alto-nivelado de la Figura 8-15 es 148 Hz. La distancia perceptual promedio entre el punto más alto y el más bajo del F0 dentro de todas las vocales acentuadas analizadas en la Tabla 8-2 fue 0,6 ST en hablantes hombres y 0,5 ST en hablantes mujeres.

Figura 8-15: Tono alto-nivelado en la vocal acentuada de [ka.pu.ti]



8.2.2 Tono alto-ascendente

La Tabla 8-3 muestra el promedio de los valores de F0 de tonos altos-ascendentes en sílabas acentuadas. Un tono alto-ascendente usualmente empieza con un F0 cercano al de un tono alto-nivelado pero los valores de F0, en vez de permanecer estables, se mantienen subiendo conforme pasa el tiempo. La columna “Punto Más Bajo” de la Tabla 8-3 muestra el valor promedio del F0 más bajo mientras la columna “Punto Más Alto” indica el valor promedio del F0 más alto en un tono alto-ascendente. La cuarta columna muestra la distancia perceptual en ST entre el F0 más alto y el más bajo. La Figura 8-12 presentó un ejemplo de un tono alto-ascendente (de un hablante hombre) en una sílaba inicial acentuada.

Tabla 8-3: Rango de F0 del tono alto-ascendente en sílabas acentuadas

	Punto más bajo	Punto más alto	Distancia en semitonos (ST)
Hablantes Hombres	146 Hz ⁽²³⁾	161 Hz ⁽²⁶⁾	1,7 ST
Hablantes Mujeres	230 Hz ⁽¹⁸⁾	253 Hz ⁽²⁰⁾	1,7 ST

La Figura 8-16 y la Figura 8-17 muestran que un tono alto-nivelado y un tono alto-ascendente no contrastan; son variantes del mismo tono alto en las sílabas acentuadas. Ambos contornos tonales corresponden a la palabra [pu.tu.ti] ‘volverse polvoriento’, pronunciada por dos hablantes mujeres. En la Figura 8-16, la hablante usa un tono alto-ascendente asociado con la sílaba acentuada mientras que en la Figura 8-17, la hablante usa un tono alto-nivelado. En la Figura 8-16,

el F0 de la sílaba acentuada empieza a 241 Hz y luego sube hasta 263 Hz (una distancia de 1,5 ST). En contraste, en la Figura 8-17, el valor más bajo del F0 en la vocal acentuada es 220 Hz y el más alto es 223 Hz (una distancia de 0,2 ST).

Figura 8-16: Tono alto-ascendente en la sílaba acentuada de [ˈpu.tu.ti] ‘volverse polvoriento’

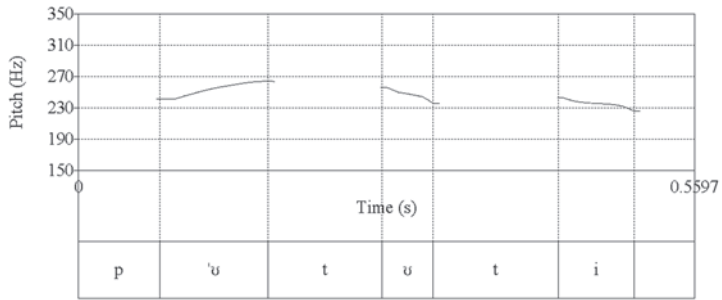
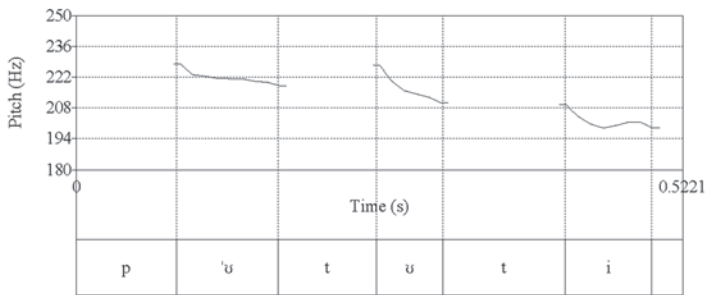


Figura 8-17: Tono alto-nivelado en la sílaba acentuada de [ˈpu.tu.ti] ‘volverse polvoriento’



8.2.3 Tono alto-descendente

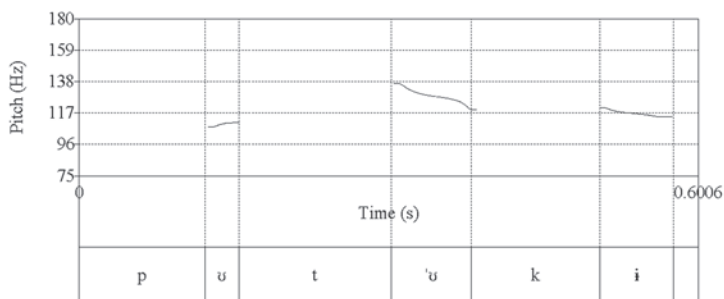
Aunque su ocurrencia es rara (solo 2% de los casos), un tono alto asociado con una sílaba que lleva acento primario puede realizarse como un tono alto-descendente. La mayoría de los casos de esta realización (casi 90%) corresponde a segundas sílabas acentuadas. En estos casos, el tono sube durante la sílaba inicial, alcanza temprano su pico durante la duración de la segunda vocal (la vocal que lleva el acento principal) y luego empieza a descender hasta que llega al final de la palabra. La Tabla 8-4 muestra los promedios de los valores más altos y más bajos del F0 en tonos alto-descendentes y su distancia perceptual promedio en semitonos.

Tabla 8-4: Rango de F0 del tono alto-descendente en sílabas acentuadas

	Punto más bajo	Punto más alto	Distancia en semitonos (ST)
Hablantes Hombres	143 Hz ⁽²⁰⁾	131 Hz ⁽¹⁹⁾	1,5 ST
Hablantes Mujeres	241 Hz ⁽¹⁹⁾	222 Hz ⁽¹⁸⁾	1,4 ST

Se muestra un ejemplo de un tono alto-descendente sobre la segunda sílaba acentuada de la palabra [pʉ.ˈtʉ.ki] ‘estar lleno’ (CMPL) en la Figura 8-18. La palabra fue pronunciada por un hablante hombre. El tono alto-descendente de la sílaba acentuada empieza en 136 Hz y desciende a 124 Hz (una distancia de 1,6 ST). Obsérvese que aunque el tono de la sílaba acentuada está cayendo, aún es relativamente más alto que los tonos de las sílabas inacentuadas. Además, como en el caso de los tonos alto-ascendente y alto-nivelado, el tono alto-descendente de las sílabas acentuadas tampoco es contrastivo. Es solo una posible realización fonética del tono alto asociado con sílabas que llevan acento primario.

Figura 8-18: Tono alto-descendente en sílaba acentuada de [pʉ.ˈtʉ.ki] ‘estar lleno’ (CMPL)



8.3 EL F0 DEL TONO BAJO DE SÍLABAS INICIALES INACENTUADAS

Mientras que las sílabas con acento primario siempre se asocian con un tono alto, las sílabas inacentuadas pueden asociarse con un tono bajo o uno alto. Las sílabas iniciales inacentuadas siempre muestran un tono bajo.⁴ En esta sección,

⁴ Uso el término ‘tono bajo’ sin ninguna implicación de su estatus fonológico sino sólo como un término conveniente para describir el comportamiento del F0 en sílabas inacentuadas. Desde un punto de vista fonológico, uno puede desarrollar un análisis en el cual el shipibo alinea las sílabas acentuadas con un tono alto. Las demás sílabas inacentuadas pueden dejarse subespecificadas con respecto a su tono. Ellas se realizarán con un tono bajo conforme las cuerdas vocales anticipan o se alejan de la realización de un tono alto.

me concentraré en el F0 de este tipo de sílabas. Recuérdese que las sílabas iniciales inacentuadas siempre son adyacentes a una sílaba con acento primario marcado por un tono alto. Esto se debe a que el acento primario no puede aparecer más allá de la segunda sílaba de una palabra. Si la sílaba inicial es inacentuada, el acento principal debe estar en la segunda sílaba de la palabra.

A nivel fonético, un tono bajo puede realizarse como un tono bajo nivelado, descendente o ascendente. Me referiré a las realizaciones de este tono como tono bajo-nivelado, tono bajo-descendente y tono bajo-ascendente. Al igual que el tono alto-descendente, podría parecer sorprendente que un tono bajo se pueda realizar como un tono bajo-ascendente. Como se mostrará en breve, hay casos de sílabas iniciales inacentuadas en las cuales el F0 sube en anticipación de la siguiente sílaba acentuada. En estos casos, el F0 de la palabra siempre alcanza su pico más alto dentro de la sílaba acentuada.

Las sílabas iniciales inacentuadas claramente prefieren el tono bajo-nivelado (72% de los casos examinados). El tono bajo-ascendente es el menos favorecido por las sílabas iniciales inacentuadas (solo 2% de los casos). La Tabla 8-5 muestra las realizaciones del tono bajo y la preferencia de sus ocurrencias. Los porcentajes se basan en el análisis de aproximadamente 2200 sílabas iniciales inacentuadas pronunciadas en palabras trisilábicas por hablantes hombres y mujeres.

Tabla 8-5: Preferencia de las realizaciones del tono bajo en sílabas iniciales inacentuadas

Tono bajo-nivelado	72%
Tono bajo-ascendente	2%
Tono bajo-descendente	26%

En las siguientes secciones, se caracterizarán en términos del F0 cada realización fonética del tono bajo conforme éste ocurre en sílabas iniciales inacentuadas.

8.3.1 Tono bajo-nivelado en sílabas iniciales inacentuadas

La Tabla 8-6 muestra el promedio del F0 (en Hercios) del tono bajo-nivelado en sílabas iniciales inacentuadas. Se usó el mismo procedimiento utilizado para obtener los promedios de F0 en las sílabas acentuadas, excepto por el área excluida a los bordes de las vocales inacentuadas. Ya que las sílabas inacentuadas son más cortas y en consecuencia las consonantes adyacentes pueden tener un

efecto mayor sobre su F0, se excluyó un área equivalente a 15% de la duración de la vocal a cada lado de la misma cuando la vocal duraba más de 50 ms, y 20% cuando la vocal era más corta que 50 ms. Se excluyeron del análisis las vocales inacentuadas que presentaron fonación no nodal.

Tabla 8-6: Promedio de F0 del tono bajo-nivelado de sílabas iniciales inacentuadas

Hablantes Hombres	119 Hz ⁽¹⁸⁾
Hablantes Mujeres	215 Hz ⁽¹⁵⁾

Debe observarse que aunque tanto el tono alto como el tono bajo se puedan realizar como nivelados (véase la Tabla 8-2 y la Tabla 8-6), los rangos de F0 de cada uno por género son bastante diferentes. La Tabla 8-7 nos permite comparar esto a través de la distancia promedio en ST entre el tono alto-nivelado que ocurre en sílabas acentuadas y el tono bajo-nivelado que ocurre con sílabas inacentuadas. La distancia entre estos tonos es mayor a un semitono tanto para hablantes hombres como para hablantes mujeres.

Tabla 8-7: Distancia perceptual entre un tono alto-nivelado y un tono bajo-nivelado

	Tono alto-nivelado (sílabas acentuada)	Tono bajo-nivelado (sílabas inicial inacentuada)	Distancia en semitonos (ST)
Hombres	138 Hz	119 Hz	2,6 ST
Mujeres	230 Hz	215 Hz	1,2 ST

Puede verse dos ejemplos de tonos bajos-nivelados en las sílabas iniciales de las palabras presentadas en la Figura 8-9 y la Figura 8-11. Los contornos tonales que se muestran en esas palabras, pronunciadas por un hablante mujer y un hablante hombre respectivamente, también nos permiten comparar entre un tono bajo-nivelado y un tono alto-nivelado. El último ocurre alineado con la sílabas acentuada segunda en cada caso.

8.3.2 Tono bajo-descendente en sílabas iniciales inacentuadas

El tono bajo de las sílabas iniciales inacentuadas también puede realizarse como un tono bajo-descendente pero esta variante fonética no es tan común en ese contexto como el tono bajo-nivelado. El tono bajo-descendente aparece en 26%

de los casos. La Tabla 8-8 presenta el promedio del rango de F0 en el tono bajo-descendente de sílabas iniciales inacentuadas. La tabla indica el valor promedio más alto y más bajo del F0 en Hercios y la distancia perceptual entre esos valores en ST. La Figura 8-5 y la Figura 8-8 presentan ejemplos del tono bajo-descendente en una sílaba inicial inacentuada. Obsérvese que el valor más alto y más bajo del F0 en un tono bajo-descendente es más bajo que los valores del F0 en un tono alto-descendente realizado en sílabas acentuadas (compárese los promedios de la Tabla 8-8 con los de la Tabla 8-4).

Tabla 8-8: Rango de F0 del tono bajo-descendente en sílabas iniciales inacentuadas

	Punto más alto	Punto más bajo	Distancia en semitonos (ST)
Hablantes Hombres	129 Hz ⁽²⁰⁾	118 Hz ⁽¹⁷⁾	1,5 ST
Hablantes Mujeres	226 Hz ⁽¹⁹⁾	206 Hz ⁽¹⁸⁾	1,6 ST

La Tabla 8-9 compara los promedios de los puntos más altos del F0 en un tono alto-descendente asociado con una sílaba acentuada (segunda columna) con aquellos de un tono bajo-descendente asociado con una sílaba inicial inacentuada (tercera columna). Tanto para hablantes hombres como mujeres, la distancia es mayor a un semitono (véase la cuarta columna).

Tabla 8-9: Distancia perceptual de los puntos más altos del F0 en tonos descendentes

	Tono alto-descendente (sílabas acentuadas)	Tono bajo-descendente (sílabas iniciales inacentuadas)	Distancia en semitonos (ST)
Hombres	143 Hz	129 Hz	1,8 ST
Mujeres	241 Hz	226 Hz	1,1 ST

La Tabla 8-10 compara los puntos más bajos del F0 en tonos descendentes asociados con sílabas acentuadas y sílabas iniciales inacentuadas. Nuevamente, los resultados indican que la distancia perceptual promedio entre el valor más bajo del F0 en un tono alto-descendente y su correspondiente en un tono bajo-descendente es mayor a un semitono.

Tabla 8-10: Distancia perceptual de los puntos más bajos del F0 en tonos descendentes

	Tono alto-descendente (sílabo acentuada)	Tono bajo-descendente (sílabo inicial inacentuada)	Distancia en Semitonos (ST)
Hombres	131 Hz	118 Hz	1,8 ST
Mujeres	222 Hz	206 Hz	1,3 ST

La comparación presentada en la Tabla 8-9 y la Tabla 8-10 muestra que el punto más alto de un tono bajo-descendente en una sílabo inicial inacentuada (129 Hz para hombres y 226 Hz para mujeres) es virtualmente el mismo que el punto más bajo de un tono alto-descendente asociado a una sílabo acentuada (118 Hz para hombres y 206 Hz para mujeres). Esto indica que el F0 de una sílabo acentuada es a propósito reajustado a un rango mucho más alto después de un tono bajo-descendente para señalar el estado de acentuado de la sílabo. Esto se puede ver claramente en la Figura 8-5 y la Figura 8-8.

8.3.3 Tono bajo-ascendente en sílabas iniciales inacentuadas

La realización de un tono bajo como ascendente en una sílabo inicial inacentuada es muy rara (solo 2% de los casos). Como en el caso del tono bajo-descendente, el tono bajo-ascendente de las sílabas inacentuadas también muestra un rango de F0 diferente al tono alto-ascendente que ocurre en las sílabas acentuadas. La Tabla 8-11 presenta los promedios de los puntos más bajos y más altos del F0 así como también su distancia perceptual en semitonos. Obsérvese que el valor más alto del F0 en un tono bajo-ascendente es muy cercano al valor más bajo del F0 en un tono alto-ascendente. Esto puede verse comparando la Tabla 8-11 y la Tabla 8-3.

Tabla 8-11: Rango de F0 de tonos bajos-ascendentes en sílabas iniciales inacentuadas

	Punto más bajo	Punto más alto	Distancia en semitonos (ST)
Hablantes Hombres	133 Hz ⁽²⁰⁾	145 Hz ⁽²²⁾	1,5 ST
Hablantes Mujeres	210 Hz ⁽³¹⁾	232 Hz ⁽³⁵⁾	1,7 ST

Como en los casos anteriores, se muestra en la Tabla 8-12 y la Tabla 8-13 la distancia perceptual entre los valores más altos y más bajos del F0 en un tono alto-ascendente asociado con sílabas acentuadas y un tono bajo-ascendente asociado con sílabas iniciales inacentuadas. En todos los casos, la distancia es mayor a un semitono.

Tabla 8-12: Distancia perceptual de los puntos más bajos del F0 en tonos ascendentes

	Tono alto-ascendente (sílabo acentuada)	Tono bajo-ascendente (sílabo inicial inacentuada)	Distancia en semitonos (ST)
Hombres	146 Hz	133 Hz	1,6 ST
Mujeres	230 Hz	210 Hz	1,6 ST

Tabla 8-13: Distancia perceptual de los puntos más altos del F0 en tonos ascendentes

	Tono alto-ascendente (sílabo acentuada)	Tono bajo-ascendente (sílabo inicial inacentuada)	Distancia en semitonos (ST)
Hombres	161 Hz	145 Hz	1,8 ST
Mujeres	253 Hz	232 Hz	1,5 ST

La Figura 8-19 presenta un ejemplo de un tono bajo-ascendente en una sílaba inicial inacentuada. El valor más bajo del F0 durante la vocal inicial de la palabra [u.'na.ma] 'ignorante' en la Figura 8-19 es 136 Hz. De allí, el tono sube a 146 Hz. La distancia perceptual entre estos valores es 1,2 ST, lo que muestra que el F0 es en realidad percibido como ascendente. Este tono bajo-ascendente puede compararse con el contorno tonal de la misma palabra pronunciada por otro hablante hombre que presenta la Figura 8-20. Esta vez, la sílaba inicial inacentuada posee un tono bajo-nivelado (su F0 es 114 Hz). Esta comparación también hace claro que el tono bajo-nivelado no contrasta con el tono bajo-ascendente.

Figura 8-19: Tono bajo-ascendente en una sílaba inicial inacentuada:
[u.'na.ma] 'ignorante'

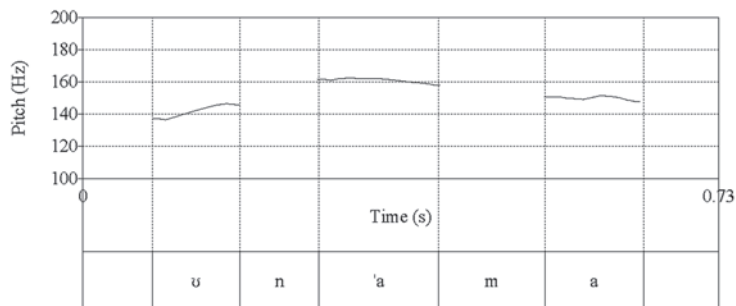
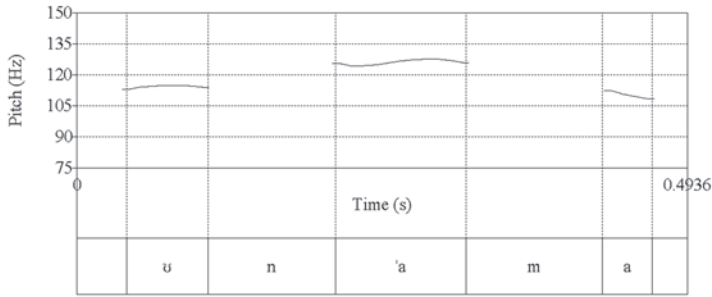


Figura 8-20: Tono bajo-nivelado en una sílaba inicial inacentuada:
[u.ˈna.ma] ‘ignorante’



8.4 EL GÉNERO COMO FACTOR CONDICIONANTE DEL TONO DE LAS SÍLABAS INACENTUADAS SEGUNDAS

A diferencia de las sílabas iniciales inacentuadas, las cuales siempre se asocian con un tono bajo (véase sección 8.3), las sílabas inacentuadas segundas pueden recibir un tono bajo o un tono alto. Un tono bajo siempre se realiza como descendente en esta posición. La Tabla 8-14 presenta los valores promedios de los puntos más altos y los más bajos del F0 en un tono bajo-descendente asociado con una sílaba inacentuada segunda. También se incluye la distancia perceptual entre los valores del F0. Los promedios de la Tabla 8-14 vienen del análisis de aproximadamente 2000 sílabas inacentuadas de segunda posición que llevan un tono bajo-descendente. Alrededor de 1500 de las vocales con un tono bajo-descendente fueron producidos por hablantes hombres y 500 por hablantes mujeres (la razón de la disparidad se explicará en breve). Todas las sílabas aparecieron en palabras trisilábicas.

Tabla 8-14: Rango de F0 de tonos bajos-descendentes en sílabas inacentuadas segundas

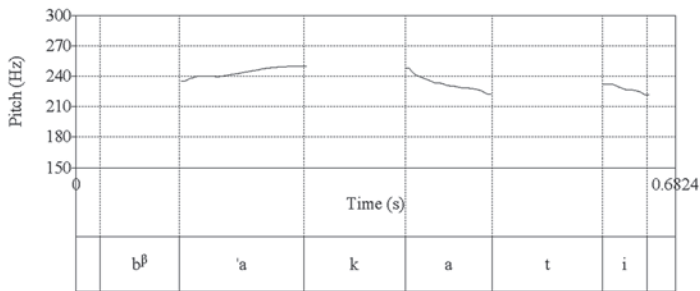
	Punto más alto	Punto más bajo	Distancia en semitonos (ST)
Hablantes Hombres	143 Hz ⁽²²⁾	129 Hz ⁽¹⁹⁾	1,8 ST
Hablantes Mujeres	237 Hz ⁽¹⁶⁾	220 Hz ⁽¹⁶⁾	1,3 ST

Los hablantes hombres consistentemente muestran un tono bajo-descendente alineado con la sílaba inacentuada que sigue a la sílaba inicial acentuada. Un ejemplo de este patrón se puede observar en la palabra [ˈka.pu.ti] ‘doblar’ en

la Figura 8-4. El valor más alto del F0 en la sílaba inacentuada segunda es 140 Hz y el más bajo es 129 Hz (la distancia perceptual es 1,4 ST).

Las hablantes mujeres también pueden usar un tono bajo-descendente en ese contexto. Esto se muestra en la Figura 8-21 a través del contorno tonal de la palabra [l'b^βa.ka.ti] 'volverse suave'. El punto más alto del F0 en sílaba inacentuada segunda es 241 Hz y, de allí, el F0 baja a 222 Hz (una diferencia de 1,4 ST).

Figura 8-21: Tono bajo-descendente en la sílaba inacentuada segunda:
[l'b^βa.ka.ti]



Sorprendentemente, las hablantes mujeres de shipibo difieren de los hablantes hombres en que típicamente usan un tono alto en la sílaba inacentuada segunda. Es decir, tanto los hablantes hombres como las mujeres pueden mostrar un tono bajo realizado como descendente asociado con esas sílabas. Sin embargo, solo las hablantes mujeres pueden además mostrar un tono alto en ese contexto. El tono alto de las sílabas inacentuadas segundas en el habla de mujeres no es una tendencia pequeña. De hecho, es una preferencia muy fuerte. Como se observa en la Tabla 8-15, el tono alto-nivelado ocurre en ese ambiente en un 64% de las aproximadamente 1500 sílabas inacentuadas segundas que las hablantes mujeres produjeron. El tono bajo-descendente descrito en la Tabla 8-14 para las hablantes mujeres (véase también la Figura 8-21) corresponde al 36% restante.

Tabla 8-15: Tono de sílabas inacentuadas segundas en hablantes mujeres

Level-High	64%
Falling-Low	36%

La Tabla 8-16 muestra el F0 promedio del tono alto-nivelado que usan las hablantes mujeres en las sílabas inacentuadas segundas. He incluido en la tabla el F0 del tono bajo-nivelado que usan en las sílabas inacentuadas iniciales y del tono

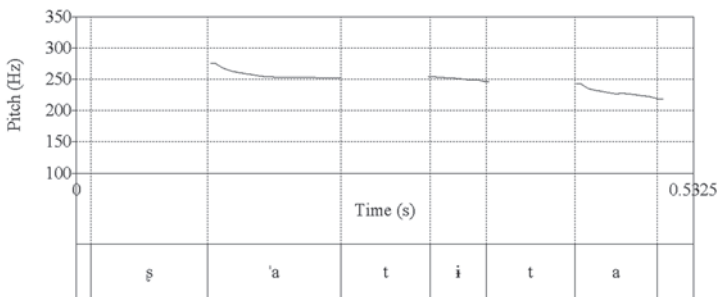
alto-nivelado que asocian a la sílabas acentuadas iniciales y segundas (véase las celdas en gris). Obsérvese que el F0 promedio del tono alto-nivelado de las sílabas acentuadas (tanto de las sílabas iniciales como de las segundas) es virtualmente el mismo que el F0 del tono alto-nivelado que aparece con las sílabas inacentuadas segundas: 230 Hz. Perceptualmente, es el mismo tono. En contraste, hay una distancia de más de un semitono (1,2 ST) entre 230 Hz y 215 Hz (el último es el F0 del tono bajo-nivelado de las sílabas inacentuadas iniciales que usan las hablantes mujeres).

Tabla 8-16: F0 del tono alto-nivelado en sílabas inacentuadas segundas (mujeres)

Sílabas Segundas Inacentuadas	231 Hz ⁽¹⁶⁾
Sílabas Iniciales Acentuadas	232 Hz ⁽²¹⁾
Sílabas Segundas Acentuadas	228 Hz ⁽¹⁹⁾
Sílabas Iniciales Inacentuadas	215 Hz ⁽¹⁵⁾

La Figura 8-22 muestra un ejemplo de una sílaba inacentuada segunda asociada con un tono alto-nivelado: [ʃa.ti.ta] ‘cortar’ (PP2). La palabra fue pronunciada por una hablante mujer. Nótese cómo el F0 casi no se mueve durante las dos sílabas iniciales. Los valores más altos y más bajos del F0 durante las vocales de la primera y segunda sílabas están menos de 0,3 ST de distancia: son perceptualmente lo mismo. Además, el contorno tonal observado en la Figura 8-22 nunca ocurre cuando el acento principal cae sobre la segunda sílaba. En ese caso, la sílaba inacentuada inicial obligatoriamente recibe un tono bajo (realizado como bajo-nivelado, bajo-descendente o bajo-ascendente – véase la sección 8.3).

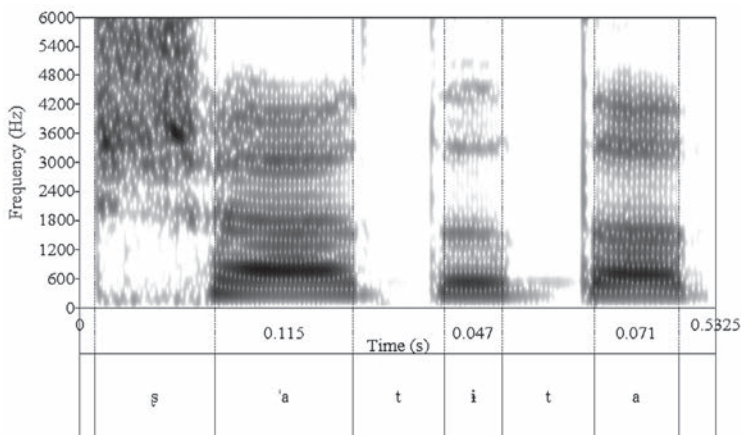
Figura 8-22: Tono alto-nivelado de sílabas inacentuadas segundas: [ʃa.ti.ta] ‘cortar’ (PP2)



El contorno tonal observado en la Figura 8-22 nos hace cuestionarnos el rol del tono perceptual como clave fonética para el acento principal. El uso de un tono alto sobre una sílaba inacentuada neutraliza la función del tono alto como clave fonética para identificar cuál de las dos sílabas iniciales lleva el acento primario. Esto no es un problema para el hablante adulto quien ya ha adquirido la fonología la lengua pero sí representa un desafío para el aprendiz de shipibo. Dado que tanto la primera como la segunda sílaba de una palabra podrían tener un tono alto, ¿cuál de ellas sería asignada como el núcleo prosódico de la palabra? Este dilema puede resolverse con la ayuda de otras claves fonéticas. El tono perceptual es una de las claves fonéticas que indica que una sílaba lleva el acento principal. Otras claves fonéticas, como la centralización de las vocales inacentuadas y particularmente la duración vocálica, juegan un rol importante.

La Figura 8-23 muestra el rol que juega la duración en determinar el estatus de acentuado de la sílaba inicial de la palabra [ˈʂa.ti.ta] ‘cortar’ (PP2).⁵ Esta palabra fue pronunciada por una hablante mujer. El contorno tonal de esta palabra se mostró en la Figura 8-22. A pesar de que las dos primeras sílabas reciben un tono alto, ellas muestran diferencias duracionales. La sílaba inicial que lleva el acento primario dura aproximadamente 65 ms más que la sílaba inacentuada segunda.

Figura 8-23: Duración vocálica en [ˈʂa.ti.ta] ‘cortar’ (PP2)



⁵ En la Figura 8-23, la sílaba final es ligeramente más larga que la segunda sílaba, aunque ambas son inacentuadas. Aquí, la diferencia duracional se debe a un fenómeno fonético de alargamiento vocálico en sílabas finales (un fenómeno común que se da en muchas lenguas). El estatus de inacentuado de la sílaba final está indicado por el tono bajo-descendente que recibe (véase la Figura 8-22).

8.5 TONO PERCEPTUAL, DURACIÓN E INTENSIDAD COMO CLAVES PARA EL ACENTO PRINCIPAL

Esta sección demuestra que en shipibo la duración es la clave fonética más fuerte para el acento principal y que cuando la duración se combina con el tono perceptual, juntos se vuelven las claves más eficientes para identificar las vocales acentuadas.

Primero, veamos qué tan exitoso es el tono perceptual, la duración y la intensidad cuando se les toma como claves fonéticas independientes para identificar las vocales con acento primario. La Tabla 8-17 muestra que la duración es la clave más fuerte, seguida del tono perceptual. Con el fin de obtener los porcentajes que se presentan en la Tabla 8-17, se analizaron aproximadamente 15000 vocales que pertenecen a palabras trisilábicas. Cada vocal se marcó apropiadamente como acentuada con acento principal o inacentuada. Entre esas vocales, hubo 5000 vocales acentuadas. Para cada vocal (acentuada o inacentuada), se obtuvo su F0 (en Hz), su duración (en ms) y su amplitud (en dB), y se identificó dentro de cada palabra las vocales con el valor más alto para cada propiedad acústica. Luego, se determinó si las vocales marcadas como portadoras del acento principal coincidían con las que tenían los valores más altos de F0, duración y amplitud.

Los resultados muestran que solas ninguna de las claves pudo encontrar todas las vocales con acento primario. La clave más exitosa es la duración. Logra predecir hasta un 80% de las vocales acentuadas. El tono perceptual (el correlato auditivo del F0) es relativamente exitoso también pero no tanto como la duración. El tono perceptual logra descubrir 68% de las vocales con acento primario. Aquí solo estoy considerando aquellos casos en los cuales la distancia perceptual en semitonos entre el F0 más alto de la vocal con el acento principal y el F0 de las vocales inacentuadas es mayor a un semitono. La clave fonética menos confiable para identificar las vocales acentuadas en shipibo es la intensidad (el correlato perceptual de la amplitud de la onda de sonido). Solo logró identificar 50% de todas las vocales con acento primario.

Tabla 8-17: Duración, Tono Perceptual e Intensidad como Claves Fonéticas para las Vocales Acentuadas (I)

Duración:	80%
Tono perceptual:	68%
Intensidad:	50%

El grado de éxito de la duración, el tono perceptual y la intensidad en identificar las vocales con acento primario se incrementa cuando se les combina. Esto se muestra en la Tabla 8-18. Cuando la duración y el tono perceptual se combinan, son completamente exitosos en determinar qué vocales llevan acento primario. Cuando la intensidad se combina con el tono perceptual o la duración, las posibilidades de descubrir las vocales acentuadas se incrementa pero ninguna de las combinaciones logra identificar todas las vocales acentuadas con un éxito mayor al 90%.

Tabla 8-18: Duración, Tono Perceptual e Intensidad como Claves Fonéticas para las Vocales Acentuadas (II)

Duración y tono perceptual:	100%
Intensidad y tono perceptual:	87%
Intensidad y duración:	88%

De este modo, la duración y el tono perceptual se perfilan como la combinación de claves fonéticas más confiable que se puede usar para identificar la posición del acento primario en shipibo. Cuando el tono perceptual no puede identificar el acento (como en los casos de palabras con el acento principal en la sílaba inicial en el habla de las hablantes mujeres - véase la Figura 8-23), la duración se vuelve decisiva.

8.6 EL F0 FUERA DE LA VENTANA DEL ACENTO PRIMARIO

Hasta ahora, he presentado el comportamiento del F0 dentro de la ventana del acento primario, es decir, las dos sílabas iniciales de la palabra. Ahora veamos el comportamiento del F0 fuera de esa ventana. Cuando una palabra no contiene ninguna sílaba cerrada, el F0 cae monótonamente después de la sílaba que lleva el acento primario.⁶ Esto se puede observar en la Figura 8-24: [b^βá.nà.jà.mà.kì]⁷ ‘no (lo) sembró’ (/b^βana/ ‘sembrar’, /-jama/ NEG, y /-kì/ CMPL). En la Figura 8-25, muestro el contorno de intensidad de la palabra de la Figura 8-24

⁶ Esto se complica por el hecho que los sufijos pueden tener su propio acento léxico y que los acentos secundarios en general no necesariamente se correlacionan con cambios del F0. La sección 8.7 discute en detalle el comportamiento del F0 en sufijos que tienen acento léxico.

⁷ En la transcripción angosta de las realizaciones fonéticas, uso el diacrítico de la tilde aguda, [´], para señalar la sílaba que lleva un tono alto y el de la tilde grave, [˘], para indicar aquellas que muestran un tono bajo. Los diacríticos [ˀ] y [ˁ] los reservo para indicar el estatus fonológico acentuado de una sílaba. El diacrítico [ˀ] se usa para señalar la sílaba núcleo de la palabra prosódica

con el fin de demostrar que la intensidad no puede usarse como una clave para determinar si hay acentos secundarios en casos como [b^βa.nà.jà.mà.kì]. Como se puede observar en la Figura 8-25, la intensidad se mantiene sin cambios durante casi toda la palabra. Los valles que se observan no están relacionados al acento sino a la caída de la energía acústica durante la producción de las consonantes obstruyentes [b^β] y [k].

Figura 8-24: F0 de ['b^βa.na.ja.ma.ki] 'no (lo) sembró'
(realizado como [b^βá.nà.jà.mà.kì])

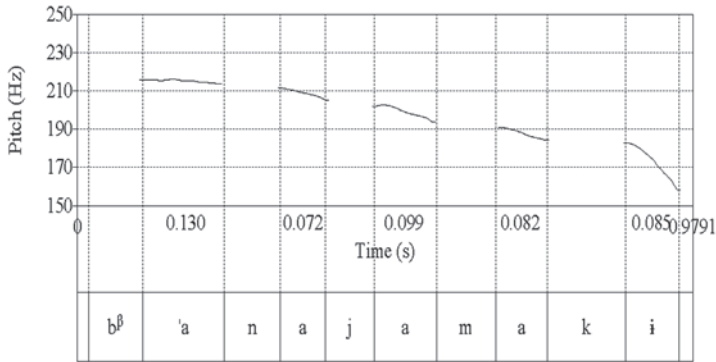
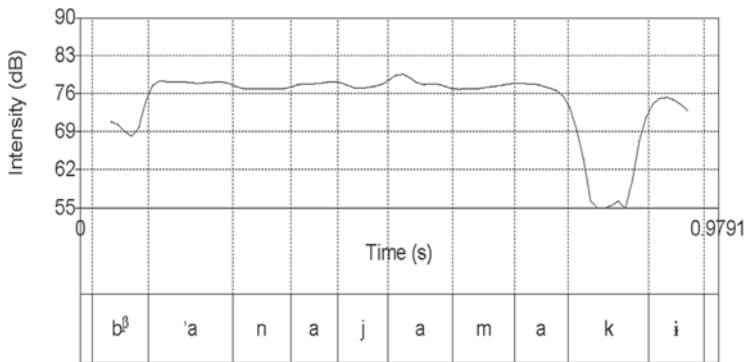


Figura 8-25: Contorno de intensidad de ['b^βa.na.ja.ma.ki] 'no (lo) sembró'



Dado que la determinación de si una sílaba es acentuada o no es hasta cierto punto un problema fonológico, la respuesta a si hay acentos secundarios en palabras como la de la Figura 8-24 dependerá de los presupuestos teóricos que

(es decir, el acento primario) y el símbolo [], para las demás sílabas núcleo que acompañan a la sílaba con el acento primario (es decir, los acentos secundarios).

tenga el investigador. Aquí presento los factores fonéticos que pueden servir de guía al analista. Claramente el comportamiento del F0 y de la intensidad en palabras como [ˈbʰa.na.ja.ma.ki] no pueden servir como claves para determinar la presencia de acentos secundarios. ¿Podría la duración servir como una clave para estos casos? En la Figura 8-24, se indica la duración de cada vocal. Como se espera, la vocal con el acento primario, [a], aquella que está asociada con el tono alto de la sílaba [bʰa], es la más larga. Es 58 ms más larga que la siguiente vocal inacentuada. Obsérvese que más allá de la ventana del acento primario, la duración de la tercera vocal, también una [a], no disminuye. Por el contrario, se incrementa en 27 ms con respecto a la vocal inacentuada previa. Después de esta vocal, la duración de la siguiente vocal disminuye nuevamente. La cuarta vocal es 17 ms más corta que la anterior. La duración de la vocal final, la cual usualmente muestra algún grado de alargamiento fonético, es muy cercana a la de la penúltima vocal.

En el caso presentado en la Figura 8-24, la tercera vocal obtiene una duración mayor que la segunda y cuarta vocal. En otras ocurrencias, la tercera vocal muestra aproximadamente la misma duración que la segunda vocal mientras que la cuarta muestra una duración más corta. En ambos casos, parece haber algo especial acerca de la tercera vocal que la protege de volverse más corta que la segunda vocal. Si este patrón duracional se toma como clave para determinar los acentos secundarios en la ausencia de un tono alto, entonces la evidencia sugiere la presencia de un ritmo trocaico que corre de izquierda a derecha a través de la palabra. De este modo, la representación fonética de la palabra en la Figura 8-24 sería [ˈbʰaˈ.na.jaˈ.ma.ki]; es decir, con el acento primario en la sílaba inicial y un acento secundario en la tercera sílaba. Aquí el símbolo [ˈ] representa una vocal que es fonéticamente más larga que las vocales vecinas.

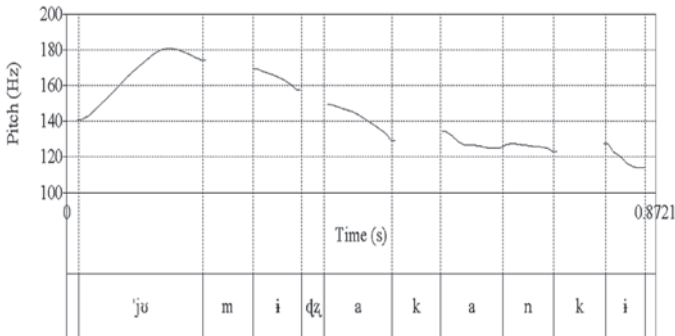
Aunque mi posición es que el patrón de duración vocálica que acaba de describirse sirve como clave para identificar el estado acentuado de las sílabas más allá de la ventana del acento primario, la evidencia que este patrón aporta para determinar los acentos secundarios no es conclusiva. Los cambios en la duración vocálica más allá de la ventana del acento primario pueden seguir dos patrones alternativos. Algunas veces, las vocales se vuelven más cortas conforme dejan atrás al acento primario. Otras veces, todas las vocales después del acento primario muestran virtualmente la misma duración. No he encontrado ninguna tendencia clara o preferencia entre estos patrones duracionales.

8.6.1 El F0 de las sílabas cerradas

A diferencia de las palabras que solo contienen sílabas abiertas más allá de la ventana del acento primario, sílabas cerradas sí afectan el comportamiento del F0. En particular, las sílabas cerradas fuera de la ventana del acento principal usualmente atraen un tono alto, el cual revela su estado de acentuado. Sin embargo, la tendencia de las sílabas cerradas a aparecer asociadas con un tono alto se da mayormente en el habla cuidada. En el habla coloquial, uno también puede observar el mismo patrón pero no es inusual encontrar sílabas cerradas que no muestran ninguna asociación con un tono alto.

Veamos, a través del examen de la Figura 8-26 y la Figura 8-27, las dos realizaciones fonéticas del tono alto cuando se encuentra asociado con una sílaba cerrada más allá de la ventana del acento primario. El F0 del tono alto en las sílabas cerradas que no llevan acento primario usualmente no sube a frecuencias más altas.⁸ Sin embargo, sabemos que hay un tono alto porque la caída del F0 se detiene. En este caso, el tono alto se realiza nivelado. Esto se muestra en la Figura 8-26, la cual presenta el F0 de la palabra [ˈjũ.mi.dʒa.kan.ki] ‘(lo) pescaron’ (/jumidʒa/ ‘percar’, /-kan/ (CON.3PL), /-ki/ CMPL). Se puede determinar el tono alto-nivelado asociado con la sílaba cerrada [kán] observando cómo el F0 cae después de la sílaba acentuada inicial y deja de caer una vez que alcanza la sílaba [kán]. Después de la sílaba cerrada, el F0 continúa su caída.

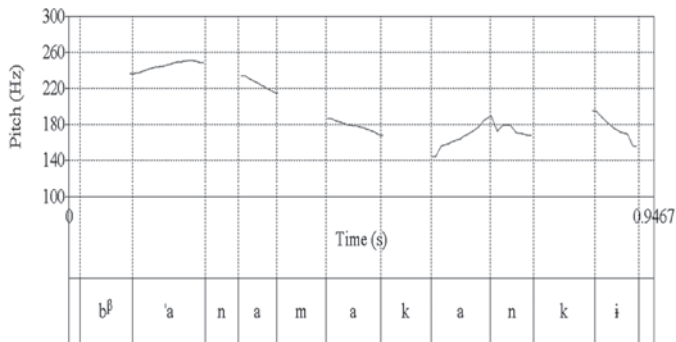
Figura 8-26: [ˈjũ.mi.dʒa.kan.ki] ‘(lo) pescaron’
(realizado como [jú.mì.dʒà.kán.kì])



⁸ Esto se debe a un fenómeno conocido como declinación, el rango tonal disminuye conforme se aleja del inicio de la palabra. Esto hace que cada tono alto subsecuente sea más bajo en su frecuencia fundamental que el tono alto anterior. De este modo, en shipibo, el tono alto más allá de la ventana del acento primario usualmente se le reconoce porque el F0 deja de caer o después de una caída muestra un incremento aunque casi nunca retorna a los valores de F0 al inicio de las palabras.

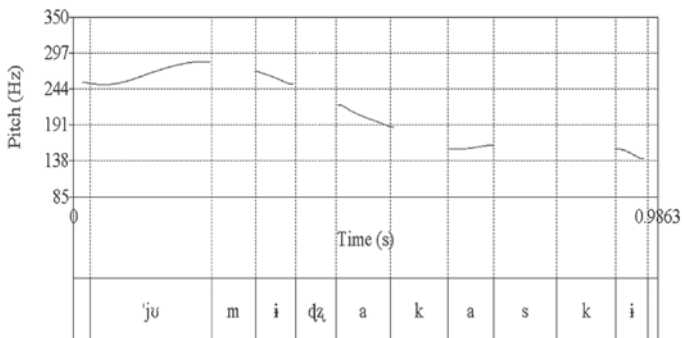
Mientras en la Figura 8-26, la sílaba cerrada obtuvo un tono alto realizado como nivelado, la Figura 8-27 muestra que el tono alto asociado con la sílaba cerrada, [kan], fuera de la ventana del acento principal también puede realizarse como un tono alto-ascendente. La Figura 8-27 muestra el contorno de F0 de la palabra [b^βa.na.ma.kan.ki] ‘(lo) sembraron’ (/b^βana/ ‘sembrar’, /-ma/ CAUS, /-kan/ (CON.3PL), /-ki/ CMPL). La subida del F0 durante la sílaba cerrada puede interpretarse como una clave fonética que indica su estado como portador del acento secundario.

**Figura 8-27: [b^βa.na.ma.kan.ki] ‘(lo) sembraron’
(realizado como [b^βá.nà.mà.kán.kì])**



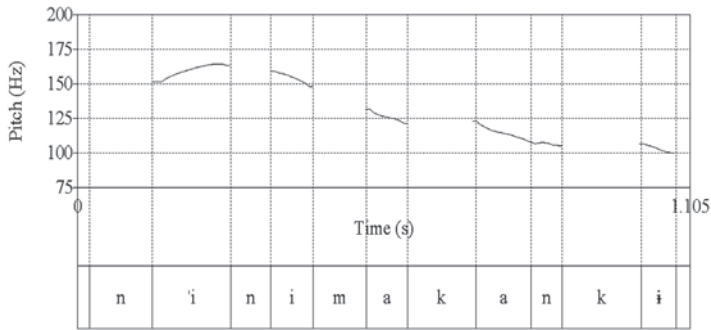
La Figura 8-28, [jũ.mi.dʒa.kas.ki] ‘quería pescar(lo)’, demuestra que las sílabas cerradas por una sibilante se comportan del mismo modo que las sílabas cerradas por una coda nasal. Ambos tipos de codas, sin importar su sonoridad, atraen un tono alto (en este caso realizado como un tono alto-nivelado). La palabra se compone de la raíz /jũmɨdʒa/ ‘pescar’ y los sufijos /-kas/ (DES) y /-ki/ (CMPL)

**Figura 8-28: /jũmɨdʒa -kas -ki/ ‘quería pescar(lo)’
(realizado como [jú.mè.dʒà.kás.kì])**



Como se mencionó anteriormente, la asociación de las sílabas cerradas con un tono alto no es obligatoria, particularmente durante el habla rápida. La Figura 8-29 muestra esto a través del F0 de la palabra [¹ni.ni.ma.kan.ki] ‘(lo) jalaron’ (/nini/ ‘jalar’, /-ma/ CAUS, /-kan/ (CON.3PL), /-ki/ CMPL). A pesar de tener una sílaba cerrada, la palabra se realiza como [ní.nì.mà.kàn.kì]. Es decir, el F0 cae después del acento principal y se mantiene cayendo hasta que alcanza el final de la palabra. Sin embargo, obsérvese que a pesar de esto, la vocal [a] de la sílaba cerrada [kan] es más larga que la vocal central que le precede y que le sigue. En la ausencia de un tono alto, interpretó este alargamiento vocálico fonético como evidencia del estatus acentuado de la sílaba cerrada.

Figura 8-29: [¹ni.ni.ma.kan.ki] ‘(le) hicieron jalar(lo)’
(realizado como [ní.nì.mà.kàn.kì])

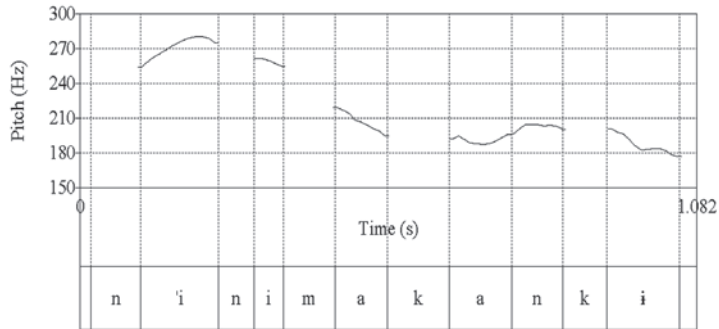


La Figura 8-30 muestra la misma palabra, [¹ni.ni.ma.kan.ki] ‘(le) hicieron jalar(lo)’. La comparación de la Figura 8-29 y la Figura 8-30 aporta evidencia adicional que la asociación de las sílabas cerradas con un tono alto fuera de la ventana del acento primario es opcional. Mientras que las sílabas cerradas más allá de esa ventana algunas veces no obtienen un tono alto, esta opcionalidad no ocurre cuando las sílabas cerradas están acentuadas dentro de la ventana del acento primario. En ese contexto, las sílabas cerradas siempre logran asociarse con un tono alto (véase el contorno tonal de la Figura 8-5).

En muchas lenguas, la habilidad de la sílaba para llevar acento secundario depende de su posición dentro de la palabra y de la distancia con otras sílabas acentuadas. Los siguientes párrafos describen el comportamiento del F0 en sílabas cerradas que aparecen en diferentes contextos conocidos por tener un efecto en la distribución de los acentos secundarios.

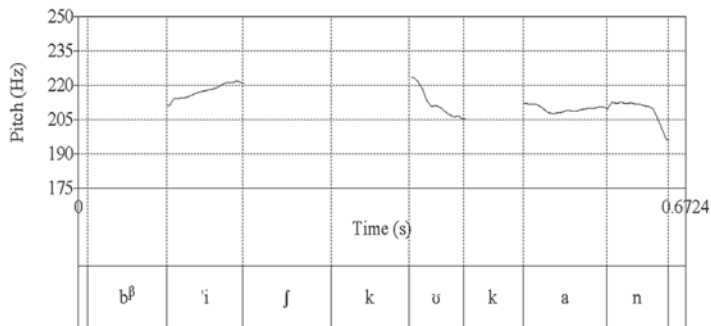
El comportamiento del F0 en la Figura 8-31 demuestra que las sílabas cerradas en shipibo no están sujetas a los efectos de *non-finalidad* o *extrametricalidad*

Figura 8-30: [¹ni.ni.ma.kan.ki] ‘(le) hicieron jalar(lo)’
(realizado como [ní.nì.mà.kán.kì])



que hacen que las sílabas finales eviten ser acentuadas o alternatively, en el caso específico de las sílabas cerradas, que se comporten como si fuesen sílabas abiertas. En shipibo, las sílabas cerradas pueden atraer un tono alto inclusive si aparecen en posición final de palabra como se muestra en la Figura 8-31 con la palabra [¹b^βiʃ.ku.kan] (/b^βiʃkuk/ ‘herida’ y /-an/ FL). La sílaba cerrada final aparece portando un acento secundario, indicado por la presencia de un tono alto-nivelado que puede identificarse observado que el F0 deja de caer durante la vocal de la sílaba final.

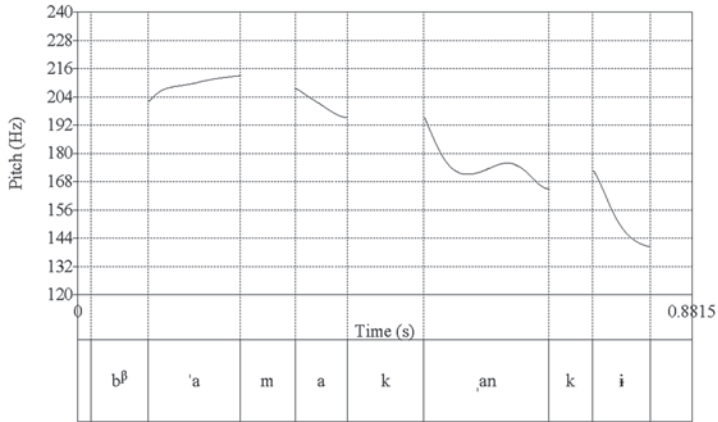
Figura 8-31: [¹b^βiʃ.ku.kan] ‘herida’ (FL), realizado como [b^βiʃ.kù.kán]



El número de sílabas de una palabra y la distancia de una sílaba cerrada con respecto al acento primario influye la habilidad de las sílabas cerradas de llevar un tono alto. Ya hemos visto que las sílabas cerradas pueden atraer un tono alto cuando ocurren una o dos sílabas lejos del acento primario (véase la Figura 8-30 y la Figura 8-31).

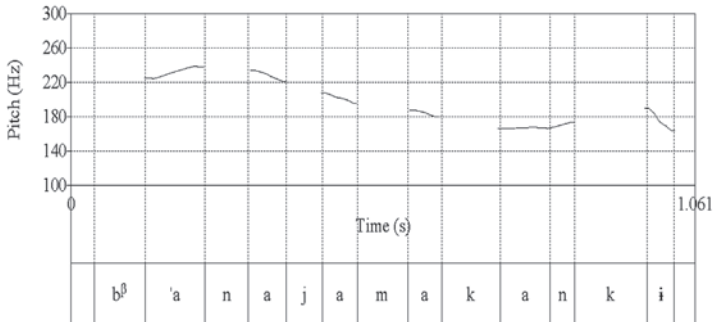
En la Figura 8-31, había una sílaba abierta que intervenía entre la sílaba con el acento primario y la sílaba cerrada que ocurre al final de la palabra. En la Figura 8-32, la sílaba cerrada también aparece una sílaba de distancia del acento principal, pero esta vez, la sílaba cerrada no está en posición final de la palabra. Obsérvese que la sílaba cerrada aún posee un tono alto.

Figura 8-32: [¹b^βa.ma,kan.ki] ‘aparecieron de pronto’
(realizado como [b^βá.mà.kán.kì])



La Figura 8-33 muestra el F0 de la palabra [¹b^βa.na.ja.ma,kan.ki] ‘no (lo) sembraron’ (/b^βana/ ‘sembrar’, /-jama/ NEG /-kan/ (CON.3PL), /-ki/ CMPL). En este caso, hay tres sílabas abiertas interviniendo entre la sílaba cerrada y la sílaba con el acento primario. El F0 deja de caer durante la sílaba cerrada, lo cual indica la presencia del tono alto. Como en la Figura 8-26, el F0 corresponde a un tono alto-nivelado.

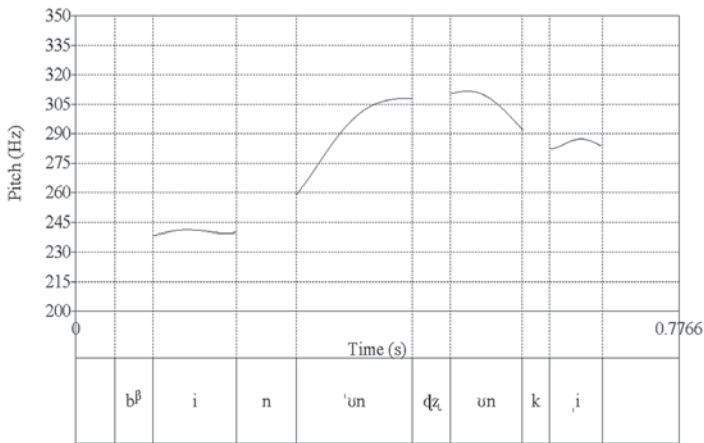
Figura 8-33: [¹b^βa.na.ja.ma,kan.ki] ‘no (lo) sembraron’
([b^βá.nà.jà.mà.kán.kì])



Finalmente, la Figura 8-34 y la Figura 8-35 muestran casos donde dos sílabas cerradas ocurren adyacentes. En la Figura 8-34, la primera sílaba cerrada pertenece a la ventana del acento primario. Cuando esta configuración ocurre, la primera sílaba cerrada siempre obtiene el acento primario: [b^βi.¹nʊn.dzʊn.₁ki]. Es decir, el Shipibo no reposiciona el acento primario a la sílaba inicial para que la segunda sílaba cerrada pueda ser acentuada: *[¹b^βi.nʊn.₁dzʊn.ki].

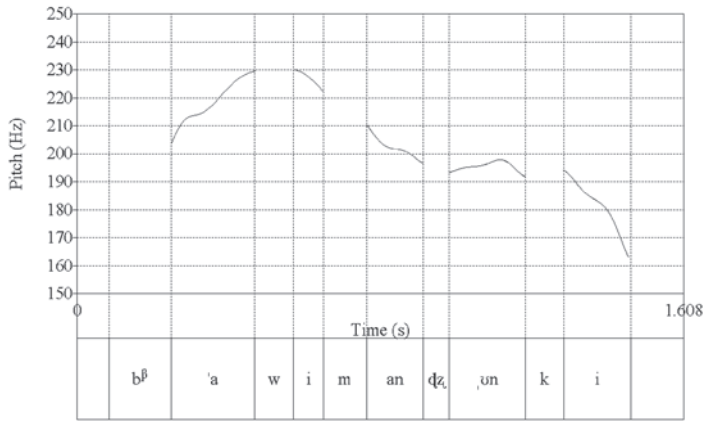
Además, obsérvese en la Figura 8-34 que el F0 no cae durante la sílaba final lo cual tomo como un indicador de la presencia de un tono alto. Esto significa que el acento secundario ha sido empujado hacia la última sílaba para evitar un choque acentual, *[b^βi.¹nʊn.₁dzʊn.ki]. El tono alto que hubiese ocurrido asociado con [dzʊn] es forzado a asociarse con la siguiente sílaba abierta. De este modo, cuando el shipibo tiene que evitar la creación de un choque acentual, opta por reposicionar los acentos que ocurren fuera de la ventana del acento primario.

**Figura 8-34: [b^βi.¹nʊn.dzʊn.₁ki] (esp. de palmera - REP)
(realizado como [b^βi.nʊn.dzʊn.kí])**



La Figura 8-35 también muestra dos sílabas cerradas adyacentes. Sin embargo, esta vez, el acento principal ocurre sobre la sílaba inicial y ambas sílabas cerradas están fuera de la ventana del acento primario. En este contexto, es la segunda sílaba cerrada la que obtiene el acento secundario. Esto puede verse en la Figura 8-35 a través del tono alto asociado con la sílaba [dzʊn]. Este comportamiento es idéntico al que se observa dentro de la ventana del acento primario cuando contiene dos sílabas cerradas. Como se mostró en la sección 8.1, en ese contexto es la segunda sílaba cerrada la que recibe el acento primario (por ejemplo, [waʃ.míɲ] ‘algodón’, véase los datos en (28)).

Figura 8-35: [l^hb^ha.wi.man.ɟzún.ki] (esp. de pescado - FL, REP)
 ([b^há.wi.màn.ɟzún.kì])



8.7 EL F0 EN SUFIJOS CON ACENTO LÉXICO

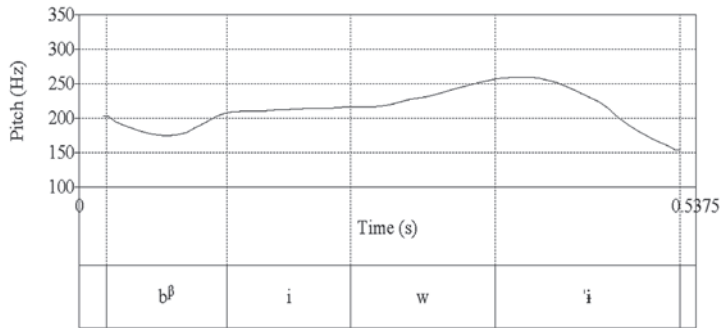
Hasta ahora, he presentado casos en los cuales el comportamiento del F0 y la asociación del tono alto con ciertas sílabas son predecibles. Sin embargo, el shipibo también muestra casos de sufijos que tienen acento léxico. Esta sección se enfoca en caracterizar el comportamiento del F0 en este tipo de sufijos.

Si el sufijo con acento léxico es monosilábico o tiene su acento léxico en la sílaba inicial (si es polisilábico), el acento asociado con el sufijo puede identificarse fácilmente porque cuando ocurre dentro de la ventana del acento primario, obligatoriamente atrae un tono alto y su vocal muestra una duración más larga. Tomemos como ejemplo el sufijo imperativo, /-^hwi/, el cual tiene un acento léxico. Cuando este sufijo se añade a una raíz monosilábica como /b^hi/ ‘traer, recoger’, el acento primario aparece sobre el sufijo, a pesar de que ambas sílabas son abiertas y el hecho que los principios generales de la asignación del acento en shipibo colocarían el acento primario sobre la sílaba inicial. Esto puede observarse en la Figura 8-36. Un tono alto aparece alineado con la vocal del sufijo imperativo durante la parte inicial de la vocal.

Antes de continuar la discusión de los sufijos con acento léxico, nótese que ya que las oraciones declarativas en shipibo se marcan con un tono descendente al final del verbo,⁹ el F0 empieza a caer en el sufijo imperativo después del pico del tono alto. Compárese este patrón con una palabra bisilábica como la que se

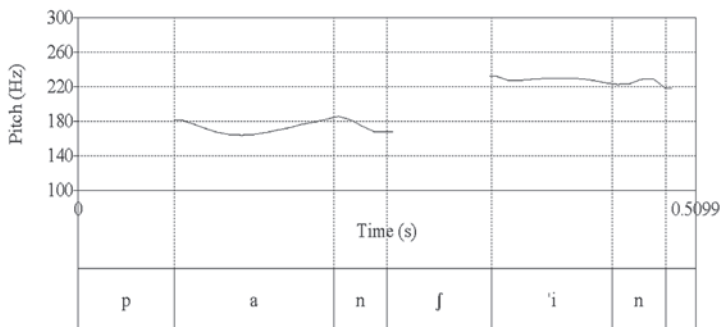
⁹ El shipibo es una lengua SOV donde el verbo ocupa la posición final de la oración. Sin embargo, el orden de palabras no es fijo, otros órdenes sintácticos son posibles.

Figura 8-36: [b^βi. 'wi] 'trae(lo)!' (realizado como [b^βɪ. 'wî])



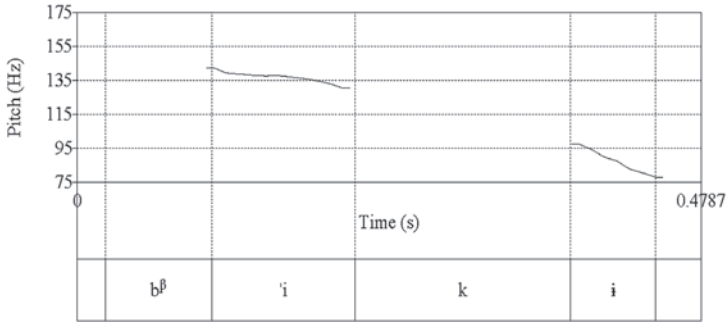
muestra en la Figura 8-37: [pan. 'ʃin] (esp. de pescado). Esta palabra se tomó de la frase ['b^βa.ki.dʒa pan. 'ʃin 'b^βi.ki] 'el niño trajo panshin' (literalmente: niño-EV panshin trajo). En este caso, la sílaba final de la palabra [pan. 'ʃin] porta el acento principal tal como lo indica el tono alto con el cual está asociado pero el tono no cae al final de la palabra como sucede en la Figura 8-36 ya que la palabra no está en posición final de una frase declarativa.

Figura 8-37: [pan. 'ʃin] (esp. de pescado) (realizado como [pàn.ʃín])



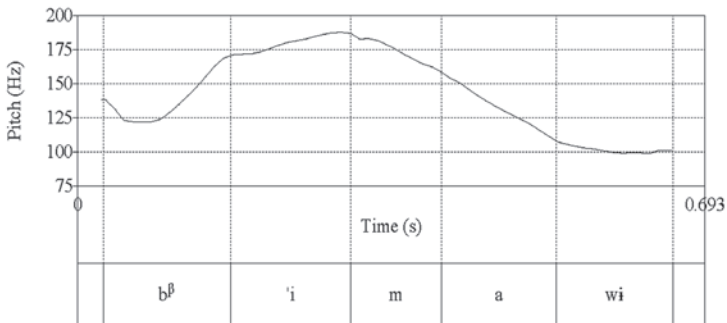
No todos los sufijos tienen acento léxico. La Figura 8-38 muestra el comportamiento de otro sufijo monosilábico, /-ki/ (CMPL), que no posee un acento léxico. Como en la Figura 8-36, el sufijo también aparece agregado a la raíz verbal /b^βi/ 'traer, recoger'. En contraste a la Figura 8-36, el acento principal cae sobre la sílaba inicial y el sufijo obtiene un tono bajo-descendente.

Figura 8-38: [^lb^βi.ki] ‘(lo) trajo’ (realizado como [b^βí.kì])



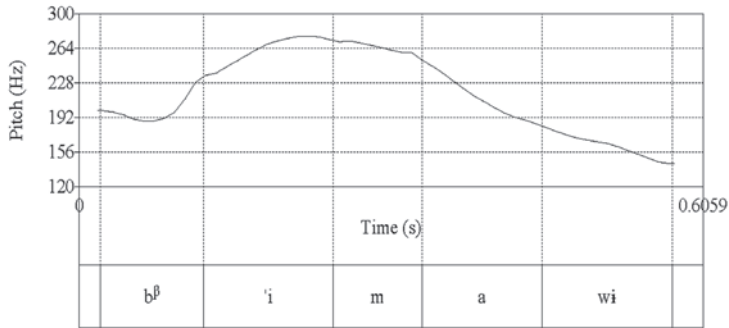
Fuera de la ventana del acento primario, los sufijos con acento léxico pueden mostrar un tono alto pero esto no siempre sucede. La Figura 8-39 muestra al sufijo /-wi/ (IMP), fuera de la ventana del acento principal: [^lb^βi.ma.wi] ‘haz(lo) que (lo) traiga!’ Además del sufijo imperativo, la palabra está compuesta de la raíz /b^βi/ ‘traer, recoger’, y el sufijo /-ma/ (CAUS). Nótese que el F0 cae como es usual después de la sílaba con el acento primario, en este caso la sílaba inicial. Sin embargo, crucialmente el F0 deja de caer durante el sufijo imperativo, lo cual interpreto como una indicación de la presencia de un tono alto.

Figura 8-39: [^lb^βi.ma.wi] ‘haz(lo) que (lo) traiga!’
(realizado como [b^βí.mà.wí])



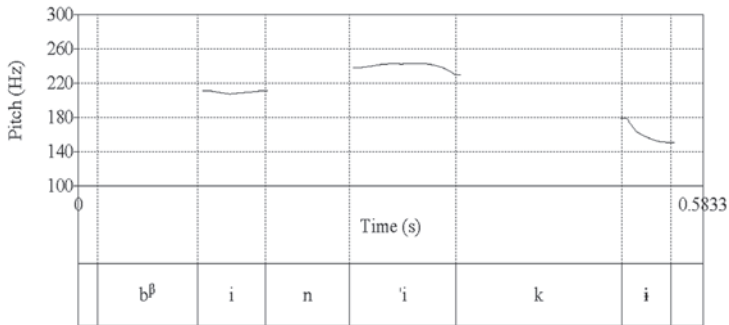
Alternativamente, fuera de la ventana del acento primario, el acento léxico del sufijo imperativo puede dejarse sin realizar. En este caso, el sufijo ocurre como un tono bajo-descendente como se muestra en la Figura 8-40. Aunque los patrones mostrados en la Figura 8-39 y la Figura 8-40 son bastante comunes para los sufijos con acento léxico fuera de la ventana del acento primario, hay una ligera tendencia a que el patrón de la Figura 8-40 ocurra más frecuentemente.

**Figura 8-40: [l^hb^βi.ma.wi] ‘haz(lo) que (lo) traiga!’
(realizado como [b^βi.mà.wì])**



Los sufijos léxicos no solo ocurren en la posición final de las palabras como en el caso del sufijo imperativo. También pueden ocurrir al interior de la palabra como lo muestra la Figura 8-41 a través del sufijo con acento léxico /-^hni/ (REM1).¹⁰ En este ejemplo, [b^βi.^hni.ki] ‘(lo) trajo hace ya mucho tiempo’, el sufijo /-^hni/ aparece después de la raíz monosilábica verbal /b^βi/ ‘traer, recoger’. Ya que ocurre dentro de la ventana del acento primario, el sufijo es capaz de llevar el acento principal de la palabra. El sufijo /-ki/ (CMPL) lo sigue.

**Figura 8-41: [b^βi.^hni.ki] ‘(lo) trajo hace ya mucho tiempo’
(realizado como [b^βi.ní.kì])**

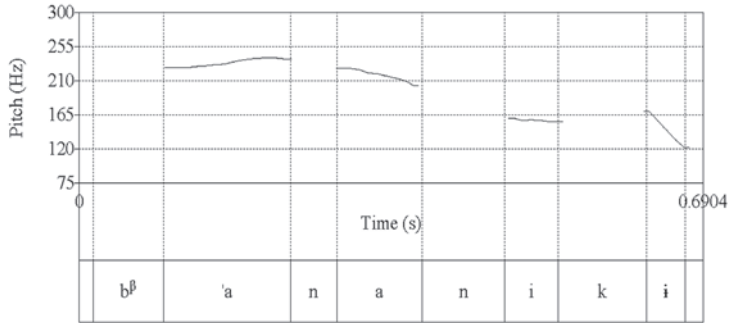


Fuera de la ventana del acento primario, este sufijo puede aparecer asociado con un tono alto, el cual generalmente se realiza como nivelado. Esto se puede

¹⁰ Para algunos de los hablantes entrevistados, el sufijo /-^hni/ (REM1) no tiene un acento léxico. Para esos hablantes, la palabra en la Figura 8-41 se pronuncia con el acento primario sobre la sílaba inicial: [b^βi.ni.ki].

apreciar en la Figura 8-42: [l^βa.na.ni.ki] ‘(lo) sembró hace ya mucho tiempo’ (/b^βana/ ‘sembrar’, /-^lni/ REM1 y /-ki/ CMPL).

Figura 8-42: [l^βa.na.ni.ki] ‘(lo) sembró hace ya mucho tiempo’
 ([b^βá.nà.ní.kì])



8.8 F0 Y LA POSICIÓN DE LOS ACENTOS SECUNDARIOS EN COMPUESTOS

Habiendo ya caracterizado el comportamiento del F0 en palabras simples, ahora podemos dirigir nuestra atención a compuestos (es decir, palabras formadas por dos o más palabras). En particular, esta sección examinará cómo el F0 puede servir como indicador de la presencia de acentos secundarios. Como en el caso de las palabras simples, cuando dos o más raíces/bases se juntan para formar un compuesto, el acento primario siempre aparece dentro de la ventana de acentuación en el borde izquierdo de la palabra (es decir, las dos sílabas iniciales).

La posición de los acentos secundarios, cuando se realizan fonéticamente por un tono alto, dependerá de dónde se encuentran los acentos primarios cuando cada palabra que forma el compuesto ocurre independientemente. Es decir, más allá de la ventana del acento primario en compuestos, las sílabas que llevarían el acento primario si las raíces del compuesto se usasen como palabras independientes llevan los acentos secundarios.

Este patrón se puede observar en los datos en (30) a (33). En cada conjunto de datos, se puede ver dónde el acento aparece cuando las palabras que forman el compuesto son independientes. El acento principal del segundo sustantivo se convierte en el acento secundario una vez que el sustantivo forma parte del compuesto.

- (30) a. [¹a.ta.pa] ‘gallina’
 b. [ʃu.b^βu] ‘casa’
 c. [¹a.ta.pa,_iʃu.b^βu] ‘gallinero’
- (31) a. [¹a.ni] ‘grande’
 b. [ka.¹pɪ] (from /kapik/) ‘caimán’
 c. [¹a.ni.ka,_ipɪ] ‘cocodrilo’
- (32) a. [ka.na] (esp. de guacamayo)
 b. [dʒu.¹nun] (/dʒunʉ -n/) ‘serpiente’ (FL)
 c. [ka.na.dʒu,_inun] (esp. de víbora amarilla venenosa - FL)
- (33) a. [kan.¹tsin] (un tipo de palmera)
 b. [pa.¹dʒan.ta,_inin] (/pa¹dʒanta-nin/) ‘plátano’ (FL)
 c. [kan.¹tsin.pa,_idʒan.ta,_inin] (un tipo de plátano - FL)

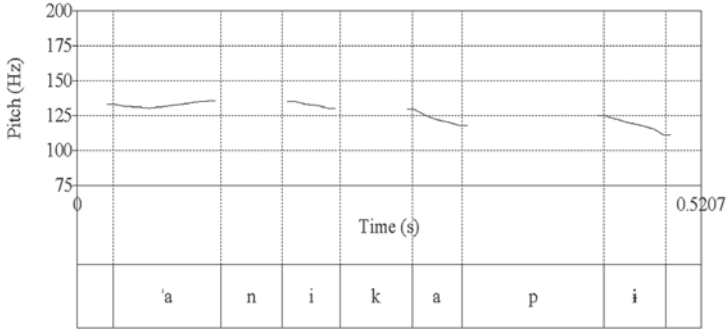
Sin embargo, el tratar de evitar los choques acentuales (Lieberman y Prince 1977) crea conflictos con el requerimiento de preservar los acentos de las palabras independientes que forman el compuesto. Este conflicto puede causar cambios en la posición final de los acentos secundarios. Las siguientes secciones discuten las diferentes estrategias que usa el shipibo para resolverlos.

8.8.1 La realización de los acentos secundarios en compuestos

La realización de los acentos secundarios en compuestos, indicada por los cambios en el F0, no es obligatoria, especialmente en el habla rápida. Esta tendencia se puede observar en contorno tonal mostrado en la Figura 8-43 para el compuesto [¹a.ni.ka,_ipɪ] ‘cocodrilo’ formado de las raíces /ani/ ‘grande’ y /kapik/ ‘caimán’.

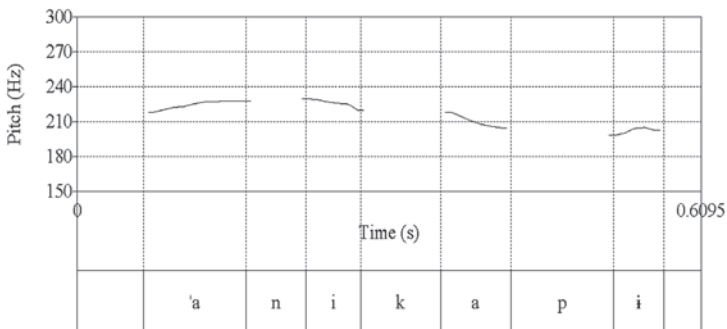
En [¹a.ni.ka,_ipɪ], esperamos un acento secundario sobre la sílaba final ya que la segunda sílaba de la palabra para ‘caimán’ lleva un acento principal cuando ocurre como una palabra independiente (/kapik/ → [ka.¹pik] → [ka.¹pɪ]). Sin embargo, como se observa en la Figura 8-43, solo la sílaba inicial, la cual lleva el acento primario del compuesto, está alineada con un tono alto. Después de la sílaba inicial, el F0 se mantiene cayendo hasta el final.

Figura 8-43: /ani kapik/ → [ˈa.ni.kɑ.ɸi] 'cocodrilo'
(realizado como [á.nì.kà.ɸì])



Sin embargo, las sílabas con acento secundario pueden reclamar su propio tono alto. Esto se da frecuentemente en el habla cuidada. La Figura 8-44 muestra el mismo compuesto, [ˈa.ni.kɑ.ɸi] ‘cocodrilo’, pero esta vez, el hablante sí alinea un tono alto con la vocal final indicando su estatus acentuado. Después del tono alto-ascendente asociado con la sílaba inicial, el F0 cae. Crucialmente, deja de caer justo antes de la sílaba final, y se vuelve un tono alto-nivelado. Aunque en la Figura 8-44, luce como que el F0 está subiendo en la vocal final, el incremento es menos de un semitono (0,6 ST), en consecuencia, lo considero como un tono alto-nivelado.

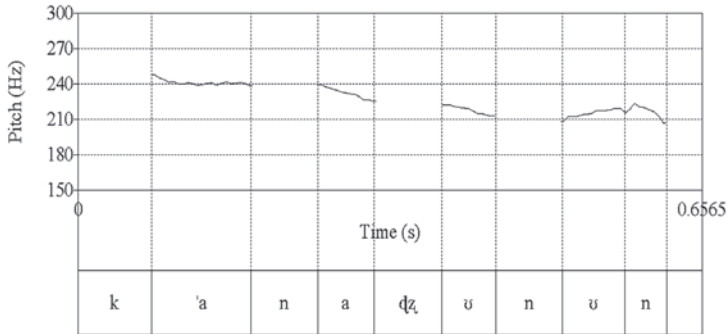
Figura 8-44: /ani kapik/ → [ˈa.ni.kɑ.ɸi] 'cocodrilo'
(realizado como [á.nì.kà.ɸì])



No solo las sílabas con codas latentes pueden llevar acentos secundarios en los compuestos (como en el caso de la Figura 8-44), sino que también lo pueden hacer las sílabas con codas nasales o codas sibilantes. Esto se presenta en la Figura

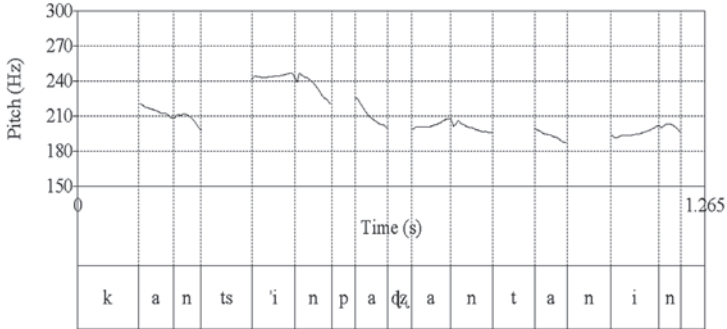
8-45, la cual muestra el contorno tonal del compuesto ['ka.na.dʒu.nun] (esp. de víbora - ['ka.na] (esp. de guacamayo), [dʒu.nun] 'serpiente, víbora', y el sufijo /-n/ FL). En este caso, la coda nasal (la de la sílaba [nun]) atrae el tono alto. Obsérvese que el patrón del F0 es idéntico al de la Figura 8-44.

Figura 8-45: /kana dʒunʉ-n/ → ['ka.na.dʒu.nun] 'esp. de víbora' (FL)
 ([ká.nà.dʒù.nún])



Antes de continuar la discusión hacia los casos de resolución de los choques acentuales, examinemos un compuesto más que muestra una mayor complejidad en términos del número y distribución de las sílabas abiertas y cerradas: [kan.'tsin.pa.dʒan.ta.nin] (un tipo de plátano - FL). La palabra contiene las raíces [kan.'tsin] (un tipo de palmera), [pa.'dʒan.ta] 'plátano' y el sufijo /-nin/ (FL). Cuatro de las seis sílabas de este compuesto son cerradas. Con la excepción de la sílaba inicial, todas las sílabas cerradas llevan acento y reciben un tono alto, que se puede identificar porque no hay una caída del F0 desde las sílabas inacentuadas precedentes. Obsérvese lo que sucede después de la sílaba con el F0 más alto, ['tsin], la cual lleva el acento primario. El F0 case por una sílaba (la sílaba [pa]) pero detiene su caída cuando llega a la vocal de la sílaba cerrada [dʒan]. Como es de esperar, el F0 reasume su caída durante la coda nasal de esa sílaba y se mantiene cayendo durante toda la sílaba [ta]. Nuevamente, tan pronto como llega a la vocal de la sílaba cerrada final [nin], el F0 detiene su caída, indicando que lleva el segundo acento secundario del compuesto.

Figura 8-46: [kan.¹tsin.pa.₁dʒan.ta.₁nin]
 (esp. de plátano - [kàn.tsín.pà.dʒán.tà.nín])



8.8.2 Choques acentuales y desacentuación

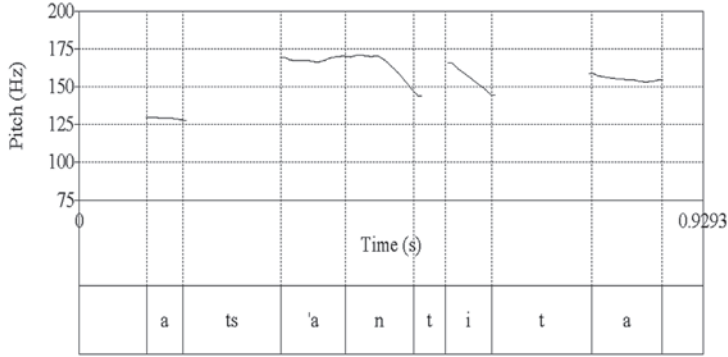
El shipibo evita la creación de choques acentuales (es decir, dos sílabas acentuadas que ocurren adyacentes). En compuestos, la posibilidad de crear un choque acentual ocurre cada vez que un sustantivo bisilábico con el acento en la segunda sílaba se une a otro sustantivo que le sigue y que tiene el acento en la sílaba inicial. El shipibo evita los choques acentuales por medio de desacentuar una de las sílabas. Si el sustantivo bisilábico con el acento sobre la segunda sílaba cae dentro de la ventana del acento primario, su segunda sílaba siempre mantiene su estatus de acentuado, la sílaba fuera de la ventana del acento primario pierde su acento (es decir, se desacentúa). Véase los datos de (34) a (37). Obsérvese que en estos casos, los compuestos pueden mostrar dos patrones acentuales. En el primer patrón, la sílaba con el acento primario y la sílaba final ocurren acentuadas. En el segundo patrón, solo la sílaba con el acento primario recibe un tono alto.¹¹

- (34) a. [ka.¹pɪ] (from /kapik/) 'caimán'
 b. [ʰdʒu.nu] 'serpiente'
 c. [ka.¹pɪ.dʒu.₁nu] ~ [ka.¹pɪ.dʒu.nu] (esp. de víbora)
- (35) a. [a.¹tsan] (from /atsa -n/) 'yuca' (FL)
 b. [ʰti.ta] 'madre'
 c. [a.¹tsan.ti.₁ta] ~ [a.¹tsan.ti.ta] (un tipo de yuca)

¹¹ Alternativamente, uno puede unificar los dos patrones en el nivel fonológico asumiendo que el acento secundario siempre está presente en la representación fonética (por ejemplo, [a.¹tsan.ti.₁ta]) pero en realización fonética, no siempre obtiene un tono alto ([à.¹tsán.tì.₁tá] or [à.¹tsán.tì.₁tà]).

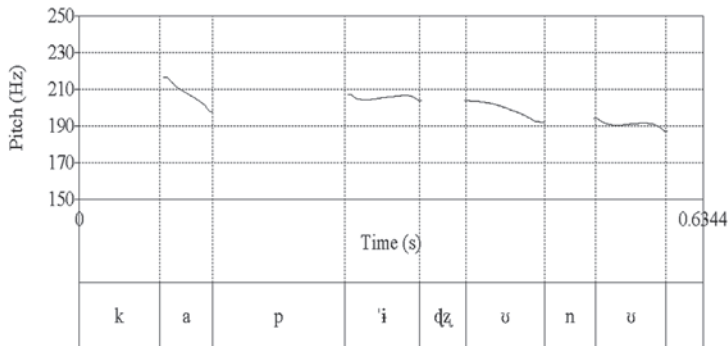
de la palabra, en la Figura 8-48, el F0 detiene su caída durante la sílaba final. En lugar de un tono bajo-descendente, la sílaba final muestra un tono alto-nivelado.

Figura 8-48: /atsa-n tita/ → [a.¹tsan.ti.ta] (tipo de yuca)



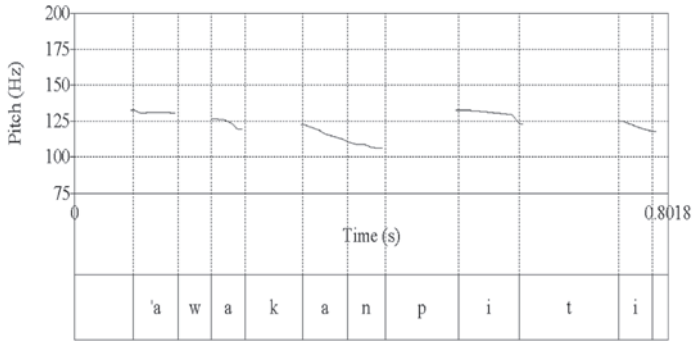
El patrón que se muestra en la Figura 8-48 no es exclusivo de este compuesto, [a.¹tsan.ti.ta], en particular. Otro compuesto similar se presenta en la Figura 8-49, [ka.¹pi.dzɯ.nu] (un tipo de serpiente), esta vez producido por una hablante mujer. Este compuesto está formado por /kapik/ ‘caimán’ (pronunciada como una palabra independiente con el acento sobre la segunda sílaba: [ka.¹pi]) y /dzɯnu/ ‘serpiente’ (pronunciada como una palabra independiente con el acento en la primera sílaba: [dʒu.nu]). Nuevamente, a fin de evitar tener dos sílabas acentuadas adyacente, la sílaba [dzɯ] se desacentúa. Como en la Figura 8-48, el primer tono alto se asocia con la sílaba que tiene el acento primario (la segunda sílaba) y después de esa sílaba, el F0 cae hasta que llega a la sílaba final. De modo similar a lo que se observó en la Figura 8-48, en ese momento, el F0 deja de caer y se convierte en un tono alto-nivelado.

Figura 8-49: /kapik dzɯnu/ → [ka.¹pi.dzɯ.nu] (un tipo de víbora)
 ([kà.pi.dzɯ.nú])



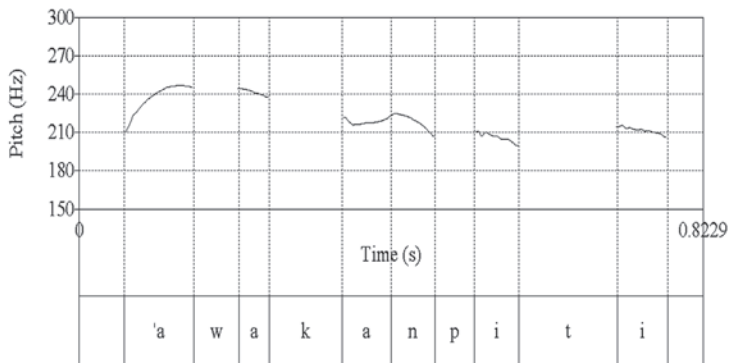
cuando el sustantivo se usa como una palabra independiente: [ˈpi.ti]. Hay una tendencia de la mayoría de los hablantes de shipibo hacia favorecer esta estrategia. Obsérvese cómo el F0 cae durante toda la sílaba [kan] en la Figura 8-51. Luego el F0 sube durante la sílaba siguiente, [pi], indicando que lleva un acento secundario.

Figura 8-51: /awak-an piti/ → [ˈa.wa.kan.pi.ti]
(*esp. de liana* - [á.wà.kàn.pí.tì])



Sin embargo, el desacentuar la sílaba cerrada no es la única manera de evitar los choques acentuales. Algunas veces, los hablantes desacentúan la sílaba inicial de la segunda palabra del compuesto, en este caso la sílaba [pi] y mueven el acento secundario a la siguiente sílaba de modo que la sílaba cerrada pueda obtener un acento secundario y un tono alto: [ˈa.wa.kan.pi.ti]. Este patrón se muestra en la Figura 8-52. Después de la sílaba inicial con el acento primario, el F0 cae hasta que llega a la sílaba cerrada donde detiene su caída. Después, el F0 empieza a caer nuevamente hasta que llega al inicio de la sílaba final durante la cual el F0 deja de caer una vez más. Compárese con el F0 de la Figura 8-51.

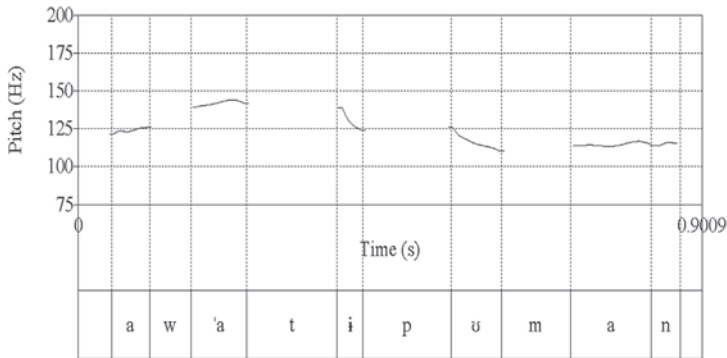
Figura 8-52: /awak-an piti/ → [ˈa.wa.kan.pi.ti]
(*esp. de liana* - [á.wà.kàn.pì.tí])



El compuesto [a.¹wa.ti.pu.₁man] (esp. de planta) nos da un buen ejemplo de un caso de desacentuación para evitar un choque acentual *[a.¹wa.₁ti.pu.₁man], pero donde se bloquea la posibilidad de mover el acento (véase la Figura 8-53). Este compuesto está formado por [a.¹wa] ‘tapir’ (/awak/) y [ti.pu.₁man] ‘tendón del cuello’ (FL). Ya que el primer sustantivo es bisilábico y tiene el acento sobre la segunda sílaba, la sílaba inicial [ti] del siguiente sustantivo [ti.pu.₁man] se desacentúa para evitar la creación de un choque acentual: [a.¹wa.ti.pu.₁man]. Sin embargo, si el acento se moviese a la siguiente sílaba, [pu], crearía otro choque acentual, esta vez con la sílaba cerrada final [man]: *[a.¹wa.ti.pu.₁man].

Este compuesto ofrece un buen contraste con el compuesto [a.wa.₁kan.pi.₁ti] ~ [a.wa.kan.₁pi.ti] (Figura 8-51 y Figura 8-52), donde la sílaba cerrada puede opcionalmente desacentuarse. En [a.¹wa.ti.pu.₁man], la sílaba cerrada nunca se desacentúa para permitir que la sílaba abierta anterior, [pu], lleve el acento secundario: *[a.¹wa.ti.₁pu.man] y así mantener el ritmo de los acentos. Esto apoya la intuición que las posiciones de los acentos en los compuestos se heredan tanto como es posible de los acentos de las raíces o bases cuando se usan como palabras independientes. De este modo, en el compuesto [a.wa.₁kan.pi.₁ti] ~ [a.wa.kan.₁pi.ti], la sílaba [pi], a diferencia de [pu] en [a.¹wa.ti.pu.₁man], puede opcionalmente desacentuar la sílaba cerrada anterior ya que es una sílaba que lleva acento cuando [pi.ti] se usa como una palabra independiente.

Figura 8-53: /awak tipum-an/ → [a.¹wa.ti.pu.₁man]
(esp. de planta - [à.wá.tì.pù.mán])



Antes de terminar esta sección, es importante comparar los diferentes grados de fuerza que muestran las sílabas cerradas cuando ocurren dentro y fuera de la ventana del acento primario. Considérese, por ejemplo, la sílaba cerrada del compuesto [a.¹tsan.ti.₁ta] ~ [a.¹tsan.ti.ta] (tipo de yuca - Figura 8-48). Esta sílaba

cerrada está dentro de la ventana del acento primario y como consecuencia, es suficientemente fuerte como para forzar la desacentuación de la siguiente sílaba como la única forma de resolver el choque acentual. Una sílaba cerrada que pertenece a esa ventana acentual nunca renuncia a su derecho de llevar el acento principal de la palabra. La forma *[¹a.tsan.₁ti.ta], con el acento primario forzado a moverse a la sílaba inicial, nunca es una alternativa para resolver un choque acentual en shipibo.

En comparación, considérese la sílaba cerrada del compuesto [¹a.wa.kan.pi.₁ti] ~ [¹a.wa.kan.₁pi.ti] (esp. de liana - Figura 8-52). Esa sílaba cerrada no está protegida por la ventana del acento primario y en consecuencia, no puede forzar la desacentuación de la siguiente sílaba para evitar el choque acentual. De este modo, algunas se desacentúa ([¹a.wa.kan.₁pi.ti]), algo que nunca ocurre con una sílaba cerrada que pertenece a la ventana del acento primario), y otras veces, logra reclamar su estatus como portadora de un acento secundario empujando otros acentos hacia la derecha ([¹a.wa.₁kan.pi.₁ti]).

8.9 RESUMEN

En este capítulo, se ha establecido que el tono alto es una clave importante para identificar las sílabas con acento primario. Se ha presentado numerosos ejemplos del comportamiento del F0 y su correlato, el tono perceptual, dentro del dominio de las palabras del shipibo. Se vio cómo los contornos tonales cambian de acuerdo a la estructura silábica de la palabra. Excepto por algunos pocos casos donde el acento principal no es predecible, éste cae sobre la segunda sílaba de las palabras cuando la sílaba es cerrada. De otro modo, cae sobre la sílaba inicial. El capítulo presentó evidencia fonética por medio de contornos tonales que cuando dos sílabas cerradas compiten por el acento principal, la sílaba cerrada inicial siempre es tratada como si fuese una sílaba abierta. Es decir, no puede atraer el acento principal y de este modo, no recibe un tono alto.

Se mostró las propiedades del F0 del tono alto en las sílabas con acento primario. Es decir, se determinó cuál era el promedio de los rangos de F0 en sílabas con acento primario tanto para hablantes hombres como para hablantes mujeres. Se estableció que la realización fonética preferida del tono alto en este tipo de sílaba es el tono alto-nivelado. El capítulo luego discutió las propiedades fonéticas del tono bajo en sílaba inacentuadas iniciales. En este tipo de sílabas, la realización preferida es un tono bajo-nivelado. Uno de los hallazgos más interesantes de este capítulo fue que las hablantes mujeres del shipibo posee una distribución ligeramente diferente del tono alto que lo hablantes hombres.

Las hablantes mujeres prefieren usar un tono alto-nivelado en sílabas inacentuadas segundas donde los hablantes hombres usan un tono bajo-descendente.

Se estudió el grado de eficiencia del tono perceptual, la duración y la intensidad como claves fonéticas que sirven para identificar las sílabas con acento primario. Se mostró que la duración es la clave más eficiente (identifica el 80% de este tipo de sílabas), seguida por el tono alto (68% de éxito en identificar las sílabas con acento primario). La clave menos eficiente fue la intensidad. Solo identifica 50% de esas sílabas. Cuando la duración y el tono perceptual se toman juntos como claves para el acento primario, se obtiene el nivel de éxito más alto. Ambos juntos pueden identificar el 100% de los casos de sílabas con acento primario.

El capítulo examinó el comportamiento del tono perceptual más allá de la ventana de acento primario. Más allá de esa ventana, las sílabas cerradas pueden atraer un tono alto indicando su condición de portadores del acento secundario. Sin embargo, el estilo del habla, en particular el habla rápida, afecta la habilidad de las sílabas cerradas para ocurrir alineadas con un tono alto. En este tipo de habla, las sílabas cerradas no siempre pueden asociarse con un tono alto. También mostré otra fuente de tonos altos en shipibo: la existencia de sufijos con acentos léxicos. Por medio de los contornos tonales, mostré que aunque este tipo de sufijo siempre atrae un tono alto dentro de la ventana del acento primario, más allá de ese ambiente, no se puede garantizar la ocurrencia de un tono alto asociado con ese tipo de sufijos.

Finalmente, el capítulo se enfocó en el F0 de los compuestos. Las posiciones acentuadas que tiene un compuesto, tal como lo indica la presencia de un tono alto, mayormente se heredan de aquellas sílabas que recibirían acento si las palabras que forman el compuesto ocurriesen independientemente. Sin embargo, con el fin de evitar los choques acentuales, las posiciones acentuales pueden alterarse ya sea por medio de la desacentuación o por medio de mover los acentos a sílabas adyacentes.

Con este capítulo he concluido la caracterización acústica de los segmentos y el acento del shipibo. El siguiente capítulo volcará nuestra atención a asuntos fonológicos. En particular, tendrá como objetivo principal bosquejar una fonología del shipibo conectando los análisis acústicos presentados con los patrones fonológicos observados en la lengua.

9. UN BOSQUEJO DE LA FONOLOGÍA DEL SHIPIBO

El propósito de este capítulo es conectar las mediciones acústicas presentadas en los capítulos anteriores con un bosquejo de la fonología de los segmentos y del acento del shipibo. Bosquejar la fonología de una lengua involucra interpretaciones tanto de hechos fonéticos como de patrones fonológicos. Con el fin de poder obtener tales interpretaciones debemos usar un marco teórico. Inclusive la propuesta fonológica más simple debe asumir un marco teórico con el cual se puedan modelar los fenómenos fonológicos que se observan. En esta sección, haré explícitas los presupuestos que usaré a fin de presentar un bosquejo de la fonología del shipibo.

Implícitamente o explícitamente, cualquier teoría fonológica tiene dos componentes básicos: un conjunto de posibles representaciones y un sistema computacional que manipula tales representaciones. Las representaciones están formadas por un conjunto finito de unidades. Para algunos lingüistas, esas unidades pueden ser segmentos (por ejemplo, un inventario de consonantes y vocales) y suprasegmentos (acento, tonos). Sin embargo, para la mayoría de fonólogos, los bloques básicos de las representaciones fonológicas son unidades más pequeñas que los segmentos. Estas unidades son los rasgos ([labial], [+sonoro], etc). En este punto, también encontramos diferentes enfoques sobre qué es un rasgo y cómo interactúa con otros rasgos. Los rasgos pueden asumirse, al estilo de SPE (Chomsky y Halle 1968), que se agrupan en matrices sin ninguna organización interna o siguiendo un enfoque de geometría de rasgos (Clements 1985), se les puede concebir con un conjunto de unidades organizadas jerárquicamente. Alternativamente, las unidades de las representaciones fonológicas también pueden pensarse como gestos que se sobreponen en el tiempo (véase fonología articulatoria o gestual - Browman y Goldstein 1989, 1992).

De nuevo, implícitamente o explícitamente, todas las teorías fonológicas aceptan la existencia de algunos niveles organizacionales por encima de los segmentos/rasgos (moras (μ), sílabas (σ), pies métricos (π), palabras prosódicas (PIPr), etc). El grado de controversia entre fonólogos con respecto a la evidencia para estos niveles organizacionales varía bastante dependiendo de la posición teórica que mantenga el analista.

Tal vez uno de los puntos más controversiales en los estudios fonológicos modernos es el sistema computacional. Algunos fonólogos asumen que las representaciones son manipuladas por reglas (este tipo de enfoque generalmente es conocido como ‘derivacional’). De este modo, una representación que posee una cierta propiedad/rasgo en un contexto dado sufre un cambio específico. El resultado luego pasa a otra regla, la cual puede realizar otros cambios. Otros fonólogos niegan la existencia de reglas y asumen que las representaciones se procesan a través de constricciones que escogen entre un conjunto de posibles resultados (candidatos) cuál es el mejor bajo ciertas condiciones (este enfoque se le conoce comúnmente como ‘no-derivacional’). La teoría más prominente de este tipo es la Teoría de la Optimicidad (Prince y Smolensky 1993). Adicionalmente, hay fonólogos que creen que la motivación de las reglas o constricciones debe ser interna a la teoría mientras otros piensan que deben ser motivadas sobre bases funcionales.

Para el propósito de este capítulo, que es presentar un bosquejo de la fonología del shipibo que pueda servir de punto de contacto entre la caracterización acústica de la lengua y los estudios fonológicos, mantendré mis presupuestos representacionales y computacionales tan simples y comúnmente aceptados como es posible. Asumiré la existencia de rasgos asociados a algún tipo de unidad temporal (por ejemplo, la mora), lo cual es una asunción bastante bien aceptada en diferentes teorías fonológicas. También asumiré que los rasgos se organizan dentro de los segmentos.¹ Usaré niveles organizacionales por encima del nivel del segmento como lo requiera el fenómeno que se esté discutiendo. Aunque evitaré el uso de reglas, sí usaré estratos fonológicos en varios puntos de la discusión.

9.1 UNA FONOLOGÍA DE LAS VOCALES DEL SHIPIBO

9.1.1 Los rasgos de las vocales

Shipibo tiene cuatro vocales fonémicas que represento con los símbolos fonéticos /i, a, u, ɨ/. Las oposiciones básicas entre las cuatro vocales del shipibo pueden

¹ En uno de los desarrollos más recientes de la Teoría de la Optimicidad se trata de eliminar la noción de geometría de rasgos (Oostendorp and Weijer [2005]).

describirse fonológicamente por medio de tres rasgos, tal como se muestra en la Figura 9-1: [\pm alto], [\pm posterior], [redondeado].² Conjuntamente con esta información fonológica disponible a la gramática, podemos asumir que los hablantes de shipibo también guardan información fonética detallada sobre las propiedades acústicas de la posición de cada vocal en el espacio vocálico.³ Esas posiciones acústicas ideales se caracterizaron en términos de los formantes F1 y F2 en la sección 2.1.

Figura 9 1: Rasgos de las vocales del shipibo

	[-posterior]	[+posterior]	[+posterior], [redondeado]
[+alto]	i	i	u
[-alto]		a	

9.1.2 La longitud vocálica

La sección 2.4 nos presentó evidencia que en la superficie el shipibo tiene dos tipos de longitud vocálica: vocales cortas (V) y vocales largas (V:). Esto nos hace preguntarnos si esto es una distinción fonológica, es decir, si la longitud vocálica es parte de las representaciones fonológicas de la lengua o no. El estatus fonológico de la longitud de las vocales largas no es un problema fácil de resolver como parece a primera vista.

Examinemos cómo funciona la longitud vocálica en shipibo. Las raíces monosilábicas que pueden aparecer sin sufijos obligatoriamente tienen una vocal larga. Este es el caso de las raíces nominales (por ejemplo, [ˈna:] ‘nido’). En shipibo, no es posible tener una raíz monosilábica nominal con un vocal corta. Esto no es particularmente sorprendente. En muchas lenguas, como el inglés, alemán, holandés, latín, bengalí, entre otras, se sabe que los monosílabos léxicos muestran vocales largas (Hayes 1980; Cole 1990; McCarthy y Prince 1991; Piggott 1993; Mester 1994; Lahiri 2000). La explicación más aceptada es que este alargamiento ocurre para cumplir con el requisito del tamaño mínimo de la palabra (McCawley

² Véase la sección 9.2.2 para más información sobre la representación en rasgos de la vocal central alta /i/.

³ De hecho, uno puede ir un paso más adelante y asumir que cuando los niños adquieren una lengua empiezan guardando información fonética detallada del input lingüístico que reciben. Parte de esa información, las oposiciones cruciales entre los segmentos que causan distinciones en el significado, se fonologiza y se vuelve accesible a sistemas cognitivos más altos (lenguaje). La información fonética restante permanece guardada en los sistemas sensorio-motores como parte de nuestra base perceptual y articulatoria. Este tipo de información fonética se activa en la realización del habla.

1968; Prince 1976, 1983; Hyman 1985; McCarthy y Prince 1986; Zec 1988; Hayes 1989; Ito 1989; Hayes 1994; Davis 1995; Hayes 1995, entre otros).

Además, como parte de este fenómeno de la talla mínima de la palabra, si una raíz monosilábica se sufixa o si tiene suficiente material fonológico para satisfacer la talla mínima de la palabra, la vocal larga no ocurre. Esto también se observa en shipibo. Ya que las raíces verbales no pueden aparecer sin sufijos y todos los sufijos obligatorios (es decir, sufijos de tiempo, de cambio de referencia, etcétera) son por lo menos monosilábicos, la palabra resultante siempre satisface el tamaño mínimo de la palabra. De este modo, en shipibo, las raíces monosilábicas verbales como las que se muestran en (38) siempre tienen una vocal corta. En (38), los verbos ocurren con el sufijo /-ti/ (INF).

- (38) a. [ʼa.ti] ‘hacer’
 b. [b^βi.ti] ‘obtener’
 c. [b^βi.ti] ‘traer’
 d. [b^βu.ti] ‘tomar’
 e. [ha.ti] ‘existir’
 f. [h^u.ti] ‘venir’
 g. [i.ti] ‘ser, suceder’
 h. [ka.ti] ‘ir, partir’
 i. [ni.ti] ‘pararse, caminar’
 j. [pi.ti] ‘comer’

Todo esto sugiere que la longitud vocálica no es subyacente en las raíces monosilábicas del shipibo ya que son derivables por medio de un requisito lingüístico que gobierna el tamaño mínimo de las palabras. Sin embargo, aquí es donde intervienen las relaciones paradigmáticas entre las palabras y las cosas se complican. Cuando una raíz nominal monosilábica se verbaliza, la raíz verbal resultante hereda la vocal larga del sustantivo. De este modo, por ejemplo, el sustantivo monosilábico [ku:] ‘pus’ se convierte en el verbo [ku:.ti] ‘supurar’ (/ku:/ ‘pus’, /-ti/ INF). La vocal de la raíz verbal es larga aunque la palabra tiene suficiente material fonológico como para satisfacer el requisito del tamaño mínimo. Éste también es el caso de todas las palabras en el paradigma que tienen la raíz [ku:]. Todas ellas heredan la vocal larga [u:] (véase los datos en (39)).

- (39) a. [ʼku:]‘pus’ (sustantivo sin sufijos)
 b. [ʼku:.ti] ‘supurar’ (/ -ti/ INF)
 c. [ʼku:.kɨ] ‘supuró’ (/ -kɨ/ CMPL)
 d. [ʼku:.wan.kɨ] ‘supuró hoy día’ (/ -wan/ PSD1, / -kɨ/ CMPL)

Los datos en (40) muestran un patrón similar al de las palabras en (39) pero esta vez el paradigma es nominal.

- (40) a. [tʰi:] ‘fuego’
 b. [tʰi:.ki] ‘fuego’ (/ -ki/ locativo)
 c. [tʰi:.dʒɨs] ‘solo el fuego’ (/ -dʒɨs/ ‘solo’)

Otra consecuencia de la preservación de las vocales largas entre palabras de un paradigma es la creación de pares mínimos cuya única diferencia es la longitud vocálica. Supongamos que hay dos raíces idénticas en términos de sus segmentos. Si una de ellas puede aparecer como una raíz sin sufijos, mostrará una vocal larga que compartirá con todas las palabras que contengan dicha raíz (tal como sucede en los datos en (39) y (40)). En contraste, la raíz que no puede aparecer sin sufijos siempre tendrá una vocal corta (véase los datos en (38)). Éste es el caso de los verbos [b^βu.ti] ‘tomar’ y [b^βu:.ti] ‘sanarse (una herida, la piel)’. La raíz verbal monosilábica para ‘sanarse (una herida, la piel)’ puede funcionar como un sustantivo: [b^βu:] ‘lugar sanado en la piel’. La vocal larga luego se impone sobre todas las formas verbales que contengan dicha raíz ([b^βu:.ti] (INF), [b^βu:.ki] (CMPL)). En contraste, la raíz monosilábica verbal /b^βu/ ‘tomar’ nunca ocurre como un sustantivo lo cual le impide obtener una vocal larga. En consecuencia, las formas verbales [b^βu.ti] y [b^βu:.ti] muestran un contraste en la longitud de la vocal inicial. Este es el único par mínimo de este tipo que conozco. Véase la Figura 9-2 y la Figura 9-3, las cuales muestran el espectrograma de estas palabras. Ambas fueron producidas por el mismo hablante. La duración de las vocales, [u] y [u:], se indican en cada espectrograma.

Figura 9-2: /b^βu -ti / → [b^βu.ti] ‘tomar’ (INF)

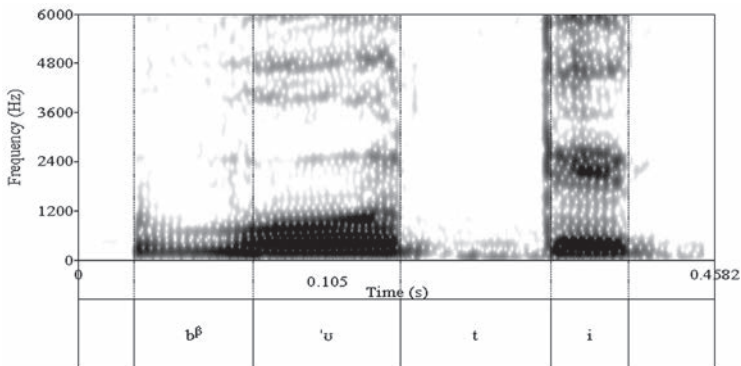
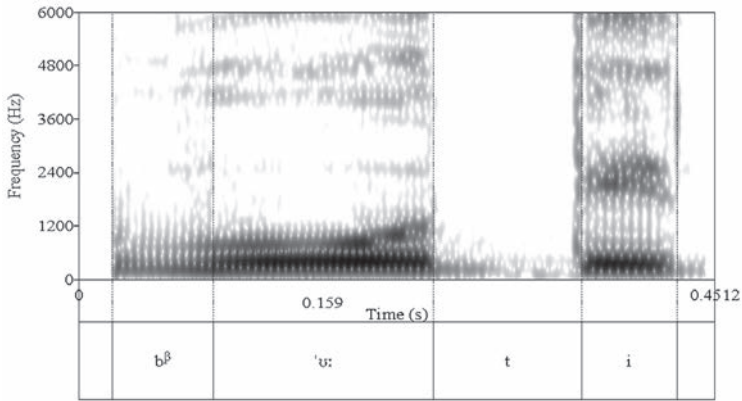


Figura 9-3: /b^βu: -ti / → [l^hb^βu:ti] ‘sanarse (una herida, la piel)’ (INF)



Ese par mínimo y el hecho que las vocales largas en las raíces monosilábicas nunca se acortan, incluso cuando ocurren sufijadas, puede tomarse como evidencia que sincrónicamente el shipibo distingue entre vocales cortas y vocales largas: /b^βu -ti / → [l^hb^βu:ti] ‘tomar’ (INF) versus /b^βu: -ti / → [l^hb^βu:ti] ‘sanarse (una herida, la piel)’ (INF).

El único caso que conozco que podría ser una excepción a la falta de alternancias entre vocales cortas y vocales largas es el prefijo [b^βu-] ‘en el cabello, a los lados de la cabeza,’ el cual viene de la raíz nominal monosilábica [l^hb^βu:] ‘cabello, pelo.’ Este prefijo ocurre, por ejemplo, en el verbo [b^βu. l^htis.ti] ‘sacar con fuerza el pelo de la cabeza’ (/b^βu-/ ‘pelo,’ /tisi/ ‘sacar, arrancar’ y /-ti/ (INF)). El sustantivo posee una vocal larga, [l^hb^βu:], pero el prefijo tiene una vocal corta, [b^βu-]. Examinemos brevemente este caso. Shipibo, como muchas otras lenguas panos, tiene un conjunto restringido de prefijos que se relacionan semánticamente a partes del cuerpo (Hall de Loos y Loos 1978). De hecho, éstos son los únicos prefijos que tiene shipibo. Cada prefijo es de una sílaba y su material fonológico viene del lado izquierdo del sustantivo del cual se deriva. La lista en (41) da algunos ejemplos es esos prefijos y de su relación fonológica con el sustantivo del cual provienen (Loriot, Lauriault *et al.* 1993).

(41) Prefijo	Glosa	Origen del Prefijo	Glosa
a. [b ^β u-]	‘en el pelo, o lados de la cabeza’	[l ^h b ^β u:]	‘pelo’
b. [b ^β a-]	‘en el sobaco, en el pecho’	[b ^β aʃ. l ^h ka]	‘sobaco’
c. [b ^β i-]	‘en el ojo, ceja, cara o superficie’	[l ^h b ^β i. dʒu]	‘ojo’
d. [tʃi-]	‘en las nalgas, cola’	[l ^h tʃi. ʃu]	‘nalga’
e. [han-]	‘en la lengua, boca’	[l ^h ha.na]	‘lengua’

f.	[hin-]	‘en el pene, cola’	[^h hi.na]	‘pene’
g.	[dʒan-]	‘en la rodilla’	[^h dʒan.b ^β u.ʃu]	‘rodilla’

Desafortunadamente, la palabra [b^βu:] es el único sustantivo monosilábico que se refiere a una parte del cuerpo y que tiene una vocal larga. En consecuencia, es imposible aseverar con certeza sobre si el acortamiento de la vocal es un patrón regular o no. Sin embargo, incluso si fuese un patrón regular, el acortamiento no necesariamente revelaría el estado subyacente de las vocales largas. Hay que tomar en cuenta que cuando una raíz se convierte en un prefijo, una parte importante de su naturaleza cambia. Es muy posible que cuando la raíz léxica [b^βu:] se vuelve el prefijo [b^βu-], de pronto está sujeta a requisitos fonológicos diferentes que obligan a los prefijos a tener vocales cortas.

El tipo de evidencia más fuerte para postular el estatus subyacente de las vocales largas del shipibo viene de la existencia de raíces polisilábicas cuyas sílabas iniciales contienen una vocal larga. Este tipo de evidencia, sin embargo, no es robusta ya que solo hay unas cuantas de esas raíces, seis para ser más preciso (véase la sección 2.4). Esto hace al shipibo un caso raro de una lengua que tiene vocales subyacentes largas pero que éstas están restringidas a raíces monosilábicas que tienen la habilidad de aparecer sin sufijos.

9.1.2.1 *Un análisis alternativo para las vocales largas*

He analizado la longitud vocálica como una propiedad que está presente en las representaciones subyacentes del shipibo. Sin embargo, hay un análisis alternativo que quita el peso de explicar la distribución de las vocales largas de las representaciones subyacentes. En vez, su distribución puede explicarse en términos del sistema computacional de la fonología. Si, como propone la Teoría de la Optimalidad (Prince y Smolensky 1993), el conocimiento fonológico puede expresarse como un conjunto de requisitos (restricciones) que se imponen sobre la buena formación de las representaciones fonéticas y que la satisfacción de esos requisitos tiene diferentes grados de importancia en cada gramática, se puede desarrollar el siguiente análisis para las vocales largas del shipibo.

Como se discutió en la sección anterior, interlingüísticamente las palabras tienden a mostrar un tamaño mínimo. Llamemos a esa tendencia *Minimalidad de la Palabra* (MP). En algunas lenguas, las palabras léxicas son mínimamente bisilábicas y en otras, las palabras léxicas más pequeñas tienen por lo menos una vocal larga. Shipibo pertenece al último grupo. Los datos en (42) muestran algunos ejemplos de raíces nominales monosilábicas. Este tipo de raíces es la fuente principal de vocales largas en shipibo.

- (42) a. [tʰi:] ‘trabajo’
 b. [bʰu:] ‘cabello, pelo’
 c. [tʰa:] ‘astilla’

También sabemos que las lenguas tienden a variar de acuerdo a si permiten o no la ocurrencia de vocales largas. Algunas lenguas, como el castellano, solo tienen vocales cortas mientras que otras, como el tailandés, distinguen entre vocales cortas y largas ([hèt] ‘champiñón’ vs. [hèt:] ‘causa’). En casos como el tailandés, la longitud vocálica está presente en las representaciones subyacentes de la lengua y se preserva en las representaciones fonéticas. Llamemos a este comportamiento de preservar la longitud subyacente de las vocales en las representaciones fonéticas, *preservación de la longitud vocálica* (PLV).

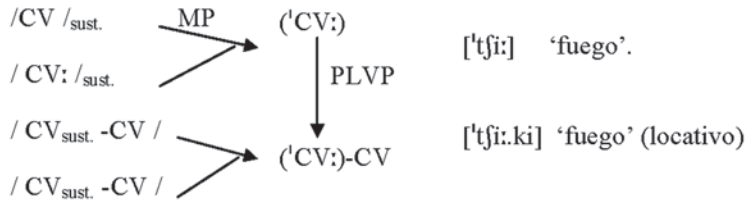
En shipibo, ambas fuerzas fonológicas, MP y PLV, ayudan a entender por qué los monosílabos muestran vocales largas. La satisfacción de MP es más importante que la satisfacción PLV. De este modo, una forma subyacente como /tʰi/ ‘fuego’ aparece con una vocal larga, [tʰi:], para poder satisfacer el requisito del tamaño mínimo de la Palabra impuesto a las palabras léxicas. Esto ocurre aunque la satisfacción de MP significa cambiar la longitud subyacente de una vocal corta a una larga y de este modo, desobedecer a PLV.

Además, como se demostró en los datos en (39) y (40), el shipibo presenta un giro inesperado al requerir la preservación de la longitud vocálica a través de las palabras de un paradigma. De este modo, si una palabra posee una vocal larga en cierto contexto, esta vocal larga se hereda a todos los miembros del paradigma. Esta condición puede pensarse como una versión especial del requisito general PLV. Llamémoslo *preservación de la longitud vocálica en paradigmas* (PLVP).

En la fonología del shipibo, la satisfacción de PLVP es tan importante como el requisito MP pero tiene un grado mayor de importancia que el PLV general. De este modo, mientras la representación subyacente de un monosílabo podría tener una vocal corta y ser forzada a presentar una vocal larga (/tʰi/ → [tʰi:] ‘fuego’), una forma sufijada del mismo monosílabo debe preservar la vocal larga a fin de satisfacer PLVP, inclusive si ya satisface el tamaño mínimo de la palabra (MP) (/tʰi -ki/ → [tʰi.ki] ‘en el fuego’). Esto se puede representar esquemáticamente en la Figura 9-4.

Obsérvese que en el análisis propuesto en la Figura 9-4, realmente no importa si subyacentemente los sustantivos monosilábicos tienen vocales largas o no. En cualquier caso, los resultados convergen a los patrones observables de la lengua. Si la vocal larga es subyacente, entonces ocurre larga. Si es corta, es forzada a ocurrir larga para satisfacer MP y al hacerlo así, las otras palabras en el paradigma

Figura 9-4: Preservación de la longitud vocálica en paradigmas



heredan la vocal larga. Bajo este análisis, la distribución de las vocales largas del shipibo está completamente codificada dentro del sistema computacional, no en las representaciones.

¿Cómo se compara este análisis al análisis representacional? La evaluación de ambos análisis no es fácil, particularmente por la existencia de las seis raíces polisilábicas que poseen vocales largas en sus sílabas iniciales (véase la sección 2.4). Sin embargo, poniendo esos seis casos de lado por el momento, se puede afirmar que la principal ventaja que tiene el análisis computacional es que explica por qué las vocales largas del shipibo muestran una distribución tan restringida dentro del léxico; es decir, aparecen en monosílabos que pueden ocurrir solos sin sufijos. En contraste, un análisis representacional encuentra difícil explicar por qué, si fuesen parte de las representaciones subyacentes, las vocales largas no pueden aparecer libremente en raíces polisilábicas.

Pero ¿cómo el análisis computacional podría lidiar con las seis raíces polisilábicas que tienen vocales largas o con el par mínimo ['b^βu.ti] 'tomar' versus ['b^βu:ti] 'sanarse (una herida, la piel)'? El análisis puede fácilmente explicar la existencia de los pares mínimos donde la longitud vocálica parece estar en juego. La raíz verbal monosilábica para 'tomar', /b^βu/, no puede aparecer como una raíz sin sufijos y en consecuencia el requisito MP siempre está satisfecho y la raíz solo muestra vocales cortas.

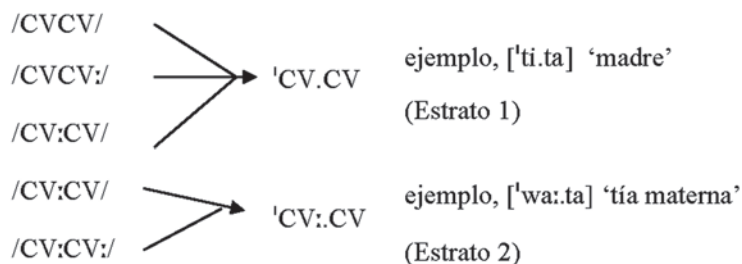
Algunas raíces verbales pueden funcionar como sustantivos sin adquirir una vocal larga. En estos casos, sin embargo, hay un proceso morfológico de sufijación. Este es el caso, por ejemplo, de la raíz /pi/ 'comer': [pi.ti] (INF), [pi.ki] (CMPL), [pi.ʃun] (CR) y otras formas similares. La vocal [i] de esta raíz siempre es corta. La raíz /pi/ puede funcionar como un sustantivo. Sin embargo, crucialmente, la operación morfológica usada para deverbalizar /pi/ es la sufijación. Se añade el nominalizador /-ti/, el cual proporciona suficiente material fonológico para satisfacer MP: [pi.ti] 'comida de carne', [pi.tin] (forma larga de [pi.ti]). De este modo, ya que la raíz /pi/ nunca aparece sin sufijos, ninguna de las palabras en su paradigma posee una vocal larga.

En contraste, la raíz verbal monosilábica para ‘sanarse (una herida, la piel)’ puede funcionar como una raíz nominal sin sufijos [ʼb^βu:] ‘lugar sanado de la piel’ y como resultado de satisfacer el requisito MP, adquiere una vocal larga que luego es heredada por todas las palabras en el paradigma. Éste es probablemente también el caso del verbo ‘trabajar’. Como sustantivo, ocurre con una vocal larga para satisfacer MP: [ʼt̪i:] ‘trabajo’. Luego, la vocal larga del sustantivo se hereda en todo el paradigma y de este modo, las formas verbales que contienen dicha raíz también muestran también una vocal larga ([ʼt̪i.ti] ‘trabajar’ (INF), [ʼt̪i.ki] ‘trabajar’ (CMPL), etc.), y como en el caso anterior, todos los sustantivos y adjetivos creados sobre la base de [ʼt̪i:] también heredan la vocal larga ([ʼt̪i.tai] ‘trabajador, obrero’, [ʼt̪i.ya] ‘trabajador’ (adjetivo), etc.)

Con respecto a las seis raíces polisilábicas con vocales largas (na: b^βa.ti ‘aumentar el líquido’ (INF), ɟza: ʂi.ti ‘aclimatarse’ (INF), ɟzu: kũn ‘tos ferina’, ɟzu: n. b^βi (esp. de árbol), wa: ta ‘tía materna’, wi: ki.ti ‘entrar en la pierna’ (INF)), se pueden tratar como casos especiales, como su minúsculo número sugiere. Pero ¿en qué sentido son especiales? Asumiendo que una gramática puede tener diferentes estratos léxicos, los requisitos fonológicos pueden diferir de un estrato léxico a otro. La idea de los estratos léxicos no es, por supuesto, una idea nueva en fonología (Kiparsky 1982a, 1985; Mohanan 1986; Kiparsky 1997, 2000). Uno de los casos más famosos de estratos léxicos en la literatura fonológica es el proceso de sonorización de compuestos del japonés (también conocido como Rendaku - Ito y Mester 1986, 1995, 2003). El japonés es bastante bien conocido por tener un léxico agrupado en cuatro estratos léxicos diferentes: el estrato Yamato (palabras nativas), el Sino-Japonés, el Mimético y el de préstamos. Rendaku ocurre característicamente en el estrato Yamato.

En el caso del shipibo, la lengua tiene un estrato léxico que contiene palabras polisilábicas con vocales largas. En ese estrato, la lengua posee una restricción que requiere la preservación de la longitud vocálica en las sílabas iniciales de las raíces (llamemos a esta restricción PLV-σ1). La satisfacción de esa restricción es más importante que satisfacer a una restricción que prohíbe la ocurrencia de vocales largas (NO-VOCALLES-LARGAS). En otras palabras, bajo la influencia de estas restricciones, las vocales largas subyacentes pueden mantener su longitud pero solo si ocurren como parte de la primera sílaba de la palabra. Las vocales largas subyacentes más allá de esa posición se acortarán ya que no están protegidas por PLV-σ1 y entonces deben obedecer a la restricción NO-VOCALLES-LARGAS. Este es el estrato 2 representado en la la Figura 9-5. En contraste, en el estrato general, las vocales largas no pueden ocurrir en ninguna sílaba de una raíz polisilábica porque la satisfacción de NO-VOCALLES-LARGAS es más importante que la preservación de la longitud vocálica (véase el estrato 1 en la Figura 9-5).

Figura 9-5: Longitud vocálica en raíces polisilábicas



La referencia a la sílaba inicial en el estrato 2 no es para nada sorprendente. La fonología del shipibo tiene una fuerte tendencia a referirse a las posiciones iniciales en general:

- las vocales largas solo pueden ocurrir en la sílaba inicial de las palabras (véase la sección 2.4).
- la centralización vocálica afecta a las vocales inacentuadas. Sin embargo, si una vocal es inacentuada pero ocurre en una sílaba inicial, muestra un grado muy bajo de centralización que la misma vocal en una sílaba inacentuada no-inicial (véase la sección 2.2).
- las sílabas iniciales de las palabras por defecto atraen el acento principal (al menos que la segunda sílaba fuera cerrada - véase la sección 8.1).
- los segmentos glotales solo ocurren en posiciones iniciales (véase la sección 2.5 y la sección 4.2). Con la excepción de algunos préstamos, la fricativa glotal, /h/, solo aparece como el arranque de las sílabas iniciales de las palabras. El caso de la oclusiva glotal subyacente, /ʔ/, es un poco más complejo pero la caracterización de su distribución también refiere a las posiciones iniciales. Solo ocurre al segmento inicial de ciertos sufijos.⁴
- el estudio de las africadas (véase la sección 4.4.3) indica que la africada bilabial sonora /b^β/ tiende a mostrar una realización fonética especial cuando ocurre después de pausas, un contexto fonológico que corresponde a la posición inicial de la frase.

⁴ Es importante mencionar que en un estudio cuidadosamente diseñado basado en los datos disponibles de alrededor de quince lenguas panos, González (en preparación) encontró que los segmentos glotales ([ʔ] y [h]) tienden a tener una distribución restringida. Aunque en algunas lenguas panos, estos segmentos glotales pueden ocurrir en posiciones finales (especialmente en posición final de palabra), en la mayoría de los casos, la distribución de estos segmentos está relacionada a las posiciones iniciales dentro de la palabra o de las raíces/morfemas.

Antes de terminar con esta sección, es importante enfatizar que bajo este análisis de restricciones, no importa cuál es el estatus subyacente de las vocales largas. En vez de codificar sus distribución en las representaciones de la lengua, la distribución se codifica en el sistema computacional, es decir, se codifica como relaciones entre restricciones fonológicas.

9.1.3 Cambios de la duración vocálica determinados por el contexto lingüístico

- El efecto de las posiciones privilegiadas sobre la duración vocálica

El estudio acústico indica diferentes comportamientos de la duración vocálica de acuerdo al contexto fonológico en el cual aparecen las vocales. Algunos de estos comportamientos se explican fácilmente por la existencia de posiciones privilegiadas (Beckman 1998). Otros, como se mostrará en breve, son más difíciles de explicar.

Empecemos por discutir los cambios en la duración vocálica determinados por el acento. Los resultados acústicos obtenidos para la duración de las vocales del shipibo en la sección 2.3 y la sección 7.1.2 indican que las vocales cortas en sílabas acentuadas tienden a ser entre 30 ms y 40 ms más largas que las vocales de sílabas inacentuadas. A pesar de estas diferencias fonéticas duracionales, estas vocales tienen la misma longitud fonológica. En términos moraicos, las vocales cortas tienen una mora ya sea en posiciones acentuadas o inacentuadas sin importar su duración fonética. Las vocales largas, en contraste, se caracterizan por poseer dos moras.

Las diferencias consistentes observadas en la duración de las vocales del shipibo de acuerdo a si están acentuadas o no, lo cual es un fenómeno fonético bastante conocido, pueden entenderse como un reflejo de si las vocales ocupan una posición lingüística de privilegio o no. Como lo describe Beckman 1998, las posiciones de privilegio (también conocidas como posiciones prominentes), incluyen las sílabas acentuadas, las posiciones iniciales y las raíces. Éstas son posiciones que disfrutan de ventajas perceptuales ya sea psicolingüísticamente o en términos de prominencia fonética.

En el plano fonético, las vocales que aparecen en estas posiciones de privilegio, en particular, en sílabas acentuadas, están de alguna forma protegidas de modo que pueden preservar su prominencia fonética. Como parte de esa protección, su duración fonética no se reduce como ocurre con las vocales de sílabas inacentuadas. Otras características fonéticas de las vocales acentuadas, también observadas en el shipibo incluyen sus resistencia a centralizarse, lo cual está acompañado de formantes más estables (véase la sección 2.2). Este comportamiento es diferente

al de las sílabas inacentuadas las cuales muestran varios grados de centralización/reducción y cuyos formantes son más propensos a ser afectados por los segmentos vecinos.

Además de las sílabas acentuadas, en algunas lenguas las posiciones iniciales pueden comportarse como posiciones de privilegio. Por ejemplo, en mongol (Svantesson 2005) y shona (Beckman 1995), el conjunto de alturas vocálicas que pueden ocurrir en las sílabas no-iniciales es un subconjunto de aquellas que pueden albergar las sílabas iniciales de palabra. En Tuva (Krueger 1977), mientras las sílabas iniciales pueden tener tanto vocales con fonación modal como vocales con fonación glotalizada, las sílabas no-iniciales solo pueden albergar vocales modales. En !Xoo Bushman (Traill 1985), se encuentra una situación similar con respecto a las consonantes clics. Las sílabas no-iniciales solo pueden contener consonantes simples, pero las sílabas iniciales pueden albergar tanto consonantes simples como clics.

El shipibo pertenece al tipo de lenguas en que las posiciones iniciales activamente se manifiestan en el nivel fonético y fonológico. Fonológicamente, por ejemplo, las posiciones iniciales del shipibo pueden albergar segmentos laríngeos, vocales largas y además son el lugar preferido, por defecto, para asignar el acento principal. Fonéticamente, las posiciones iniciales también muestran un efecto interesante en shipibo. Las sílabas iniciales de palabra atenúan el grado de centralización de las vocales inacentuadas (véase la sección 2.2). Es decir, las vocales en las sílabas iniciales inacentuadas muestran casi los mismos valores para F1 y F2 como lo hacen las vocales acentuadas. En contraste, las vocales inacentuadas en las sílabas no-iniciales muestran una mayor tendencia a moverse hacia el centro del espacio vocálico.

- La duración de las vocales largas y el tamaño de las palabras

Mientras las diferencias encontradas con respecto a la duración vocálica y los grados de centralización puede explicarse como los efectos que tienen las posiciones privilegiadas sobre la realización de las vocales, existe un patrón también relacionado con la duración vocálica que no puede explicarse de ese modo: las vocales largas tienden a sufrir alargamiento vocálico cuando ocurren en palabras monosilábicas. Esto no sucede cuando ocurren en palabras polisilábicas. Esta generalización se basa en el análisis de la duración vocálica presentado en la sección 2.4.

El estudio acústico mostró que en sustantivos monosilábicos que aparecen solos sin sufijos, las vocales largas tienden a durar alrededor de 210 ms. Sin embargo, cuando la palabra que alberga la vocal larga no es monosilábica (ya sea una raíz monosilábica seguida de por lo menos un sufijo o una raíz polisilábica que

contiene una vocal larga), la duración de la vocal larga disminuye por aproximadamente 30 ms. Es decir, una vez que una vocal larga aparece en una palabra que es por lo menos bisilábica, la duración de las vocales largas no disminuye más. Tanto las vocales largas en palabras bisilábicas y trisilábicas muestran la misma duración (aproximadamente 180 ms). Véase las ilustraciones en la sección 2.4.1.

En un trabajo tradicional de documentación, el trabajador de campo probablemente solo diría que el shipibo muestra vocales largas y vocales cortas. Sin un estudio acústico, el alargamiento vocálico que sufren las vocales largas en shipibo cuando ocurren en palabras monosilábicas habría pasado desapercibido. Pero esto no es todo. Este alargamiento es otra propiedad fonética que no se sigue automáticamente de afirmar que una lengua tiene vocales cortas y largas. El húngaro también tiene vocales cortas y largas pero sus vocales largas no sufren alargamiento cuando ocurren en las mismas condiciones que las vocales largas del shipibo. Examinemos este caso en mayor detalle.

White y Mady 2008a, b llevaron a cabo un estudio acústico de la duración de las vocales cortas y largas del húngaro hablado en Budapest. Estudiaron la influencia que la estructura prosódica tiene sobre la duración de las vocales acentuadas en sílabas iniciales de palabra. En particular, testaron la duración de vocales cortas y largas de ese tipo en palabras monosilábicas, bisilábicas y trisilábicas que ocurrían en posiciones medias y finales dentro de frases.

Ya que el estudio del shipibo presentado en este libro no incluye palabras en posición final dentro de una frase, solo me concentraré en los resultados de White y Mady para las palabras con vocales largas que ocurren en posición media en una frase dado que esos resultados son comparables a aquellos del shipibo.⁵ El experimento realizado por ellos para el húngaro es bastante similar al que se realizó para el shipibo. En ambos experimentos, las vocales largas aparecen acentuadas en la sílaba inicial de las palabras y fueron medidas en palabras monosilábicas, bisilábicas y trisilábicas.

White y Mady encontraron que en posición media dentro de una frase, las vocales largas del húngaro muestran la misma duración sin importar el número de sílabas de la palabra que las contiene. Esto contrasta con el alargamiento fonético encontrado en las vocales largas del shipibo cuando aparecen en palabras

⁵ Para las palabras que ocurrieron en posición final de frase, White y Mady encontraron que las vocales largas del húngaro sufrieron un alargamiento dramático cuando ocurrían en palabras monosilábicas. El alargamiento disminuyó conforme el número de sílabas disminuía. De este modo, en la posición final de frase, las vocales largas de palabras monosilábicas fueron más largas que aquellas de palabras bisilábicas, y las vocales largas de palabras bisilábicas fueron más largas que aquellas que aparecieron en palabras trisilábicas.

monosilábicas. El contraste entre el shipibo y el húngaro se muestra esquemáticamente en la Tabla 9-1

Tabla 9-1: Vocales largas acentuadas al inicio de palabras en posición media de frase

	¿Alargamiento fonético de las vocales?	
	<u>Shipibo</u>	<u>Húngaro</u>
En palabras monosilábicas	Sí	No
En palabras bisilábicas	No	No
En palabras trisilábicas	No	No

Éste es ciertamente un contraste inesperado. Con respecto al shipibo, ya que las vocales largas siempre aparecen acentuadas y en la sílaba inicial de la palabra, entonces, desde una perspectiva fonológica, las posiciones privilegiadas no pueden explicar las asimetrías duracionales que acaban de describirse. El alargamiento fonético observado en las vocales largas de palabras monosilábicas no puede atribuirse simplemente a factores en el estilo o velocidad del habla. Las palabras se produjeron usando un habla normal y bajo las mismas condiciones segmentales y prosódicas.

Una posible explicación es que el alargamiento fonético observado en las vocales largas de las palabras monosilábicas del shipibo se debe a su ocurrencia como el segmento final de la palabra. Se sabe que los segmentos finales tienden a presentar alargamiento. En contraste, las vocales largas de palabras bisilábicas y trisilábicas no están en posición final de palabra y por lo tanto el alargamiento no ocurre. El problema con esta explicación es que no puede explicar por qué las vocales largas del húngaro en palabras monosilábicas no sufren alargamiento dado que también ocurren en la posición final de palabra. El segundo problema es que el contexto de posición final de palabra no parece ser la causa del alargamiento vocálico en otras lenguas. Como White y Mady han demostrado para el húngaro, el contexto que realmente causa el alargamiento segmental es la posición final de *frase*. Pero dado que las palabras elicitadas del shipibo no aparecieron en posición final de frase, el alargamiento observado aún queda sin resolver.

Me gustaría entretener la idea que la explicación a este fenómeno fonético puede entenderse como un reflejo de ciertas propiedades fonológicas del shipibo. En particular, pienso que la explicación reside en el tamaño de la palabra que alberga a la vocal larga. Como puede verse en la Tabla 9-1, el alargamiento

de las vocales largas del shipibo deja de ocurrir una vez que la palabra es por lo menos bisilábica. En otras palabras, el alargamiento de las vocales largas en palabras monosilábicas se relaciona a la falta de suficiente material segmental para alcanzar la bisilabidad. Desde este punto de vista, el alargamiento fonético de las vocales largas en las palabras monosilábicas indican que el shipibo trata de algún modo de alcanzar la longitud abstracta de una palabra bisilábica estirando la duración de las vocales de los monosílabos dado que en el nivel fonológico, la lengua no puede recurrir a la inserción de segmentos como hacen otras lenguas como Axininca, Lardil o Tamil.

Pero ¿por qué la bisilabidad? En los estudios de la teoría métrica, los fonólogos han observado que las lenguas agrupan sílabas para formar constituyentes prosódicos jerárquicamente más altos: los pies métricos. En shipibo, hay evidencia independiente que la lengua hace lo posible para obtener bisilabidad a varios niveles. Primero, la ventana del acento primario, localizada en el margen izquierdo de la palabra, está formada por las dos sílabas iniciales de la palabra. Shipibo impone esta ventana de dos sílabas sin importarle la estructura de las dos sílabas iniciales. Éstas pueden ser ambas abiertas o cerradas o una abierta y la otra cerrada. El objetivo importante para el shipibo parece ser el lograr un pie métrico bisilábico donde el acento primario pueda asignarse. Pero la bisilabidad no solo opera en el nivel puramente fonológico. También interactúa con la morfología. La alomorfía mostrada por ciertos sufijos depende de si las raíces o bases a las cuales se juntan tienen un número par o impar de sílabas. Es decir, en términos métricos, este tipo de alomorfía depende de si la raíz termina en una sílaba que pertenece a un pie métrico o no.

Además de jugar un rol en la asignación del acento primario y en condicionar la alomorfía de ciertos sufijos, la bisilabidad también tiene un profundo impacto en el tamaño preferido de las raíces. Las raíces del shipibo son abrumadoramente bisilábicas. Muy pocas raíces son trisilábicas (la mayoría de ellas son préstamos de otras lenguas o palabras con una estructura morfológica antigua que se ha perdido). Las raíces monosilábicas son aún más raras. De hecho, los estudios diacrónicos sugieren que en algún momento, la bisilabidad, como el tamaño ideal de las raíces, era un requisito tan fuerte en el proto-pano que causó la apócope de la vocal final de las raíces trisilábicas. Éste es el origen de los segmentos latentes de shipibo (véase la sección 9.2.3). La consonante final de raíces como /isak/ ‘pájaro’ originalmente fue el arranque de una tercera sílaba: */isaka/. Debido al proceso de apócope, la consonante /k/ quedó como el segmento final de la palabra. Ya que este segmento no es una sibilante o una nasal no se le permite aparecer en posición de coda. La consonante oclusiva se convirtió en un segmento latente, capaz de afectar la posición del acento primario

([i.'sak] → [i.'sa]) pero incapaz de aparecer al menos que le siga una vocal (/isak -an/ → ['i.sa.kan] 'pájaro' FL).

Toda esta evidencia indica que la bisilabicidad es una fuerza activa en la fonología del shipibo. Pero tal vez lo más interesante desde el punto de vista teórico es que la bisilabicidad es independiente del requisito que las palabras sean mínimamente bimoraicas (es decir, la restricción MP que fuerza las vocales a ser largas en palabras monosilábicas). En shipibo, mientras la bisilabicidad es el tamaño deseable para una raíz, la bimoracidad es el tamaño mínimo obligatorio de las palabras. Ambas fuerzas fonológicas están activas en el misma lengua.

La independencia de la bisilabicidad y de la bimoracidad también se puede observar en otras lenguas. El tamil es una de esas lenguas pero a diferencia del shipibo, la bimoracidad es el tamaño mínimo obligatorio de las raíces y la bisilabicidad es el tamaño deseado para las palabras. De este modo, las raíces monosilábicas en tamil deben ser por lo menos bimoraicas (por ejemplo, [ma:n] 'venado', [pal] 'diente') pero además, el tamil puede opcionalmente alcanzar la bisilabicidad para sus palabras monosilábicas insertando la vocal [ɨ]. Cuando esto ocurre, las raíces monosilábicas no pierden su bimoracidad. De este modo, las palabras para 'venado' y 'diente' en tamil pueden también ocurrir como [ma:nɨ] y [pallɨ] (Vijayakrishnan 2002). Estas palabras satisfacen la bisilabicidad pero al mismo tiempo las raíces satisfacen la bimoracidad.

Obsérvese que en tamil, la bisilabicidad tiene la opción de alcanzar al reino de la fonología y de este modo inducir epéntesis en monosílabos. Sin embargo, el shipibo, que no puede recurrir a la epéntesis de vocales para satisfacer la bisilabicidad, manifiesta su insatisfacción con palabras de menos de dos sílabas al tratar de estirar la duración fonética de los segmentos de estas palabras.

Pero entonces, ¿por qué el húngaro no muestra alargamiento de sus vocales largas en palabras monosilábicas cuando éstas ocurren en posición media de frase? A diferencia del shipibo y del tamil, el húngaro no muestra ninguna necesidad de lograr la bisilabicidad como el tamaño mínimo de sus palabras, aunque sí parece tener un requisito débil de palabras mínimamente bimoraicas (Grimes 2005, 2009). Bimoracidad, como una restricción que gobierna el tamaño mínimo de las palabras en húngaro es tan débil que la lengua inclusive permite la existencia de algunas pocas palabras monosilábicas con vocales cortas (*fa* 'árbol' y *ma* 'hoy'). Sin embargo, la mayoría de sus palabras monosilábicas tienen una vocal larga ([fo:] 'sal') o tienen una vocal corta segunda de una consonante, la cual aporta la segunda mora a la palabra ([ʃok] *sok* 'mucho'). De este modo, esta diferencia con el shipibo sugiere que la ausencia de la bisilabicidad como una fuerza activa en la fonología del húngaro puede ser la explicación del por qué sus vocales largas no sufren ninguna presión de alargarse en palabras monosilábicas. No hay

necesidad en húngaro de estirar fonéticamente la duración de los segmentos de las palabras monosilábicas ya que la bisilabidad es un requisito inactivo en la fonología de la lengua.

9.2 UNA FONOLOGÍA DE LAS CONSONANTES DEL SHIPIBO

9.2.1 Los rasgos de las consonantes

Ahora volquemos nuestra atención a las consonantes. Empezaré el bosquejo de la fonología de las consonantes caracterizándolas por sus rasgos fonológicos. Estos se muestran en la Figura 9-6. Cualquier caracterización fonológica de las consonantes del shipibo debe ser capaz de distinguirlos de la clase de las vocales. Esto tradicionalmente se logra por medio del rasgo binario [\pm consonántico]. Los segmentos [+consonántico] son aquellos en los cuales existe una constricción radical en la cavidad oral. Una constricción radical quiere decir que el pase del aire a través de la boca se detiene completamente o es suficientemente angosto como para producir fricción. En contraste, los segmentos [-consonánticos] son aquellos como las vocales en las cuales no ocurre tal constricción. Desde la perspectiva acústica, la diferencia entre segmentos [-consonánticos] y [+consonánticos] se observa en los altos niveles de energía que se encuentra en los espectrogramas-espectros de los segmentos [-consonánticos] y en los bajos niveles de energía acústica de los segmentos [+consonánticos].

Con respecto al modo de articulación, el rasgo [\pm continuo] distingue entre oclusivas y todos los otros segmentos. Segmentos [-continuos], [p, t, k, m, n], son aquellos en los cuales el aire que viene de los pulmones no puede escapar a través de la boca debido a que hay alguna oclusión en alguna parte de la cavidad oral. En términos acústicos, las oclusivas no muestran energía acústica (oclusivas sordas) o si la muestran, la energía se concentra en las frecuencias bajas (oclusivas sonoras, nasales). Los segmentos [+continuos] incluyen, en el caso del shipibo, las fricativas y las glides. Además, las fricativas sibilantes, [s, ʃ, ʂ], pueden agruparse juntas como la clase de las [+estridentes]; es decir, consonantes cuya articulación involucra la producción de fricción en la cavidad oral. Desde el punto de vista acústico, el rasgo [+estridente], como se mostró en el capítulo 4 para la fase de fricción de las fricativas y africadas, se caracteriza por la presencia de energía acústica aleatoria.

Utilizo el rasgo [liberación retrasada] para diferenciar las africadas de las oclusivas. El rasgo [liberación retrasada] aparece como [l.r.] en la Figura 9-6. Este rasgo se refiere a la fase de fricción que se encuentra en las consonantes africadas.

Figura 9-6: Rasgos de las consonantes del shipibo

[+consonántico]	Labial		Coronal		Dorsal	Laríngeo
		[+anterior]	[+anterior]	[-anterior]		
			[+distr.]	[-distr.]		
[-continuo]	p	t			k	
[+continuo, +estridente]		s	ʃ	ʂ		
[l. r.] ↙ [-son] ↘ [+son]		ts	tʃ			
		b ^β		dʒ		
[+sonante] ↙ [-cont], ↘ [nasal]	m	n				
			j		w	
[glotis constreñida]						ʔ
[glotis distendida]						h

Las africadas del shipibo son la única clase de segmentos obstruyentes que explotan la binaridad del rasgo [±sonoro]. Todas las oclusivas y fricativas son [-sonoras]. Como se demostró en el análisis acústico presentado en el capítulo 4, solo las africadas pueden distinguirse entre [-sonoras] ([ts, tʃ]) y [+sonoras] ([b^β, dʒ]).

Las nasales y las glides forman una sola clase bajo el rasgo [+sonante] (es decir, segmentos cuya producción no involucra la creación de presión que se incremente en la cavidad oral. Este tipo de presión en la cavidad oral hace difícil la vibración espontánea de las cuerdas vocales, lo cual es una característica de los segmentos sonantes). Dentro de esta clase, las nasales y las glides pueden distinguirse por el rasgo binario [±continuo] y el rasgo privativo [nasal]. Las nasales, por supuesto, llevan el rasgo [nasal] debido a que el aire sale a través de la nariz y son [-continuas] dado que durante su articulación el aire no puede escapar por la boca. En

contraste, las semi-consonantes [j] y [w] son [+continuas] ya que el aire fluye continuamente a través de la boca. De este modo, clasifico a las consonantes sonantes, [m, n, j, w], como [+consonánticas, +sonantes], a las obstruyentes como [+consonánticas, -sonantes] y a las vocales como [-consonánticas, +sonantes] (véase también Kenstowicz 1994).

Retornando a los rasgos laríngeos, además del rasgo [±sonoro], el shipibo usa los rasgos [glotis constreñida] (*constricted glottis*) y [glotis distendida] (*spread glottis*). Aunque la Figura 9-6 no lo muestra, estos tres rasgos son dependientes del nudo laríngeo en la representación de los segmentos. El primero caracteriza a la oclusiva glotal, [ʔ], un segmento durante el cual su articulación involucra sostener las cuerdas vocales muy juntas de modo que el aire se detenga momentáneamente. En shipibo, las oclusivas glotales frecuentemente no se realizan como una única cerrazón glotal sino como una secuencia de ellas, lo cual resulta en fonación glotalizada que se extiende sobre el final de la vocal precedente y el inicio de la siguiente vocal. En contraste, el segmento [glotis distendida] del shipibo, [h], involucra sostener las cuerdas vocales aparte causando el fenómeno comúnmente conocido como aspiración (la fricción producida por el aire al rozar con las cuerdas vocales). La energía acústica generada por la aspiración, aunque es ruido, tiende a concentrarse en las mismas frecuencias en que la energía acústica se concentra en las vocales que rodean la consonante [h].

Con respecto al punto de articulación, el shipibo usa los articuladores: labios, corona y dorso de la lengua, cuyos gestos se expresan fonológicamente a través de los rasgos privativos [labial], [coronal] y [dorsal]. Las consonantes coronales se pueden distinguir más por medio de dos rasgos binarios: [±anterior] y [±distribuido]. Los segmentos clasificados como [+anterior] se refieren a segmentos dentales y alveolares ([t, s, ts, n]), mientras que [-anterior] se refiere a aquellos que son post-alveolares y retroflejos ([ʃ, ʂ, tʃ, dʒ]). Dentro del último grupo, [-anterior], podemos distinguir entre [ʃ, tʃ] y [ʂ, dʒ]. Los segmentos retroflejos son [-distribuido] ya que la superficie de la corona no entra en un contacto completo extendido con el paladar. En contraste a las consonantes retroflejas, las alveo-palatales [ʃ, tʃ] son [+distribuidas].

9.2.1.1 *Sobre el estatus subyacente de las semi-consonantes /j, w/*

Hay una pregunta importante que debe ser respondida sobre las semi-consonantes, /j, w/: ¿cómo es posible que el Shipibo las distinga subyacentemente de las vocales altas? Esta es una pregunta legítima dado que tradicionalmente se asume que en el nivel subyacente los segmentos [j] y [w] son las vocales altas /i/ y /u/, los cuales ocurren como glides cuando no logran ocupar el núcleo de una sílaba debido

a principios fonológicos generales (silabificación y acento). De este modo, por ejemplo, obsérvese la alternancia en una lengua como el castellano entre [i] y [j] como se encuentra entre las palabras *policía* y *policiaco*. En sustantivos, por defecto, el acento en castellano cae en la penúltima vocal (Hualde 2005; Oltra-Massuet 2005). En el caso de /polisia/, la penúltima vocal es la /i/ que sigue a la sibilante /s/: [po.li.'si.a]. Sin embargo, una vez que el sustantivo [po.li.'si.a] se vuelve el adjetivo /polisiako/, la vocal /a/ es ahora la penúltima vocal y en consecuencia, el acento principal cae sobre ella: [po.li.'sja.ko]. Ya que la vocal /i/ deja de ser la portadora del acento principal y está adyacente a otra vocal, se convierte en la glide palatal [j].

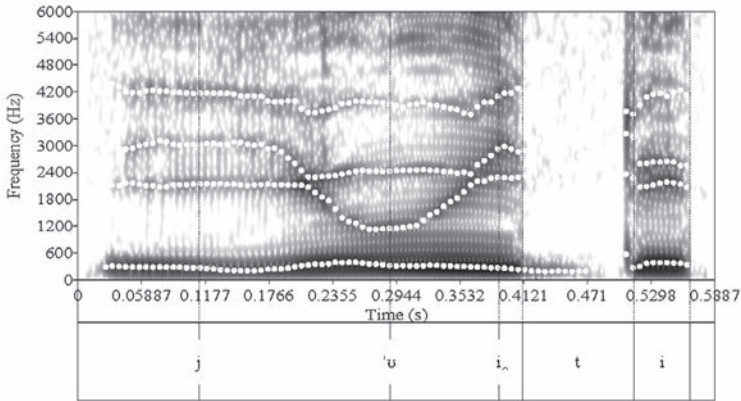
Sin embargo, inclusive en lenguas como el castellano donde las vocales altas y las glides están mayormente en distribución complementaria, existen algunos casos en los cuales su distribución no es predecible, casos en los cuales la naturaleza de ser una vocal o una glide debe especificarse léxicamente. Examinemos dos ejemplos de este fenómeno en castellano. Comparemos las palabras [ˈdwe.lo] ‘duelo’ y [du.'e.to] ‘dueto’ (Hualde 2005). Ambos tienen el acento primario en la penúltima vocal [e] y son segmentalmente muy similares, pero la última palabra no puede ocurrir como *[ˈdwe.to] y la primera no puede ser *[du.'e.lo]. En este ejemplo, la vocal /u/ en ‘dueto’ no puede convertirse en una glide. Veamos ahora otro par de palabras del castellano: [se.kre.'ta.rja] ‘secretaria’ y [se.kre.ta.'ri.a] ‘secretaría’. En este par, la /j/ en ‘secretaria’ no puede convertirse en una vocal, de otro modo, el acento caerá sobre ella ya que se volvería la penúltima vocal de la palabra.

Los datos del castellano que acaban de examinarse muestran que las lenguas pueden en realidad codificar la distinción entre vocales altas y glides en la representación subyacente de las palabras. Y esto es exactamente lo que encontramos en el shipibo con respecto a las vocales /i, u/ ([-consonánticas, +sonantes, +continuas]) y las semi-consonantes /j, w/ ([+consonánticas, +sonantes, +continuas]). Encontramos evidencia para el estatus subyacente de /j/ y /w/ en la comparación de palabras como en (43.c-d). Si el segmento inicial de la raíz /juí/ ‘hablar (con alguien)’ fuese la vocal /i/, */iui/, esperaríamos erróneamente que la palabra apareciera como *[iwi] ya que en shipibo las vocales altas obligatoriamente se convierten en glides en contextos intervocálicos (véase la sección 6.2.3). Este fenómeno también se puede ver en (43.a-b) al observar el comportamiento de la vocal /i/ en la raíz /pui/ ‘defecar’. Además, el shipibo sí tiene una palabra [iwi] que significa ‘rayo’, lo cual indica que si la raíz /juí/ ‘hablar (con alguien)’ fuese en realidad */iui/, podría aparecer como *[iwi]. Nada en la lengua lo prohibiría.

- (43) a. /pui -ti/ → [ˈpu.i.ti] ~ [ˈpu̠i.ti] ‘defecar’
 b. /pui -a/ → [ˈpu.ja] ‘defecar’ (PP2)
 c. /ju̠i -ti/ → [ˈju̠.i.ti] ~ [ˈju̠̠i.ti] ‘hablar (con alguien)’
 d. /iwi/ → [ˈi.wi] ‘rayo’

Además de la evidencia fonológica, la sección 6.2.1 mostró que las realizaciones fonéticas de las vocales altas como glides en Shipibo tienden a mostrar diferencias con respecto a la realización de las semi-consonantes subyacentes /j/ y /w/. Las glides subyacentes tienden a realizarse con cierta cantidad de fricción, especialmente /j/. Por ejemplo, el espectrograma en la Figura 9-7 muestra una ocurrencia de la realización de la palabra /ju̠i -ti/ → [ˈju̠.i.ti] ‘hablar (con alguien)’. La glide palatal [j] se realiza fonéticamente larga y con un claro timbre consonántico. La constricción en la cavidad oral para [j] se vuelve suficientemente angosta como para producir una ligera cantidad de fricción mientras en aire pasa por la cavidad oral (es decir, la /j/ subyacente se realiza como [j]).

Figura 9-7: Espectrograma de /ju̠i -ti/ → [ˈju̠.i.ti] ‘hablar (con alguien)’
 ([ˈju̠̠i.ti])



En contraste a las glides subyacentes, cuando las vocales altas subyacentes se realizan como glides, muestran más variaciones en su realización. Véase los datos en (22). Éstas pueden aparecer como las vocales [i, u], las semi-vocales [j, w] o las semi-consonantes [j, w]. En el último caso, también podrían mostrar algo de fricción. La sección 6.2.2 mostró por medio de espectrogramas estas tres posibles realizaciones para la vocal inicial /i/ en la palabra en (22.g): /iuti/ ‘tomar, dirigir’ (INF).

- (44) a. /ia/ → [ᵀia] ~ [ᵀja] ~ [i.ᵀa] ía ‘piojo’
 b. /ia -ti/ → [ᵀia.ti] ~ [ᵀja.ti] ~ [i.ᵀa.ti] íáti ‘llevar algo en el
 hombro’ (INF)
 c. /iu/ → [ᵀiu.ti] ~ [ᵀju.ti] ~ [i.ᵀu.ti] íoti ‘tomar, dirigir’ (INF)

9.2.1.2 Las restricciones fonotácticas sobre la adyacencia de segmentos

El shipibo presenta dos tipos de restricciones fonotácticas que involucran a las vocales altas. La primera restricción completamente prohíbe las secuencias *[ji], *[wʊ], *[ij] y *[ʊw] tanto al interior de morfemas como en la frontera entre morfemas. Dada la gran similitud entre las vocales altas y las glides, este fenómeno se le conoce en la literatura como *Principio del Contorno Obligatorio* (OCP - Leben 1973; Goldsmith 1976; McCarthy 1986).

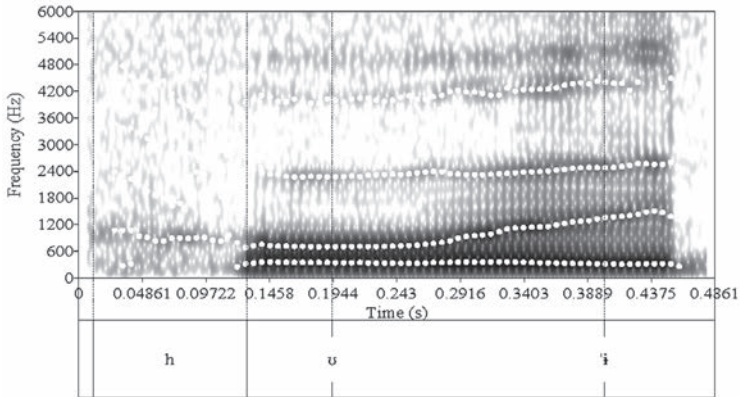
Los siguientes párrafos muestran un número de contextos en los cuales se esperaría que ocurran las secuencias *[ji], *[wʊ], *[ij] y *[ʊw] que prohíbe el OCP. Empecemos por examinar algunos contraejemplos potenciales al OCP en el interior de los morfemas. El diccionario de Loriot, Lauriault *et al.* 1993 reporta la existencia de cuatro palabras que contienen las secuencias: [ij] y [ʊw]. Siguiendo la representación ortográfica dada en el diccionario, esas palabras son: *johué* ‘hola’, *johuéati* ‘saludar (a alguien)’, *showáwa* (esp. de pájaro), y *síyo* ‘sello’. La última palabra es un préstamo del castellano [ˈse.ʝo] ‘sello’. La ortografía de estas palabras sugiere que deberían pronunciarse como [hʊ.ᵀwᵀi], [hʊ.ᵀwᵀi.ʔa.ti], [ʃʊ.ᵀwa.wa] y [ˈsi.ᵀju], respectivamente. Sin embargo, pronto veremos que esta expectativa es incorrecta.

Desafortunadamente, los archivos de audio hechos para esta investigación no incluyeron la palabra *johuéati* ‘saludar (a alguien)’, la cual está formada por la raíz *johué* ‘hola’ y el sufijo *-ati* (es decir, /-ʔati/ VZ1). Sin embargo, sí se grabó la palabra para ‘hola’, la cual contiene la secuencia [ʊw] en discusión. Todos los hablantes entrevistados concordaron en pronunciar *johué* como [hʊ.ᵀᵀi], sin una semi-consonante intervocálica [w].⁶ El espectrograma de la Figura 9-8 presenta una realización típica de esta palabra. No hay ningún rastro de una semi-consonante labial-dorsal. Los valores de F1 y F2 durante la vocal [ʊ] se mantienen estables. Como se espera, el F2 sube conforme se acerca a la vocal

⁶ Uno de los hablantes shipibos entrevistados ocasionalmente pronunció la palabra *johué* ‘hola’ como [hʊ.ᵀwᵀi]; es decir, con la glide intervocálica [w]. La glide pudo ser claramente observada en los espectrogramas. Sin embargo, se le preguntó si tenía alguna preferencia por la pronunciación [hʊ.ᵀwᵀi] o [hʊ.ᵀᵀi], la entrevistada respondió que ya que sabía que la palabra se escribía con ‘hu’ (es decir, el diágrafo para el sonido [w]), debería pronunciarse como [hʊ.ᵀwᵀi]. Interpreto esto como una indicación que la hablante en este caso estaba modificando su habla para acomodar la representación escrita.

central [ɨ]. En este ejemplo, la vocal [ʊ] se ha realizado como la vocal media [o]. Eso explica por qué su F1 es ligeramente más alto que el F1 de [ɨ]. Además, nótese que no hay ningún cambio observable en la cantidad de energía acústica durante la producción de [ʊ] que pueda indicar la presencia de una semi-consonante.

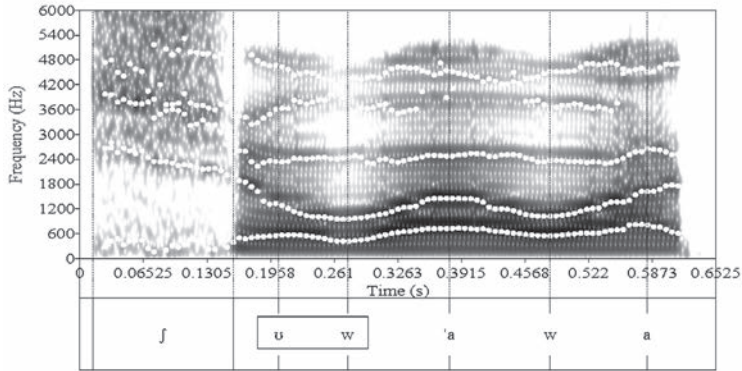
Figura 9-8: Espectrograma de la palabra [hʊ.ʔi] ‘hola’



La palabra *showáwa* (esp. de pájaro), [ʃʊ.ʔwa.wa] sí contiene la semi-consonante [w]. Sin embargo, la mayoría de los hablantes pronuncian la primera vocal de la palabra, [ʊ], con diferentes grados de centralización. Desafortunadamente el contexto en el cual aparece la /ʊ/ hace difícil establecer con certeza la causa de la centralización. Podría ser un efecto fonético que refleja la insatisfacción de OCP en el nivel fonológico o podría ser la influencia de la sibilante palatal precedente, [ʃ], que afecta la altura y grado de posterioridad de la vocal [ʊ]. La Figura 9-9 muestra el espectrograma de [ʃʊ.ʔwa.wa], pronunciado por una hablante mujer. La vocal [ʊ] se realiza como una vocal media central. Esto puede observarse en su relativamente alto F2 (más de 1200 Hz) y F1 (cerca de 650 Hz).

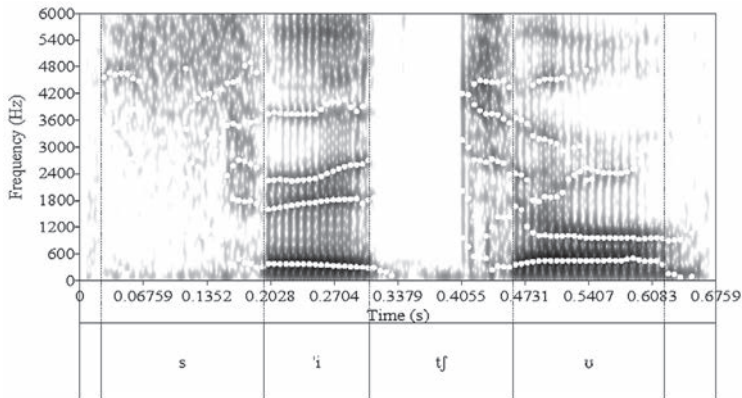
La presencia de la labial-dorsal [w] se indica por la caída de F1 y F2 así como también por la clara disminución de la energía acústica (debido a la constricción más angosta de [w] en la cavidad oral). Esto se ve en el espectrograma de [w] por medio de las manchas blanquizas localizadas entre 1500 Hz y 2500 Hz. La palabra [ʃʊ.ʔwa.wa] es probablemente la única palabra en shipibo que tiene la secuencia [ʊw].

Figura 9-9: Espectrograma de la palabra [ʃü.¹wa.wa] (esp. de pájaro)



En cuanto a la palabra *síyo* ‘sello’, se encontró pocas ocurrencias de esta palabra pronunciada como [ʃi.jü]. La mayoría de los hablantes usan una africada ([ʃi.tʃü] o [ʃi.dʒü] – véase el espectrograma en la Figura 9-10). Algunos usan una sibilante palatal sonora ([ʃi.ʒü]) mientras otros la pronuncian como [ʃi.ü].

Figura 9-10: Espectrograma de la palabra [ʃi.chü] ‘sello’

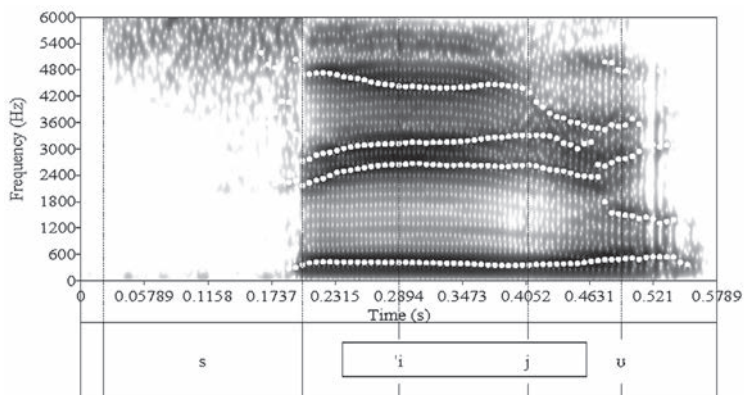


Solo un hablante consistentemente usó la semi-consonante [j] en la palabra para ‘sello’. El espectrograma de la Figura 9-11 muestra esto: [ʃi.jü].⁷ Como en el caso de la semi-consonante [w] de la Figura 9-9, se puede apreciar una disminución

⁷ La discontinuidad de los formantes durante la realización de la vocal [ü] se debe a la presencia de una oclusiva glotal que el hablante introdujo al final de la palabra (probablemente porque la palabra se pronunció en aislamiento).

de la energía acústica durante la articulación de la semi-consonante palatal [j]. El valor de F2 permanece estable durante la mayor parte de la duración de [i] y [j]. F1 baja ligeramente conforme se acerca a [j]. Hacia la mitad de [j], ambos formantes empiezan a moverse en la dirección de los valores necesarios para la vocal [u].

Figura 9-11: Espectrograma de la palabra [ˈsi.ju] ‘sello’



De este modo, he presentado evidencia que el shipibo evita tener las secuencias [ji]/[ij] y [wʊ]/[ʊw] al interior de morfemas. Mientras las secuencias *[ji] y *[wʊ] están completamente prohibidas, las secuencias [ij] y [ʊw] solo se encuentran marginalmente en dos palabras (una de ellas, un préstamo del castellano). En las fronteras de morfemas, esas secuencias nunca aparecen. Ninguna raíz termina en [j] o [w] de modo que las secuencias *[ji] y *[wʊ] nunca tienen la posibilidad de ocurrir en la frontera de morfemas. Cuando una raíz o base termina en /i/ o /u/ y el siguiente sufijo empieza en /j/ o /w/, las semi-consonantes no aparecen (/i + j/ → [i] y /u + w/ → [u]).⁸ Esto puede analizarse como coalescencia de ambos segmentos o como elisión de la glide. No he encontrado evidencia que pueda decidir entre estas dos opciones.

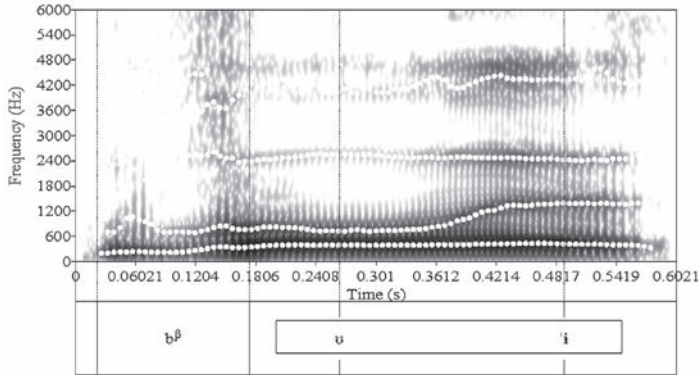
El ejemplo en (45) presenta este fenómeno para la secuencia /u + w/. Cuando la raíz verbal /b^βu/ recibe el sufijo /-^lwi/ (IMP), la forma resultante es [b^βu.^li] ‘tómalo!’ La semi-consonante /w/ del sufijo imperativo no se realiza. Esto puede observarse en el espectrograma de la Figura 9-12.⁹

⁸ El símbolo ‘+’ indica una frontera entre morfemas.

⁹ Las palabras y frases mostradas en los espectrogramas de la Figura 9-9 a la Figura 9-12 fueron pronunciadas en aislamiento, lo cual explica por qué los segmentos, particularmente en los espectrogramas de la Figura 9-9 y la Figura 9-10, aparecen con una duración algo más larga que la duración esperada.

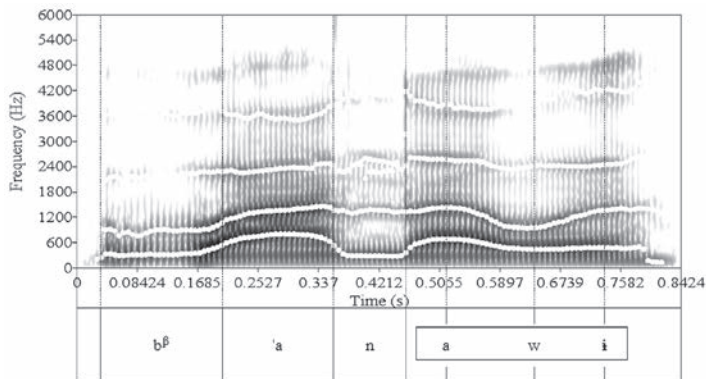
- (45) /b^βu -^lwi/ → [b^βu.^li]
 tomar -IMP.
 '¡Tómalo!'
- (46) /b^βana -^lwi/ → [b^βa.na.wi]
 sembrar -IMP.
 '¡Siémbralo!'

Figura 9-12: Espectrograma de /b^βu -^lwi/ → [b^βu.^li] '¡Tómalo!'



En contraste, cuando el sufijo /-^lwi/ (IMP) aparece después de una raíz o base que termina en cualquier vocal excepto /u/, la semi-consonante /w/ del sufijo imperativo sí se realiza. El espectrograma en la Figura 9-13 muestra el ejemplo de (46). El sufijo imperativo aparece después de la raíz verbal /b^βana/ 'sembrar'.

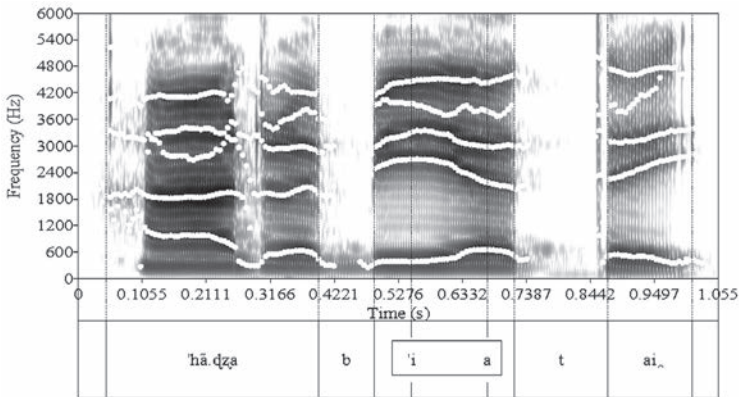
Figura 9-13: Espectrograma de /b^βana -^lwi/ → [b^βa.na.wi] '¡Siémbralo!'



El ejemplo en (47) muestra que el mismo patrón ocurre para /i+j/. En esa oración, tenemos la raíz verbal /b^βi/ ‘traer’ seguida de los sufijos /-jat/ (FUT1) y /-ai/ (PP1). Ya que la raíz verbal termina en la vocal /i/, la semi-consonante palatal [j] del sufijo /-jat/ no se realiza. Véase el espectrograma de la Figura 9-14.

- (47) /han -dʒa b^βi -jat -ai/ → [ˈhan.dʒa ˈb^βi.a.ta.i]
 Él -EV traer -FUT1 -PP1
 ‘él (lo) traerá mañana’
- (48) /han -dʒa b^βana -jat -ai/ → [ˈhan.dʒa ˈb^βa.na.ja.ta.i]
 Él -EV traer -FUT1 -PP1
 ‘él (lo) sembrará mañana’

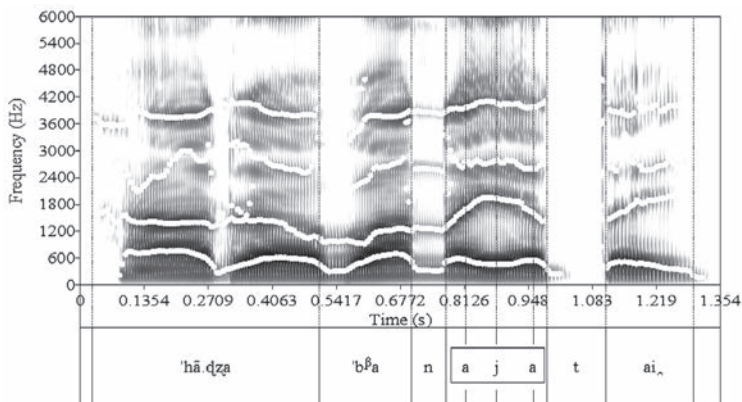
Figura 9-14: Espectrograma de [ˈhan.dʒa ˈb^βi.a.ta.i]
 ‘él (lo) traerá mañana’



Ahora, si se añade el sufijo /-jat/ después de una raíz que termina en cualquier vocal excepto /i/, como /b^βana/ ‘sembrar’, entonces sí aparece la semi-consonante /j/. Esto puede observarse en el espectrograma de la Figura 9-15.

El shipibo tiene otro caso que restringe la adyacencia de segmentos con ciertas características. Se trata de las secuencias *[ʃi]. Mientras no hay palabras con esta secuencia en el shipibo, sí existen palabras con la secuencia invertida, [iʃ], por ejemplo, [ˈtʃi.su] ‘nalgá’ y [ˈtʃi.su.ku] ‘demasiado corto’. Obsérvese que el evitar la secuencia *[ʃi] no está motivado sincrónicamente por la fonología de la lengua. No es el punto de articulación de la sibilante retrofleja la causa que el shipibo no acepte la secuencia *[ʃi] ya que existen muchas palabras con la secuencia [dʒi], la cual involucra la africada retrofleja y la vocal alta frontal: [ˈdʒi.dʒu] (esp. de mono), [dʒi.s.ˈb^βi] ‘soga’, [ˈpi.dʒu] (esp. de pescado). De este modo, hay una

Figura 9-15: Espectrograma de [ʰan.dʒaː ʰb̥a.na.ja.ta.i]
‘él (lo) sembrará mañana’



diferencia entre el evitar tener adyacentemente vocales altas y glides y el evitar la secuencia *[ʃi]. El último parece ser un residuo diacrónico que se está debilitando.

9.2.2 Los desafíos de un análisis fonológico de la labialización del shipibo

En esta sección, discuto los problemas que el fenómeno de labialización del shipibo coloca a los análisis fonológicos. La solución que se propone nos obligará a revisar las representaciones fonológicas de sus vocales. La generalización de este fenómeno es como sigue: las consonantes labiales, [p, b̥, m], se labializan cuando son seguidas por la vocal central [i] (por ejemplo, [ʰp̥i.ku.dʒu] (esp. de hormiga), [ʰb̥wi.pu.ti] ‘cubrir los ojos de alguien’, [ʰm̥i.ti.ki] ‘muñeca’).

El principal desafío que enfrenta un análisis fonológico es cómo explicar el hecho que una vocal central *no-redondeada* pueda causar labialización. Un análisis fonológico debe hacerse tomando en cuenta los resultados obtenidos en el capítulo 3. En particular, los análisis acústicos basados en los patrones de duración y la transición de formantes mostraron que la labialización no involucra la secuencia de las vocales [ui]. Además, se identificó a la vocal central [i] como el único segmento capaz de causar este fenómeno. En el capítulo 3 se demostró que la vocal [u] del shipibo, cuya articulación sí involucra el redondeamiento de los labios, extrañamente no puede causar la ocurrencia de labialización (e.g. [pu.tu.ti] ‘volverse polvoriento’). De modo similar, se presentó evidencia que ninguna de las consonantes labiales es la causa de la labialización. La labialización no ocurre cuando esas consonantes aparecen adyacentes a otras vocales ([pi.ti] ‘comer’).

El fenómeno de labialización descrito acústicamente en la sección 3.2.4 no es único al shipibo. Puede encontrarse en otras lenguas panos como el cashibo y el

capanhua (Loos 1969; Shell 1985), lo cual sugiere que el origen del fenómeno debe haber ocurrido antes de que cada una de las lenguas panos modernas hayan aparecido. Dado que no hay ningún registro de cómo era la lengua en ese tiempo, solo podemos utilizar lo que sabemos de las lenguas panos a nivel sincrónico y en particular los datos que el shipibo nos da. Sin embargo, la carencia de datos diacrónicos no debe ser un obstáculo para analizar este fenómeno sincrónicamente. Después de todo, un niño que aprende shipibo debe adquirir el fenómeno de labialización sin tener acceso a las etapas anteriores de la lengua.

Loos (1969:183) es el primero en proponer un análisis fonológico de este fenómeno para el capanhua, una lengua pano muy cercana al Shipibo. El análisis propuesto por Loos 1969 caracteriza el fenómeno de labialización en capanhua como ‘aplanamiento’ (‘flattening’), lo que interpreto como un tipo de disimilación que involucra los rasgos [plano] ([flat]) y [grave]. Los sonidos especificados como [+planos], Jakobson, Fant *et al.* 1963 los definen acústicamente como un ‘cambio descendente de un conjunto de formantes’ y en particular de F2. Éste es el efecto acústico obtenido cuando los labios sobresalen al redondearse como en la vocal [u] o durante la labialización de las consonantes [p^w], [b^{βw}] y [m^w]. La especificación [+grave] se usa para referirse a sonidos en los cuales la energía acústica se concentra en las partes más bajas del espectrograma (Jakobson, Fant *et al.* 1963). Las consonantes labiales, [p, b^β, m], y las vocales posteriores, [a, i, u], son [+grave].

Bajo el análisis de Loos 1969, las consonantes labiales, que son [+graves], se convierten en [+planas] (es decir, labializadas) cuando son seguidas por una vocal [-plana, +grave] (es decir, las vocales centrales no-redondeadas). A fin de específicamente distinguir a la vocal [i] de la otra vocal central [a], el análisis propuesto por Loos 1969 usa el rasgo [difuso]. Las vocales bajas son [-difusas] y las altas son [+difusas].

Aunque este análisis logra identificar a los segmentos correctos involucrados en el fenómeno y excluir a todos los otros segmentos, enfrenta un serio problema. Obsérvese que la disimilación (‘aplanamiento’) es causado por los rasgos [+grave] de los segmentos adyacentes [p] y [i].¹⁰ Sin embargo, la adyacencia no deseada de estos segmentos y de sus rasgos [+grave] no se arregla cambiando el valor del rasgo [grave] en uno de los segmentos sino que en vez de ellos se inserta la labialización (el rasgo [+plano]) entre ambos segmentos como si esta adición separaría a la consonante bilabial de la vocal central [i] y en consecuencia, también

¹⁰ Supuestamente dentro del dominio de la sílaba, de otro modo la disimilación ocurriría en la secuencia hetero-silábica [i.p] y como se mostró en los análisis espectrográficos del capítulo 3, eso es incorrecto.

separaría sus rasgos [+grave]. Sin embargo, la sección 3.2.4 presentó evidencia fonética basada en la duración y en el comportamiento del formante F2 que la labialización no es equivalente a la adición de un nuevo segmento. Es decir, la labialización no es un segmento que pueda separar a una consonante bilabial como [p] y a la vocal [i̠].

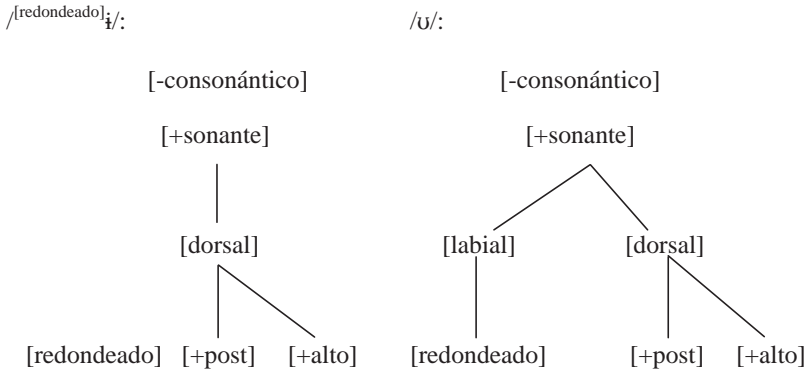
Un análisis alternativo, basado en rasgos articulatorios, trata la labialización del shipibo como un rasgo flotante, [redondeado], de la vocal [i̠]: / [redondeado] i̠/. Dado que el rasgo [redondeado] solo puede asociarse con un articulador labial, debe entonces encontrar un segmento labial que lo albergue. Esto se presenta esquemáticamente en la Figura 9-16. De este modo, las consonantes coroneales y velares se excluyen como posibles lugares donde el rasgo [redondeado] puede asociarse, excepto si una articulación labial secundaria se les añadiese o si el articulador principal de la consonante cambiase a labial. Estas dos últimas opciones no son permitidas por la fonología del shipibo. En contraste, las consonantes labiales ya vienen con el articulador correcto para el rasgo flotante [redondeado]. Como también ha sido reportado en otras lenguas con rasgos flotantes (Zoll 1994, 1996), el rasgo flotante [redondeado] es gobernado por la direccionalidad; es decir, el rasgo flotante, en el caso del shipibo, debe buscar por un lugar donde albergarse hacia el lado izquierdo y dentro del dominio de la sílaba. Si no encuentra el articulador correcto, el rasgo no se realiza fonéticamente.

Figura 9-16: El rasgo flotante [+redondeado] de la vocal [i̠]

/p ^[labial] [redondeado]i̠/	→	[p ^w i̠]
/t ^[coronal] [redondeado]i̠/	→	[t̠i̠]
/k ^[dorsal] [redondeado]i̠/	→	[k̠i̠]

Si este análisis es correcto, tiene importantes consecuencias para la representación subyacente de las vocales. Recuérdese que ya que la vocal /u/ involucra el redondeamiento de los labios, también está especificada como [redondeada] (véase la sección 9.1.1). Si es así, entonces ¿cómo las vocales /i/ y /u/ se diferencian subyacentemente dado que ambas son [+posterior] y ahora ambas llevan el rasgo [redondeado]? La respuesta reside en la disponibilidad subyacente del articulador labial.

Figura 9-17: Representación subyacente de las vocales /i/ y /u/ (modificado)



La articulación de las vocales involucra el movimiento del dorso de la lengua. Esto se presenta en el nivel fonológico por medio del rasgo [dorsal]. Dos tipos de rasgos son dependientes de [dorsal]. Éstos son [\pm posterior] y [\pm alto], los cuales codifican la información sobre la altura y grado de posterioridad de las vocales. Las vocales posteriores como [u] que involucran redondeamiento de los labios tienen un articulador adicional codificado por el rasgo [labial] y su rasgo dependiente [redondeado]. De este modo, hay una diferencia crucial entre las vocales posteriores redondeadas y centrales en términos de articuladores: las vocales como [u] tienen el articulador [labial] disponible y las vocales como [i] no lo tienen. Véase el gráfico de la Figura 9-17.

En shipibo, aunque la vocal central no tienen un articulador [labial], sí posee el rasgo [redondeado] sin ninguna asociación a un articulador es decir, un rasgo flotante. El rasgo [redondeado] es flotante porque la vocal no posee un articulador [labial] con el cual se pueda asociar el rasgo. Si tuviese uno, el rasgo [redondeado] se asociaría con él, pero entonces la vocal central sería la vocal posterior redondeada (véase la Figura 9-17). Cuando la vocal central sigue a una consonante labial, /p, b^β, m/, logra encontrar un lugar donde asociar su rasgo flotante ya que estas consonantes poseen el articulador [labial], lo cual dispara causa el fenómeno de labialización.

Interlingüísticamente, lenguas como el chaha (McCarthy 1983; Archangeli y Pulleyblank 1994; Rose 1994; Zoll 1996), una lengua semítica de Etiopía, se conocen por sus rasgos flotantes. En esta lengua, por ejemplo, la labialización expresa el objeto de tercera persona singular en los verbos y se codifica por medio del rasgo flotante [redondeado]. Ese rasgo flotante se pega a la consonante del verbo que esté más a la derecha y que permita ser labializada (es decir, las

consonantes labiales y dorsales). A diferencia del shipibo, sin embargo, el rasgo flotante del chaha puede viajar dentro de la raíz verbal para alcanzar a la consonante labializable más cercana. Si la raíz no contiene ninguna consonante labial o dorsal, entonces, el rasgo flotante no ocurre. Los datos en (49) muestran ejemplos de cómo la labialización funciona como un rasgo flotante en chaha.

(49) a.	/dænæg/	→	[dænæg]	‘golpear’
b.	/dænæg ^[redondeado] /	→	[dænæg ^w]	‘golpearlo’
c.	/sʷæfær/	→	[sʷæfær]	‘cubrir’
d.	/sʷæfær ^[redondeado] /	→	[sʷæf ^w ær]	‘cubrirlo’
e.	/mæsær/	→	[mæsær]	‘parecer’
f.	/mæsær ^[redondeado] /	→	[m ^w æsær]	‘parecerlo’
g.	/sædæd/	→	[sædæd]	‘cazar’
h.	/sædæd ^[redondeado] /	→	[sædæd]	‘cazarlo’

9.2.3 Segmentos latentes

Uno de los aspectos más fascinantes de la fonología del shipibo es la existencia de segmentos latentes (también conocidos como segmentos fantasmas). Los segmentos latentes del shipibo son consonantes que están presentes en la representación subyacente pero que solo se pueden realizar si el siguiente segmento es una vocal. No pueden realizarse como codas porque no pertenecen al conjunto de codas permitidas (las sibilantes [s, ʃ, ʂ] y la nasal [n]). Sin embargo, aunque no se realicen, sí afectan la asignación del acento como lo hace cualquier otra coda.

Veamos ejemplos del comportamiento de un segmento latente. Shipibo tiene varias estrategias para cambiar la transitividad de los verbos. Una de esas estrategias es añadir el sufijo /-t/ (INTVZ) después de la raíz verbal. Este sufijo toma un verbo transitivo y lo convierte en uno intransitivo. Además del sufijo /-t/, los datos en (50) y (51) contienen las raíces verbales transitivas /tʃʊpi/ ‘abrir, desplegar’ y /nʊʃi/ ‘rasgar (la ropa)’, y los sufijos: /-ki/ (CMPL), /-ti/ (INF), /-ai/ (PP1) y /-a/ (PP2). Nótese en (50.a) y (51.a) que el sufijo /-t/ no puede aparecer pero que el acento primario está en la segunda sílaba como si ésta tuviese una cosa. El shipibo no permite que la consonante /t/ ocurra en la representación fonética como coda, pero el acento, al momento de ser asignado, la puede ver. En (50.b) y (51.b), el sufijo que sigue a /-t/ es una vocal, en consecuencia, /-t/ puede realizarse como un arranque y por supuesto, la segunda sílaba esta vez en verdaderamente abierta. El acento puede ver esto y siguiendo las reglas generales de acentuación, cae sobre la primera sílaba.

- (50) a. /tʃupɪ -t -ki/ → [tʃu.ˈpɪt.ki] → [tʃu.ˈpɪ.ki] ‘abrirse, desplegarse’
 (INTVZ, CMPL)
 b. /tʃupɪ -t -a/ → [ˈtʃu.pɪ.ta] ‘abrirse, desplegarse’
 (INTVZ, PP2)
- (51) a. /nuʃi -t -ti/ → [nu.ˈʃit.ki] → [nu.ˈʃi.ti] ‘rasgarse (la ropa)’
 (INTVZ, INF)
 b. /nuʃi -t -ai/ → [ˈnu.ʃi.ta.i] ‘rasgarse (la ropa)’
 (INTVZ, PP1)

Ahora veamos los datos en (52) y (53), los cuales tienen los mismos morfemas que en (50) y (51) excepto por el sufijo /-t/ (INTVZ). En (52) y (53), no hay ningún segmento latente y por lo tanto el acento primario cae sobre la sílaba inicial (ya que la segunda es abierta). A diferencia de (50.b) y (51.b), en (52.b) y (53.b), no aparece una consonante [t] entre la raíz y la vocal que sigue. Esto descarta la posibilidad que el segmento [t] observado en (50.b) y (51.b) pertenezca al sufijo /-a/ o que sea insertado para evitar que dos vocales entren en contacto.

- (52) a. /tʃupɪ -ki/ → [ˈtʃu.pɪ.ki] ‘abrir, desplegar’ (CMPL)
 b. /tʃupɪ -a/ → [ˈtʃu.pɪ.a] ‘abrir, desplegar’ (PP2)
- (53) a. /nuʃi -ti/ → [ˈnu.ʃi.ti] ‘rasgar (la ropa)’ (INF)
 b. /nuʃi -ai/ → [ˈnu.ʃi.a.i] ‘rasgar (la ropa)’ (PP1)

Además de la evidencia fonológica que acaba de presentarse, también existe evidencia fonética de la existencia de segmentos latentes en shipibo. Esta evidencia viene de los resultados acústicos obtenidos en la sección 2.3 y la sección 7.1.2. En el capítulo 2, vimos que las vocales acentuadas de segundas sílabas tienden a ser más largas (por alrededor de alrededor de 15 ms) que las vocales acentuadas de sílabas iniciales. De este modo, por ejemplo, la vocal [i] acentuada de [si.ˈki.ti] ‘quebrarse’ (/sikɪ -t -ti/ → [si.ˈkɪt.ti] → [si.ˈki.ti]) tiende a ser más ligeramente más larga que la vocal acentuada [i] de [ˈsi.ki.ti] ‘quebrar’ (/sikɪ -ti/ → [ˈsi.ki.ti]).

La Figura 9-18 y la Figura 9-19 presentan los espectrogramas de los verbos [ˈsi.ki.ti] ‘quebrar’ y [si.ˈki.ti] ‘quebrarse’.¹¹ Ambas palabras tienen la misma

¹¹ Recuérdese que los resultados presentados en el capítulo 2, sección, 2.3, corresponden a la medición de vocales cortas en sílabas abiertas. Éstas ocurrieron ya sea en la primera o segunda sílaba de palabras trisilábicas. En la mayoría de los casos, la tercera sílaba de cada palabra correspondió a un sufijo que seguía a una raíz bisilábica. Sin embargo, dado que el acento en shipibo aparece sobre la segunda sílaba si ésta es cerrada (García-Rivera 1994; Elias-Ulloa 2000a), se decidió considerar

secuencia segmental y fueron pronunciadas por una hablante mujer. Ambas tienen la raíz /siki/ ‘quebrar’ y aparecen en sus formas infinitivas indicadas por el sufijo /-ti/. El último, el verbo intransitivo ‘quebrarse’, tiene además el sufijo /-t/ (INTVZ). Las palabras solo se diferencian por la posición del acento. Estos dos espectrogramas permiten ilustrar las diferencias en la duración entre vocales acentuadas en sílabas iniciales y segundas. Mientras en la Figura 9-18, la vocal acentuada [i] dura 102 ms, en la Figura 9-19, la misma vocal acentuada, esta vez en una segunda sílaba, seguida en el nivel subyacente por un segmento latente, tiene 20 ms más de duración.

Figura 9-18: /siki -ti/ → [ˈsi.ki.ti] ‘quebrar’ (INF)

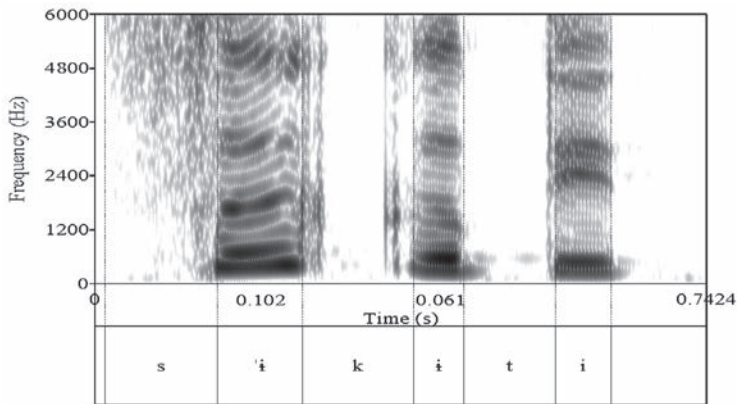
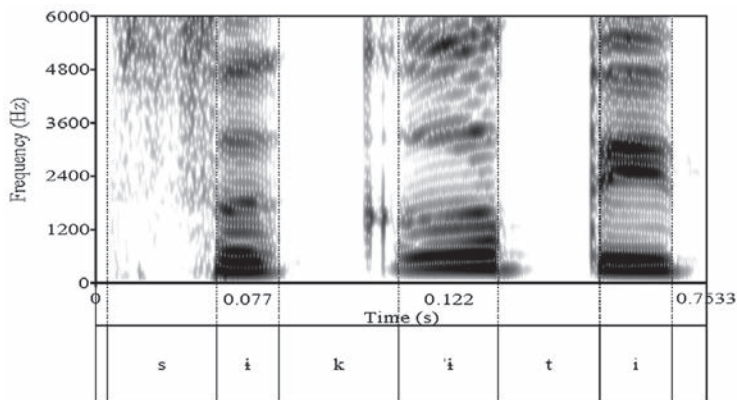


Figura 9-19: /siki -t -ti/ → [si.ˈki.ti] ‘quebrarse’ (INTVZ, INF)



palabras con consonantes latentes en la segunda sílaba. Esta decisión hizo posible incluir palabras con el acento en la segunda sílaba pero que en la superficie aparecieran como segundas sílabas abiertas.

El segundo tipo de evidencia fonética que apoya la existencia de segmentos latentes en shipibo viene de la sección 7.1.2, donde las mediciones de las vocales acentuadas de sílabas cerradas por una sibilante mostraron tener virtualmente la misma duración sin importar si ocurren en sílabas iniciales o segundas sílabas. Este hallazgo contrasta con los resultados de la sección 2.3. Por último, el capítulo 7 también mostró que la duración de las vocales acentuadas seguidas por una coda latente ([ma.¹ta.ki] ‘quitarse (algo) de la cabeza’ (INTVZ, CMPL)) es bastante similar a la duración de las vocales acentuadas seguidas de una coda sibilante ([ma.¹tas.ki] ‘golpear en la cabeza’ (Esos patrones fonéticos juntos indican que las vocales acentuadas seguidas de una coda latente presentan la misma duración que las vocales acentuadas de sílabas cerradas. Si no hubiese un segmento latente en la segunda sílaba de palabras como [ma.¹ta.ki], esa asimetría duracional sería difícil de explicar.

9.2.3.1 *Los segmentos latentes y la asignación del acento*

Los segmentos latentes del shipibo tienen dos propiedades que un estudio fonológico debe enfrentar: (i) son capaces de interactuar con la asignación del acento y (ii) todos han sido neutralizados con respecto a sus puntos de articulación. Empezaremos por abordar el primer problema.

Con respecto a los segmentos latentes del shipibo, el desafío más difícil que enfrenta una teoría fonológica con un sistema computacional no-derivacional es explicar la influencia que esos segmentos ejercen sobre la asignación del acento. Para que la fonología del shipibo sepa dónde asignar el acento, los segmentos latentes deben ser silabificados como codas en algún nivel representacional. Ya que la silabificación es completamente predecible, no se le considera parte de las representaciones subyacentes. Además, ya que los segmentos latentes no ocurren en las representaciones fonéticas, esto sugiere que deben ser codas en algún nivel intermedio de representación, lo cual es exactamente lo que una teoría no-derivacional trata de evitar (/inpɪt -ki/ → [im.¹pɪt.ki] → [im.¹pɪ.ki] ‘agujerear la sien (de alguien)’ (CMPL) - véase los datos en (54)).

¿Qué tipo de análisis podría proponer una teoría no-derivacional para tratar los casos de segmentos latentes y su interacción con la silabificación y la asignación del acento? Bajo un análisis más abstracto, se puede asumir que los segmentos latentes son parte de las representaciones fonéticas pero que no logran obtener una realización fonética. Una representación fonética es el resultado (output) del componente fonológico de una gramática y como tal, es un concepto lingüístico abstracto. Esta representación es interpretada por nuestro sistema motores, los cuales están a cargo de implementar su realización articulatoria. De este modo,

bajo esta perspectiva, la representación fonética de /inpit -ki/ en Shipibo es [im. 'pit.ki].¹² Este análisis tendría que asumir que el requisito de la *Condición de Coda* (es decir, la restricción que solo las sibilantes y la nasal [n] puedan recibir en posición de coda una expresión fonética en Shipibo) se aplica en la interfase entre las representaciones fonéticas y los sistemas motores de modo que una palabra con una representación fonética como [im. 'pit.ki] últimamente se realice como [im. 'pi.ki].

Una segunda alternativa sería recurrir a estratos fonológicos dentro del componente fonológico. En cada nivel, los requisitos fonológicos presentan diferentes grados de importancia. De este modo, uno puede asumir que el shipibo tiene dos estratos. Una forma como /inpit -ki/ primero va al nivel en que la *Condición de Coda* no tiene ningún poder para eliminar una representación como [im. 'pit.ki]. En ese mismo nivel, se asigna el acento. Luego, la forma va al segundo nivel en el cual las representaciones fonéticas posibles basadas en la forma [im. 'pit.ki] compiten para ser seleccionadas como el resultado del componente fonológico. De este modo, en este segundo nivel, tendremos por lo menos compitiendo las formas *[im. 'pit.ki], [im. 'pi.ki] y *[im. 'pi.ti.ki]. A diferencia del primer nivel, esta vez el requisito de la *Condición de Coda* sí tiene el poder de eliminar aquellas formas que contengan una coda que no sea sibilante o nasal. Esto elimina a *[im. 'pit.ki]. Una vez esta forma es separada de la competencia, la fonología debe decidir entre el candidato con elisión, [im. 'pi.ki], o aquel con inserción vocálica, *[im. 'pi.ti.ki]. Shipibo escoge el candidato con elisión.¹³

9.2.3.2 El punto de articulación por defecto de los segmentos latentes

El segundo problema que un análisis fonológico debe enfrentar es que el shipibo ha neutralizado el punto de articulación de sus segmentos latentes. Estos segmentos pueden ocurrir como las consonantes finales de raíces y sufijos. Con la excepción de los verbos 'ser, estar, suceder' y 'hacer, decir' y las formas largas de los sustantivos y adjetivos que discutiré en breve, los segmentos latentes siempre aparecen por defecto como la oclusiva coronal [t].

¹² Esta alternativa no-derivacional ciertamente se acerca a un modelo conocido como *Contención* (*Containment*) dentro de la Teoría de la Optimicidad, un modelo originalmente propuesto por Prince and Smolensky (1993), Prince and Smolensky (2004) y que luego fue reemplazado por la *Teoría de la Correspondencia* (*Correspondence Theory* - McCarthy and Prince [1995]).

¹³ Bajo esta perspectiva, el sistema computacional de la fonología mantiene su característica no-derivacional dentro de cada nivel fonológico, una idea originalmente propuesta por McCarthy and Prince (1993b) y adoptada por el modelo Strata OT (Kiparsky (2000), Bermúdez-Otero (2003)).

Los datos en (54) contienen la raíz verbal /inpit/ ‘agujerear la sien (de alguien)’. El segmento final de esta raíz es un segmento latente /t/. Diacrónicamente, esta raíz tiene su origen en un antiguo proceso de prefijación que ya no está activo en la lengua. Este proceso agrega un prefijo que refiere a una parte del cuerpo a una raíz verbal. En el caso de la raíz /inpit/, el prefijo es /in-/ ‘en la sien’ y el verbo transitivo original al cual se pegó el prefijo era: /pika/ ‘hacer huecos’ (esta raíz verbal aún puede encontrarse sin el prefijo). Ya que el resultado era una raíz trisilábica, [in.pi.ka], la tercera sílaba se apocopó (véase más sobre este proceso en la sección 9.1.3), lo cual resultó en la raíz /inpik/. Conforme el tiempo pasó, todas las oclusivas finales de estas formas verbales, sin importar su punto de articulación original, se neutralizaron en el mismo articulador: [coronal]. De este modo, la raíz verbal ‘agujerear la sien (de alguien)’ terminó siendo /inpit/.

Obsérvese que la /t/ final de /inpit/ no puede interpretarse como el sufijo /-t/ (INTVZ), ya que el verbo resultante tiene una lectura transitiva. Además, no puede argüirse que la oclusiva coronal en (54.b) se inserta para evitar que las vocales heteromorémicas estén adyacentes. La oclusiva coronal no ocurre si la raíz no posee un segmento latente: /nuʃi -a/ → [¹nu.ʃi.a] ‘rasgar (la ropa)’ (PP2). Véanse también los datos en (52) y (53).

- (54) a. /inpit -ki/ → [im.¹pi.ki] ‘agujerear la sien (de alguien)’ (CMPL)
 b. /inpit -a/ → [¹im.pi.ta] ‘agujerear la sien (de alguien)’ (PP2)

Los segmentos latentes pueden también aparecer como las consonantes finales de sufijos y en este contexto también siempre ocurren como coroneles. Los datos en (55) muestran un ejemplo de un segmento latente que ocurre como la consonante final del sufijo /-jat/ (FUT1). Este segmento latente se comporta de manera idéntica a lo que hemos visto de los segmentos latentes al final de las raíces; es decir, es capaz de hacer que una sílaba se comporte como cerrada y de este modo, puede atraer el acento. Cuando logra aparecer como arranque de sílaba, el acento regresa a la sílaba inicial. Además de /-jat/, hay otros sufijos en shipibo que tienen un segmento latente. Todos poseen una oclusiva coronal final. Algunos ejemplos de estos sufijos incluyen: /-ʔib^βat/ ‘ayer, el otro día’, /-pakit/ ‘río abajo’, /-ʔinat/ ‘ría arriba’, /-b^βait/ ‘todo el día’, etc.

- (55) a. /pi -jat -ai/ → [¹pi.ja.ta.i] ‘comer’ (FUT1, PP1)
 b. /pi -jat -kan -ai/ → [pi.¹ja.ka.na.i] ‘comer’ (FUT1, CON.3PL, PP1)

Como en el caso de las raíces, no todos los sufijos terminan en un segmento latente. Por ejemplo, los sufijos /-ma/ (CAUS), /-jama/ (NEG), /-jodʒa/

(INTENS), /-dʒɪb^βi/ ‘otra vez’ no muestran ningún segmento latente cuando aparecen seguidos por una vocal (/maʃka -dʒɪb^βi -ai/ → [ˈmaʃ.ka.dʒɪ.b^βi.a.i] ‘(él/ella) se desmaya otra vez’.

A fin de obtener un mejor entendimiento del comportamiento de los segmentos latentes del shipibo, examinemos brevemente cómo se comporta este tipo de segmentos en otras lenguas. Encontramos casos de segmentos latentes en lenguas como lardil, yawelmani, castellano latinoamericano, francés, italiano (Miogliola), variedades no-róticas del inglés, entre otros. En lardil (Wilkinson 1988; Goldsmith 1993), solo las consonantes corales se permiten en posición de coda. De este modo, la palabra para ‘corto’, cuya representación subyacente es /putuk/, aparece en lardil como [putu] cuando no tiene sufijos. El segmento latente ocurre cuando se agrega a la raíz un sufijo como /-aɾ/, tradicionalmente conocido como futuro: [putukaɾ]. En variedades del castellano latinoamericano, como es el caso del castellano peruano, ciertas consonantes finales de palabras (como /b/, /d/ y /x/) generalmente no se pronuncian en el habla coloquial: /fakultad/ → [fa.kul.ˈta] *facultad*, /relox/ → [re.ˈlo] *reloj*. Pero una vez que son seguidas por un sufijo que empieza en una vocal, como en las formas plurales, esas consonantes reaparecen obligatoriamente: [fa.kul.ˈta.ðes], [re.ˈlo.xes].

Un aspecto interesante de los casos del lardil y del castellano peruano es que los segmentos latentes pueden mostrar diferentes puntos de articulación. Esto contrasta con el shipibo en el cual las consonantes latentes siempre aparecen como la oclusiva coronal [t] en las formas verbales. Sin embargo, el shipibo no es el único caso en que los segmentos latentes se han neutralizado al punto de articulación coronal. En el dialecto románico ligurino (gallo-italiano) de Miogliola, Italia (Ghini 2001, 2002), todos los segmentos latentes también son consonantes corales. Sin embargo, a diferencia del shipibo, las consonantes corales en Miogliola pueden ser las oclusivas /t, d/ o las líquidas /l, r/. De este modo, la palabra en el Miogliola para ‘escalera’ tiene un segmento latente final: /katsyt/. Cuando se pronuncia en aislamiento, la palabra ocurre como [katsý] pero una vez que se le añade el sufijo derivacional /-eɲ/, aparece el segmento latente /r/: [katsyréɲ] ‘escalera’ (diminutivo).

La neutralización del punto de articulación de las consonantes a coronal es muy común interlingüísticamente. Aunque no es un caso de segmentos latentes, el vasco posee uno de los casos más interesantes de este tipo de neutralización. Palabras como [ogɪ] ‘pan’, [sagu] ‘ratón’ y [idi] ‘buey’, cuando reciben un morfema derivacional, pierden su vocal final y la consonante que precede a esa coda se convierte en coronal: [ot-apur] ‘migaja de pan’, [sat-or] ‘topo’, [it-zain] ‘guía de bueyes’ (Hualde 1991; Rice 1996). Además del vasco, hay lenguas como el Lardil (Wilkinson 1988; Goldsmith 1993) y el Carrier (Walker 1979; Kenstowicz

1994) que neutralizan todas sus codas a consonantes coronales, y otras como el Axininka Campa que tiene la consonante coronal [t] como su segmento epentético por defecto (Payne 1981; McCarthy y Prince 1993a). Estos patrones son parte de una tendencia interlingüística conocida como el *Síndrome Coronal* (Kenstowicz 1994; De Lacy 2006), por el cual las lenguas tienden a recurrir a coronal como el punto de articulación por defecto en diferentes patrones fonológicos.

En shipibo, en el caso de las codas y las consonantes finales de las raíces, esta tendencia es bastante fuerte. Todas las codas deben ser coronales: [s, ʃ, ʂ, n]. Y las consonantes finales de las raíces deben pertenecer al conjunto lícito de codas (es decir, [s, ʃ, ʂ, n]) o ser [t]. Si la consonante final es [t], se convierte en un segmento latente. Véase los datos en (56).

- (56) a. /isin/ → [i.'sin] 'enfermedad'
 b. /mɨntsɨs/ → [mɨn.'tsɨs] 'uña'
 c. /yawɨʃ/ → [ya.'wiʃ] 'armadillo'
 d. /witaʃ/ → [wi.'taʃ] 'pierna'
 e. /ɨnpɨt -kɨ/ → [ɨm.'pɨ.kɨ] 'agujerear la sien' (CMPL)
 f. /pi -jat -kan -ai/ → [pi.'ja.ka.na.i] 'comer' (FUT1, CON.3PL, PP1)

En contraste, en lenguas como lardil o castellano, la preservación del punto de articulación subyacente de los segmentos latentes es más importante que respetar el síndrome coronal. De este modo, en el lardil y el castellano, los segmentos latentes ocurrirán con diferentes puntos de articulación cuando logran aparecer como arranques de una vocal.

9.2.3.3 *Los segmentos latentes dorsales*

Así como los segmentos latentes de formas verbales y sufijos, el shipibo también tiene segmentos latentes en raíces no-verbales (sustantivos y adjetivos). En contraste a los casos que acaban de presentarse, los segmentos latentes de raíces nominales y adjetivales no se han neutralizado a la oclusiva coronal [t] sino a la oclusiva dorsal [k].

Antes de presentar los segmentos latentes [k] de sustantivos y adjetivos, es importante mencionar que existen dos raíces verbales que tienen [k] como segmento latente. Éstas son raíces verbales que pueden usarse como sufijos verbales y auxiliares en construcciones sintácticas (véase Valenzuela 2003b). Son los verbos /ʔik/ 'ser, suceder' y /ʔak/ 'hacer, decir'.¹⁴ Cuando son seguidos por los sufijos

¹⁴ Estas raíces verbales son el origen de los sufijos /-ʔati/ (VZ1) y /-ʔiti/ (VZ2).

/-ti/ (INF) o /-ki/ (CMPL), sus segmentos latentes no aparecen. Véase los datos en (57.a-b). Sin embargo, cuando son seguidos por un sufijo como /-ai/ (PP1), sus segmentos latentes sí aparecen. Véase los datos en (57.c-d).

- (57) a. /ʔik -ti/ → [ʔi.ti] ‘ser, suceder’ (INF)
 b. /ʔak -ti/ → [ʔa.ti] ‘hacer, decir’ (INF)
 c. /ʔik -ai/ → [ʔi.kai] ‘ser, suceder’ (PP1)
 d. /ʔak -ai/ → [ʔa.kai] ‘hacer, decir’ (PP1)

En el caso de sustantivos y adjetivos, el segmento latente dorsal [k] no es la excepción sino la norma. Esto puede observarse en los datos en (58). Como en los datos en (50) y (51) para los verbos, la asignación del acento en (58) depende de si el segmento latente es considerado una coda en algún nivel representacional o si es el arranque de la vocal que le sigue. La vocal del sufijo /-(a)n/ aparece en paréntesis para indicar su naturaleza epentética. El timbre de esta vocal depende de las características del punto de articulación de la consonante precedente. Cuando es precedida por [k], como en el caso de los sustantivos con segmentos latentes, el timbre de la vocal es más cercano a [a]. Sin embargo, cuando la consonante precedente es una sibilante coronal, como en (59.a-d), la vocal del sufijo se realiza como una vocal alta frontal o como una vocal central ([i] o [ɨ]).

- (58) a. /putuk/ → [pʊ.'tuk] → [pʊ.'tʊ] ‘lleno, repleto’
 b. /putuk -(a)n/ → [pʊ.tʊ.kan] ‘lleno, repleto’ (FL)
 c. /ʃanuk/ → [ʃa.'nuk] → [ʃa.'nʊ] (esp. de víbora)
 d. /ʃanuk -(a)n/ → [ʃa.nʊ.kan] (esp. de víbora - FL)
 e. /kiʃak/ → [ki.'ʃak] → [ki.'ʃa] ‘labio’
 f. /kiʃak -(a)n/ → [ki.ʃa.kan] ‘labio’ (FL)

Obsérvese que el patrón del acento en los datos en (59.a-d) es idéntico al que se mostró en los datos en (58). Una vez más esto muestra la similaridad de los patrones fonológicos entre los segmentos latentes codas y las codas regulares. Los datos en (59) también sirven para descartar la posibilidad que el segmento latente [k] que aparece en (58) sea parte de la representación subyacente del sufijo /-(a)n/. Finalmente, obsérvese que la presencia o ausencia de un segmento latente puede crear pares mínimos falsos basados en la posición del acento primario. Compárese (58.a) con (59.g).

- (59) a. /tʃuɔɖzi/ → [tʃu.ˈɖzi] ‘duro’
 b. /tʃuɔɖzi/-(a)n/ → [ˈtʃu.ɖzi.ɲin] ‘duro’ (FL)
 c. /kintʃiʃ/ → [kin.ˈtʃiʃ] ‘de labio negro’
 d. /kintʃiʃ/-(a)n/ → [ˈkin.tʃi.ʃin] ‘de labio negro’ (FL)
 e. /tita/ → [ˈti.ta] ‘madre’
 f. /tita/-(a)n/ → [ti.ˈtan] ‘madre’ (FL)
 g. /putu/ → [ˈpu.tu] ‘polvo’
 h. /putu/-(a)n/ → [pu.ˈtun] ‘polvo’ (FL)

Colocando de lado los casos del segmento latente /k/ en los verbos /ʔik/ ‘ser, suceder’ y /ʔak/ ‘hacer, decir’, los cuales son excepciones, y basados en los datos presentados en este sección, uno puede concluir que el shipibo ha desarrollado dos modos diferentes de lidiar con el punto de articulación de los segmentos latentes. En el caso de las consonantes finales de los verbos y sufijos, el shipibo ha tomado coronal como el punto de articulación por defecto. Sin embargo, quiere que sus segmentos latentes en raíces nominales y adjetivales no sean coronales. En términos de articuladores, no ser coronal significa ser labial o dorsal.¹⁵ El shipibo ha optado por marcar sus segmentos latentes en raíces nominales y adjetivales con la opción [dorsal], [k]. De hecho, esto es interesante porque el capanahua, una lengua pano cercana al shipibo, tomó la opción [labial].

Los datos en (60) nos permiten comparar la palabra para ‘arcilla’ tanto en shipibo como en capanahua. Los ejemplos en (60.a-b) son las formas básicas sin sufijos. En ambas lenguas, el segmento latente final hace que la segunda sílaba sea cerrada y en consecuencia que atraiga al acento primario. Una vez que se agrega el sufijo /-(a)n/ (FL), el segmento latente aparece como [k] en shipibo y como [p] en Capanahua.

- (60) a. /mapuk/ → [ma.ˈpuk] → [ma.ˈpu] (shipibo)
 b. /mapup/ → [ma.ˈpup] → [ma.ˈpu] (capanahua)
 b. /mapuk/-(a)n/ → [ma.ˈpu.kan] (shipibo)
 c. /mapup/-(a)n/ → [ma.ˈpu.pan] (capanahua)

En contraste a los sustantivos y adjetivos donde el shipibo y el capanahua difieren con respecto al punto de articulación de sus segmentos latentes, los datos en (61) muestran que ambas lenguas coinciden en neutralizar sus segmentos latentes de las formas verbales al punto coronal: [t]. Los ejemplos contienen el

¹⁵ Véase la sección 9.2.4 para más información sobre el punto de articulación no-coronal en las consonantes finales de raíces nominales y adjetivales.

verbo ‘lavar ropa’: /patsat/ en Shipibo y /paʔtsat/ en capanahua. En (61.a), el verbo shipibo recibe el sufijo /-ki/ (CMPL) mientras en (61.b), el del capanahua recibe /-kin/ (infinitivo). Ambos sufijos empiezan en una consonante, en consecuencia, el segmento latente no puede ocurrir pero el acento primario es atraído a la segunda sílaba. En (61.c-d), los verbos reciben el sufijo de participio imperfecto, el cual es /-ai/ en shipibo y /-i/ en capanahua. Ya que ambos sufijos empiezan en una vocal, los segmentos latentes aparecen como arranques y el acento principal cae sobre la sílaba inicial.

- (61) a. /patsat -ki/ → [pa.'tsat.ki] → [pa.'tsa.ki] (shipibo)
 b. /paʔtsat -kin/ → [pa.'tsat.kin] → [paʔ.'tsa.kin] (capanahua)
 c. /patsat -ai/ → ['pa.tsa.tai] (shipibo)
 d. /paʔtsat -i/ → ['paʔ.tsa.ti] (capanahua)

El marcaje morfológico de los segmentos latentes en raíces nominales y adjetivales con el rasgo [dorsal] se puede analizar a través de un requisito específico que ha desarrollado el shipibo. Llamemos a este requisito *evitar coronalidad en raíces no-verbales*. Mientras en shipibo, se satisface este requisito específico a las raíces nominales y adjetivales adoptando el punto [dorsal], en Capanahua, se le satisface adoptando el punto [labial].

Este marcaje morfológico no contiene ningún significado asociado con él, lo cual podría parecer extraño, pero no lo es. Lo que el shipibo le hace a los segmentos latentes en raíces nominales y adjetivales no es diferente de lo que, por ejemplo, el castellano le hace a las raíces verbales. El castellano tiene tres sufijos que no aportan ningún significado o rasgos a los verbos a los que se pegan, tradicionalmente conocidos como *vocales temáticas*: /-al/, /-el/, /-i/. Cada verbo en castellano es arbitrariamente asignado a uno de estos sufijos que no hacen otra cosa que marcar, clasificar a cada verbo como perteneciente a un paradigma verbal el cual sigue un tipo específico de conjugación (por ejemplo, /abl -a -r/ ‘hablar’, /ped -i -r/ ‘pedir’, /ca -e -r/ ‘caer’).

El marcaje morfológico de las raíces no-verbales también explica el comportamiento de las consonantes latentes en raíces nominales que convierten en verbos. Todas las palabras en los datos en (55) contienen la misma raíz. En (62.a-b), las palabras son sustantivos. Obsérvese que en (62.b) el segmento latente de la raíz aparece como [k], el arranque de la siguiente vocal. En leyendas o cuentos míticos, los hablantes del shipibo pueden aceptar convertir el sustantivo para ‘caimán’ en un verbo ‘convertirse en un caimán’. Esta forma también ha sido reportada en Lorient, Lauriault *et al.* 1993. El verbo denominizado en (62.c-e) mantiene su segmento latente final. Esto se puede ver por la posición del acento

en (62.c). Sin embargo, ya que la raíz es ahora un verbo, el punto de articulación del segmento latente de pronto cambia a coronal [t]. Véase los datos en (62.d-e). La existencia de este tipo de alternancia hace difícil proponer una sola representación subyacente y un análisis derivacional. En contraste, si el análisis no es derivacional como se propone aquí, entonces la alternancia se sigue del conflicto entre tener [coronal] como el punto de articulación por defecto para las codas y segmentos latentes en general y evitar [coronal] en los segmentos latentes de raíces nominales y adjetivales.

En las representaciones subyacentes de la raíz en (62), los segmentos latentes aparecen como coronales. Sin embargo, nada depende de que sea así. Ya que el evitar la coronalidad en raíces nominales y adjetivales es un requisito impuesto sobre las representaciones fonéticas, el segmento latente en (62) podría haber sido representado subyacentemente por /k/, como en (58), y los resultados aún serían los mismos. En este caso, el segmento latente /k/ ocurrirá como [t] cuando la raíz que lo contiene es un verbo ya que la restricción *Evitar Coronalidad en Raíces No-Verbales* no tiene ningún control sobre las raíces verbales. Las consonantes finales de las raíces verbales están gobernadas por el *Síndrome Coronal*.

- (62) a. /kapit/ → [ka.¹pi] ‘caimán’
 b. /kapit-(a)n/ → [¹ka.pi.kan] ‘caimán’ (FL)
 c. /kapit-ma-ki/ → [ka.¹pi.ma.ki] ‘convertirse en caimán’ (CAUS, CMPL)
 d. /kapit-ai/ → [¹ka.pi.ta.i] ‘convertirse en caimán’ (PP1)
 e. /kapit-a/ → [¹ka.pi.ta] ‘convertirse en caimán’ (PP2)

9.2.3.4 *El punto de articulación de las nasales finales de raíces*

Aunque las nasales finales de raíz no son segmentos latentes, ni como codas ni como arranques, el comportamiento de su punto de articulación se sigue del análisis propuesto para los segmentos latentes.

En shipibo el requisito de tener [coronal] como el punto de articulación por defecto gobierna tanto el punto de articulación de las codas ([s, ʃ, ʒ, n]) y el punto de articulación de los segmentos finales de los morfemas. Esto fuerza a los segmentos latentes a ocurrir como la consonante coronal [t], como se mostró en los datos en (54), (55) y (56). Esta misma fuerza también hace que las codas nasales y las nasales finales de morfemas (ya sea en posición de coda o de arranque) aparezcan como [n]. En los datos en (63.a-b), podemos ver una consonante nasal que ocurre como coda y en tal posición su punto de articulación por defecto es coronal. Las codas nasales son por defecto coronales sin importar si son parte de

una raíz o de un sufijo. En (63.c-d), podemos ver que las nasales finales de raíz y las nasales finales de sufijos que pueden se silabifican como arranques también deben ser coronales. De este modo, la única diferencia entre un segmento latente como /t/ y /n/ es que éste último puede ocurrir tanto como coda y arranque mientras que el primero solo puede aparecer en posición de arranque.

- (63) a. /kũndzũn/ → [kũn.ˈdʒũn] (esp. de árbol)
 b. /tita -(a)n/ → [ti.ˈtan] ‘madre’ (FL)
 c. /ʃinan -ai/ → [ˈʃi.na.na.i] ‘pensar’ (PP1)
 d. /pi -kan -ai/ → [ˈpi.ka.na.i] ‘comer’ (CON.3PL, PP1)

Como en el caso de los segmentos latentes en raíces nominales y adjetivales, el punto de articulación de las consonantes nasales también muestra un comportamiento diferente en este tipo de raíces. Véanse los datos en (64). En el caso de los segmentos latentes en este tipo de raíces, la fonología evita la coronalidad. Al decidir entre los puntos no-coronales, el shipibo opta por [dorsal] de modo que sus segmentos latentes aparecen como [k] cuando se silabifican como arranques. El evitar la coronalidad también gobierna el punto de articulación de las nasales finales de raíces nominales y adjetivales cuando se silabifican como arranques. Sin embargo, ya que la nasal velar */ŋ/ no pertenece al inventario consonántico subyacente de la lengua, el shipibo fuerza a las nasales finales de las raíces nominales y adjetivales a ser [labiales], lo cual es la otra forma de evitar ser [coronal].¹⁶ La sección 5.5.3 ya presentó datos que mostraban este comportamiento en la nasal final de raíces nominales. Repito un subconjunto de esos datos en (14). Las nasales indicadas en negritas son nasales finales de raíz. En la primera columna, aparecen como arranques que preceden al sufijo /-(a)n/ (FL). Las mismas raíces sin el sufijo se muestran junto a la glosa. Obsérvese la alternancia [n] ~ [m] entre las formas cortas y largas de las raíces.

- (64) a. [ˈpĩn.pĩ.man] ‘mariposa’ (forma larga de [pĩn.ˈpĩn])
 b. [ˈtʊʃ.mʊ.man] ‘ampolla’ (forma larga de [tʊʃ.ˈmʊn])
 c. [ˈjʊ.na.man] ‘ahumado’ (forma larga de [jʊ.ˈnan])
 d. [ˈha.kʊ.man] ‘bueno’ (forma larga de [ha.ˈkʊn])

¹⁶ Capanahua resolvió este conflicto escogiendo [labial] de modo que sus segmentos finales de raíz tanto en sustantivos como adjetivos aparecen como [p] y [m] cuando se silabifican como arranques de sílaba.

El evitar la creación de un segmento nasal velar */ŋ/ es un comportamiento fonológico conocido como *Preservación de la Estructura* (*Structure Preservation*) y ha sido encontrado en muchas otras lenguas (véase Kiparsky 1982b, 1985). La Preservación de la Estructura inhibe procesos fonológicos que crean segmentos que no son parte del inventario subyacente de una lengua. Las codas fricativas [s, ʃ, ʂ] no muestran la alternancia que muestran las nasales. Están protegidas debido a que las sibilantes no pueden ser labiales o dorsales. Es decir, su rasgo [+estridente] debe ser preservado en la representación fonética. En consecuencia, solo pueden ser coronales. Además, incluso si el rasgo [+estridente] no necesitara ser preservado, las fricativas como *[f] o *[x], que podrían evitar la coronalidad en las consonantes finales de raíces nominales y adjetivales, son descartadas por el requisito de la Preservación de la Estructura.

Finalmente, como en el caso de la alternancia [t] ~ [k] en (62), esta vez los datos en (65) muestran la alternancia [n] ~ [m]. Todas las palabras en (65) contienen la raíz /tuʃmʊn/ ‘ampolla’. Cuando su consonante nasal final ocurre como coda en (65.a) y (65.c), está sujeta al requisito que todas las codas deben ser coronales. Cuando la nasal final de la raíz en (65.d), un verbo, aparece como arranque, el punto de articulación de la nasal es coronal ya que la coronalidad también gobierna el punto de articulación de las consonantes finales de morfemas. Sin embargo, cuando la nasal final de la raíz en (65.d) ocurre como arranque, su punto de articulación no está más gobernado por la necesidad de tener [coronal] como el punto de articulación por defecto de las consonantes. En vez de ello, ya que la raíz es esta vez un sustantivo, la fonología del shipibo evita la coronalidad forzando a la nasal final que aparezcan como [labial]: [m].

- (65) a. /tuʃmʊn/ → [tuʃ.ˈmʊn] ‘ampolla’
 b. /tuʃmʊn -an/ → [ˈtuʃ.mʊ.man] ‘ampolla’ (FL)
 c. /tuʃmʊn -ti/ → [tuʃ.ˈmʊn.ti] ‘ampollarse’ (INF)
 d. /tuʃmʊn -a/ → [ˈtuʃ.mʊ.na] ‘ampollarse’ (PP2)

Para resumir hasta aquí, hay tres requisitos importantes que la fonología del shipibo debe satisfacer en todos los contextos. Estos tres requisitos son: *codas coronales* (es decir, todas las codas deben ser consonantes coronales), *preservación de la estructura* (es decir, no se puede crear segmentos que no son parte del inventario segmental subyacente de la lengua) y *evitar coronalidad* (es decir, las consonantes finales de raíces nominales y adjetivales no deben ser coronales). Además, el shipibo tienen un cuarto requisito que solo puede satisfacerse si se satisfacen primero los tres requisitos que acaban de numerarse. Este cuarto requisito es *consonantes finales*

coronales de morfemas (el punto de articulación por defecto de las consonantes finales de raíces y sufijos es coronal).

9.2.4 Problemas fonológicos que rodean a las nasales

Hay dos problemas fonológicos que presentan las nasales en shipibo. El primer problema se relaciona al comportamiento de su punto de articulación, el cual ya se discutió en la sección 9.2.3.4. El segundo problema, que cubre esta sección, es si las vocales nasalizadas deberían analizarse como una secuencia de una vocal seguida de una consonante nasal coda. Esta es una pregunta recurrente en los estudios del shipibo en particular y de las lenguas pano en general. Los resultados obtenidos en la sección 5.2 apoyan el análisis que el shipibo no tiene vocales nasales fonémicas. La respuesta a la misma pregunta para las otras lenguas pano debe esperar por un estudio acústico.

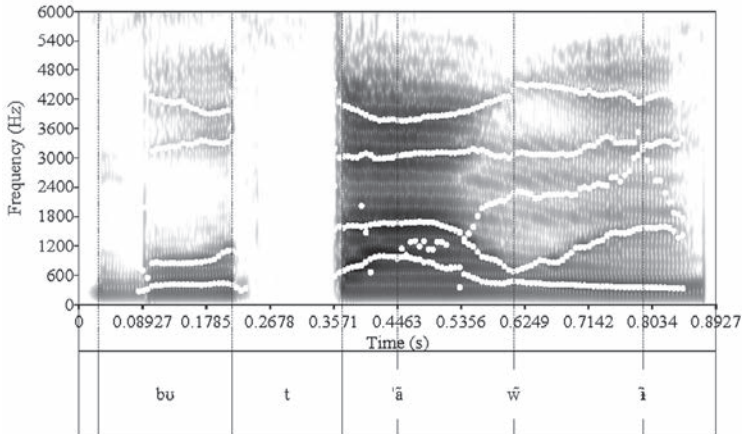
La sección 5.2.1 presentó la duración de una vocal seguida de una coda nasal y la duración de las codas nasales. En promedio, una vocal seguida de una coda nasal dura 120 ms en una sílaba acentuada y 90 ms en una sílaba inacentuada. Las codas nasales duran en promedio 60 ms ya sea en sílabas acentuadas o inacentuadas. Sobre la diferencia fonética entre vocales nasalizadas y vocales nasales fonémicas, es importante mencionar el trabajo hecho por Everett (en prensa), quien ha demostrado acústicamente para el karitiâna, una lengua tupí hablada en el Brasil, que las vocales nasales nunca muestran una fase oral mientras que las vocales nasalizadas sí tienden a tener una fase oral. Un contraste interesante entre el shipibo y el karitiâna con respecto a las vocales nasalizadas es que en karitiâna, la nasalización puede ocurrir circun-oralizada (es decir, la nasalización se concentra en el medio de la vocal pero sus bordes se mantienen orales). Sin embargo, la circun-oralización, aunque es una posible realización de las vocales nasalizadas en karitiâna, no es la realización más común. Como en shipibo, la realización más común de las vocales nasalizadas es primero mostrar una fase oral seguida de una fase nasal. De nuevo, este es otro caso donde dos lenguas, como el shipibo y el karitiâna, que poseen el mismo tipo de segmento (en este caso, las vocales nasalizadas), pero esos no muestran en mismo comportamiento fonético.¹⁷

Ahora examinemos un caso típico de una vocal que se ha fusionado completamente con su coda nasal sin dejar ninguna huella de la consonante nasal. Esto ocurre en shipibo cuando una coda nasal es seguida por una glide, como por

¹⁷ Everett (en preparación) reporta que la circun-oralización de las vocales nasalizadas es inusual tipológicamente. Además del karitiâna, las lenguas Jê, kaingang (Wiesemann (1972), Wetzels (1995) D'Angelis (1999)) y apinayé (Callow (1962), Burgess and Ham (1968)) también son conocidas por tener circun-oralización.

ejemplo en /bU -tan -wi/ → [bU.¹tan.wi] ‘anda y tómallo!’, realizado fonéticamente como [bU.¹tã.wĩ]. Una ocurrencia de esta palabra se presenta en el espectrograma de la Figura 5-17. Aunque no es posible hacer mediciones exactas de la vocal nasalizada cuando ocurre seguida de una glide, bajo las mediciones más conservadoras, la duración de la vocal acentuada de la sílaba [tã] no debe ser menos de 165 ms. Esta duración es bastante larga considerando las mediciones promedio obtenidas en la sección 5.2.1. Una vocal acentuada [a] seguida de una coda nasal dura aproximadamente 126 ms en una segunda sílaba. La explicación para los 40 ms extra que dura esta vocal es la presencia de una coda nasal que la sigue en el nivel representacional: [bU.¹tan.wĩ].

**Figura 9-20: /bU -tan -wi/ → [bU.¹tan.wi] ‘anda y tómallo!’
(realizado como [bU.¹tã.wĩ])**



Además de la evidencia fonética obtenida en este estudio, hay también evidencia fonológica de que las codas nasales corresponden a segmentos fonológicos independientes. El primer tipo de evidencia fonológica viene del comportamiento del acento. Las sílabas cerradas le quitan el acento primario a las sílabas iniciales. Compárese los datos en (66.a) y (66.b): en el último, el acento cae sobre la segunda sílaba porque es cerrada por una sibilante. Las codas nasales se comportan de un modo similar a las codas sibilantes. Atraen el acento sin importar si llegan a realizarse como un segmento independiente (véase (66.c)) o si coalescen completamente con los segmentos vecinos (véase (66.d)) y el espectrograma de la Figura 5-17).

- (66) a. /b^βU -ti/ → [b^βU.ti] ‘tomar’ (INF)
 b. /b^βU -kas -ti/ → [b^βU.¹kas.ti] ‘querer tomar’ (DES, INF)
 c. /b^βU -tan -ti / → [b^βU.¹tan.ti] ‘ir y tomar’ (/tan/ ‘ir
 y regresar’, INF)
 d. /b^βU -tan -wi/ → [b^βU.¹tan.wi] ([b^βU.¹tã.wĩ]) ‘anda y tómalolo!’
 (/tan/ ‘ir y regresar’, IMP)

El segundo tipo de evidencia fonológica para analizar las vocales completamente nasalizadas como una secuencia de una vocal seguida de una coda nasal viene de la resilabificación. Las codas nasales se resilabifican como arranques y cuando esto sucede claramente se separan de la vocal precedente y aparecen como segmentos completamente independientes. Véase los datos en (65). Este comportamiento puede verse claramente en los espectrogramas presentados en la sección 5.3 donde las nasales ocurren en posición de arranque.

9.3 LA INTERACCIÓN ENTRE SÍLABA Y ACENTO

En esta sección final del bosquejo de la fonología del shipibo, discutiré la sílaba y su rol en la asignación del acento. Empecemos por examinar la estructura silábica.

9.3.1 La estructura de la sílaba del shipibo

Los datos en (60) muestran ejemplo de las diferentes estructuras silábicas del shipibo. Cada sílaba obligatoriamente tiene una vocal como su núcleo. Esa vocal puede ser corta o larga. La distribución de las vocales largas se discutió en la sección 9.1.2. Además del núcleo, la sílaba puede opcionalmente tener un arranque y una coda. Que una consonante ocupe la posición de arranque o coda de la sílaba se ve reflejado por las claves fonéticas. De este modo, en la sección 7.1.1 se encontró que las nasales y las sibilantes muestran una duración más larga cuando ocurren como arranques. Además, el estudio acústico también obtuvo otra importante generalización relacionada a la duración de los segmentos. En contraste a las vocales cuya duración cambia de acuerdo al estado de acentuado de la sílaba que las contiene, (véase la sección 2.3), la duración de las consonantes en posición de arranque o coda tiende a ser más resistente a este factor. El estudio encontró que las duraciones de los arranques y codas en las sílabas iniciales y segundas con o sin acento son muy similares. Esta asimetría entre consonantes y vocales con respecto al acento sirve como clave fonética a los hablantes del shipibo para recuperar la estructura prosódica de las palabras y frases.

- (67) a. V : [ʼa.ta.pa] ‘gallina’
 b. CV : [ʼti.ta] ‘madre’
 c. CVC : [tʃa.ʼdʒas] (esp. de pájaro)
 d. CVCC : [ʼpʊŋʂ.ti.ti] ‘cortar el brazo’ (INF)
 e. VC : [ʼis.ku] (esp. de pájaro)
 f. VCC : [ʼins.pi] ‘lado de la cabeza’
 g. CV: : [ʼni:] (esp. de lagartija)
 h. CV:C : [ʼki:n] ‘deseo’

Aunque los arranques y las codas son opcionales en shipibo, hay una tendencia a que las sílabas tengan arranques. Además, el shipibo solo permite que una consonante ocupe la posición de arranque de la sílaba. Cualquier consonante puede ocupar esa posición. Sin embargo, en las fronteras morfélicas, cuando uno de los morfemas termina en una consonante y la siguiente empieza en una vocal, las consonantes que ocurren como arranques están restringidas a las consonantes [t, k, m, n, s, ʃ, ʂ]. Esto se debe a lo que se discutió en la sección 9.2.3: las consonantes finales de morfema sufren restricciones respecto a su punto de articulación. Cuando ocurren como arranques, deben ser coronales: [t, n, s, ʃ, ʂ]. Sin embargo, las consonantes finales de morfemas en raíces nominales y adjetivales tratan de evitar ser coronales cuando ocurren como arranques. Logran hacer esto para las oclusivas y nasales pero no para las sibilantes. De este modo, los arranques que vienen de consonantes finales de raíces nominales y adjetivales pueden ser: [k, m, s, ʃ, ʂ].

Solo la nasal coronal [n] y las sibilantes coronales [s, ʃ, ʂ] pueden aparecer como codas. Además, el shipibo permite que las codas oclusivas se comporten como cualquier otra coda con respecto a la interacción con el acento (véase la sección 9.2.3), aunque no se realicen fonéticamente. El shipibo prefiere codas simples pero tiene algunos pocos casos de codas complejas. Es importante señalar que las sílabas con codas complejas y aquellas con vocales largas solo pueden aparecer al inicio de palabras (véase la sección 2.4, la sección 7.2 y la discusión sobre las posiciones prominentes en la sección 9.1.3). En estos casos, las codas pueden tener un máximo de dos consonantes donde la consonante que sigue al núcleo debe ser una nasal y la segunda consonante, una sibilante (la única excepción es la palabra [ʼunh.nu] ‘allá muy lejos’ en la cual la coda es de la primera sílaba tienen una fricativa laríngea como la segunda consonante del margen silábico complejo - véase la sección 4.2 y la sección 7.2).

9.3.1.1 Las glides y la silabificación

La sección 6.2.2 examinó y presentó evidencia fonética de cómo el shipibo evita acentuar las vocales iniciales de palabras que empiezan con dos vocales adyacentes. De modo similar, en ese mismo capítulo (véase la sección 6.2.3) discutió y dio evidencia fonética de cómo las vocales altas obligatoriamente se convierten en semi-consonantes cuando ocurren en contextos intervocálicos. En esta sección, discutiré la interacción fonológica entre estos dos patrones. Llamemos al primero *evitar vocales acentuadas iniciales de palabra* y al segundo, *evitar vocales altas en contextos intervocálicos*. El evitar tener vocales altas en ambientes intervocálicos es más importante de satisfacer en shipibo que el evitar colocar el acento primario en las vocales altas iniciales de las palabras que empiezan con dos vocales.

Los datos en (68) presentan ejemplos de este conflicto. En la forma infinitiva del verbo /iʊ -ti/ ‘tomar, dirigir’ en (68.a), la vocal /ʊ/ no tiene ninguna oportunidad de convertirse en una semi-consonante ya que no aparece en un contexto intervocálico. En este caso, el acento puede evitar caer sobre la vocal inicial moviéndose sobre la vocal [ʊ]. Uno puede asumir que la representación fonética de la forma infinitiva de este verbo es [i.ʊ.ti] y que la sílaba inicial inacentuada [i] puede mostrar varios grados de semi-consonantismo en su realización: [ᶯi.ʊ.ti] ~ [ʲi.ʊ.ti] ~ [i.ʊ.ti].

- (68) a. /iʊ -ti/ → [ᶯi.ʊ.ti] ~ [ʲi.ʊ.ti] ~ [i.ʊ.ti] ‘tomar, dirigir’ (INF)
 b. /iʊ -a/ → [ʲi.wa] ‘tomar, dirigir’ (PP2)
 c. /ʊi -ti/ → [ᶯʊi.ti] ~ [ʲwi.ti] ~ [ʊ.ʲi.ti] ‘quemar (cerámica cruda)’ (INF)
 d. /ʊi -a/ → [ʲʊ.ja] ‘quemar (cerámica cruda)’ (PP2)

Las cosas, sin embargo, son diferentes cuando cambiamos el sufijo /-ti/ (INF) por el sufijo /-a/ (PP2), como en (68.b): /iʊ -a/. En este caso, los dos patrones entran en conflicto: el evitar acentuar la sílaba inicial quiere que el acento primario caiga sobre la segunda vocal, /ʊ/. Al mismo tiempo, la fonología de la lengua requiere que la vocal /ʊ/ aparezca como una semi-consonante ya que ocurre en un contexto intervocálico, y como una semi-consonante, este segmento, no puede llevar acento. En este caso, el shipibo claramente prefiere satisfacer el segundo requerimiento. La vocal /ʊ/ es forzada a ocurrir como una semi-consonante labiodorsal [w], [ʲi.wa]. El acento entonces cae sobre la sílaba inicial de la vocal. Las alternativas en las cuales el acento evita la vocal inicial, *[ᶯi.ʊ.a] ~ *[ʲi.ʊ.a] ~ *[i.ʲʊ.a], son consistentemente rechazadas por todos los hablantes.

La interacción entre convertir las vocales altas intervocálicas en semi-consonantes y evitar acentuar las vocales altas iniciales crea una alternancia notable dentro del paradigma de los verbos que tienen como sus dos segmentos iniciales las vocales subyacentes /i/ y /u/. Cuando se logra evitar el acento sobre la vocal inicial de la palabra, esta vocal se convierte en una glide (por ejemplo, la /i/ inicial de la forma infinitiva /iʉ -ti/ ‘tomar, dirigir’ → [ʰjʉ.ti]). Sin embargo, cuando la segunda vocal acaba en una posición intervocálica, se le fuerza a ser una semi-consonante y esta vez el segmento inicial ocurre como una vocal (por ejemplo, la /u/ de la forma /iʉ -a/ ‘tomar, dirigir’ (PP2) → [ʰi.wa]). De este modo, en la superficie, observamos las formas alternantes [ʰjʉ.ti] y [ʰi.wa], donde ambas comparten la misma raíz subyacente /iʉ/ ‘tomar, dirigir’.

La alternancia también nos da evidencia que apoya la afirmación que la representación subyacente de las palabras que empiezan con una vocal alta /i/ o /u/ (véase la sección 6.2.2) tiene dos vocales subyacentes iniciales. Primero, obsérvese que las semi-consonantes iniciales de palabra, [w] y [j], no muestran la variación descrita para las vocales /i/ y /u/ iniciales de palabras que además empiezan con dos vocales (por ejemplo, /iʉ -ti/ → [ʰjʉ.ti] ~ [ʰjʉ.ti] ~ [i.ʰu.ti]). Las semi-consonantes subyacentes siempre ocurren como [j] y [w]. De este modo, las palabras para ‘hacha’ /jami/ o para ‘mezquino, tacaño’ /juʌʃi/ siempre se realizan como [ʰja.mi] y [ʰju.a.ʃi], donde el segmento inicial tiene un timbre consonántico característico. En contraste, el verbo para ‘cargar (algo) en los hombros’, /ia -ti/, pueden realizarse ya sea como [ʰja.ti], [ʰja.ti] o [i.ʰa.ti]. Esto indica que el segmento subyacente inicial de este verbo no es la semi-consonante [j].

Obsérvese que si la representación subyacente de la palabra para ‘mezquino, tacaño’ fuese /iʉʌʃi/, donde el segmento inicial es la vocal /i/ en lugar de la semi-consonante /j/, entonces uno esperaría incorrectamente que la /u/ intervocálica ocurriese como la semi-consonante [w]: *[ʰi.wa.ʃi].

La segunda pieza de la evidencia que indica que la representación subyacente de las palabras que empiezan con una vocal alta /i/ o /u/ tiene dos vocales iniciales subyacentes viene de la habilidad de cada una de esas vocales para aparecer bajo ciertas condiciones como un segmento vocálico. La vocal subyacente inicial obligatoriamente ocurre como un segmento vocálico cuando la segunda está en un ambiente intervocálico y en consecuencia, se le fuerza a ser una semi-consonante (por ejemplo, /iʉ -a/ → [ʰi.wa] ‘tomar, dirigir’ (PP2)). En contraste, una secuencia similar de segmentos, pero con una glide subyacente inicial, no se convertirá en una vocal. En este caso, aparecerá como [jua], por ejemplo, /juak -(a)n / → [ʰju.a.kan] ‘comida’ (FLC cuando la segunda vocal subyacente no se encuentra en un ambiente intervocálico, ambas vocales subyacentes ocurren como vocales en la representación fonética con el acento sobre la segunda vocal con el

fin de evitar acentuar la vocal inicial (/iʊ -ti/ → [i.¹ʊ.ti] ‘tomar, dirigir’). De esta representación fonética (el resultado abstracto del componente fonológico de la gramática), se sigue la variación observada en la realización de la vocal inicial en palabras como [i.¹ʊ.ti]. Dado que la vocal inicial está inacentuada y adyacente a una vocal acentuada en la representación fonética, está sujeta a sufrir diferentes grados de semi-consonantismo (debilitamiento). Algunas veces se convierte en un sonido breve y transitorio pero sin perder su timbre vocálico ([¹ʊ.ti]). Otras veces, se realiza con un timbre más consonántico ([¹jʊ.ti]) o en otras ocasiones, se realiza completamente como una vocal ([i.¹ʊ.ti]). De este modo, el shipibo no tiene diptongos en el nivel representacional pero puede mostrarlo en la realización del habla.

La lista en (69) presenta las representaciones subyacentes y fonéticas de las palabras que empiezan con una vocal alta. Recuérdesse que en cada una de las representaciones fonéticas dadas en (69), la vocal alta inicial puede mostrar diferentes realizaciones. De este modo, como se determinó en la sección 6.2.2, por ejemplo, la representación fonética de [i.¹a] ‘piojo’ puede realizarse como [¹ja] ~ [¹ja] ~ [i.¹a].

(69) a.	/ia/	→	[i. ¹ a]	‘piojo’
b.	/iamaʃ/	→	[i. ¹ a.maʃ]	‘es decir, así’
c.	/ia -ti/	→	[i. ¹ a.ti]	‘cargar (algo) en los hombros’ (INF)
d.	/iʊ -ti/	→	[i. ¹ ʊ.ti]	‘tomar, dirigir’ (INF)
e.	/ʊi/	→	[ʊ. ¹ i]	‘inusualmente grande’
f.	/ʊi /	→	[ʊ. ¹ i]	‘lluvia’
g.	/ʊin -a/	→	[ʊ. ¹ i.na]	‘ver’ (PP2)
h.	/ʊipua /	→	[ʊ. ¹ i.pʊ.a]	(esp. de caracol)
i.	/ʊiʃʊdʒʊ/	→	[ʊ. ¹ i.ʃʊ.dʒʊ]	(esp. de pájaro)
j.	/ʊi -ti/	→	[ʊ. ¹ i.ti]	‘quemar (cerámica cruda)’ (INF)

9.3.1.2 Las vocales altas y la reestructuración fonética de las sílabas

Las representaciones fonéticas ofrecidas en (69) también indican que la fonología del shipibo no se da cuenta de las realizaciones fonéticas como glides que la vocal inicial puede mostrar durante la producción de esas palabras. La asignación del acento puede usarse como un test fonológico para apoyar esta aseveración. Examinemos la representación fonética [i.¹a.maʃ] (que viene de la representación subyacente /iamaʃ/ ‘es decir, así’). Esta representación fonética puede realizarse como [¹ja.maʃ] ~ [¹ja.maʃ] ~ [i.¹a.maʃ]. Si la semi-vocalización y la semi-consonantización observadas en las dos primeras realizaciones,

[¹ja.maʃ] ~ [¹ja.maʃ], fuesen creadas por la fonología de la lengua, entonces la palabra sería bisilábica con la segunda sílaba cerrada. El acento en shipibo es atraído por las segundas sílabas cerradas (por ejemplo, [jü.¹nan] ‘ahumado’, [ja.¹wiʃ] ‘armadillo’). Siguiendo, este patrón, uno erróneamente esperaría que el acento cayese sobre la segunda sílaba de la palabra: *[ja.¹maʃ].

El acento realmente cae sobre la sílaba inicial: [¹ja.maʃ]. Sin embargo, la posición del acento se sigue de la representación fonética: [i.¹a.maʃ], la cual tiene tres sílabas. Como se mencionó antes, el acento en shipibo evita caer sobre vocales altas iniciales de palabras que empiezan con dos vocales subyacentes. Como consecuencia, el acento aparece sobre el segundo mejor segmento silábico de la izquierda disponible, es decir, la vocal [a] de [i.¹a.maʃ].

Durante la producción de esta palabra, la vocal inicial está sujeta a sufrir varios grados de semi-consonantización debido a su estatus inacentuado y la adyacencia a una vocal acentuada (es decir, [¹ja.maʃ] ~ [¹ja.maʃ] ~ [i.¹a.maʃ]). La sílaba cerrada final nunca logra atraer el acento ya que desde un punto de vista fonológico, está fuera de la ventana del acento primario de la lengua.

Un fenómeno similar de reestructuración de sílabas que convierte las vocales altas en glides crea sílabas con un arranque seguido de un núcleo complejo. En este caso, las vocales altas solo pueden manifestarse como las semi-vocales: [j, ɥ]. Por ejemplo, la palabra para ‘moho’, /dʒapui/ usualmente se realiza como [¹dʒa.pu̯j], con la vocal final ocurriendo como la semi-vocal [j]. Como en el caso anterior, este fenómeno ocurre fuera del reino de la fonología. Primero, aunque la segunda sílaba [¹dʒa.pu̯j] ocurre con un núcleo complejo [u̯j], el cual debería hacer que la sílaba sea pesada y de este modo, atraer el acento primario, pero no lo hace así. Hasta donde le concierne a la fonología del shipibo, la representación de esta palabra es [¹dʒa.pu.i], con una segunda sílaba ligera abierta de modo que el acento primario cae sobre la sílaba inicial.

Independientemente del comportamiento de la asignación de acento, la reestructuración de las sílabas por medio de convertir las vocales altas en glides tampoco afecta la selección de los alomorfos que dependen del número de sílabas de la palabra. Como se mencionó en los capítulos anteriores, los sustantivos del shipibo tienen dos formas: las formas cortas y las formas largas. Las formas largas corresponden a la base de un sustantivo seguida de un sufijo /-(a)n/ o /-nin/ (FL). La selección del alomorfo se hace de acuerdo al número de sílabas en la base. Cuando la base tiene un número par de sílabas, se añade el sufijo /-n/ (por ejemplo, [¹ti.ta] ‘madre’ / [ti.¹tan] (forma larga de ‘madre’)). De otro modo, la gramática escoge /-nin/ (por ejemplo, [¹a.ta.pa] ‘gallina’ / [¹a.ta.pa.nin] (forma larga de ‘gallina’)). La palabra para ‘moho’, a pesar de ser realizada como una palabra de dos sílabas, [¹dʒa.pu̯j], debido a que la vocal final se convirtió en una semi-vocal, solo

puede aceptar el alomorfo /-nin/, [¹dʒa.pu̯.nin], lo cual indica que la fonología aún considera que la base tiene tres sílabas: [¹dʒa.pu.i].

9.3.2 El acento primario

Los capítulos 2 y 8 nos presentaron una caracterización acústica de las claves fonéticas que revelan el estado acentuado de una sílaba: la duración vocálica y el tono perceptual (pitch). En términos fonológicos, el acento, el hecho que una sílaba sea prominente dentro de la estructura prosódica, es mayormente predecible en shipibo. Como ya se mencionó, el acento principal ocurre sobre la segunda sílaba de una palabra si ésta es cerrada, de otro modo, ocurre sobre la sílaba inicial. Con raras excepciones (como por ejemplo, la mayoría de las formas vocativas de palabras trisilábicas que llevan el acento primario sobre la sílaba final), el acento principal no puede aparecer más allá de la segunda sílaba de una palabra. Es decir, la primera y la segunda sílabas de las palabras forman la ventana acentual del acento primario.

En fonología, podemos capturar los límites del acento primario dentro de las dos sílabas iniciales de las palabras por medio de la noción de pie métrico, el agrupamiento de sílabas en constituyentes prosódicos más altos. Los pies métricos en shipibo son bisilábicos (véase también la sección 9.1.3). De este modo, el pie métrico del acento primario del shipibo se alinea con el lado izquierdo de las palabras. Dentro de ese pie, el acento primario se alinea por defecto con la sílaba izquierda al menos que la segunda sílaba sea cerrada. Esta sensibilidad que muestra el acento a las sílabas cerradas se conoce como *Weight-by-Position* (WBP - Prince 1990; Hayes 1995). WBP requiere que la coda de las sílabas cerradas lleven una mora (es decir, una unidad fonológica de tiempo y peso). Esta mora extra hace que la sílaba cerrada sea bimoraica (pesada). En contraste, las sílabas abiertas con vocales cortas son monomoraicas (ligeras), solo tienen una mora llevada por la vocal. Cuando una sílaba dentro del pie métrico principal es bimoraica en shipibo, aleja al acento primario de su posición por defecto. Esto explica los patrones acentuales observados en (70). Los paréntesis indican los pies métricos y el símbolo μ representa la mora.

- (70) a. /tita/ → (¹ti_μ.ta_μ) ‘madre’
 b. /tita -(a)n/ → (ti_μ.¹tan_{μμ}) ‘madre’ (FL)
 c. /jawif/ → (ja_μ.¹wif_{μμ}) ‘armadillo’
 d. /jawif -(a)n/ → (¹ja_μ.wi_μ)(_μfan_{μμ}) ‘armadillo’ (FL)

Obsérvese que cuando ocurren dos sílabas abiertas dentro del pie métrico principal, como en (70.a), es la sílaba inicial la que obtiene el acento. Esto sugiere que cuando dos sílabas estructuralmente idénticas compiten por el acento primario, el shipibo escoge por defecto aquella que está más a la izquierda. Sin embargo, sorprendentemente, cuando dos sílabas cerradas compiten por el acento primario, el shipibo siempre se lo da a la segunda sílaba. Esto puede observarse en los datos en (71).

- (71) a. (wa_μʂ.¹mɪn_{μμ}) ‘algodón’
 b. (kɪ_μm.¹pɪʂ_{μμ}) ‘mal aliento’
 c. (mi_μs.¹pan_{μμ}).ti_μ ‘preparar tamales’ (INF)
 d. (pʊ_μn.¹wɪʂ_{μμ}).ti_μ ‘arañar el brazo de alguien’ (INF)

Siguiendo la idea que las sílabas cerradas iniciales se comportan como sílabas abiertas, Elias-Ulloa 2006 analiza el patrón acentual en (71) como casos donde las sílabas cerradas muestran peso moraico variable. Es decir, en (71), mientras que la sílaba cerrada inicial es monomoraica (la vocal solo lleva una mora), la segunda sílaba es bimoraica (tanto la vocal como la coda llevan una mora cada una). Obsérvese que los datos en (71) no son casos de acento léxico. Una vez que la coda de la segunda sílaba se silabifica como un arranque, se vuelve abierta y el acento regresa a la sílaba inicial. Véase (72). Esto sigue exactamente el mismo patrón observado en (70.c-d).

- (72) a. /waʂmɪn -(a)n/ → (¹waʂ.mɪ)(,man) ‘algodón’ (FL)
 b. /kɪmpɪʂ -(a)n/ → (¹kɪm.pi)(,ʂan) ‘mal aliento’ (FL)

¿Qué inhibe una coda moraica en las sílabas cerradas iniciales? Hay varias fuerzas métricas que hacen sentir su influencia aquí. El shipibo requiere que las sílabas bimoraicas sean acentuadas. Éste es un fenómeno conocido como *Weight-to-Stress Principle* (WSP - Prince 1990; Hayes 1995). De este modo, por ejemplo, si la sílaba cerrada inicial en (71.b) fuese bimoraica, *(kɪm_{μμ}.¹pɪʂ_{μμ}) ‘mal aliento’, ese requerimiento no estaría satisfecho. Una alternativa para satisfacer WSP sería que cada sílaba cerrada formase su propio pie métrico y en consecuencia cada uno apareciese acentuado: *(¹kɪm_{μμ})(,pɪʂ_{μμ}). El problema con esta solución es que causaría un choque acentual (Lieberman y Prince 1977). Como se discutió en el capítulo, un choque acentual ocurre cuando sílabas acentuadas aparecen adyacentes. Esta alternativa es eliminada por el shipibo, el cual siempre cumple con evitar la creación de choques acentuales.

Pero, entonces, ¿por qué no se le permite a la coda de la sílaba inicial llevar una mora pero sí se le permite a la segunda sílaba cerrada? La respuesta a esta pregunta reside en *Grouping Harmony* (GH), es decir, la tendencia a evitar pies métricos que tengan una sílaba bimoraica seguida de una monomoraica: ($\sigma_{\mu\mu}.\sigma_{\mu}$). Muchas lenguas tratan de evitar este tipo de pie métrico (Prince 1990; Hayes 1995; Kager 1999). Por ejemplo, el árabe del Cairo evita el pie ($\sigma_{\mu\mu}.\sigma_{\mu}$) por medio de dejar que las sílabas monomoraicas formen su propio pie métrico (Mitchell 1960; McCarthy 1979; Prince 1990:9; Kenstowicz 1994:291-2; Hayes 1995:67-71, entre muchos otros). En esta lengua, el agrupamiento de sílabas en pie métricos corre de izquierda a derecha y el pie métrico principal se localiza tan cerca como sea posible al lado derecho de la palabra. Obsérvese en (73.b) que si la sílaba bimoraica, [qa:], formase su propio pie con la siguiente sílaba, lo cual crearía un pie ($\sigma_{\mu\mu}.\sigma_{\mu}$), predeciríamos equivocadamente que el acento principal caería sobre la sílaba inicial. El pie ($\sigma_{\mu\mu}.\sigma_{\mu}$) se evita dejando que la sílaba [qa:] forme su propio pie métrico, lo cual se refleja en el hecho que el acento primario cae sobre la segunda sílaba.

- (73) El árabe del Cairo evita el pie métrico ($\sigma_{\mu\mu}.\sigma_{\mu}$)
- | | | | |
|----|--|--|------------|
| a. | (_i fa _μ .ja _μ)(^l ra _μ .tu _μ).hu _μ | | ‘su árbol’ |
| b. | (_i qa: _{μμ})(^l hi _μ .ra _μ) | *(^l qa: _{μμ} .hi _μ).ra _μ | ‘Cairo’ |

Otras lenguas como el tonga (Churchward 1953; Mester 1991; Hayes 1995) evitan la ocurrencia del pie métrico ($\sigma_{\mu\mu}.\sigma_{\mu}$) ajustando el peso moraico de sus sílabas. Tonga requiere que el pie métrico principal se alinee con el lado izquierdo de la palabra. A fin de satisfacer este requisito, el tonga fuerza a la vocal larga de la sílaba inicial a acortarse (a volverse monomoraica) de modo que pueda agruparse junto con la segunda sílaba sin crear un pie métrico ($\sigma_{\mu\mu}.\sigma_{\mu}$). Véase (74.a). Una vez que hay suficientes sílabas en el lado derecho para formar el pie métrico principal, la vocal larga de la sílaba inicial aparece con su vocal larga subyacente. Shipibo se comporta exactamente igual al tonga (véase los datos en (71)), excepto que en vez de ajustar el peso moraico de sus sílabas con vocales largas, ajusta el peso de sus sílabas cerradas.

- (74) El tonga evita el pie métrico ($\sigma_{\mu\mu}.\sigma_{\mu}$)
- | | | | | |
|----|-------------|---|---|------------------------|
| a. | /si:βi/ | → | (^l si _μ .βi _μ), *(^l si: _{μμ} .βi _μ) | ‘exceder’ |
| b. | /si:βi -ta/ | → | (_i si: _{μμ})(^l βi _μ .ta _μ) | ‘exceder’ (transitivo) |

9.3.3 El acento secundario

Tradicionalmente se ha asumido que el shipibo carece de acentos secundarios. Sin embargo, uno de los resultados más importantes que ha arrojado el estudio acústico de la frecuencia fundamental y su correlato, el tono perceptual, presentado en el capítulo 8 (sección 8.6) es que la lengua sí tiene acentos secundarios. Debo colocar una advertencia en este punto. La propuesta de la existencia de acentos secundarios en una lengua es un asunto fonológico y uno particularmente relacionado a qué considera el analista que cuenta como evidencia para proponer su existencia. En este estudio, afirmo la existencia de acentos secundarios en el shipibo basado en evidencia fonética, es decir, en el comportamiento del F0 (tono perceptual) y en alguna medida, en el comportamiento de la duración.

La evidencia más robusta viene del comportamiento del F0 y las sílabas cerradas que ocurren más allá del pie métrico principal (es decir, la ventana del acento primario). En ese contexto, las sílabas cerradas tienden a aparecer alineadas con un tono alto. La presencia de un tono alto no es obligatoria, particularmente en el habla rápida. Sin embargo, tomo la habilidad de las sílabas cerradas de ocurrir asociadas con un tono alto fuera del pie métrico principal como una indicación de su estado acentuado.

Los datos en (70.d) nos presentaron un ejemplo de una sílaba con un acento secundario. Algunos datos adicionales se presentan en (75). He marcado las sílabas acentuadas que reciben un tono alto con un acento agudo [´]. Los datos en (70) y (72) muestran una sílaba cerrada que ocurre como la sílaba inicial o segunda de una palabra. Los datos en (75) muestran una sílaba cerrada a dos, tres y cuatro sílabas de distancia de lado izquierdo de la palabra.

- (75) a. /b^βama -kan -ki/ → (¹b^βá.ma)(,kán.kì) ‘aparecieron de repente’ (CON.3PL, CMPL)
- b. /jumidza -kan -ki/ → (¹jú.mi)(dza.,kán).kì‘(lo) pescaron’ (CON.3PL, CMPL)
- c. /b^βana -jama -kan -ki/ → (¹b^βá.na)(,ja.ma)(,kán.kì) ‘no (lo) sembraron’ (NEG, CON.3PL, CMPL)

Ahora examinemos los datos en (76). En la sección anterior, mostré que cuando dos sílabas cerradas compiten por el acento dentro del pie métrico principal, solo una de ellas lo obtiene con el fin de evitar la creación de un choque acentual. Los datos en (76.a) muestran lo que sucede cuando dos sílabas cerradas adyacentes compiten por su estatus acentuado pero una pertenece al pie métrico principal y la otra pertenece a un pie métrico secundario. En este contexto, la sílaba del pie métrico principal obtiene el acento primario y la sílaba cerrada adyacente se queda inacentuada. El acento secundario se mueve a la sílaba que sigue a la segunda sílaba cerrada. Sin embargo, hay bastante variación con respecto a si los hablantes asignan un tono alto a esa sílaba. Tomo la posibilidad de que la sílaba pueda aparecer con un tono alto como una indicación que un acento secundario ha sido alejado del acento principal para poder evitar un choque acentual. El contorno tonal de (76.a) y (76.b) se muestran en la sección 8.6.1.

Los datos en (76.b) también nos brindan evidencia de que, como ocurre dentro del pie métrico principal, cuando dos sílabas cerradas se agrupan para formar un pie métrico, la primera sílaba es forzada a ser monomoraica. Obsérvese que la sílaba cerrada [dʒún] obtiene un tono alto, lo cual indica su estatus acentuado, mientras que la sílaba cerrada precedente, [màn], recibe un tono bajo. Esto es exactamente el patrón que se encontró en el pie métrico principal (por ejemplo, [wì.ˈtáʃ] ‘pierna’, [wìʃ.ˈtín] ‘estrella’), lo cual quiere decir que los mismos requisitos métricos que dieron forma al pie principal también gobiernan la forma de los pies métricos secundarios del shipibo.

- (76) a. /b^βinun -dʒunki/ → (b^βi.ˈnún)(dʒun.ᵛkí) (esp. de palmera
- REP)
- b. /b^βawin -(a)n -dʒunki/ → (ˈb^βá.wi)(man.ᵛdʒún).kí (esp. de pescado
- FL, REP)

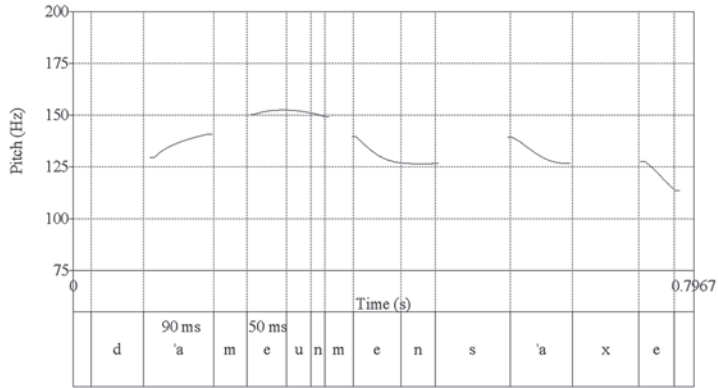
A diferencia de las sílabas abiertas dentro del pie métrico principal, las cuales siempre muestran un tono alto si ocurren acentuadas, más allá del pie métrico principal, las sílabas abiertas no muestran ningún tono alto, al menos que posean un acento léxico (véase la sección 9.3.4). Esto hace difícil la determinación de si llevan acento o no. Sin embargo, las sílabas abiertas pueden mostrar, aunque no obligatoriamente, cambios fonéticos en la duración de sus vocales, lo cual puede tomarse como una clave de su estatus como portadores de acento secundario. En la sección 8.6, presenté la palabra [ˈb^βa.na.ja.ma.kí] ‘(no) lo sembró’ y mostré que las vocales de la primera y tercera sílaba muestran una duración mayor a la duración de las vocales de las otras sílabas. Si estas diferencias duracionales

indicasen la presencia de acentos secundarios, entonces esta palabra tendría la siguiente estructura métrica: (¹b^βa.na)(ja.ma).kì.

Sin embargo, obsérvese que ya que las vocales tienden a mostrar diferentes duraciones intrínsecas (por ejemplo, la vocal [a] es intrínsecamente más larga que las otras vocales que ocurren en el mismo contexto, véase los capítulos 2 y 7), este método de encontrar acentos secundarios comparando la duración de las vocales en sílabas abiertas adyacentes solo puede usarse con palabras cuyas sílabas tienen las mismas vocales, como en la palabra (¹b^βa.na)(ja.ma).kì. Este método no puede aplicarse en una palabra como [¹tʃí.ʂà.jù.dzà.kì] ‘(le) dolió muchísimo’. La vocal [à] de la sílaba [dzà] es probablemente más larga que la vocal [ù] de [jù] pero no hay forma de saber si esto se debe al acento o a la duración intrínseca de la vocal [a]. Para los casos como [¹tʃí.ʂà.jù.dzà.kì], asumo que las sílabas se agrupan en pies bisilábicos y que por defecto, ya que en estos casos ambas sílabas son abiertas, la primera ocupa una posición prominente en el estructura prosódica: (¹tʃí.ʂà)(jù.dzà).kì. La sílaba [jù] no recibe un tono alto ya que el shipibo solo asocia el tono alto con sílabas que llevan el acento primario y con sílabas cerradas que llevan acento secundario.

Cierro esta sección comparando el comportamiento del tono alto de sílabas acentuadas del shipibo con el comportamiento que encontramos en castellano. Colocando de lado los casos de sílabas abiertas fuera del pie métrico principal, uno podría tener la impresión a primera vista que no es sorprendente que las sílabas acentuadas del shipibo se alineen con un tono alto. Sin embargo, este alineamiento no es necesariamente una propiedad de las sílabas acentuadas. Por ejemplo, en castellano, las sílabas acentuadas usualmente se asocian con un tono bajo y el tono alto ocurre en la sílaba postónica, excepto si la palabra ocurre al final de una frase o tiene prominencia discursiva (Navarro Tomás 1957; Sosa 1991; Prieto, van Santen *et al.* 1995; Face 2001; Hualde 2003, entre otros). De este modo, por ejemplo, en una frase como [¹dà.mé un.men.¹sá.xè] ‘dame un mensaje!’ (véase el contorno tonal en Figura 9-21), la primera palabra, [¹dà.mé], tiene una sílaba acentuada que lleva un tono bajo mientras que la siguiente sílaba inacentuada aparece con un tono alto. En la segunda palabra, ya que ocurre al final de la frase, la sílaba acentuada lleva un tono alto mientras que la siguiente sílaba lleva un tono bajo. A pesar de que se les asigne un tono bajo a las sílabas acentuadas, los hablantes de castellano pueden identificar el estatus acentuado de las diferentes sílabas por la duración de las vocales.

Figura 9-21: Un ejemplo del alineamiento tonal en las sílabas acentuadas del castellano



9.3.4 Casos de acento léxico

El shipibo tiene casos de acento léxico, Son pocos comparados a los casos donde la posición del acento es robustamente predecible. La mayoría de los casos de acento léxico se restringen a préstamos y algunos sufijos. El capítulo 8 mostró esto al examinar el comportamiento impredecible del tono perceptual en préstamos como [ˈis.kũn] ‘nueve’ (sección 8.1.1) y sufijos como el imperativo /-ˈwĩ/ (sección 8.7). En esta sección, me gustaría entretener la idea que los casos de acento léxico del shipibo corresponden a situaciones donde la asociación de un tono alto a una vocal dada Shipibo ha sido codificada como información subyacente.

Este enfoque contrasta con uno en que el acento léxico se entiende como un diacrítico (es decir, [ˈ]) que requiere que ciertas sílabas ocupen una posición fuerte dentro de la estructura prosódica de la palabra. En la práctica, el uso de este diacrítico equivale a decir que la lengua guarda la estructura prosódica en sus representaciones subyacentes (al menos, la estructura de la sílaba marcada para ocupar una posición prosódica nuclear). En contraste, al afirmar que los casos de acento léxico son asociaciones subyacentes de tonos altos a vocales, el análisis evita asumir que la estructura prosódica es fonémica. Además, solo dice que ciertas vocales tienen un tono alto asociado con ellas y pone el destino de esa asociación en las manos de los principios generales que gobiernan la asignación del acento en la lengua.

Antes de continuar identifiquemos las fuerzas fonológicas que manejan la asociación de los tonos altos a las sílabas acentuadas. El capítulo 8 mostró que las sílabas con acento primario siempre se asocian con un tono alto. Fuera del pío métrico principal, las sílabas cerradas pueden mostrar, dependiendo del tipo de habla, un tono alto. En el caso de las sílabas abiertas, muestran un tono alto

obligatorio si llevan el acento primario. Más allá del pie métrico principal, las sílabas abiertas podrían mostrar diferencias duracionales que revelan si portan algún acento secundario o no pero no ocurren asociadas con un tono alto (al menos que tengan acento léxico). Llamemos al requisito que asocia las sílabas con acento primario a un tono alto: *alineamiento de tono alto al acento primario*. La atracción del tono alto a sílabas cerradas se debe a la bimoracidad de estas sílabas. Llamemos a ese requisito: *alineamiento de tono alto a sílaba bimoraica*. Los casos de acento léxico, entendidos como la asociación impredecible de un tono alto a ciertas vocales indica que para el shipibo es importante preservar esas asociaciones subyacentes en las representaciones fonéticas, es decir, esas asociaciones no pueden ser elididas o reasociadas con otras sílabas. Llamemos a ese requisito: *preservación del tono alto*.

En el caso de una palabra como /b^βana -ki/ → ('b^βa_μ.na_μ).kɨ_μ 'lo sembró' (CMPL), los requisitos métricos que gobiernan la posición del acento, discutida en la sección 9.3.2, determinarán que las dos sílabas iniciales de la palabra forman el pie métrico principal, y asignarán a esa sílaba inicial a una posición prominente, es decir, la portadora del acento primario. El *alineamiento de tono alto al acento primario* eliminará la posibilidad que la sílaba con el acento primario ocurra con un tono bajo o como sucede en castellano (véase la Figura 9-21), que el tono alto aparezca en una sílaba postónica: *('b^βà_μ.na_μ).kɨ_μ, *('b^βà_μ.ná_μ).kɨ_μ. Este requisito métrico asegurará que la sílaba con el acento primario aparezca con un tono alto: ('b^βá_μ.na_μ).kɨ_μ.

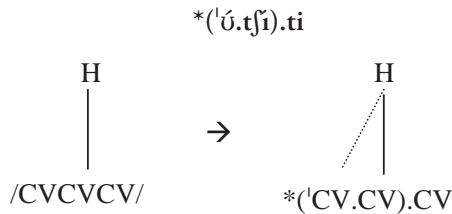
En una palabra como /witaʂ-(a)n/ → ('wi_μ.ta_μ)(,ʂan_μ) 'pierna' (FL), la sílaba cerrada será bimoraica para satisfacer WBP y como tal, se le permite formar su propio pie métrico para satisfacer WSP. El requisito *Alineamiento de Tono Alto a Sílaba Bimoraica* se encargará de asociar un tono alto a esa sílaba en la representación fonética de la palabra: ('wí_μ.ta_μ)(,ʂán_μ). Cuando la palabra se pronuncie, el tono alto de ese acento secundario podría realizarse, aunque alternativamente podría dejarse sin realizar en el habla rápida. En contraste, en un palabra como /b^βana -jama -ki/ → ('b^βa_μ.na_μ)(,ja_μ.ma_μ).kɨ_μ 'no (lo) sembró' (NEG, CMPL), la sílaba que lleva el acento secundario es abierta y por lo tanto, monomoraica. En este contexto, el requisito *alineamiento de tono alto a sílaba bimoraica* no tiene la fuerza para asegurar una asociación con un tono alto: ('b^βá_μ.nà_μ)(,jà_μ.mà_μ).kɨ_μ.

En el caso de palabras con acento léxico como /utʃíti/ 'perro' (originalmente un préstamo del ashéninka-asháninka [o.'tsi.tsi] ~ [o.'tsi.ti]), la representación subyacente tiene un tono alto. Como es usual, el shipibo agrupa las dos sílabas iniciales en el pie métrico principal. A pesar que la segunda sílaba es abierta, hay varias razones por las cuales el shipibo no puede designar la sílaba inicial de *['u. tʃí.ti] 'perro' como el portador del acento primario cuando hay una tono alto subyacente en la segunda sílaba. Examinemos estas razones.

Primero, el requisito *alineamiento de tono alto al acento primario* descarta la forma *('ù.tʃí).ti. Obsérvese que aunque se preserva la asociación subyacente del tono alto a la segunda sílaba, la sílaba con el acento primario tiene un tono bajo. Además, para poder satisfacer el requisito *preservación del tono alto*, el tono alto subyacente no puede ser elidido o asociado a otra sílaba. Esto elimina la alternativa *('ú.tʃí).ti ya que el tono alto subyacente se ha desasociado de la segunda sílaba y se ha movido a la primera sílaba de modo que el acento primario aparece con un tono alto como lo requiere el *alineamiento de tono alto al acento primario*.

Tercero, la *preservación del tono alto* también prohíbe la propagación del tono alto subyacente ya que esto significaría que en la representación fonética, el tono alto se asocia con vocales a las cuales no está asociado en el nivel subyacente. La propagación del tono alto se representa esquemáticamente en la Figura 9-22. La H representa un tono alto.

Figura 9-22: Prohibición de la propagación de los tonos altos subyacentes:



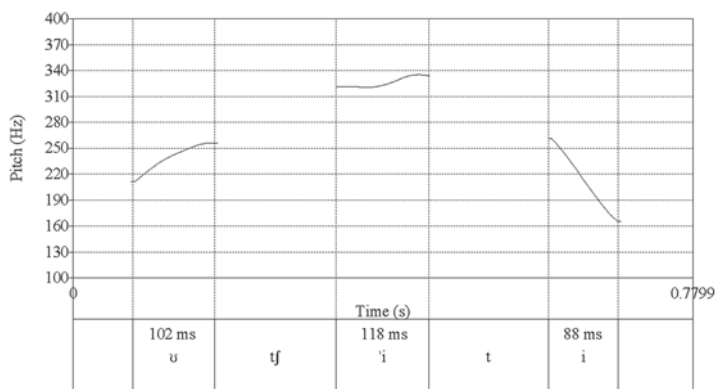
Una alternativa que satisface el requisito *preservación del tono alto*, coloca el acento en la sílaba inicial y además tiene la sílaba con el acento primario asociado con un tono alto es insertar un nuevo tono alto como se muestra esquemáticamente en la Figura 9-23. Sin embargo, esto coloca dos tonos altos adyacentes uno al otro, *('ú.tʃí).ti, una configuración que se sabe muchas lenguas tratan de evitar. Este fenómeno se conoce como el *principio del contorno obligatorio* (OCP - Leben 1973; McCarthy 1986; Odden 1986, 1988; Yip 1988).

Figura 9-23: OCP prohíbe tonos altos adyacentes: *('ú.tʃí).ti



La única alternativa que le queda al shipibo es entonces mover el acento primario a la segunda sílaba de la palabra: (ù.'tʃi).tì. Esto satisface el requisito de que el portador del acento primario esté asociado con un tono alto. Ya que el tono alto subyacente no se elide o comparte con otras sílabas y no aparece adyacente a otros tonos altos, la forma satisface los requisitos *preservación del tono alto* y *OCP*. El costo de satisfacerlos es desobedecer el requisito que quiere que el acento primario aparezca por defecto alineado con el lado izquierdo del pie métrico principal. Un ejemplo de la realización del contorno tonal de la palabra shipiba (ù.'tʃi).tì ‘perro’, la cual tiene un tono alto subyacente, se presenta en la Figura 9-24.

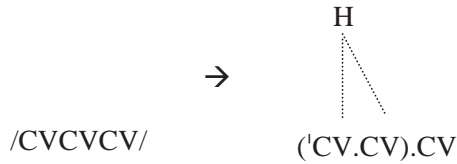
Figura 9-24: Contorno tonal de la palabra (ù.'tʃi).tì ‘perro’



Con respecto a la prohibición de propagar un tono alto (véase la Figura 9-22) y OCP en la Figura 9-23, uno podría preguntarse si este análisis es correcto. Después de todo, la sección 8.4 mostró que en shipibo, las hablantes mujeres, a diferencia de los hablantes hombres, tienden a colocar un tono alto tanto sobre la sílaba inicial acentuada como también sobre la sílaba inacentuada que sigue. Esto se muestra esquemáticamente en la Figura 9-25. Sin embargo, estos casos no involucran una transgresión de OCP, sino más bien la propagación hacia la derecha de un tono alto dentro del dominio del pie métrico principal. A diferencia de la Figura 9-22, esta propagación no desobedece el requisito de *preservación del tono alto* ya que no hay ningún tono alto subyacente.

Además de evitar la necesidad de asumir estructura prosódica subyacente, este análisis tiene la ventaja de explicar el comportamiento diferente que se observa entre las sílabas abiertas con y sin un tono alto subyacente. Fuera del pie métrico principal, las sílabas abiertas que no tienen un tono alto subyacente no reciben

Figura 9-25: /b^βana -ki/ → (l^βá_μ.ná_μ).kì_μ ‘(lo) sembró’
(CMPL) - hablantes mujeres



un tono alto, ni siquiera si ocupan una posición acentuada en la estructura prosódica: /b^βana -jama -ki/ → (l^βá_μ.ná_μ)(jâ_μ.mâ_μ).kì_μ ‘(lo) sembró’ (NEG, CMPL) (aunque, como se vio en la sección 8.7, sí podrían mostrar efectos duracionales).

En contraste, las sílabas abiertas que contienen vocales que tienen un tono alto subyacente sí tienen la opción de aparecer asociadas con un tono alto fuera de la ventana del pie métrico principal: /b^βana -ní -ki/ → (l^βá_μ.ná_μ)(ní_μ.kì_μ) ‘(lo) había sembrado ya hace bastante tiempo’ (REM1, CMPL). En contraste, en un análisis donde los casos de acento léxico se analizan como estructuras prosódicas guardadas subyacentemente, queda como un misterio por qué una palabra como, (l^βá_μ.ná_μ)(ní_μ.kì_μ), puede mostrar un tono alto en la sílaba abierta que lleva el acento secundario pero la sílaba abierta que ocupa una posición de portadora de acento secundario en la palabra, (l^βá_μ.ná_μ)(jâ_μ.mâ_μ).kì_μ, no puede ocurrir asociada con un tono alto. Después de todo, ambas sílabas abiertas ocupan la misma posición prominente y ya que bajo este análisis alternativo, no hay ningún tono subyacente, predice erróneamente que ambos tipos de sílabas abiertas deberían comportarse de modo similar.

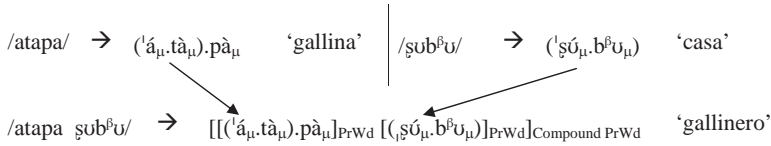
9.3.5 Choques acentuales y la preservación del tono alto del acento primario en compuestos

El patrón acentual que muestran los compuestos en el shipibo no se sigue de concatenar dos o más palabras y agrupar sus sílabas en pies métricos. Si ese fuese el caso, la distribución de los acentos en un compuesto como (l^á_μ.tâ_μ).pâ_μ.(sù. b^βù_μ), el cual está formado de las palabras (l^á.ta).pa ‘gallina’ y (l^{sù}.b^βù) ‘casa’, sería: *(l^á_μ.tâ_μ)(pâ_μ.sù_μ).b^βù_μ, como sucede con la palabra (l^βá_μ.ná_μ)(jâ_μ.mâ_μ).kì_μ. La sección 8.8 caracterizó el comportamiento del F0 en compuestos y encontró que la distribución de los acentos en este tipo de palabras se hereda del patrón acentual de cada una de las palabras que forma el compuesto.

Otra característica de los compuestos es que las sílabas que llevarían el acento primario, si ocurriesen como palabras independientes, no solo ocupan

una posición acentuada sino que también preservan su tono alto. En el caso del compuesto (¹á_μ.tà_μ).pà_μ.(₁šú_μ.b^βü_μ) ‘gallinero’, mostrado en la Figura 9-26, la sílaba [₁šú] puede aparecer asociada con un tono alto porque la palabra (¹šú_μ.b^βü_μ) ‘casa’, como una palabra independiente, tiene un tono alto sobre su sílaba acentuada inicial. La habilidad de esta sílaba abierta de mostrar un tono alto contrasta con el comportamiento de las sílabas abiertas fuera del pie métrico principal en palabras independientes, las cuales no llevan un tono alto al menos que el tono sea subyacente. Obsérvese que la *Preservación del Tono Alto en Compuestos* es similar a cómo las vocales largas se heredan y preservan a través de las palabras en un paradigma (véase la sección 9.1.2.1).

Figura 9-26: Preservación de tonos altos asociados con el acento primario a través de palabras



Analizo los compuestos del shipibo como palabras prosódicas independientes que se agrupan para formar un constituyente prosódico que llamaré *Palabra Prosódica de Compuestos*. Dentro de cada palabra prosódica interna al compuesto, el patrón acentual trata de ser tan similar como le sea posible al patrón acentual que se encuentra cuando la palabra ocurre independiente. Sin embargo, ya que el shipibo evita la creación de choques acentúales, los patrones acentúales dentro del compuesto pueden sufrir ajustes. La lengua reposiciona o como último recurso, elide uno de los acentos que podría crear el choque acentual.

Considérese, por ejemplo, la sílaba cerrada del compuesto (a.¹tsan)(ti.₁ta) (tipo de yuca), que contiene las palabras (a.¹tsan) ‘yuca’ (FL) y (¹ti.ta) ‘madre’. La sílaba cerrada está dentro del pie métrico principal y como consecuencia es suficientemente fuerte como para forzar la desacentuación de la siguiente sílaba como la única forma de resolver el choque acentual. Una sílaba cerrada que pertenece al pie métrico principal nunca pierde su derecho a llevar el acento primario de la palabra cuando compite con una sílaba que está fuera de ese pie métrico. De este modo, la forma *[¹a.tsan.₁ti.ta], con el acento principal movido a la sílaba inicial, nunca es una alternativa en el shipibo para resolver un choque acentual.

En comparación, consideremos la sílaba cerrada del compuesto (¹a.wa)(₁kan) (pi.₁ti) ~ (¹a.wa).kan.(₁pi.ti) (esp. de liana). Este compuesto se estructura de la

palabra (¹a.wa)_(,)kan ‘tapir’ (FL) y (¹pi.ti) ‘comida’. La sílaba cerrada no está protegida por el pie métrico principal y por lo tanto, no puede forzar las desacentuación obligatoria de la siguiente sílaba para evitar el choque acentual. De este modo, esa sílaba cerrada algunas veces queda desacentuada, (¹a.wa).kan._(,)(pi.ti), algo que nunca ocurriría con una sílaba cerrada que pertenece al pie métrico principal. Otras veces, sin embargo, logra reclamar su estatus como portador de acento secundario, (¹a.wa)_(,)kan(pi._(,)ti). Cuando esto sucede, el acento de (¹pi.ti) se disloca a la siguiente sílaba para evitar la creación de un choque acentual: *(¹a.wa)_(,)kan)_(,)(pi.ti). De este modo, los requisitos de tener una sílaba cerrada bimoraica y el evitar los choques acentuales son las únicas fuerzas que juntas pueden opcionalmente predominar sobre el requisito de *preservación del tono alto en compuestos*.

Los únicos casos en que el shipibo recurre a la elisión del tono alto de una sílaba en un compuesto es cuando la reposición del acento crea otro choque acentual. Este caso puede verse en el compuesto (a.¹wa).ti._(,)(pu._(,)man) (un tipo de planta - FL), que contiene las palabras (a.¹wa) ‘tapir’ y (¹ti.pu)_(,)(man) ‘tendón (del cuello)’ (FL). Si la sílaba [ti] conservase su acento en el compuesto, entonces crearía un choque acentual: *(a.¹wa)_(,)(ti.pu)_(,)(man). Si el acento se moviese una sílaba hacia adelante, también crearía un choque acentual: *(a.¹wa)(ti._(,)pu)_(,)(man).

9.3.6 El género y la adquisición de la distribución de los tonos altos

El último aspecto que me gustaría discutir sobre la fonología del shipibo es el hallazgo reportado en la sección 8.4 concerniente a la diferente distribución del tono alto entre hablantes hombres y mujeres. En palabras con el acento sobre la sílaba inicial, las hablantes mujeres tienden a asociar un tono alto tanto con la sílaba acentuada y con la sílaba postónica mientras que en las mismas palabras, los hablantes hombres solo asocian un tono alto con la sílaba acentuada y un tono bajo con la sílaba postónica.

Las diferencias basadas en género que se encuentran en la distribución del tono alto del shipibo no quieren decir que hombres y mujeres tengan diferentes fonologías. Más bien, el problema debe entenderse como un problema de uso del conocimiento lingüístico: los hablantes hombres y mujeres deben adquirir el mismo conocimiento lingüístico que una sílaba inacentuada segunda puede tener un tono bajo-descendiente o un tono alto-nivelado. Después de todo, los hablantes hombres y mujeres deben lograr comunicarse y entenderse mutuamente. Los hablantes hombres deben aprender a esperar un tono alto-nivelado sobre una sílaba inacentuada segunda cuando hablan con mujeres y las

hablantes mujeres deben aprender a no esperar ese patrón cuando hablan con un hablante hombre.¹⁸

Las diferencias lingüísticas basadas en género no son desconocidas en la literatura. Por ejemplo, Trudgill 1974 reporta que en Gros Ventre, una lengua amerindia que pertenece a la familia algonquina, los hablantes hombres y mujeres difieren en el punto de articulación de sus consonantes palatalizadas: los hombres usan oclusivas dentales palatalizadas (por ejemplo, [djatsa] ‘pan’) mientras que las mujeres usan oclusivas velares palatalizadas (por ejemplo, [kjatsa] ‘pan’). Otro ejemplo de una diferencia lingüística basada en género viene del castellano argentino (Fontanella de Weinberg 1978, 1992; Lipski 1994; Chang 2008). Esta variedad de castellano tiene un fenómeno de ensordecimiento que fue inicialmente asociado con el habla de mujeres educadas de Buenos Aires. Este fenómeno vuelve la fricativa sonora [ʒ] en la fricativa sorda [ʃ]. Hoy en día, sin embargo, el fenómeno se ha propagado igualmente entre hombres y mujeres (Chang 2008). En el caso del shipibo, la diferencia de género es de naturaleza supra-segmental (es decir, afecta el tono asociado con las sílabas segundas inacentuadas) en lugar de rasgos segmentales como el punto de articulación o rasgos laríngeos como la sonoridad.



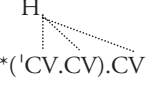

Examinemos, desde una perspectiva fonológica, el caso del tono alto que se propaga a la sílaba postónica. La propagación del tono alto a la siguiente sílaba inacentuada en palabras que tienen acento primario inicial es interesante porque nos hace preguntarnos por qué solo se propaga hasta la segunda sílaba y no más allá de ese punto y también por qué en casos donde el acento primario cae sobre la segunda sílaba, el tono alto no se propaga a la sílaba inicial inacentuada. Estas configuraciones se ilustran esquemáticamente en la Figura 9-27.

Analizó la restricción en la propagación del tono alto más allá de la segunda sílaba de una palabra como un reflejo del dominio en el cual ocurre este fenómeno: el pie métrico principal. De este modo, los hablantes shipibos, tanto hombres como mujeres, saben que su fonología puede asociar un tono alto con las sílabas del pie métrico que contiene el acento primario o restringir el tono alto solo a la sílaba designada como la portadora del acento primario. Las hablantes mujeres prefieren usar la primera alternativa mientras que los hablantes hombres prefieren la segunda. Ambos, hombres y mujeres, saben que la propagación no puede ir más allá del dominio del pie métrico principal, lo cual descarta la estructura en

¹⁸ Una pregunta que quedará abierta y fuera del foco de este libro es cuándo los niños que aprenden shipibo se dan cuenta que la sociedad les impone distinciones relacionadas a qué tono deben usar en las sílabas inacentuadas segundas. Es decir, cuándo un niño, por ejemplo, se da cuenta que usar un tono alto-nivelado sobre una sílaba inacentuada segunda es una característica del habla femenina y que la sociedad pide que él use en ese contexto un tono bajo-descendente.

(c) de la Figura 9-27. Sin embargo, el pie métrico principal como dominio de propagación del tono alto no es el único factor que está involucrado. La direccionalidad de la propagación también juega un rol importante. El shipibo permite que la propagación del tono alto se dé hacia la derecha pero no hacia la izquierda. Éste es un tipo bastante común de propagación tonal (véase Flemming 2008, Akinlabi y Liberman 2001). Esto elimina la estructura en la estructura en (d) de la Figura 9-27. De este modo, mientras el tono alto de la sílaba con acento

Figura 9-27: Patrones de los tonos altos dentro de pie métrico principal

a.	En el habla de mujeres (/b ^β ana -ki/ → ('b ^β á _μ .ná _μ).kà _μ 'lo sembró' (CMPL)).	
b.	En el habla de hombres y mujeres (/matas -ki/ → (mà _μ .tá _μ s _μ).kà _μ 'lo golpeó en la cabeza' (CMPL)).	
c.	Esta configuración no ocurre en shipibo.	
d.	Esta configuración no ocurre en shipibo.	

primario puede ser compartido con la siguiente sílaba inacentuada, el tono alto de una sílaba acentuada segunda no puede compartirse con la sílaba anterior.

Es importante mencionar que interlingüísticamente lenguas como bambara, una lengua manding de Mali (Rialland y Badjimé 1989; Weidman y Rose 2006), se comporta como el shipibo en el sentido que su fonología puede referirse al pie métrico como el dominio fonológico en el cual se pueden propagar los tonos. Sin embargo, a diferencia del shipibo, en bambara, la propagación de los tonos no se limita al pie métrico principal. Palabras del bambara como /^HL kulukutu/ 'una pelota' o /^{LH} garijɛgɛ/ 'una oportunidad' ocurren como (kú.lú)(kù.tù) y (gà.rì)(jé.gé), respectivamente. Obsérvese que el cada pie métrico muestra un tono desplegado sobre sus dos sílabas. De este modo, el tono alto de /^HL kulukutu/ se propaga sobre las dos primeras sílabas iniciales (las cuales corresponden al pie métrico inicial) y el tono bajo se extiende sobre las siguientes dos sílabas (las cuales corresponden al segundo pie métrico de la palabra).

9.4 RESUMEN

Informado por los resultados de los análisis acústicos, este capítulo ha identificado, además de la composición de rasgos de las vocales y consonantes, los fenómenos fonológicos más característicos del shipibo. He bosquejado sus propiedades y determinado las fuerzas fonológicas que están en juego. Éstas no son exclusivas al shipibo sino que están motivadas interlingüísticamente como ha demostrado la discusión de varias lenguas que no están relacionadas al shipibo. En el caso de las vocales, las posiciones prominentes en la estructura prosódica, en particular, las sílabas acentuadas y las sílabas iniciales, han sido cruciales en entender las correlaciones que encontraron los análisis acústicos del shipibo entre duración y centralización vocálica y la posición que ocupan las vocales dentro de las palabras. Las posiciones prominentes protegen las vocales de los fenómenos de debilitamiento. Esto explica por qué las vocales en posiciones no-prominentes son más vulnerables a sufrir acortamiento o moverse hacia el centro del espacio vocálico. Pero las posiciones prominentes del shipibo van más allá de dar claves fonéticas asimétricas en el comportamiento de las vocales sobre las cuales los aprendices de shipibo puedan recrear la estructura prosódica de su lengua. Las posiciones prominentes también ofrecen santuario a segmentos que no se les permite ocurrir en otros sitios. De este modo, por ejemplo, las sílabas iniciales del shipibo son los únicos lugares que pueden albergar vocales largas, segmentos laríngeos e incluso estructuras silábicas con codas complejas.

Al examinar las mediciones de la duración de las vocales largas, este capítulo pudo identificar claramente los roles importantes que juegan en la fonología del shipibo los requisitos del tamaño mínimo y de la preservación de la longitud vocálica. En particular, se ha argumentado en el capítulo que en el shipibo la búsqueda de al menos un tamaño mínimo bimoraico en las palabras es obligatorio mientras que al mismo tiempo, la lengua aspira a que sus raíces sean mínimamente bisilábicas. Se ha sugerido que la insatisfacción de la bisilabicidad en palabras monosilábicas es la razón de por qué el shipibo alarga fonéticamente sus vocales largas en este tipo de palabras. Este alargamiento fonético se detiene una vez que la palabra es mínimamente bisilábica.

Tal vez uno de los fenómenos más interesantes relacionados a las vocales largas es su distribución bastante restringida y ‘casi’ predecible. Siempre ocurren acentuadas y en sílabas iniciales (es decir, en posiciones prominentes). Son obligatorias en raíces monosilábicas que tienen la habilidad de ocurrir sin sufijos. Además, pueden aparecer en las sílabas iniciales de algunas pocas raíces polisilábicas. El requisito de la preservación de la longitud vocálica garantiza que una vez que una palabra ha obtenido una vocal larga, todas las formas de ese paradigma deben heredar

la vocal larga. Este patrón contrasta marcadamente con lo que se encuentra en otras lenguas en las cuales las vocales largas aparecen en todo el léxico, no solo en monosílabos o sílabas iniciales. En este capítulo, he bosquejado un análisis derivacional y otro no-derivacional para explicar la distribución de las vocales largas.

Con respecto a las consonantes, además de discutir su composición de rasgos y presentar evidencia para el estatus subyacente de las glides [j] y [w], me he concentrado en segmentos latentes y nasales. En el capítulo, he demostrado que hay evidencia fonética y fonológica que indica la existencia de segmentos latentes en el shipibo. La evidencia fonológica se basa en el comportamiento del acento y la aparición de los segmentos latentes cuando logran silabificarse como arranques de sílaba. La evidencia fonética que se encontró en los capítulos 2 y 7 da apoyo adicional a la evidencia fonológica sobre la existencia de segmentos latentes. La evidencia fonética viene de los patrones duracionales encontrados en las vocales acentuadas. En los contextos donde una coda latente sigue a una vocal acentuada, esa vocal muestra el mismo comportamiento duracional que una vocal en una sílaba cerrada acentuada. Además, en este capítulo se ha examinado una característica fascinante de los segmentos latentes: la neutralización de sus puntos de articulación. He argumentado que con la excepción de los casos de las consonantes latentes en raíces nominales y adjetivales, en las cuales los segmentos latentes se han neutralizado al punto de articulación dorsal, coronal es por defecto el punto de articulación preferido de los segmentos latentes en shipibo.

He extendido el análisis del punto de articulación propuesto para los segmentos latentes a la alternancia que muestran las consonantes nasales cuando ocurren como segmentos finales de las raíces. Como en el caso de los segmentos latentes, el punto de articulación por defecto de las nasales finales de morfemas es coronal. Sin embargo, en raíces nominales y adjetivales, el shipibo las obliga a evitar el rasgo [coronal]. El shipibo satisface este requisito escogiendo el punto de articulación [dorsal] en el caso de los segmentos latentes oclusivos, pero ya que la lengua no posee un segmento nasal dorsal, las nasales finales de raíces nominales y adjetivales toman la especificación [labial], la cual es el otro modo de evitar ser [coronal]. Además, basado en las medidas de duración y la evidencia fonológica (el comportamiento del acento y la resilabificación), se ha mostrado que las codas nasales pueden coalescer completamente con las vocales precedentes pero desde la perspectiva fonológica la lengua sigue considerando esa fusión como una secuencia de una vocal seguida de una coda nasal.

Con respecto a la sílaba y al acento, el capítulo ha presentado las principales propiedades estructurales de la sílaba del shipibo y ha discutido cómo la sílaba puede ser reestructurada por la fonética por medio de convertir a las vocales altas en glides. Se ha identificado las fuerzas métricas que determinan la posición del

acento. La ventana del acento primario ha sido capturada en términos de un pie métrico bisilábico alineado con el lado izquierdo de la palabra prosódica. Dentro de este pie métrico, el acento se posiciona por defecto en la sílaba izquierda. Esta asignación por defecto del acento es anulada cuando la segunda sílaba es cerrada, bimoraica, y en consecuencia capaz de alejar el acento de la sílaba inicial. Al discutir la interacción de la sílaba y el acento, la característica más notable que se encontró fue la habilidad de las sílabas cerradas del shipibo para ajustar su peso moraico con el fin de satisfacer a las fuerzas que gobiernan las formas de los pies métricos.

Además, basados en los resultados acústicos obtenidos en el capítulo 8, en este capítulo se ha tomado la habilidad de las sílabas cerradas fuera del pie métrico principal de aparecer con un tono alto como evidencia de que llevan acentos secundarios. Se ha identificado las fuerzas métricas que alinean los tonos altos con las sílabas que llevan el acento primario y con las sílabas cerradas que ocurren más allá del pie métrico principal. Esto es significativo porque hasta ahora se ha asumido que el shipibo es una lengua que no tenía acentos secundarios. He bosquejado un análisis fonológico que además de explicar la asignación del acento en los patrones predecibles puede también lidiar con el acento léxico. En este enfoque, los tonos altos se guardan como propiedades subyacentes de ciertos morfemas.

El capítulo también ha examinado la asignación del acento en los compuestos, donde el patrón acentual se hereda de las palabras que forman el compuesto. El evitar la creación de choques acentuales juega un rol importante en estos casos, ya que ajusta la distribución de las sílabas acentuadas. Finalmente, he examinado dos problemas fonológicos concernientes a la propagación de los tonos altos en habla de las hablantes mujeres: el dominio y la direccionalidad de la propagación. Se identificó al pie métrico principal como el dominio en el cual opera el fenómeno y además se encontró que la lengua solo permite que la propagación de los tonos altos ocurra hacia la derecha.

10. CONCLUSIONES

El principal objetivo de este libro ha sido el caracterizar las propiedades acústicas de las vocales y consonantes del shipibo en diferentes contextos. Además, he discutido en detalle el comportamiento de la frecuencia fundamental (F0) dentro del dominio de la palabra y he presentado un bosquejo de la fonología de la lengua. Ahora la pregunta es: ¿qué hemos logrado con la obtención de todo este conocimiento?

Primero, hemos obtenido una documentación acústica completa del sistema de sonidos de una lengua amazónica, un tipo de conocimiento que usualmente solo está disponible para lenguas dominantes como el castellano o el inglés. Segundo, como parte de la tarea de documentar el sistema de sonidos de la lengua, hemos obtenido una idea más precisa de cómo el sistema de sonidos del shipibo es exactamente. Tomemos, como un ejemplo para aclarar este punto, la palabra shipiba *bari* ‘sol’.

Antes de este estudio, la transcripción ortográfica de orientación fonológica sugería que la palabra podría corresponder a la representación fonológica /βari/. Tradicionalmente se dice que la letra ‘b’ del Shipibo corresponde a un sonido similar a la consonante [v] pero en lugar de ser labiodental, es bilabial. Al mismo tiempo, las propiedades articulatorias del sonido representado por la letra ‘r’ han sido generalmente ignoradas. Algunas veces, se ha hecho la sugerencia de que es una consonante vibrante muy similar a la [r] del castellano en la palabra [toro] ‘toro’. Además, de acuerdo a las convenciones ortográficas, las cuales han sido hasta ahora lo más cercano a una descripción fonética/fonológica del inventario segmental de la lengua, se asume que una palabra que carece de tilde debe tener el acento primario localizado en la sílaba inicial. De este modo, uno debería esperar

que la palabra *bari* ‘sol’ corresponda fonéticamente a [ʼβa.ri]. Tradicionalmente no se tenía ninguna caracterización de la realización fonética del acento en shipibo.

Basado en los resultados del estudio acústico del shipibo presentado en los capítulos anteriores, ahora sabemos que /βari/ no es la representación fonológica correcta de la palabra shipiba para ‘sol’, sino que debe ser /b^βaɖzi/,¹ donde la consonante inicial es una africada bilabial sonora, /b^β/, y la intervocálica es una africada retrofleja sonora, /ɖz/. Éste es un hallazgo que ha permitido mejorar la caracterización y nuestro entendimiento del shipibo. Pero basados en el estudio acústico reportado en este estudio, podemos afirmar que sabemos mucho más que eso.

Este estudio ha revelado que el shipibo tiene una tendencia por la cual las africadas sonoras disminuyen la duración de sus fases de cerrazón cuando ocurren intervocálicamente y como parte de una sílaba inacentuada. Esto significa que las africadas sonoras /b^β/ y /ɖz/ tienden a realizarse como [β] y [z] cuando ocurren rodeadas de vocales. Por lo tanto, la palabra /b^βaɖzi/ ‘sol’ se realiza fonéticamente como [ʼb^βa.zi], aunque también es posible la forma alternativa [ʼb^βa.ɖzi].

Con respecto a la interacción entre segmentos y prosodia, ahora sabemos que las vocales inacentuadas no-iniciales tienden a sufrir varios grados de centralización vocálica. Por lo tanto, podemos ir en más detalle sobre la representación fonética del shipibo. Para la palabra /b^βaɖzi/ ‘sol’ que estoy usando como ejemplo, esto significa que la vocal /i/ se realiza más como una vocal laxa [ɪ] ya que ocurre inacentuada y en una sílaba no-inicial: [ʼb^βa.zɪ] ~ [ʼb^βa.ɖzɪ]. El estudio también encontró que la falta de acento no solo centraliza las vocales pero que también afecta la duración de las vocales. Las vocales inacentuadas se realizan con una duración más corta que las vocales acentuadas. De este modo, en una transcripción fonética mucho más angosta, podemos representar la palabra shipiba para ‘sol’ como [ʼb^βa.zɪ̃] ~ [ʼb^βa.ɖzɪ̃], donde el diacrítico [̃] indica una vocal con una duración más corta comparada a la duración de su contraparte acentuada.

También podemos ahora incluir el comportamiento del tono perceptual. Este estudio encontró un patrón interesante basado en el género de los hablantes de shipibo. Cuando una palabra que tiene el acento primario sobre la sílaba inicial, los hablantes hombres asocian un tono alto con la sílaba acentuada y un tono bajo con la sílaba postónica. Sin embargo, las hablantes mujeres tienden a asociar un tono alto tanto con la sílaba acentuada como con la postónica. De este modo, la palabra para ‘sol’ lucirá como [ʼb^βá.zɪ̃] ~ [ʼb^βá.ɖzɪ̃] si es pronunciada por un

¹ También sabemos que hay hablantes de shipibo para quienes la palabra para ‘sol’ se representa mejor en el nivel fonológico como /baɖzi/. Aquellos hablantes consistentemente carecen de la africada bilabial sonora y en su lugar muestran una oclusiva bilabial sonora, /b/.

hablante hombre pero será [ʰb^βá.zɪ̃] ~ [ʰb^βá.dzɪ̃] si es una hablante mujer. En el último caso, la centralización vocálica y las diferencias en la duración de las vocales juegan un rol importante en servir como claves fonéticas para identificar la sílaba inicial de palabra como la portadora del acento primario.

De este modo, no solo hemos obtenido una caracterización mucho más completa de cómo se realiza en el mundo físico el sistema de sonidos del shipibo, sino que también hemos logrado una caracterización que va más allá de las transcripciones basadas en el AFI (Alfabeto Fonético Internacional). La caracterización acústica brinda substancia a los futuros estudios fonológicos de esta lengua y una base más sólida para poder hacer análisis comparativos con otras lenguas.

En realidad, la documentación acústica del shipibo que ofrece este libro nos permite hacer comparaciones significativas y más precisas con las propiedades acústicas encontradas en otras lenguas. Al inicio del libro mencioné el caso de las vocales del italiano y del yoruba. Ambas lenguas poseen el mismo inventario vocálico, /i, e, ε, a, ə, o, u/, pero aún así los timbres precisos de sus vocales /e/ y /o/ muestran diferencias consistentes. En yoruba, estas vocales están mucho más cerca en el espacio vocálico a las vocales /i/ y /u/ que a las vocales /ε/ y /ɔ/. En contraste, todas las siete vocales del italiano aparecen igualmente separadas. Este fenómeno sugiere, como muchos fonetistas argumentan, que cuando adquirimos un sistema de sonidos, además de la información categórica/fonológica, también guardamos información fonética. En otras palabras, los hablantes de italiano y yoruba no solo saben que sus lenguas tienen como parte de sus inventarios vocálicos las categorías fonológicas /e/ y /o/, también saben que en su lengua ambas vocales buscan alcanzar un valor ideal para su F1, el formante que se correlaciona con la altura vocálica. Esos valores ideales pueden diferir de lengua a lengua.

Casos como este tienen un impacto profundo en la documentación de lenguas. Las propiedades fonéticas de los sonidos no deben verse más como propiedades que son universales, y en consecuencia, asumidas que son las mismas para todas las lenguas que se dice comparten un conjunto de segmentos. A fin de entender y documentar las características únicas de cada lengua, los estudios acústicos son un componente necesario de la documentación de lenguas.

Retornemos al shipibo y hagamos una simple comparación interlingüística de algunas de sus propiedades acústicas. Como una muestra, me concentraré en la altura de las vocales /i/ y /u/ del shipibo. Este estudio determinó que los hablantes hombres colocan como objetivo centrar la altura de sus vocales /i/ alrededor de 326 Hz. ¿Cómo este valor de F1 para /i/ se compara con otras lenguas? Para el castellano hablado en Madrid, Bradlow 1995 determinó que el F1 de la vocal /i/ en hablantes hombres es 286 Hz. En un estudio diferente, Pérez Silva, Acurio Palma *et al.* 2008 encontraron un valor similar para el F1 de la vocal /i/ del

castellano hablado en Lima: 275 Hz (hombres). Olive, Greenwood *et al.* 1993 encontró que el valor ideal de F1 para la vocal /i/ en hablantes hombres de inglés americano (Pittsburg, Pensilvania) es 280 Hz, un valor bastante similar al que se encontró para el castellano.

Además del shipibo, incluyamos en el comparación dos lenguas más que no tengan origen europeo: el bora y el quechua del Cusco. Ambas son lenguas indígenas habladas en el Perú pero son lingüística y geográficamente inconexas una con otra y con el shipibo. El bora es una lengua amazónica que pertenece a la familia lingüística witoto. Tiene seis vocales fonémicas: /i, e, a, o, ɨ, ʉ/. En un estudio acústico de las vocales del bora, Parker 2001 determinó que el F1 ideal de la vocal /i/ en esta lengua para hablantes hombres es 308 Hz.

En contraste al inglés americano que tiene doce vocales (/i, ɪ, e, ε, æ, ɑ, ɔ, o, ʊ, u, ʌ, ɚ/), al castellano que tiene cinco vocales (/i, e, a, o, u/), al bora que tiene seis vocales (/i, e, a, o, ɨ, ʉ/), y al shipibo que tiene cuatro (/i, a, ʊ, ɨ/); el quechua del Cusco, una lengua andina que pertenece a la familia lingüística quechua, solo tiene tres vocales, tradicionalmente representadas por los símbolos fonéticos: /i, a, u/. Pérez Silva, Acurio Palma *et al.* 2008 encontraron que la altura de la vocal /i/ en quechua del Cusco es mucho más baja que la altura de la misma vocal en castellano de Lima (variedad no-andina). En promedio, el valor ideal del F1 de la vocal /i/ en quechua para hablantes hombres es 365 Hz.² Compárese este valor al promedio 280 Hz encontrado para el inglés y el castellano (peruano y peninsular), o a los 308 Hz que se encontró para el bora o a los 326 Hz para el shipibo.

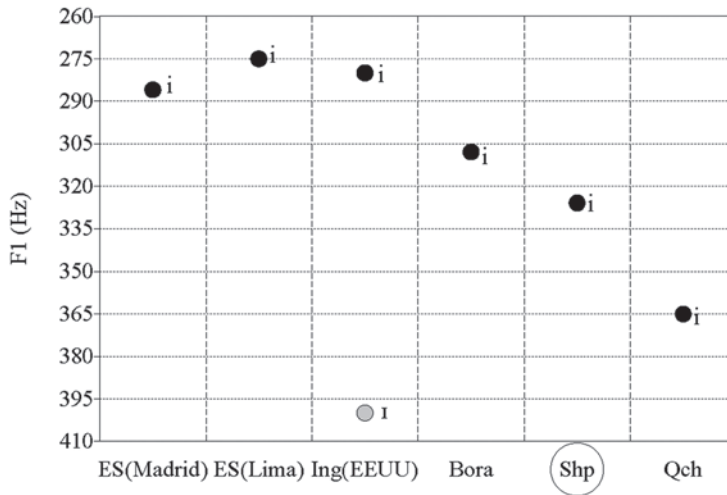
La comparación de los diferentes valores de F1 para la vocal /i/ en castellano (Madrid y Lima), inglés americano, bora, shipibo y quechua cusqueño se muestra en la Figura 10-1. El gráfico nos permite ver rápidamente que el castellano y el inglés americano tienen una altura muy similar para la vocal /i/ y que la altura más baja para esta vocal la posee el quechua. La altura ideal de la vocal /i/, expresada por el valor de su F1, en shipibo y en bora se posiciona a mitad de camino entre el castellano/inglés y el quechua.

En la Figura 10-1, he incluido la posición de la vocal /i/ del inglés americano (400 Hz, as reported by Olive, Greenwood *et al.* 1993). Obsérvese que aunque la vocal alta frontal del quechua tradicionalmente se describe y representa por el

² Las vocales /i/ y /u/ en Quechua sufren un proceso fonológico por el cual tienden a realizarse como [ɛ] y [ɔ] respectivamente cuando están en contacto con la consonante uvular /q/, un fenómeno que no se encuentra en las otras lenguas usadas en la comparación. El promedio de los valores de F1 de las vocales /i/ y /u/ que estoy reportando del trabajo de Pérez Silva, Acurio Palma and Bendejú Araujo (2008) excluyen dichos alófonos. Si esos alófonos fuesen considerados, el promedio de F1 para la vocal /i/ sería 410 Hz y para la vocal /u/ sería 486 Hz, lo cual colocaría a estas vocales en una posición mucho más baja en el espacio vocálico.

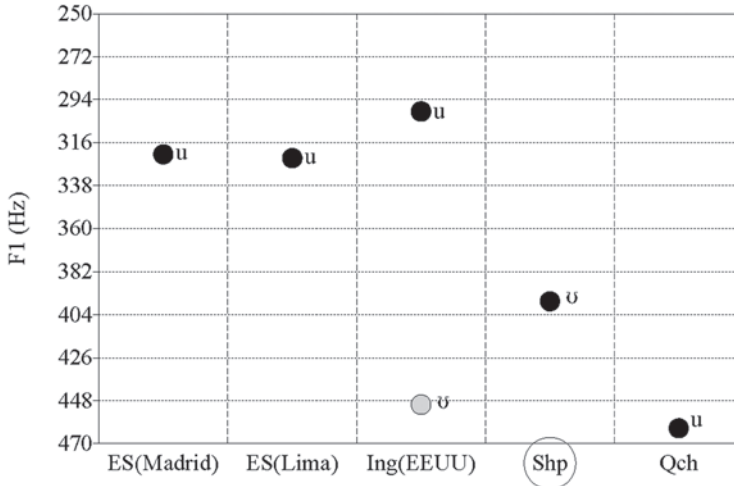
símbolo fonético /i/, su valor ideal de F1 indica que en realidad es bastante cercana a la vocal /i/. En el caso del shipibo, represento su vocal frontal con el símbolo /i/, en vez de /ɪ/, dada su mayor cercanía al F1 de la vocal /i/ del castellano. Sin embargo, ahora ya debe ser claro que más allá de las convenciones usadas para escoger un símbolo fonético para representar una vocal, no debe asumirse que las vocales ocupan la misma posición en el espacio vocálico de una lengua a otra aunque se les represente con el mismo símbolo.

Figura 10-1: Comparación interlingüística del F1 de la vocal /i/



El comportamiento ilustrado arriba no es una rara coincidencia, una propiedad extraña de comparar una vocal en diferentes lenguas. Podemos encontrar un patrón idéntico en estas lenguas comparando el valor de F1 de otra vocal como /u/ (voy a excluir al bora de esta comparación ya que no posee la vocal /u/). Esto se muestra en la Figura 10-2. Nuevamente, el castellano y el inglés muestran el mismo comportamiento, ambas lenguas tienen una vocal /u/ bastante alto. El quechua coloca su vocal /u/ en una posición mucho más baja, comparable a la vocal /u/ del inglés. La vocal posterior redondeada del shipibo aparece nuevamente a mitad de camino entre la /u/ del castellano/inglés y la /u/ del quechua, sin embargo, esta vez es mucho más cercana a la vocal /u/ del inglés. Para el shipibo, decidí representar esta vocal con el símbolo fonético /u/, en vez del símbolo /ʊ/, para resaltar la observación acústica y también la impresión perceptual que la vocal posterior redondeada del shipibo posee un timbre, en términos de altura vocálica, más bajo que la vocal /u/ del castellano.

Figura 10-2: Comparación interlingüística del F1 de la vocal /u/



Uno puede argumentar que ya que el quechua solo tiene tres vocales, la dispersión de cada vocal con respecto a la altura se espera, por supuesto, que sea mayor y en consecuencia, la posición promedio de cada vocal sería más baja en el espacio de vocales que en lenguas como el inglés e inclusive el castellano que tienen más vocales y por lo tanto cada vocal muestra una dispersión mucho menos. Aunque ésta es una tendencia observada en muchas lenguas, debemos recordar que lenguas como italiano y yoruba tienen el mismo inventario vocálico y aún así la posición y grado de dispersión de sus vocales siguen patrones diferentes.

Desde una perspectiva interlingüística, tanto la Figura 10-1 como la Figura 10-2 subrayan la importancia de documentar las propiedades acústicas de lenguas indígenas. Si no hubiese estudios acústicos de la estructura de formantes de las vocales en lenguas como quechua, bora y shipibo, podríamos pensar erróneamente, basado solo en los datos de lenguas como el castellano o el inglés, que los valores de los formantes de las vocales no muestra diferencias notables de una lengua a otra. Esta presuposición no solo ocurriría con respecto a los valores de los formantes de las vocales sino que también para cualquier otra propiedad acústica de los sonidos lingüísticos. Por ejemplo, en el capítulo 9, discutí la duración de las vocales del shipibo en comparación al húngaro (White y Mady 2008a, b). Mientras ambas lenguas poseen vocales largas, el shipibo alarga fonéticamente su duración en palabras monosilábicas pronunciadas dentro de una frase mientras en húngaro, en el mismo contexto, las vocales largas no se alargan fonéticamente. Sin un estudio acústico, solo sabríamos que ambas lenguas poseen vocales largas

y probablemente asumiríamos equivocadamente que las vocales largas de las dos lenguas se comportan de modo similar en los mismos contextos.

Como lingüista preocupado por la documentación de lenguas y con las teorías lingüísticas, espero haber podido convencer al lector que las observaciones impresionísticas que tradicionalmente se codifican en los símbolos fonéticos no pueden usarse para recuperar y entender las propiedades fonéticas reales de los sistemas de sonidos de las lenguas. Ésta es una asunción muy difundida que se basa en el supuesto erróneo que las características articulatorias que convencionalmente se le asigna a los símbolos fonéticos basta para poder recuperar cualquier otro aspecto de las propiedades físicas de los sonidos de una lengua. Si nuestra tarea es documentar lenguas y/o contribuir al avance teórico de campos que dependen de la información fonética/fonológica, entonces la documentación acústica es esencial a nuestra empresa, ya sea que seamos lingüistas de campo, lingüistas teóricos o ambos.

Me gustaría terminar este libro reflexionando con el lector sobre las posibilidades que se abren ante los lingüistas cuando uno puede hacer comparaciones interlingüísticas entre estudios que documentan las diferentes propiedades acústicas de las lenguas. La disponibilidad de estudios acústicos hace posible encontrar fenómenos lingüísticos que van más allá de lo que una simple transcripción fonética puede decirnos sobre una lengua. Imagine todos los estudios comparativos, o las revisiones de los que ya existen, que podrían hacerse si existiesen más trabajos documentando las propiedades acústicas de lenguas poco estudiadas. Aún más, imagine cómo las teorías lingüísticas tendrían que cambiar para acomodar y encajar mejor los resultados de esos estudios acústicos y los fenómenos que están aún por descubrirse con esos estudios. Como mencioné al inicio de este libro, afortunadamente la documentación acústica de lenguas está ahora emergiendo como una tendencia fuerte, un cambio en el paradigma de la documentación de lenguas el cual estoy seguro traerá nuevos avances en nuestro entendimiento de las similitudes y diferencias entre las lenguas.

BIBLIOGRAFÍA

- AKINLABI, Akinbiyi y Mark LIBERMAN
2001 «Tonal complexes and tonal alignment». En Min-Joo Kim y Uri Strauss (editores). *Actas de NELS 31*. Amherst: Graduate Students Linguistics Association Publications, pp. 1-20.
- ALEMANY, Agustín
1906 *Vocabulario de Bolsillo. Castellano-Shipibo*. Lima: Tip. del Colegio Apostólico de P. F. del Peru.
- ARCHANGELI, Diana y Douglas PULLEYBLANK
1994 *Grounded Phonology*. Cambridge, MA: MIT Press.
- BECKMAN, Jill
1995 «Shona height harmony: Markedness and positional identity». En Beckman, J., S. Urbanczyk y L. Walsh (editores). *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Amherst, Mass., GLSA Publications, pp. 53-75.
1998 «Positional Faithfulness». Tesis doctoral. Amherst: University of Massachusetts.
- BERMÚDEZ-OTERO, Ricardo
2003 «The acquisition of phonological opacity». Rutgers Optimality Archive, 593, <http://roa.rutgers.edu/view.php?id=1126>
- BLACK, H. Andrew
1992 «South American verb second phenomena: evidence from Shipibo». En H.A. Black y J. McCloskey (editores). *Syntax at Santa Cruz*, vol. 1, pp. 35-63.
- BOERSMA, Paul y David WEENINK
2008 *Praat: Doing Phonetics by Computer*. (Version 5.0.06 - 5.1.05) [Programa de computadora]. <http://www.praat.org/>

- BRADLOW, Ann
 1995 «A comparative acoustic study of English and Spanish vowels». *Journal of the Acoustic Society of America*, vol. 97, pp. 1916-24.
- BRENZINGER, Matthias
 2007 *Language Diversity Endangered*. Berlin: Walter de Gruyter.
- BROWMAN, Catherine y Louis GOLDSTEIN
 1986 «Towards an articulatory phonology». *Phonology Yearbook*, vol. 3, pp. 219-252.
 1989 «Articulatory gestures as phonological units». *Phonology*, vol. 6, pp. 201-251.
 1990 «Tiers in articulatory phonology, with some implications for casual speech». En J. Kingston y M. Beckman (editores). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, pp. 341-397.
 1992 «Articulatory phonology: an overview». *Phonetica*, vol. 49, pp. 155-180.
 1995 «Dynamics and articulatory phonology». En R.F. Port y T. van Gelder (editores). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, pp. 175-193.
- BURGESS, E. y P. HAM
 1968 «Multilevel conditioning of phoneme variants in Apinayé». *Linguistics*, vol. 41, pp. 5-18.
- CALLOW, J.C
 1962 «*The Apinayé Language*». Tesis doctoral. London: University of London.
- CAMACHO, José
 2007 «The case of the case agreement». *Conferencia IV Encuentro de Gramática Generativa*. Argentina.
 En prensa-a «On case concord: The syntax of switch-reference clauses in Shipibo». *Natural Language & Linguistic Theory*.
 En prensa-b «On the asymmetry of root vs. embedded clauses. Evidence from Shipibo second-position clitics». En A. M. D. Sciullo y V. Hill (editores). *The Syntax of Clauses and Language Interfaces*. Amsterdam: John Benjamins.
- CAMACHO, José y José ELÍAS-ULLOA
 2004 «The syntactic structure of evidentiality in Shipibo». Rutgers University y Stony Brook University, <http://www.rci.rutgers.edu/~jcamacho/publications/publications.htm>
 En prensa «Null subjects in Shipibo switch-reference systems». En J. Camacho, R. Gutiérrez-Bravo y L. Sánchez. (editores). *Information Structure in Indigenous Languages of the Americas Syntactic Approaches*. Berlin: Walter de Gruyter.
- CHANG, Charles B.
 2008 «Variation in palatal production in Buenos Aires Spanish». En M. Westmoreland y J. A. Thomas (editores). *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 54-63.

- CHOMSKY, Noam y Morris HALLE
1968 *The Sound Pattern of English*. New York: Harper & Row.
- CHURCHWARD, C. Maxwell
1953 *Tongan Grammar*. Oxford: Oxford University Press.
- CLEMENTS, G. N.
1985 «The geometry of phonological features». *Phonology Yearbook*, vol. 2, pp. 225-252.
- COLE, Jennifer Fitzpatrick
1990 «The minimal word in Bengali». En A. Halpern (editor). *The Proceedings of the West Coast Conference on Formal Linguistics 9*. Stanford: Stanford Linguistic Association, pp. 157-170.
- D'ANGELIS, Wilmar da Rocha
1999 «Gradient versions of pre- and circum-oralized consonants in Kaingang (Brazil)». En *Proceedings of the 14th International Congress of Phonetic Sciences*. Berkeley: University of California at Berkeley, pp. 1043-6.
- DAVIS, Stuart
1995 «Geminate Consonants in Moraic Phonology». En *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. Stanford, CA: Center for the Study of Language and Information, pp. 32-45.
- DE LACY, Paul V.
2006 *Markedness: Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press.
- DECKER, T. Newell y Thomas D. CARRELL
2004 *Instrumentation: An Introduction for Students in the Speech and Hearing Sciences*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- DELATTRE, P., A. LIBERMAN y F. COOPER
1955 «Acoustic loci and transitional cues for consonants». *Journal of Acoustical Society of America*, vol. 27, N° 4, pp. 769-773.
- DISNER, Sandra
1983 *Vowel Quality: The relation between universal and language-specific factors*. *UCLA Working Papers in Phonetics*, vol. 58. Department of Linguistics, UCLA.
1984 «Insights on vowel spacing». En Maddieson, I. (editor). *Patterns of Sound*. Cambridge: Cambridge University Press, pp. 136-155.
- DOCHERTY, Gerard
1992. *The timing of voicing in British English obstruents*. New York: Foris Publications.
- ELÍAS-ULLOA, José
2000a «*El Acento en Shipibo*». Tesis de licenciatura. Lima: Universidad Nacional Mayor de San Marcos.
2000b «Silabas pesadas ¿monomoraicas?: reflexionando sobre el peso silábico en Shipibo ». En *Actas de I Congreso de lenguas Indígenas de Sudamérica*. Lima: Universidad Ricardo Palma, pp. 135-149.

- 2001 «Un análisis métrico del acento en Shipibo». *Lenguaje y Sociedad*, vol. 3, pp. 27-46.
- 2005 «Variable syllable weight and quantity-insensitive allomorphy in Shipibo». En *Actas de NELS 35*, Connecticut University: GLSA Publications, pp. 171-186.
- 2006 «*Theoretical Aspects of Panoan Metrical Phonology: Disyllabic Footing and Contextual Syllable Weight*». Tesis doctoral. New Brunswick: Rutgers University.
- 2009 «The Distribution of Laryngeal Segments in Capanahua». *International Journal of American Linguistics*, vol. 75, pp. 185-211.
- EVERETT, Caleb
- 2008 «Locus Equation Analysis as a Tool for Linguistic Fieldwork». En *Language Documentation and Conservation*, vol. 2, N^o 2, pp. 185-211, University of Hawai'i Press (<http://nflrc.hawaii.edu/ldc/>).
- En prensa «Variable velic movement in Karitiâna». *International Journal of American Linguistics*.
- FACE, Timothy L.
- 2001 «Focus and early peak alignment in Spanish intonation». *Probus*, vol. 13, pp. 223-46.
- FANT, Gunnar
- 2004 *Speech Acoustics and Phonetics: Selected Writings*. Boston: Kluwer Academic.
- FAUST, Norma
- 1990 *Lecciones para el aprendizaje del idioma shipibo-conibo*. Yarinacocha: Instituto Lingüístico de Verano.
- FLEMMING, Edward
- 2008 *The Grammar of Coarticulation*. Cambridge, MA: MIT, Department of Linguistics.
- FONTANELLA DE WEINBERG, Maria Beatriz
- 1978 «Un cambio lingüístico en marcha: las palatales del español bonaerense». *Orbis*, vol. 27, pp. 215-247.
- 1992 *El Español de América*. Madrid: Editorial MAPFRE.
- FRY, Dennis Butler
- 1979 *The Physics of Speech*. Cambridge (UK): Cambridge University Press.
- GARCÍA-RIVERA, Fernando
- 1994 *Aspectos de la Fonología del Shipibo*. Lima: Universidad Nacional Mayor de San Marcos.
- GHINI, Mirco
- 2001 *Asymmetries in the Phonology of Miogliola*. Berlin: Mouton de Gruyter.
- 2002 «From prosody to place: The development of prosodic contrasts into place of articulation in the history of Miogliola». En P. Fikkert y H. Jacobs (editores). *Development in Prosodic Systems*. Berlin: Mouton de Gruyter, pp. 419-56
- GOLDSMITH, John
- 1976 *Autosegmental Phonology*. Cambridge (MA): MIT Press.

- 1993 «Harmonic phonology». En J. Goldsmith (editor). *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago: University of Chicago Press, pp. 21-60.
- GONZÁLEZ, Carolina
 2003 «*The Effect of Stress and Foot Structure on Consonantal Processes*». Tesis doctoral. University of Southern California.
 2005 «Phonologically-conditioned allomorphy in Panoan: Towards an analysis». *UCLA Working Papers in Phonology*, vol. 6, N° 11, 39-56.
 2007 «The phonology and morphology of Panoan languages». Conferencia BIDE 2007 Summer School. Universidad de Deusto, Bilbao.
 En preparation *The Patterning of Glottals in Panoan Languages*. Tallahassee, Florida State University.
- GRIMES, Stephen
 2005 «Extrametricity and minimal word length in Hungarian». Conferencia Midcontinental Workshop on Phonology 11. Ann Arbor, Michigan.
 2009 «*Quantitative Studies in Hungarian Phonotactics and Syllable Structure*». Tesis doctoral. Indiana University.
- HÁJEK, John
 1997 *Universals of Sound Change in Nasalization*. Oxford: Blackwell.
- HALL DE LOOS, Betty y Eugene LOOS
 1978 «La estructura semántica y fonológica de los prefijos verbales en Capanahua». En Eugene Loos (editor). En *Estudios Panos I*, Yarinacocha, Peru, pp. 63-132.
- HARRINGTON, J. y S. CASSIDY
 1999 *Techniques in Speech Acoustics*. Dordrecht: Kluwer Academic Publishers.
- HAYES, Bruce
 1980 *A Metrical Theory of Stress Rules*. Cambridge, MA: MIT.
 1989 «Compensatory lengthening in moraic phonology». *Linguistic Inquiry*, vol. 20, pp. 253-306.
 1994 «Weight of CVC can be determined by context». En J. Cole, y C. Kisseberth (editores). *Perspectives in Phonology*. Stanford: CSLI, pp. 61-80.
 1995 *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.
- HEWLETT, Nigel y J. Mackenzie BECK
 2006 *An Introduction to the Science of Phonetics*. Mahwah, N.J.: L. Erlbaum Associates.
- HIMMELMANN, Nikolaus y Robert LADD
 2008 «Prosodic description: An introduction for fieldworkers». *Language Documentation and Conservation*, vol. 2, N° 2, pp. 244-274.
- HUALDE, José Ignacio
 1991 *Basque Phonology*. New York: Routledge.

- 2003 «El modelo métrico y autosegmental». En P. Prieto (editor). *Teorías de la entonación*. Barcelona: Ariel, pp. 155-184.
- 2005 *The Sounds of Spanish*. Cambridge, UK: Cambridge University Press.
- HUFFMAN, Marie
 1990 *Implementation of Nasal: Timing and Articulatory Landmarks*. UCLA Working Papers in Phonetics 75. Los Angeles: University of California.
- HUFFMAN, Marie. y R. KRAKOW
 1993 *Nasals, Nasalization and the Velum*. Phonetics and Phonology 5. San Diego: Academic Press.
- HYMAN, Larry
 1985 *A Theory of Phonological Weight*. Dordrecht: Foris.
- ITO, Junko
 1989 «A prosodic theory of epenthesis». *Natural Language and Linguistic Theory*, vol. 7, pp. 217-259.
- ITO, Junko y Armin MESTER
 1986 «The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility». *Linguistic Inquiry*, vol. 17, pp. 49-73.
- 1995 «Japanese phonology». En J. Goldsmith (editor). *Handbook of Phonological Theory*. Oxford: Blackwell, pp. 817-838.
- 2003 «Lexical and postlexical phonology in Optimality Theory: Evidence from Japanese». *Linguistische Berichte: Resolving Conflicts in Grammars*, vol. 11, pp. 183-207.
- JAKOBSON, Roman, C. GUNNAR, M. FANT y MORRIS HALLE
 1963 *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge MA: The MIT Press.
- JOHNSON, Keith
 2003 *Acoustic and Auditory Phonetics*. Oxford: Blackwell.
- KAGER, René
 1999 *Optimality Theory*. Cambridge: Cambridge University Press.
- KARJIGI, Veena, Preeti RAO y K. SAMUDRAVIJAYA
 2007 «Investigation of Acoustic Attributes of Marathi Unvoiced Stops for Classification». En *Actas de Frontiers of Research on Speech and Music*. Mysire: India, pp. 110-115
- KENSTOWICZ, Michael
 1994 *Phonology in Generative Grammar*. Oxford: Blackwell.
- KIPARSKY, Paul
 1982a «From cyclic phonology to lexical phonology». En H. van der Hulst y N. Smith (editores). *The Structure of Phonological Representations*. Dordrecht: Foris, pp. 131-175.
- 1982b «Lexical phonology and morphology». En I. S. Yang (editor). *Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 3-91.

- 1985 «Some consequences of Lexical Phonology». *Phonology*, vol. 2, pp. 85-138.
- 1997 «LP and OT». Conferencia del *LSA Summer Linguistic Institute*. Ithaca: Cornell University.
- 2000 «Opacity and cyclicity». *The Linguistic Review*, vol. 17, pp. 351-367.
- KRUEGER, John Richard
 1977 *Tuvan Manual: Area Handbook, Grammar, Reader, Glossary, Bibliography*. Bloomington, Ind.: Indiana University.
- LADEFOGED, Peter
 1996 *Elements of Acoustic Phonetics*. Chicago: University of Chicago Press.
 2003 *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.
 2005 *Vowels and Consonants: An Introduction to the Sounds of Languages*. Oxford: Blackwell.
 2006 *A Course in Phonetics*. Boston, MA: Thomson, Wadsworth.
- LADEFOGED, Peter y Ian MADDIESON
 1996 *The Sounds of the World's Languages*. Oxford: Blackwell.
 2001 «Vowels of the World's Languages». En C. Kreidler (editor). *Phonology: Critical Concepts in Linguistics*. London: Taylor & Francis.
- LAHIRI, Aditi
 2000 «Phonology: Structure, representation and process». En L. Wheeldon (editor). *Aspects of Language Production*. Philadelphia: Psychology Press, pp. 165-226.
- LAHIRI, Aditi, L. GEWIRTH y Sheila BLUMSTEIN
 1984 «A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study». *Journal of Acoustical Society of America*, vol. 76, Nº 2, pp. 391-404.
- LAURIAULT, James
 1948 «Alternating mora-timing in Shipibo». *International Journal of American Linguistics*, vol. 14, pp. 22-24.
- LEBEN, William R.
 1973 *Suprasegmental Phonology*. Cambridge MA: Massachusetts Institute of Technology.
- LEE, Seunghun J.
 2008 «*Consonant-Tone Interaction in Optimality Theory*». Tesis doctoral. New Brunswick: Rutgers University.
- LENAERTS, Marc
 2006 «When inter-ethnic botanical borrowing does not rely on obvious efficacy: Some questions from Western Amazonia». *Ethnobotany Research & Applications*, vol. 4, pp. 133-146.
- LEWIS, M. Paul (editor)
 2009 *Ethnologue: Languages of the World*. Dallas, Texas: SIL International.

- LIBERMAN, Mark y Alan PRINCE
 1977 «On stress and linguistic rhythm». *Linguistic Inquiry*, vol. 8, pp. 249-336.
- LIPSKI, John M.
 1994 *Latin American Spanish*. London: Longman Group.
- LODGE, Ken
 2009 *A Critical Introduction to Phonetics*. London: Continuum.
- LOOS, Eugene E.
 1969 *The phonology of Capanahua and its Grammatical Basis*. Oklahoma: University of Oklahoma.
- LORIOT, James y Barbara HOLLENBACH
 1970 «Shipibo paragraph structure». *Foundations of Language*, vol. 6, pp. 43-66.
- LORIOT, James, Erwin LAURIAULT y Dwight DAY
 1993 *Diccionario shipibo-castellano*. Yarinacocha, Perú: Ministerio de Educación: Instituto Lingüístico de Verano.
- MARPLE, L.
 1987 *Digital Spectral Analysis with Applications*. New Jersey: Prentice-Hall.
- MCCARTHY, John
 1979 «On stress and syllabification». *Linguistic Inquiry*, vol. 10, pp. 443-466.
 1983 «Consonantal morphology in the Chaha verb». En M. Barlow, D. Flickinger y M. Wescoat (editores). *The Proceedings of the West Coast Conference on Formal Linguistics 2*. Stanford: Stanford Linguistic Association, pp. 176-188.
 1986 «OCP effects: Gemination and antigemination». *Linguistic Inquiry*, vol. 17, pp. 207-263.
- MCCARTHY, John y Alan PRINCE
 1996 *Prosodic Morphology*. New Brunswick, NJ: Rutgers University Center for Cognitive Science.
 1991 *Prosodic minimality*. Conferencia The Organization of Phonology. Champaign: University of Illinois.
 1993a «Generalized alignment». En G. Booij y J. v. Marle (editores). *Yearbook of Morphology*. Dordrecht: Kluwer, pp. 79-153.
 1993b *Prosodic Morphology I: Constraint Interaction and Satisfaction*. New Brunswick: Rutgers University Center for Cognitive Science.
 1995 «Faithfulness and Reduplicative Identity». En J. Beckman, L. Walsh Dickey y S. Urbanczyk (editores). *University of Massachusetts Occasional Papers in Linguistics 18*. Amherst, Mass.: GLSA Publications, pp. 249-384.
- MCCAWLEY, James D.
 1968 *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.
- MESTER, Armin
 1991 *Some Remarks on Tongan stress*. Santa Cruz, CA: UCSC.
 1994 «The quantitative trochee in Latin». *Natural Language and Linguistic Theory*, vol. 12, N° 1, pp. 1-61.

- MITCHELL, T. F.
1960 «Prominence and syllabication in Arabic». *Bulletin of the School of Oriental and African Studies*, vol. 23, pp. 369-89.
- MOHANAN, K. P.
1986 *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- MOSELEY, Christopher
2007 *Encyclopedia of the World's Endangered Languages*. New York: Routledge.
- NAVARRO TOMÁS, Tomás
1957 *Manual de pronunciación española*. New York: Hafner Publishing Co.
- ODDEN, David
1986 «On the role of the Obligatory Contour Principle in phonological theory». *Language*, vol. 62, pp. 353-383.
1988 «Anti Antigemination and the OCP». *Linguistic Inquiry*, vol. 19, pp. 451-75.
- OLIVE, Joseph, Alice GREENWOOD y John COLEMAN
1993 *Acoustics of American English Speech: A Dynamic Approach*. New York: Springer-Verlag.
- OLTRA-MASSUET, Isabel
2005 «Stress-by-structure in Spanish ». *Linguistic Inquiry*, vol. 36, 43-84.
- OOSTENDORP, Marc van y Jeroen MAARTEN VAN DE WEIJER
2005 *The Internal Organization of Phonological Segments*. Berlin: Mouton de Gruyter.
- PARKER, Steve
2001 «The Acoustic Qualities of Bora Vowels». *Phonetica*, vol. 58, pp. 179-95.
- PAYNE, David L.
1981 *The Phonology and Morphology of Axininca Campa*. Arlington, TX: The Summer Institute of Linguistics and University of Texas at Arlington.
- PÉREZ SILVA, Jorge, Jorge ACURIO PALMA y Raúl BENDEZÚ ARAUJO
2008 *Contra el prejuicio lingüístico de la motosidad: Un estudio de las vocales del castellano andino desde la fonética acústica*. Lima: Instituto Riva-Agüero, Pontificia Universidad Católica del Perú.
- PIGGOTT, Glyne
1993 «Satisfying the minimal word». *McGill Working Papers in Linguistics*, vol. 8, N° 2, pp. 194-233.
- POZZI-ESCOT, I.
1998 *El Multilingüismo en el Perú*. Lima: Biblioteca de la Tradición Oral Andina.
- PRIETO, P., J. VAN SANTEN y J. HIRSCHBERG
1995 «Tonal alignment patterns in Spanish». *Journal of Phonetics*, vol. 23, pp. 429-451.
- PRINCE, Alan
1976 *'Applying' Stress*. Amherst: University of Massachusetts.

- 1983 «Relating to the grid». *Linguistic Inquiry*, vol. 14, N° 1, pp. 19-100.
- 1990 «Quantitative consequences of rhythmic organization». En M. Ziolkowski, M. Noske y K. Deaton (editores). *Parasession on the Syllable in Phonetics and Phonology*. Chicago: Chicago Linguistic Society, pp. 355-398.
- PRINCE, Alan y Paul SMOLENSKY
 1993 *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick, Rutgers University Center for Cognitive Science.
- QI, Yingyong
 1989 «Acoustic Features of Nasal Consonants». Tesis doctoral. Ohio State University.
- RAPHAEL, L., Gloria BORDEN y Katherine HARRIS
 2007 *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. New York: Williams & Wilkins.
- RIALLAND, Annie y Mamadou BADJIMÉ
 1989 «Réanalyse des tons du bambara: des tons du nom à l'organisation générale du système ». *Studies in African Linguistics*, vol. 20, pp. 1-28.
- RICE, Keren
 1996 «Default variability: the coronal-velar relationship». *Natural Language and Linguistic Theory*, vol. 14, pp. 493-543.
- ROSE, Sharon
 1994 «The historical development of secondary articulation in Gurage». En *Actas de Annual Meeting of the Berkeley Linguistics Society 20*, pp. 112-124.
- SÁNCHEZ, Liliana , José CAMACHO y José ELÍAS-ULLOA
 En prensa «Shipibo-Spanish: Differences in residual transfer at the syntax/morphology and the syntax/pragmatics interfaces ». En T. Ionin y M.-L. Zubizarreta (editores). *Second Language Research. Special issue: Selective L1 Influence and Retreat from Negative Transfer*. vol. 25, N° 2.
- SHELL, Olive
 1985 *Estudios panos 3: Las lenguas pano y su reconstrucción*. Yarinacocha: Instituto Lingüístico de Verano.
- SILVERMAN, Daniel Doron
 2006 *A Critical Introduction to Phonology: Of Sound, Mind and Body*. London: Continuum.
- Solís Fonseca, Gustavo
 2003 *Lenguas de la amazonía peruana*. Lima: Visual Service.
- SOSA, Juan Manuel
 1991 *Fonética y fonología de la entonación del español hispano-americano*. Amherst: University of Massachusetts.
- STEVENS, Kenneth y Sheila BLUMSTEIN
 1978 «Invariant cues for place of articulation in stop consonants». *JASA*, vol. 64, pp. 1358-1368.

- 1981 «The Search for Invariant Acoustic Correlates of Phonetic Features». En P. Eimas y J. Miller (editores). *Perspectives on the Study of Speech*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- SVANTESSON, Jan-Olof
2005 *The Phonology of Mongolian*. Oxford: Oxford University Press.
- TRAILL, A.
1985 *Phonetic and Phonological Studies of !Xóó Bushman*. Hamburg: Helmut Buske.
- TRUDGILL, Peter
1974 *Sociolinguistics: An Introduction*. Harmondsworth: Penguin.
- UCLA DEPARTMENT OF LINGUISTICS
2007 UCLA Phonetics Lab Archive. Los Angeles, CA: UCLA. <http://archive.phonetics.ucla.edu/>.
- VALENZUELA, Pilar
2002a «Causativization and transitivity in Shipibo-Konibo». En M. Shibatani (editor). *The Grammar of Causation and Interpersonal Manipulation*. Philadelphia: John Benjamins, pp.417-483.
2002b *Relativization in Shipibo-Konibo: A Typologically-oriented Study*. Munchen: Lincom Europa.
2003a «Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan». En A. Aikhenvald y R. M. W. Dixon (editores). *Studies in Evidentiality*. Philadelphia: John Benjamins, pp. 33–61.
2003b «*Transitivity in Shipibo-Konibo Grammar*». Tesis doctoral. Portland: University of Oregon.
- VALENZUELA, Pilar, Luis MÁRQUEZ-PINEDO y Ian MADDIESON
2001 «Shipibo». *Journal of the International Phonetic Association*, vol. 31, N° 2, pp. 281-285.
- VEENA KARJIGI, P. R. y K. SAMUDRAVIJAYA
2007 «Investigation of Acoustic Attributes of Marathi Unvoiced Stops for Classification». Conferencia: Frontiers of Research on Speech and Music. Mysire, India.
- VIJAYAKRISHNAN, K.
2002 «Disyllabic trochee in Bangla: Punjabi and Tamil ». En *Actas de Generative Linguistics in the Old World – Asia*. Taiwan (<http://glow.ling.nthu.edu.tw/6.htm>)
- VON DEN STEINEN, Karl
1904 *Diccionario Sipibo. Castellano-deutsch-sipibo. Apuntes de gramática. Sipibo-castellano*. Berlin: D. Reimer.
- WALKER, Richard
1979 «Central Carrier phonemics». En D. Zimmerly (editor). *Contributions to Canadian Linguistics*. Ottawa: National Museums of Canada, pp. 93-107.

- WEIDMAN, Scott y Sharon ROSE
 2006 «A Foot-Based Reanalysis of Edge-in Tonal Phenomena in Bambara». En *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Somerville MA: Cascadilla Proceedings Project, pp. 426-434.
- WETZELS, W. Leo
 1995 «Estrutura Silábica e Contornos Nasais em Kaingáng». En Leo Wetzels (editor). *Estudos Fonológicos das Línguas Indígenas Brasileiras*. Rio de Janeiro: UFRJ, pp. 265-296.
- WHITE, Laurence y Katalin MADY
 2008a «Heads you lose: Prosodic structure and timing in Hungarian». Conferencia *Laboratory Phonology 11*. Victoria University of Wellington, New Zealand.
 2008b «The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian». Conferencia *Speech Prosody 2008*, Campinas, Brasil.
- WIESEMANN, Ursula
 1972 *Die phonologische und Grammatische Struktur der Kaingang-Sprache*. The Hague: Mouton.
- WILKINSON, Karina
 1988 «Prosodic structure and Lardil phonology». *Linguistic Inquiry*, vol. 19, Nº 2, pp. 325-334.
- YIP, Moira
 1988 «The Obligatory Contour Principle and phonological rules: A loss of identity». *Linguistic Inquiry*, vol. 19, pp. 65-100.
- ZEC, Draga
 1988 «*Sonority Constraints on Prosodic Structure*». Tesis doctoral, Stanford: Stanford University.
- ZOLL, Cheryl
 1994 *anchors Away: A Unified Treatment of Latent Segments and Floating Features*. University of California, Berkeley.
 1996 «*Parsing below the Segment in a Constraint-based Framework*». Tesis doctoral. Berkeley: University of California.