

INTRODUCCIÓN A LA LINGÜÍSTICA

**Curso para investigadores de
lenguas indígenas de Bolivia**

Patricia Alandia Mercado (ed.)

Joshua Birchall

Luca Ciucci

Gillian Gallagher

Gabriel A. Gallinate

Myriam Lapierre

Neil Myler

Andrey Nikulin

Jesse Stewart

Adam J.R. Tallman



PÁGINA Y SIGNOS



Funproeib
Andes

INTRODUCCIÓN A LA LINGÜÍSTICA

Curso para investigadores de
lenguas indígenas de Bolivia



PÁGINA Y SIGNOS



Funproeib
Andes

Primera edición, junio 2023

PÁGINA Y SIGNOS es una publicación de la Carrera de Lingüística Aplicada a la Enseñanza de Lenguas de la Facultad de Humanidades y Ciencias de la Educación de la Universidad Mayor de San Simón.

© Carrera de Lingüística Aplicada a la Enseñanza de Lenguas
Plazuela Sucre, acera sud
Telefax: (591-4) 4258803, 4544108, interno 260-261
E-mail: paginaysignos@umss.edu
<https://www.facebook.com/paginaysignos>
Cochabamba, Bolivia

Patricia Alandia Mercado (ed.)

Autores

Joshua Birchall
Luca Ciucci
Gillian Gallagher
Gabriel A. Gallinate
Myriam Lapierre
Neil Myler
Andrey Nikulin
Jesse Stewart
Adam J.R. Tallman

Editora responsable

Patricia Alandia Mercado
patr.alandia@umss.edu

Comité editorial de esta obra

Gladys Camacho
Silvana Campanini Tejerina
Gabriel Gallinate

DIAGRAMACIÓN

Talleres Gráficos “KIPUS”

DISEÑO DE LA PORTADA

Talleres Gráficos “KIPUS”

Depósito Legal: 2-1-2143-2023
ISBN: 978-9917-0-2681-5

Esta publicación se realiza gracias al auspicio de FUNPROEIB Andes-Fundación para la Educación en Contextos de Multilingüismo y Pluriculturalidad.

Impreso en Talleres Gráficos “Kipus” Telfs.: 4116196 – 4237448
Cochabamba-Bolivia

Índice

Presentación	7
Capítulo 1: Documentación y descripción lingüística	9
<i>Adam J.R. Tallman</i>	
Capítulo 2: Fonología	49
<i>Gillian Gallagher</i>	
Capítulo 3: Fonética	62
<i>Gabriel A. Gallinate & Myriam Lapierre</i>	
Capítulo 4: Morfología	100
<i>Andrey Nikulin</i>	
Capítulo 5: Sintaxis	130
<i>Neil Myler</i>	
Capítulo 6: Lexicografía y elaboración de diccionarios	141
<i>Joshua Birchall</i>	
Capítulo 7: Contacto de lenguas y variación	160
<i>Jesse Stewart</i>	
Capítulo 8: Lingüística histórica	180
<i>Luca Ciucci</i>	
Capítulo 9: Lengua y adquisición	204
<i>Gillian Gallagher</i>	

Cómo citar: Alandia Mercado, P. (2023). Presentación. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 7-8), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11106424>

PRESENTACIÓN

Este libro es un proyecto concebido y gestionado por la Dra. Gladys Camacho, Directora de Linguistics Summer School Bolivia (LSSB), una organización sin fines de lucro creada para promover la investigación lingüística a favor de las lenguas indígenas de Bolivia.

Hace dos años, en el contexto de los seminarios de lingüística promovidos por el LSSB, y ante la ausencia de material didáctico contextualizado en la realidad boliviana, Gladys Camacho convocó a investigadores de distintos países para proponer la redacción de una introducción a la lingüística dirigida a la descripción de las lenguas indígenas. Nueve lingüistas respondieron al llamado, sin embargo, los siguientes pasos ya no pudieron ser dados, por lo que solicitaron su edición y publicación a Página y Signos.

El Comité editorial de la revista Página y Signos, con gran satisfacción, aceptó acoger este proyecto con el objetivo común de proveer, a estudiantes y profesores de las carreras de Lingüística, a hablantes de lenguas indígenas dedicados a su estudio, a los investigadores de las lenguas en general y de las lenguas de Bolivia en particular, de un material único en el medio, para la formación teórica y sobre todo metodológica en el campo de la lingüística descriptiva. La Fundación para la Educación en Contextos de Multilingüismo y Pluriculturalidad, Funproeib Andes, se ha unido en este propósito apoyando su publicación.

Introducción a la lingüística: curso dirigido a investigadores de las lenguas de Bolivia no es una introducción en el sentido estricto, pues, si bien sienta las bases de las disciplinas necesarias para la descripción lingüística, la complejidad propia de esta tarea exige el conocimiento de una serie de categorías lingüísticas que no podrían ser cubiertas en una obra de esta naturaleza. No obstante, es un material valioso para acompañar procesos de enseñanza y aprendizaje, guiar los pasos de investigadores nóveles y enriquecer los trabajos de investigadores más experimentados.

Este Curso presenta nueve capítulos que, si bien han sido pensados como partes de una estructura formativa, están dotados de autonomía, con características discursivas propias de cada uno de los autores. El orden de los capítulos ha sido definido en el proceso de edición, sin embargo, no debe entenderse como una ruta obligatoria, pues esta dependerá de los intereses y necesidades de los lectores, tanto dentro de procesos de enseñanza y aprendizaje como en sus búsquedas individuales.

Cada capítulo presenta una estructura temática elegida por su autor, con niveles diferentes de profundidad y complejidad teórica, considerados necesarios para la comprensión del funcionamiento de las lenguas. Sin embargo, el aporte que singulariza

este trabajo es el cúmulo de experiencias de investigación compartidas, que proveen recorridos bibliográficos y metodológicos de utilidad invaluable, fundamentales para todo investigador.

Las abreviaturas presentadas por los autores en cada capítulo se encuentran reunidas al final y, a continuación, se presenta un apretado glosario para acompañar la lectura de los distintos capítulos.

Patricia Alandia Mercado

Editora responsable

Cómo citar: Tallman, A. J.R. (2023). Documentación y descripción lingüística. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 9-48), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11106976>

CAPÍTULO 1

DOCUMENTACIÓN Y DESCRIPCIÓN LINGÜÍSTICA

*Adam J.R. Tallman**

0. Introducción

Este capítulo presenta una introducción básica a la documentación y descripción lingüística. La primera sección introduce el concepto de documentación lingüística con el objetivo general de definirla. La segunda sección describe los métodos de documentación. La tercera sección resume la relación entre la descripción, documentación y la tipología o lingüística comparativa.

1.1. ¿Qué es y por qué hacer documentación lingüística?

Una descripción puede ser definida como un registro duradero y multiusos de una lengua (Lehmann, 2001; Himmelmann N. P., 1998; 2006; 2012); (Woodbury, 2003; 2007). Entonces, debemos establecer qué significa ‘duradero’, ‘multiusos’ y ‘una lengua’ en esa definición. Como describe Himmelmann (2006, p. 16), hay varios sentidos intercalados del concepto de ‘una lengua’.

(1)

- a. Un sistema cognitivo
- b. Fenómeno social e histórico
- c. Praxis social y política

La lengua puede ser entendida como un sistema cognitivo del individuo. Con ese tipo de abstracción no existen ‘lenguas’ como castellano, quechua o ese ejja como tales, sino sistemas cognitivos más o menos homogéneos o heterogéneas entre

* Trabaja en la Universidad Freidrich-Schiller, Jena (Alemania). Es investigador adscrito del Museo Nacional de Etnografía y Folklore de La Paz (Bolivia). Tiene un doctorado por la Universidad de Texas en Austin. Sus investigaciones se centran en la descripción y documentación de las lenguas de las Américas, en especial de las familias Algonquina, Pano y Tacana, y en la aplicación de métodos cuantitativos a la descripción y tipología de lenguas.

sí. Desde esa perspectiva, el interés principal es cómo caracterizar mecánicamente las reglas y restricciones necesarias que capturen y predigan las intuiciones de un hablante sobre qué es una oración gramatical o no gramatical, y **qué es una oración** adecuada o no adecuada, según el contexto lingüístico relacionado (ej. Uriagereka, 1998). Tomar una perspectiva cognitiva de la lengua no necesariamente significa rechazar los otros sentidos. De hecho, la relación entre el sistema cognitivo de un individuo y los otros sentidos intercalados es la pista para llegar a diferencias teóricas entre varios campos de la lingüística. Por ejemplo, unos postulan que las intuiciones sobre la gramaticalidad de una oración provienen de escenarios culturales típicos en vez de ser la producción mecánica de un sistema estructural únicamente lingüístico. El campo de la *etnosintaxis* investiga las relaciones entre intuiciones gramaticales y el contexto cultural y social del hablante (véase Evans, 2002). Otro sentido de la lengua radica en sus dimensiones sociales y culturales. Por ejemplo, se puede investigar la manera en la que la estructura de la gramática y el uso de la lengua sirven como índice o símbolo de la identidad de un hablante, de un oyente o de una comunidad o etnia (ej. Webster, 2008; Epps, Webster, Woodbury, 2017; Eckert, 2018). Finalmente, la lengua puede ser entendida como un medio de la praxis política (ej. Cameron, 1995; Eagleton, 2008; Carlucci, 2015). La documentación lingüística involucra todas estas diferentes miradas sobre el fenómeno de ‘la lengua’, por ello, la documentación es un proyecto, por necesidad, amplio, interdisciplinario, que es a la vez científico y político.

¿Qué significa que el registro debe ser ‘**multiusos**’? Significa que la documentación lingüística tiene que ser útil para diferentes grupos y propósitos. Principalmente, tiene que ser utilizable por la comunidad de hablantes. Por eso, al inicio, los enfoques de un proyecto de documentación tienen que estar dirigidos y responder, lo mejor posible, a los intereses de la comunidad de hablantes. En segundo lugar, se puede identificar a otros grupos u organizaciones interesados, como agencias nacionales (ej. IPELC, en el caso de Bolivia) e internacionales, que se interesan en la educación y en la planificación o normalización lingüística. En tercer lugar, tiene que ser utilizable por los investigadores de varias disciplinas (antropólogos, lingüistas, historiadores). Las categorías de interesados citados no son mutuamente excluyentes. Idealmente hay gente que se involucra en los tres diferentes grupos. Por ejemplo, los miembros de los Institutos de Lengua y Cultura de Bolivia (ILC) son normalmente miembros de la comunidad de hablantes, trabajan para el Gobierno y son investigadores interdisciplinarios. De hecho, los proyectos de documentación más exitosos tienen gente que responde a esas diferentes categorías.

El **registro duradero** está hecho con una perspectiva de larga duración, que va más allá de intereses y metas actuales. El registro tiene que estar organizado y archivado de manera tal que los datos puedan ser usados por gente que no estuvo

involucrada en el proceso de documentación. Eso trae la pregunta de cómo se puede conocer los intereses de generaciones del futuro. Para responder a esta pregunta, podemos hipotetizar basándonos en nuestra experiencia de uso de documentos lingüísticos del pasado: ¿Qué falta en los documentos históricos para que tengan un uso más general?

A veces los registros del pasado carecen de **metadatos**. Consideremos el siguiente fragmento de una gramática del lingüista Gilberto Prost, del Instituto Lingüístico de Verano (Prost, 1965, p. 15).

Figura 1.

Gramática de Prost (1965, p.15)

Chácobo

15

La cláusula dependiente

C1 110 = nika ci - -hiwí ha ? aki ci 'así - -j con-palo él

cuando-golpeó j' (cuando él lo golpeó así con el palo).

C1 120 = haci - -şobo ki--bişo ci 'entonces--casa a--ha-

biendo-traído j' (luego habiendo traído a la casa).

C1 130 = ?ani--şıta?á ci ki?a 'río--después-de-cruzar j

rel' (después de cruzar el río).

C1 140 = hini- -şokobo ? aşino ci 'en-agua--niños mientras -

bañan j' (mientras los niños se bañan en el agua).

En este fragmento no tenemos información sobre quién era el hablante, de dónde venía, ni cómo el lingüista adquirió la información. Imaginemos que intentamos reelicitar la misma información, y el hablante nos dice que está mal. ¿Qué podríamos concluir? Que tal vez hay diferentes variantes de chácobo. O, tal vez, no hemos podido construir el contexto apropiado de la oración. ¿Esas oraciones son parte de un mito? ¿Son traducciones directas del castellano? ¿Fueron escuchadas por el lingüista en contextos naturales? Concluir generalizaciones sobre el chácobo, basadas en esas oraciones, sería problemático, porque nos hace falta mucha información del contexto.

El registro de Prost también carece de la grabación original. Solo se sabe cómo pronunciar de una forma inexacta. Si quisiéramos discutir las posiciones de

los acentos que ha puesto Prost, no podríamos hacerlo (por ejemplo, ¿qué significa fonéticamente el acento sobre *hiwí* ‘en el árbol’ en CI 110? (Según Tallman, 2018b y Tallman y Elias-Ulloa, 2020 debe haber una clase de acento sobre la primera sílaba de *nika* ‘así’; también falta un acento sobre la segunda sílaba de *akí* ‘cuando hizo’). Como no hay una grabación de Prost, no podemos hacer la comparación para ver si hay una diferencia verdadera o una diferencia de signos diacríticos simplemente.

Faltan también traducciones detalladas. No hay una traducción directa de una lengua a otra. La única manera de llegar a una aproximación es dando diferentes traducciones posibles de diferentes contextos, en diferentes niveles de análisis e, idealmente, de diferentes hablantes (véase Woodbury (2007) sobre la noción de ‘traducción densa’). Consideremos la traducción de la oración de CI 120 ‘luego *habiendo* traído a la casa’. ¿Qué parte del chácobo lleva el significado *habiendo* que aparece en la traducción? De hecho, *bi* significa ‘traer’ o ‘venir’ si el sujeto es plural, y *so* <š> significa un evento anterior, pero ¿cómo llegamos al pasado perfecto progresivo del castellano *habiendo*? Si uno pregunta a un hablante nativo, se da cuenta de que esa oración puede ser traducida como ‘Y después de traer a la casa’, ‘Y después de venir a la casa’, ‘Y después que han traído a la casa’, etc. (Tallman, 2018b, 2019; Tallman & Stout, 2016, 2018). La traducción dada es una entre varias posibles, pero Prost da la impresión de que hay una relación isomórfica entre la semántica del castellano y la semántica del chácobo.

En el registro puede faltar también una anotación detallada o multiperspectiva (‘anotación densa’). La anotación lingüística a veces necesita muchos detalles que reflejen el análisis preferido de un lingüista entre la multitud de análisis posibles. Podemos complementar la definición de Himmelmann para decir que una documentación debe ser **multiperspectiva**, es decir, debe reflejar diferentes interpretaciones posibles de la praxis lingüística (*observable linguistic behavior*) de una comunidad de hablantes. Observar fenómenos desde diferentes perspectivas puede ayudar a la creación de una documentación más rica y detallada. Somos seres humanos y todos tenemos la tendencia de ignorar o filtrar aspectos de la realidad para que estén conformes con nuestras presunciones. Un buen descriptivista reconoce ese aspecto del conocimiento e invita a diferentes miradas sobre el fenómeno.

La noción amplia de documentación, resumida arriba, proviene de los motivos principales para tener una documentación. Hay tres respuestas a la pregunta sobre por qué tener una documentación lingüística. Un registro duradero y multiusos puede servir a la revitalización de una lengua. También una documentación bien hecha puede beneficiar a la pedagogía, sobre todo en el desarrollo de currículos indígenas y anticolonialistas, ya que debe incluir etno-historias de la comunidad de hablantes,

en lugar de que sean impuestos desde el Estado. Desde la mirada científica, la documentación es importante por la verificabilidad. Sin una documentación amplia, los datos de estudio para mostrar cualquier punto teórico no serían reproducibles, y las generalizaciones no serían replicables.

Un proyecto de documentación involucra a la comunidad de hablantes que deben ser participantes en el proceso y en las decisiones (Woodbury, 2003; Henke & Berez-Kroeker, 2016). Por eso, gran parte de mi propio trabajo con el chácobo y el pacahuara consistía en la documentación de mitos tradicionales y no tanto en la producción de materiales pedagógicos. Es importante destacar que una documentación tiene que estar bien organizada para que tenga múltiples usos. Cualquier organización tiene que estar basada en una clasificación de los datos. Una documentación bien organizada presupone una descripción lingüística del fenómeno. Pero la descripción misma depende de una documentación amplia del fenómeno. Entonces hay una relación dialéctica entre la documentación y la descripción. Los métodos de documentación que revisaremos abajo son a la vez métodos de descripción.

1.2. Métodos de documentación

Las herramientas básicas del lingüista documentalista pueden ser divididas en tres (véase también Chellilah & de Reuse, 2011; Bown, 2008).

- (2) a. Herramientas de grabación y anotación:
notas; grabaciones en foto, audio y vídeo; ELAN; FLEEx; Toolbox; Praat
- b. Herramientas de organización y archivo:
sesiones, ficheros, sistema de nombrar, Arbil, ficheros de IMDI
- c. Clases o contextos de comportamientos lingüísticos observables:
elicitación, experimentos, entrevistas, habla natural (cuentos, conversaciones), observación participante

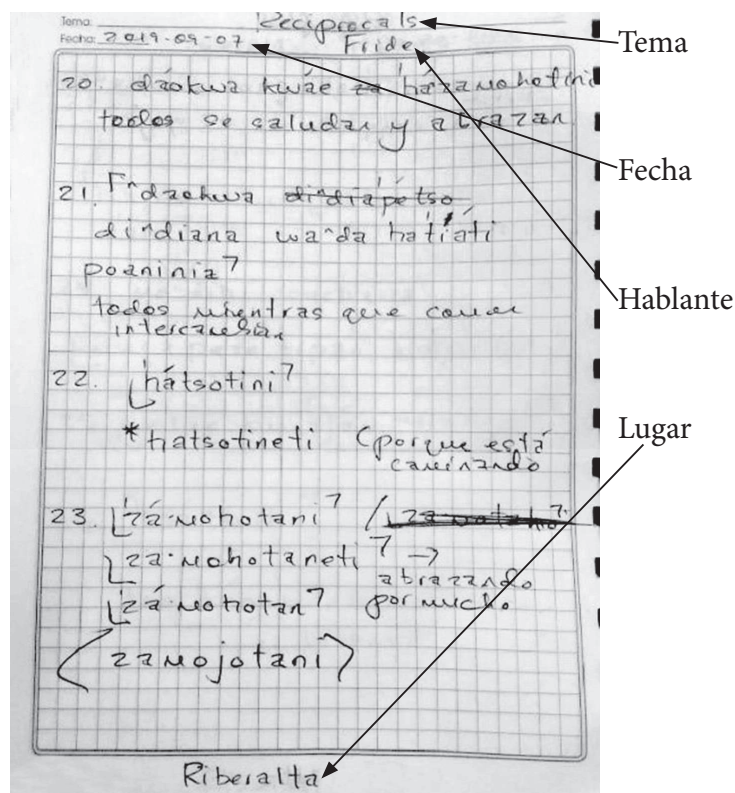
1.2.1. Herramientas de grabación, anotación, organización y archivo

1.2.1.1. Notas, datos y metadatos

Hacer buenas notas físicas todavía es muy importante, sobre todo en el inicio de un proyecto de documentación o descripción de una lengua no conocida. Debe anotarse la fecha, el lugar y el hablante y elaborarlas con notas etnográficas para capturar todo el contexto de la grabación.

Figura 2

Notas hechas en Riberalta trabajando con los araonas, 2019.9.07

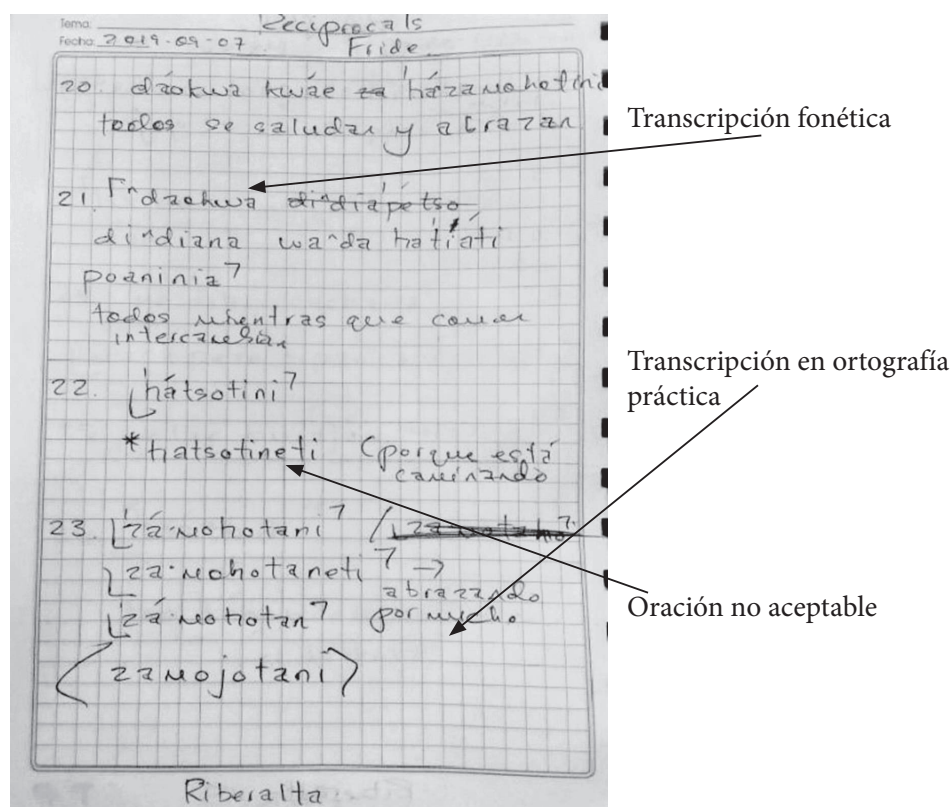


En la **Figura 2**, se ve un ejemplo de notas que tomé en 2019 trabajando con los araonas de Bolivia. Tiene la fecha ('2019.9.07'), el tema sobre el que quisimos estudiar ('Recíprocos'), el hablante (Fride Matahua) y el lugar donde estábamos trabajando ('Riberalta'). Es importante escribir las notas con un bolígrafo en lugar de un lápiz, porque dura más tiempo. Además, es importante usar un cuaderno de tapa dura para escribir notas. Cuando se esté en el terreno, aconsejaría también guardar los cuadernos en una bolsa plástica (idealmente de ziplock) para evitar daños causados por el agua.

En mis propias notas uso convenciones diferentes según si estoy escribiendo una transcripción en el Alfabeto Fonético Internacional, en una ortografía práctica, y también si una oración es aceptable o no. La transcripción fonética está marcada por paréntesis cuadrados: []. La transcripción en la ortografía práctica está en paréntesis angulares: < >. Las oraciones no aceptables están marcadas con * para 'completamente no-aceptable' y '?' para 'más o menos aceptable'.

Figura 3

Notas hechas en Riberalta, trabajando con los arañas. 2019.9.07



1.2.1.2. Organización de archivos

Las notas y los ficheros de grabaciones deben ser organizados de forma tal que sea fácil encontrar la grabación original. Por ejemplo, la grabación que contiene los ejemplos que escribí se encuentra en proyecto_Araona_2019 > 2019 > 2019-9-07. Cada trabajo lingüístico está organizado por sesiones con la fecha. Aunque puede parecer tedioso dar tanta organización a los datos, es muy importante hacerlo lógicamente al inicio de un proyecto de documentación, porque el lingüista documentalista va a manejar miles de ficheros que, si no están organizados, corre el riesgo de perderlos o de pasar mucho tiempo buscándolos. Para tener metadatos más organizados, se puede utilizar Arbil u otros sistemas de creación de ficheros de IMDI. Ese programa relaciona las sesiones con un sistema de grabación y organización de metadatos. Los metadatos tienen más información de la que incluí en las notas.

Figura 4

Foto de organización en Arbil

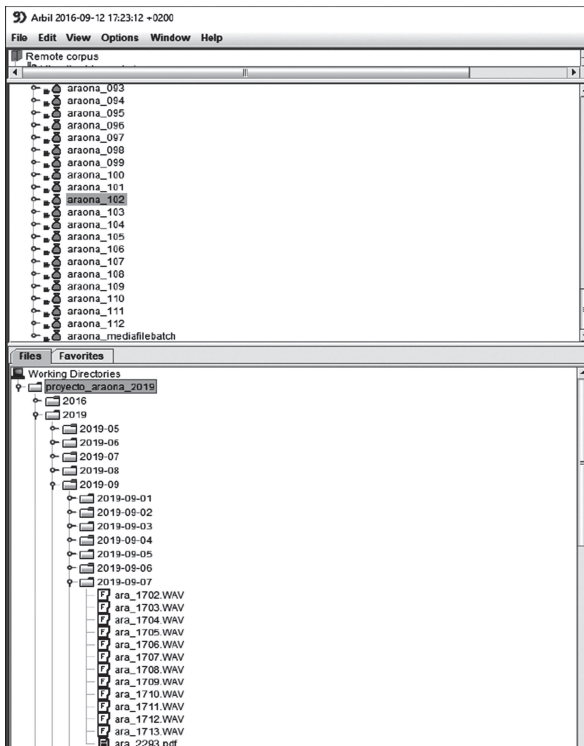
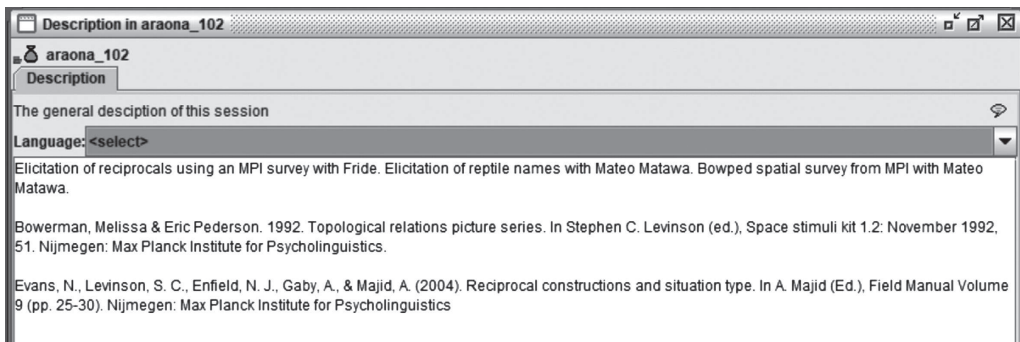


Figura 5

Una foto del contenido de una sesión de Arbil (araona_102)



Para archivos como *Endangered Language Archive (SOAS)* (ELAR) en Londres, es un requisito organizar los ficheros con un creador de IMDI.¹ En la **Figura 5** se indica exactamente qué hice en esa fecha y los materiales que usé para la elicitación.

¹ <https://www.soas.ac.uk/elar/>

1.2.1.3. Textos: grabación, transcripción y traducción densa

Otra parte importante de la documentación es la producción de textos. Normalmente, el inicio de la producción de textos desde la perspectiva documentalista consiste en grabar el habla natural. En el pasado, todos los textos fueron hechos por dicción, pero ya que ahora es más fácil hacer registros de audio, es siempre preferible tener una grabación de audio, sino se pierde mucha información de la presentación (Tedlock, 1977; 1983). Últimamente los investigadores aprecian cada vez más los aspectos visuales de la presentación, de la interacción y de la obra lingüística, y, por eso, lo ideal es grabar todo lo posible en vídeo también (Seyfeddinipur, 2011).

Cuando puedo, grabo un vídeo con el audio desde diferentes direcciones. La foto en la **Figura 6** muestra mi intento de tener una grabación amplia que involucre todo el contexto del habla natural. Chanito Matawa está entrevistando a Shani Matawa sobre los diferentes nombres de los meses del año y sus significados en araona. Durante la grabación, varios hablantes interrumpieron para dar sus opiniones y recuerdos sobre el tema. La foto muestra seis micrófonos (un micrófono AT 3010 anexado a la cámara, otro anexado a un árbol con un atril especial, dos en el suelo al frente de los hablantes y un micrófono Shure de cabeza para cada hablante). Con eso se cubre todo el contexto, con grabaciones afines de los dos hablantes. Además, saqué fotos de los alrededores de la comunidad mientras estábamos grabando. Los temas de presentación son del dominio de los hablantes. Mientras tanto, si es posible, intento explicar cómo usar los equipos cuando los estoy alistando.

Figura 6

Chanito Matawa (izquierda) entrevistando a Shani Matawa (derecha) en la comunidad de Barero, río Manupari (departamento de La Paz)



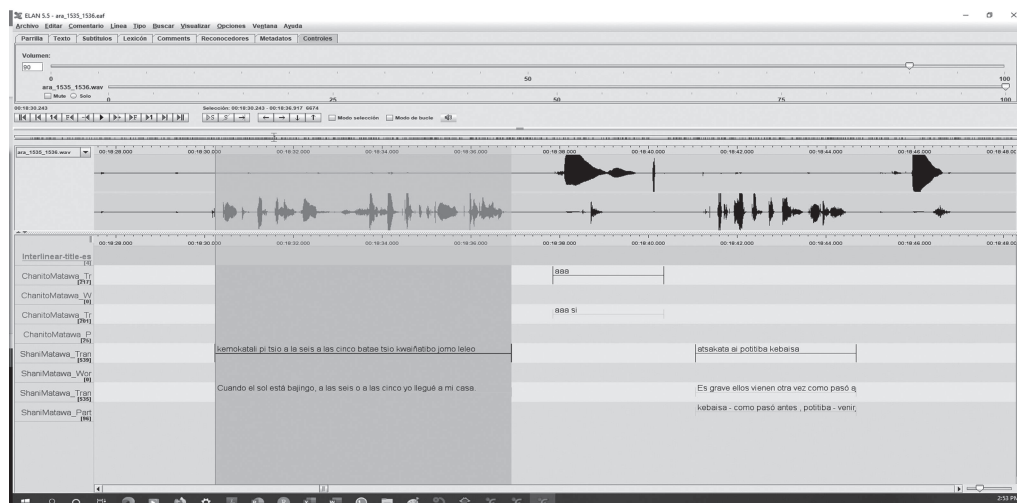
Siempre debe intentarse tener por lo menos dos hablantes involucrados en una grabación, no solo para grabar conversaciones, sino también para mitos, leyendas, etc. La razón es porque, en muchas culturas, la praxis narratológica tiene aspectos dialógicos importantes (Brody, 1996; Martin, 2000). Todos los ficheros grabados tienen que estar guardados en la misma sesión, organizada por fecha. Los metadatos pueden ser escritos en notas físicas o directamente en la computadora si uno tiene acceso a la electricidad.

Después de guardar las grabaciones en audio y vídeo, es importante nombrar los ficheros con un sistema lógico de códigos (ej. ara_0071, ara_0072, etc.). El fichero que contiene la transcripción y traducción de la grabación idealmente debería tener el mismo código que el fichero de audio o vídeo. Por ejemplo, una transcripción de la grabación ara_0070.wav sería guardado como ara_0070.eaf. Los ficheros *eaf* son producidos por el programa de ELAN, que, hoy en día, es el más comúnmente usado por lingüistas y antropólogos en la creación de transcripciones y traducciones.

ELAN es un programa que vincula transcripciones y traducciones con un audio. Se tiene que escoger una grabación y una plantilla .etf (Tallman, 2018b; Hellwig et al., 2020). La **Figura 7** presenta una foto de ELAN.

Figura 7

ELAN de ara_1535_1536 kemokatali pi tio a la seis a las cinco batae tsio kwaiñaibo jomo leleo; 'Cuando el sol está bajingo, a las seis o a las cinco yo llegué a mi casa'.



La **traducción densa** o “Thick translation” (Woodbury, 2007) se refiere a la traducción elaborada con notas y traducciones en diferentes niveles de composicionalidad, como se ve en la **Figura 8**. Por ejemplo, cuando uno elicit

traducciones de una oración, también tiene que preguntar traducciones al nivel de cada palabra, como se ve en la Figura 8. Se anota también aspectos del contexto.

Figura 8

Ilustración de la traducción densa en de un texto araona

ShaniM [539]	atsakata ai potitiba kebaisa
ShaniM [0]	
ShaniM [535]	Es grave ellos vienen otra vez como pasó antes.
ShaniM [001]	kebaisa - como pasó antes , potitiba - venir otra vez , (contexto: hablando de los misioneros)

El nivel más fino de la descripción es el nivel de morfemas (o formativos). Las traducciones o *glosas* en ese nivel son importantes para la documentación. Actualmente la herramienta más común para hacer glosas interlineares es FLEx (FieldWorks Language Explorer). FLEx es un programa que crea una base de datos léxicos (como un diccionario) relacionada con textos. Hay formatos .etf de ELAN especiales que se pueden usar para exportar un texto .eaf hecho en ELAN a FLEx (Hellwig et al., 2020).

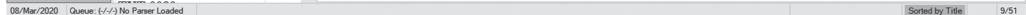
Figura 9

Parte de las entradas léxicas de FLEx del proyecto araona

The screenshot shows the FieldWorks Language Explorer (FLEx) interface. On the left, a sidebar contains a 'Lexicon' panel with a tree view showing various entry types like 'Headword', 'Morph Type', etc. The main window displays a list of lexical entries under the 'Entries' tab. The entries are organized in a table with columns for 'Headword', 'Lexeme Form', 'Glosses', and 'Grammatical Info'. The entry 'nadetse dos' is selected, and its details are shown on the right. The details include the Lexeme Form 'nadetse', Morph Type 'tema', Citation Form 'Spm', and a Gloss 'dos'. The entry is also linked to a 'Sense 1' gloss 'dos' and a 'Definition' 'cielo año'. The interface includes a menu bar at the top with options like File, Send/Receive, Edit, View, Data, Insert, Format, Tools, Parser, Window, and Help. The status bar at the bottom shows the date '20/Nov/2019', the queue status 'Queue: (-/-) No Parser Loaded', and the sorting criteria 'Sorted by Lexeme Form' with a count of '1257/1977'.

Como se ve en la **Figura 10**, cada morfema en el léxico puede vincularse con las formas que se encuentran en el texto. La Figura 10 muestra un texto con glosas interlineares hechas en FLEx. Cada glosa aparece también en el léxico, ya que uno puede construir un diccionario /léxico mientras continúa con el análisis de un texto.

Un texto ara_0049 del proyecto araona mostrando la glosa interlinear de kweda-ta-o ‘perseguir-3a-desapr’



1.2.2. Comportamientos lingüísticos observables

Los transductores pueden ser divididos sobre un continuo entre los más controlados y los más naturales. Elicitación y experimentación son el transductor más

controlado. Consiste en traducir una oración o palabra o frase en la lengua de interés por medio de la pregunta ‘¿Cómo se dice ____?’; describir una situación vista en foto o vídeo; dar juicios de gramaticalidad sobre oraciones hipotéticas o repetir oraciones de la lengua de interés. Los transductores menos controlados son las oraciones de observación de participante. Son oraciones que uno escucha al observar a los hablantes, sin incitación. La ventaja principal de la elicitación y experimentación es que el lingüista tiene control máximo sobre el contenido. También se puede conseguir evidencia negativa, por ejemplo, cuáles oraciones inaceptables son útiles para probar generalizaciones. Mientras que la elicitación ofrece un cierto grado de control al investigador y la posibilidad de concentrarse en unos temas específicos, está asociada también con una falta de validez ecológica que puede corromper los comportamientos observables lingüísticos.

Si la elicitación está basada en traducciones de una lengua de contacto (ej. castellano a bésiro), hay la posibilidad de que salgan más oraciones con calcos de la lengua de contacto que oraciones de la lengua de interés, y también hay la posibilidad de que estructuras importantes de la lengua no aparezcan en los datos de elicitación porque no corresponden a las traducciones más obvias de ciertas construcciones de la lengua de contacto (Mithun, 2001). Para tomar un ejemplo sencillo, en chácobo el objeto normalmente precede al verbo. Pero cuando uno pregunta cómo traducir de una oración sujeto-verbo-objeto del castellano, muchos hablantes dan un orden que corresponde al castellano. El orden menos marcado será sujeto-objeto-verbo (Tallman, 2018b). El orden sujeto-verbo-objeto es posible en chácobo, pero normalmente estaría marcado con una pausa después del verbo. En (3), se tiene una oración calcada.

- (3) <camano’ tépasquë joni> (más natural *camano’ joni tépasquë*)
 kamano=’ típas=ki honi
 tigre=ERG matar=DEC:PAS hombre
 ‘El tigre mató al hombre.’

Con los calcos, hay también la posibilidad de importar estructuras que son completamente agramaticales en la lengua de interés, y que suenan gramaticales solo en el contexto de elicitación (Mithun, 2001). Mientras que la elicitación es una técnica necesaria para orientarse y llenar lagunas del conocimiento y acercarse a traducciones de ‘palabras’ fuera de un contexto, no debe ser la metodología por defecto del lingüista descriptivo. Un estudio cercano de los datos en contextos ecológicamente válidos es fundamental.

El otro extremo son los datos adquiridos por observación. Esos son ecológicamente válidos, pero presentan el problema de que el investigador no suele

tener conocimiento de todo el contexto de su producción, entonces la interpretación puede ser degradada por sus propios sesgos. Con la observación natural, no se puede controlar aspectos del contexto para ver los límites de las reglas que emergen al estudiarlas, y existe la posibilidad de que ciertas construcciones no ocurran en el corpus por razones estadísticas (Matthewson, 2004). Por eso, la mayoría de los lingüistas documentalistas combinan diferentes metodologías que se pueden organizar a lo largo de un continuo de control (véase también Bonhemeyer, 2015). Se describe el continuo de comportamientos observables lingüísticos en la **Tabla 1**. El documentalista no solo aprecia las ventajas de cada metodología, sino también sus trampas, ya que los comportamientos observables lingüísticos (articulaciones e interpretaciones) son un producto de la realidad lingüística y su método de adquisición (el transductor de la realidad), al mismo tiempo (véase Wimsatt, 2005; Candea, 2018).

Tabla 1

Continuo de métodos de adquisición de observables lingüísticos

Clase	Tarea	Control
Elicitación experimentos	y Traducir (lengua de contacto de/a lengua del contacto) Juicios de aceptabilidad / gramaticalidad Experimentos de producción (fonética) Experimentos de percepción (fonética) Elicitación inducida por contexto / foto / video ...	Controlado
Guiones gráficos	...	↕
Entrevistas dinámicas	...	
Habla natural	Mitos / leyendas comunes Cuentos, diarios personales, Conversaciones	
		Ecológico
Observación participante	Vivir en la comunidad con cuaderno	

Podemos vincular el continuo de comportamientos observables lingüísticos con la noción de traducción densa ('Thick translation') introducida por Woodbury (2007). Como se señaló anteriormente, la traducción densa es necesaria en la documentación porque no hay una traducción perfecta para una palabra, frase u oración. Solo hay aproximaciones refractadas de diferentes perspectivas según ciertos contextos y según ciertos métodos de observación. La traducción es una heurística necesaria de investigación y documentación, pero una traducción sola nunca puede ser una reflexión isomórfica de la forma o significado de interés.

Para abordar una solución a ese problema, la traducción densa implica adquirir traducciones a diferentes niveles de composicionalidad, traducciones de diferentes perspectivas (ej. Diferentes hablantes), traducciones elaboradas en o con varios contextos, y traducciones basadas en un conocimiento y experiencia etnográficos. Investigar un fenómeno de interés por medio de las diferentes clases y tareas de los observables lingüísticos resumidos en la Tabla 1 es la manera más segura de documentar y de acercarse a un entendimiento de la realidad lingüística.

1.2.2.1. Más allá de la elicitación

Para comenzar, se aconseja preguntar por una lista de palabras en la lengua de contacto (si la hay), basada en la investigación de un lingüista que trabaja o trabajaba en la misma región, para conseguir una lista de palabras en el castellano regional. Por ejemplo, yo usaba las notas de Antoine Guillaume sobre la lengua cavineña para ayudarme en la construcción de una lista de palabras, no solo para aprender el castellano beniano, sino también para tener ideas de qué elicitar. Palabras como 'microondas' o 'castillo' probablemente no son las primeras cosas que se quiere elicitar en la Amazonia boliviana, por ejemplo.

También es importante comprobar las transcripciones con varios hablantes. Las palabras que se elicite con un hablante tienen que ser verificadas con otro y viceversa. Cuando se tiene una base de palabras elicidadas con un hablante, se puede verificar y elaborar las traducciones para preguntar a otro hablante y retraducirlas en la lengua de contacto.

Más allá de una elicitación sencilla, hay otros métodos más sofisticados. Por ejemplo, los manuales de campo y materiales de estímulo del Instituto de Max Plank (L&C Field Manuals and Stimulus materials) son muy útiles.²

La **Figura 11** muestra una captura de pantalla de un vídeo de *Put projet* (poner objetos) de Bowerman et al. (2004), que tiene varios vídeos que uno puede usar para elicitar oraciones en la lengua de interés sin usar la lengua de contacto. En el vídeo de abajo, acerca del cual escribí en mis notas (4), un hombre sentado lanza un libro sobre el piso. Los vídeos también vienen con instrucciones de cómo elicitar.

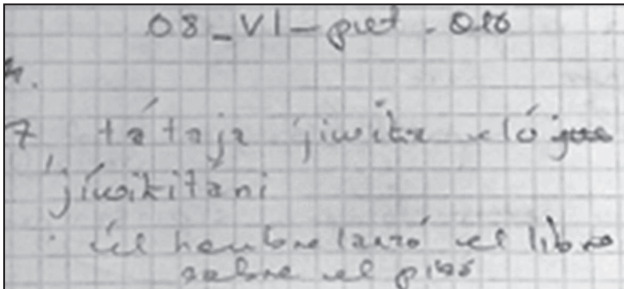
2 <http://fieldmanuals.mpi.nl/>

Figura 11

Video de Bowerman et al. (2004) del proyecto de poner ('put project')



(4)



tata-ja jiwika eló ji-wiki-tā-ni
 hombre-ERG libro sobre_piso lanzar-MOVIENDO:P-3A-SENTAR
 'El hombre (sentado) lanzó el libro sobre el piso.'

Video 08_V1_put_010 (Bowerman et al. 2004), 2019-05-17, Tsimi Matawa

Otro método, que va más allá de la elicitación básica, es *la elicitación inducida por contexto* (Matthewson, 2004). Consiste en crear contextos imaginarios y preguntar cuáles oraciones serían (no)apropiadas, (no)lícitas en esos contextos. Normalmente se usa esta metodología para investigar la semántica y la pragmática (Bochnak & Matthewson, 2015).

Después de tener un conocimiento aproximado de la semántica de una construcción o morfema, se busca saber detalles más sutiles viendo si se pueden crear contextos donde el hablante insista en que dicha construcción no es apropiada. Por ejemplo, en chácobo, los verbos se marcan por varios morfemas que expresan la distancia temporal (Tallman & Stout, 2016, 2018; Tallman, 2018b).

- (5) a. <caníquē>
 ka=ní=ki
 ir=PASREM=DEC:PAS
 ‘Él/ella se fue (hace un año).’
 ‘Él/ella se ha/había ido (hace un año).’
- b. <cayaméquē>
 ka=yamí(t)=ki
 ir=PASDIST=DEC:PAS
 ‘Él/ella se fue (hace un mes).’
 ‘Él/ella se ha/había ido.’
- c. <caitáquē>
 ka=ʔitá=ki
 ir=PASREC=DEC:PAS
 ‘Él/ella se fue (ayer o anteayer).’
 ‘Él/ella se ha/había ido.’
- d. <cáquē>
 ka=ki
 ir=DEC:PAS
 ‘Él/ella se fue (hoy día o no se sabe cuándo).’
 ‘Él/ella se ha/había ido.’
- e. <cátsiqui>
 ka=tsi=ki
 ir=AHORA=DEC:NOPAS
 ‘Él/ella se va (ahora).’
- f. <casharíki>
 ka=fari=ki
 ir=CRAS=DEC:NOPAS
 ‘Él/ella irá (mañana).’
- g. <caxéquē>
 ka=ʃí=ki
 ir=REMP=DEC:NOPAS
 ‘Él/ella irá (no sé cuándo).’

Entre las preguntas que uno podría hacer sobre los marcadores de distancia temporal están: ¿Qué hace el hablante cuando falta información sobre cuándo pasó

la acción? ¿Cuál es la relación entre el conocimiento del hablante y el marcador de distancia temporal? ¿Qué marcador usa el hablante cuando no tiene información sobre cuándo pasó el evento, por ejemplo?

La manera más fácil de crear los contextos es leer literatura sobre semántica, que usa elicitación inducida por contexto y crear el contexto análogo. Por ejemplo, creé contextos para la elicitación basada en Cable (2013). Para saber cuál oración usaría un hablante cuando falta información sobre cuándo pasó el evento, construí el contexto siguiente basado parcialmente en contextos en Cable (2013) y en los asuntos reales de los chácobo mismos.

(6) **Contexto:** Tu suegro tiene una moto que te dio a vos. Vos no sabes cuándo él la compró, pero recién se fregó. En la tienda donde vos compraste tu moto, dicen que pueden arreglar tu moto gratis por un año después de comprar una moto. Vos quieres que ellos arreglen tu moto gratis, pero no sabes si se puede porque no sabes cuándo la moto fue comprada. Solamente los dueños de la tienda tienen esta información.

Después de que el hablante de chácobo entendió el contexto, le pregunté cuál sería la oración más apropiada en chácobo, como traducción de ‘Mi suegro compró esta moto de aquí, y yo quiero que ustedes la arreglen.’, según el contexto en (6). Caco Moreno (profesor y autor chácobo) dice que (7a-c) no son apropiadas según el contexto en (6) porque implican que el hablante sabe cuándo su suegro compró la motocicleta. Entonces, cuando a uno le falta información sobre cuándo pasó el evento modificado por el marcador de distancia temporal, se usa la forma hodierna, que es la opuesta a la del patrón que fue documentado en Gikuyu (Bantu) por Cable (2013).

- (7) a. #<něaca noho raisí moto copiniquě, ebaxrí jiahuatí mětsa ni mato>
 niaka noʔo raisi= moto kopi=ní=ki
 aquí 1sg:gen suegro=ERG moto comprar=REM=DEC:PAS
 ibaʂri hia-wa=ti mitsa ni ma-to
 1SG=BENEF=AUG arreglar-TR=PURP POSS INTER 2PL-EPEN
- b. #<něaca noho raisí moto copiyaměquě, ebaxrí jiahuatí mětsa ni mato>
 niaka noʔo raisi= moto kopi=yamí(t)=ki
 aquí 1sg:gen suegro=ERG moto comprar=DIST=DEC:PAS
 i=baʂ=rí hia-wa=tí mitsa ni ma-to
 1SG=BENEF=AUG arreglar-TR=PURP POSS INTER 2PL-EPEN
- c. #<něaca noho raisí moto copihitaquě, ebaxrí jiahuatí mětsa ni mat>o
 niaka noʔo raisi= moto kopi=?itá=ki
 aquí 1SG:GEN suegro=ERG moto comprar=REC=DEC:PAS
 ibaʂri hia-wa=ti mitsa ni ma-to
 1SG=BENEF=AUG arreglar-TR=PURP POSS INTER 2PL-EPEN

- d. <nēaca noho raisi moto copiquē, ebaxrí jiahuatí mētsa ni mato>
 niaka noʔo raisi= moto kopí=ki
 aquí 1sg:gen suegro=ERG motocicleta comprar=DEC:PAS
 ibaʃri hia-wa=ti mitsa ni ma-to
 1SG=BENEF=AUG arreglar-TR=PURP POSS INTER 2PL-EPEN
 ‘Mi suegro compró esta moto de aquí, y yo quiero que ustedes la arreglen.’

A veces, los hablantes también ofrecen contextos que corresponden a oraciones elicítadas. Es importante escribir los contextos que son ofrecidos. Por ejemplo, Caco Moreno dio el contexto en (8b) para diferenciarlo mejor del (8a).

- (8) a. <caca baha bahahuayaquē noho ēhuatí>
 kaka baʔa~ baʔa-wa=yá=ki
 canasto fabricar~ fabricar-TR=PERF=DEC:PAS
 noʔo iwatí
 1SG:GEN abuela
 ‘Mi abuela ya estaba fabricando varios canastos.’
 b. <caca bahahuahi noho ēhuati iquē>
 kaka baʔa-wa=?i noʔó
 canasto fabricar-TR=CONCUR:S 1SG:GEN
 iwatí i=ki
 abuela AUX.ITR=DEC:PAS
 ‘Estaba fabricando canastos.’
Contexto: Cuando uno pasa y pregunta qué estaba haciendo.

1.2.2.2. Guiones gráficos

Los guiones gráficos (*storyboards*) de construcción dirigida consisten en representaciones pictóricas de historias, que se les pide a los consultados que describan con sus propias palabras. Se utilizan para reducir la influencia del metalenguaje en la obtención de construcciones específicas (Burton y Matthewson, 2015). Los datos de los guiones gráficos tienen una naturalidad intermedia entre los datos de obtención y el habla natural. Permiten al lingüista obtener construcciones específicas dentro de un contexto de discurso más natural, que no dependa de la casualidad en la elicitación o de problemas en la interpretación de los datos obtenidos. Desarrollé guiones gráficos a partir de los de Totem Field <http://www.totemfieldstoryboards.org/stories/>. Utilicé y modifiqué los guiones existentes para investigar la semántica del tiempo, el aspecto, el número y la modalidad.

El cuento ‘Adivina’ cuenta la historia de una mujer que busca a una adivina para preguntarle con quién debería casarse de entre dos posibles pretendientes, un hombre alto y un hombre gordo; se muestra en la **Figura 12**.

Figura 12

Guiones gráficos de la historia ‘Adivina’



El hablante araona Fride Matawa dio las siguientes traducciones para los guiones de arriba:

- (9) *Dea* *abaopoa* *po-nae* *ja-jemi-ki-jae*
 hombre alto 3-con I NTRC-coger/casar-INTRC-querer
e-a-ta-ni
 EPEN-AUX.TR-3A-SENTADO
 ‘El hombre alto quiso casarse con ella sentado.’
- (10) *Dea* *peada-ja* *mabaopo-a* *pea* *jicho-ta-ni*
 hombre otro-ERG gordo-ERG uno molestar-3A-SENTADO
ja-jemi-ki-jae *e-a-ta-ni* *wada-dae*
 INTRC-casar-INTRC-querer EPEN-AUX.TR-3A-SENTADO 3-igual
dipa *isha*
 así otra.vez
 ‘El otro hombre que es chaparro la molesta, quiero estar con vos, dice igual.’

La historia ‘Adivina’ elige construcciones de modalidad y probabilidad. Cuando la mujer protagonista pregunta a la adivina, le da diferentes posibilidades de su futuro según el hombre que escoja. La traducción da ejemplos de modalidad que podemos relacionar con el contexto del cuento en vez de confiar solo en la traducción en castellano.

Figura 13

Guiones 10 (corresponde a (11)) y 12 (corresponde a (12)) de la ‘Adivina’ (‘fortune teller’)



La oración dada en (12) sugiere que en araona no hay un marcador condicional dedicado (como *si* en castellano más la forma subjuntiva), solo que la prótasis está marcada con el marcador de ‘enfoque’ *-mo*.

- (11) *Kwipa tso mia ba-me joda dea nae*
 como ANT 2SG ver-CAUS ese hombre con
ema ja-jemi-ki-tso
 1SG INTRC-casar-INTRC-ANT
 ‘Cómo vos decís si me caso con ese hombre?’ (corresponde a guion 10)

- (12) *Joda dea abao nae ja-jemi-ki-tso-mo bakwa*
 ese hombrealto con INTRC-coger-INTRC-ANT-FOC hijo
apamo e-a-ni-mi
 mucho EPEN-hacer-SENTADO-2SG
 ‘Si te casas con ese hombre alto, vas a tener muchos hijos.’ (corresponde a guion 12)

Los guiones de Totem Field evocan contextos que pueden ser culturalmente incoherentes para unas comunidades de hablantes. Por eso, el investigador tiene que adaptar un poco las historias para que sean comprensibles. Un área de investigación importante es el desarrollo de *storyboards* culturalmente específicos. Marine Vuillermet ha desarrollado un guion gráfico que se enfoca en las construcciones de movimiento asociado, que es apropiado para un contexto amazónico.

El uso del guion gráfico de Vuillermet y Desnoyers (2013) era importante para establecer que el araona tiene marcadores de movimiento asociado del objeto, que no fueron documentados o bien descritos antes (Tallman, 2020), aunque morfemas parecidos fueron descritos para otras lenguas Tacana (Guillaume, 2008; Vuillermet, 2013). Se puede observar claramente que *-wiki* indica el movimiento del capibara ‘zododo’, objeto de *pisa* ‘cazar’ en (14).

Figura 14

Guion 10 a partir del guion gráfico de Vuillermet y Desnoyers (2013) ‘Yendo a cazar / a hunting story’



- (13) *Ainimetsio* *zododo* *po-bea-yoa*
 repentinamente capibara ir-venir-deambulando
tsabaja-ta-iki
 escuchar-3A-PAS
 ‘Repentinamente él escuchó que venía el capibara.’
- (14) *Datsio* *pisa-ta-ni* *betakata*
 entonces tirar-3A-SENTADO dos
pisa-wiki-ta-ni *tso* *jododo*
 cazar-YENDO.OBJ-3A-SENTADO ANT correr
 ‘Tiró a dos, (pero) tiró a uno que se fue corriendo.’

Desde el contexto de las lenguas bolivianas, un área importante de investigación futura es el desarrollo de *storyboards* de construcciones específicas que sean culturalmente apropiadas para las lenguas indígenas bolivianas.

1.2.2.3. Entrevistas dinámicas

Las entrevistas dinámicas involucran a dos o más participantes, donde uno de ellos hace preguntas relacionadas con las actividades culturales comunes de interés de la lengua de la comunidad de hablantes (ej. ¿Cuándo fue la última vez que fue a pescar? ¿Qué pescó?). Estas preguntas están destinadas a provocar narrativas a diferentes distancias del tiempo del discurso. (¿Qué estabas haciendo hace un mes? ¿Qué estabas haciendo hace dos semanas? ¿Qué planeas hacer mañana / en una semana?), y preguntas que están destinadas a obtener las perspectivas de los oradores (Cuenta una historia que escuchaste que no creas que es verdad y explica

por qué. ¿Alguna vez has visto un fantasma? ¿Cómo supiste que era un fantasma?). Todas las preguntas y respuestas tienen que estar hechas en la lengua. Las preguntas pueden ser sobre historias que ya habían sido transmitidas a mis consultores entrevistadores anteriormente. Se instruyó a los consultores entrevistadores para que intenten mantener un diálogo con los otros oradores a fin de grabar un discurso más conversacional, pero que provoque construcciones específicas. Si bien tales entrevistas estaban destinadas a capturar un discurso más interactivo, debe tenerse en cuenta que es difícil separar los registros del habla interaccional y los monólogos en mi corpus. La idea básica de una entrevista dinámica es provocar ciertas construcciones de interés, pero en un contexto a la vez ecológicamente válido. El entrevistador recibe instrucciones específicas para tratar de guiar la conversación en ciertas direcciones, pero no se mantiene ningún control sobre el discurso resultante, como ocurre con los guiones gráficos. Las entrevistas dinámicas son, por lo tanto, un paso más hacia el habla natural en comparación con los guiones gráficos.

Un ejemplo de los resultados de una entrevista dinámica está dado en (15). Aquí Paë Yaquë Roca está guiando la narrativa de Jëma Chávez (su madre). Primero le pregunta qué hizo ayer en (15a), que condiciona la aparición del marcador del pasado reciente =*ɣita* en (15b-e); se ve que en (15f) Jëma cambia a la construcción ‘hodierna’ que no tiene un marcador de distancia temporal. Ese diálogo muestra precisamente la frontera entre el pasado reciente y el pasado hodierno (Tallman & Stout, 2016; 2018; Tallman, 2018).

- (15) a. PAË YAQUË:
 <jisa caí jahuë mi aitaána naa bari caitahacaton>
hisa kai=’ hawi mi a(k)=ita=ɣa=na
 mira madre=VOC qué 2SG hacer=REC=NLZ=EPEN
naa bari ka=ita=ɣa=kato=’no
 DEM1 día/sol ir=REC=NLZ=NLZ=SPAT
 ‘Mira mamá, ¿qué hiciste ayer o uno de estos días que pasaron?’
- b. JËMA CHÁVEZ:
 <xobo mihpa ano ëmë biitaquë>
ɣobo mi-ɣ-ipa a(k)=no i-mí
 casa 2SG-EPEN-padre hacer=CONCUR 1SG-RFLX
bi=ita=ki
 coger=REC=DEC:PAS
 ‘Cuando tu papá estaba haciendo la casa, yo misma estaba recogiendo (frutas de patujú).’

- c. <jaquí ě maní raahitaquē>
ha=kí i maní raa=?ita=ki
 3=DAT 1SG patujú mandar=REC=DEC:PAS
 ‘Yo estaba mandando frutas de patujú.’
PAĚ YAQUE:
- d. <japá>
ha-pá
 3-MIR
 ‘Verdad, ¿y entonces!?’
- e. JĚMA CHÁVEZ:
 <Jatsi no quēyohitaquē>
hatsi no kiyo=?ita=ki
 entonces 1PL terminar=REC=DEC:PAS
 ‘Y terminamos.’
- f. <quēyotaxó tsi no oxashináquē>
kiyo-ta(n)=só tsi no oša=finá=ki
 terminar-PNCT=PREVIO:A ENL 1PL dormir=NOCHE=DEC:PAS
 ‘Al momento que terminamos, dormimos.’
- g. <oxahá huēayohá tsi jahuē atēquēhiní noquirí naa nohó xobo pístiaparí quēyotēquēno>
oša=?á wia=yo=?á tsi hawi
 dormir=NLZ amanecer=CMPL=NLZ ENL qué
a(k)=tiki(n)=?iní no-ki=rí náa
 hacer=otra.vez=INTER:NOPAS 1PL-EPEN=AUG DEM1
no?ó sobo pístia=parí kiyo=tiki(n)=no
 1SG:GEN casa pequeño=primero terminar=otra.vez=DES
 ‘Cuando dormimos, amanece; “¿Qué hacemos otra vez?” Después teníamos que terminar la casa.’
- h. <mani raayóquē>
maní raa=yó=ki
 patujú mandar=CMPL=DEC:PAS
 ‘Él dejó todo el patujú (hoy día).’ TXT 093:007

Para hacer una entrevista dinámica, el entrevistador tiene que ser hablante nativo o hablar con fluidez. A la vez, necesita comprender la construcción a la que se dirige. Eso solo puede pasar con un cierto grado de formación en lingüística, que puede surgir con el tiempo a través de sesiones de elicitación. Por ejemplo,

las entrevistas que hizo Paë Yaquë (visto en la **Figura 15**) con la comunidad de Cachuelita fueron hechas después de sesiones de elicitación intensivas en las que Paë y yo desarrollamos una apreciación de la complejidad de los temas relevantes que rodean el tiempo, el aspecto y la distancia temporal en chácobo y pachahuara.

Figura 15

Paë Yaquë Roca haciendo entrevistas dinámicas con comunarios de Cachuelita (2015)




1.2.2.4 . Habla natural

Una vez que uno se ha orientado, el tipo de datos más crucial para un registro documental son los datos del habla natural. Se reconocen diferentes géneros (**Tabla 2**) de habla en diferentes comunidades de habla. En general, estos varían de acuerdo con lo formulados o estructurados que sean. El arte verbal tiende a tener las estructuras retóricas más modeladas (rimas, líneas, estrofas, etc.), pero parece haber una variación intercultural según en qué grado se dan. Además, a menudo se asume que los géneros de habla más formulados o interpretativos son monológicos. Sin embargo, esto tampoco es cierto, ya que las interacciones entre el hablante y el oyente pueden estar muy estructuradas en ciertos tipos de géneros de habla, como en la narración de mitos.

Los géneros en la Tabla 2 son puntos de partida para la investigación. Los géneros de habla son culturalmente específicos y la clasificación debe estar basada en las diferencias que están reconocidas por la comunidad de hablantes.

Tabla 2*Grados de estructuración retórica según los géneros de habla*

Canciones / Poesía	 Grado de estructuración / formulacidad retórica
Mitos tradicionales	
Instrucciones tecnológicas / culturales / gastronómicas	
Cuentos folklóricos	
Narrativas experimentadas	
Conversaciones “libres” / chismes	
Contextos de habla de baja concentración	

Para dar una ilustración de la especificidad cultural de ciertos géneros de habla, podemos considerar la narración de mitos en chácobo. Los hablantes no memorizan los mitos, sino que los recuerdan. Para contar mitos es crucial un “oyente de mitos”, que metodológicamente responde (*‘back channels’*) con *ë* después de cada tres grupos de pausa proporcionados por el narrador principal. El narrador responde con *jm*. Esto se puede ver en el ejemplo (16) a continuación. El narrador de mitos es Ona Ortiz y el oyente de mitos es Boca Chávez.

- (16) ONA ORTIZ:
- a. < tóa picaná ca ...>
‘Entonces cuando les comieron.’
 - b. <jarohá átsa tsi naarisqui noa,>
“‘Eso es! Necesitamos yuca!’”
 - c. <i quiha ja caníquë>
‘Ellos dijeron.’
- BOCA CHÁVEZ:
- d. <ë!>
‘Y!’
- ONA ORTIZ:

- e. <jm!>
'Jm!'
- f. <tóa átsa tsi náarisqui noa,>
“‘Es yuca lo que nos falta!’”
- g. <jěnáhua no huaná>
“¿Qué hacemos?”
- h. <i tsi quiha / ja caníquē>
'Ellos dijeron.'
BOCA CHÁVEZ:
- i. <ě!>
'Y!'
ONA ORTIZ:
- j. <jm!>
'Jm!'

Es muy importante documentar conversaciones más libres. Las conversaciones tienen una estructura y hay ciertos morfemas y construcciones que solo pueden ser entendidos en el contexto de una conversación (véase Zariquiey, 2018 para un ejemplo interesante de kakataibo). Para entender las fórmulas que distinguen géneros de habla, tienen que ser comparadas entre sí.

Hay unos contextos donde la gente no habla mucho —cazar, pescar, cosechar, cocinar—. De todas formas, es importante grabar esos contextos, aunque el habla es de baja concentración. Documentar contextos de habla de baja concentración es importante para la documentación de gestos. También, después de grabar, uno puede mostrar los vídeos a los hablantes para que den una descripción de lo que está pasando en la lengua.

1.3. Descripciones lingüísticas

La sección 1.2.1. resumió la tecnología de recopilación, anotación y organización de datos. Toda anotación y organización lingüística depende de un análisis que presupone una organización descriptiva de los hechos. La sección 1.2.2. resumió diferentes técnicas de elicitación y tipos de discurso que pueden recopilarse en función de cómo y hasta qué punto el contexto comunicativo es manipulado por el investigador y / o controlado por los hablantes. Todos los métodos descritos anteriormente están motivados por cuestiones teóricas que surgen cuando se quiere describir un idioma. Como sostiene Himmelmann, la descripción gramatical es un aspecto importante de la documentación del lenguaje.

Himmelman (2007) propone una estructura para documentar lenguas que contiene un análisis descriptivo, una etnografía, un diccionario y una descripción gramatical como partes asociadas.

Tabla 3

Formato ampliado de documentación lingüística (Himmelman 2007, p. 38)

Datos primarios	Aparato	
grabaciones / registros de comportamiento lingüístico observable y conocimiento metalingüístico	Por sesión	Para toda la documentación
	Metadatos	Metadatos
	Anotación	Recursos de acceso general
	transcripción	introducción
	traducción	convenciones ortográficas, convenciones de glosa, índices vínculos a otros recursos
	otras glosas y comentarios lingüísticos y etnográficos	Análisis descriptivo Etnografía Gramática descriptiva Diccionario

Contra el argumento de que la descripción no es importante en la documentación, Himmelman explica: "...la producción de una gramática descriptiva es una parte necesaria de la documentación lingüística porque sin un análisis lingüístico ciertos aspectos del sistema lingüístico no estarían documentados" (Himmelman, 2006, p. 23).

Al no intentar generalizar los datos y organizarlos en una gramática descriptiva, el investigador podría perder de vista posibles lagunas en el registro documental. Una documentación lingüística tiene que ser producida con una anotación que presuponga una clasificación y descripción de los datos. El documentalista, por tanto, debe abordar cuestiones como: ¿Cómo se definen conceptos como 'verbos', 'sustantivos', 'afijos', 'raíces'? ¿Cómo sabemos que tenemos marcadores que codifican conceptos del tiempo (pasado, presente, futuro), aspecto (progresivo, habitual, imperfectivo), etc.?

En otras palabras, el documentalista tiene que navegar por la tarea arriesgada de

mapear conceptos desde la lingüística general hasta las categorías expresadas en un idioma en particular. Siempre que un investigador hace esto, existe un doble riesgo de tergiversar las categorías indígenas o de extender de manera inapropiada categorías de la lingüística general más allá de sus límites intuitivos. Y si el investigador elude la tarea de abordar conceptos de la lingüística general, se le puede acusar de ofuscación, exotización o ignorancia. El lingüista descriptivo y documentalista está siempre destinado a disgustar a alguien con sus decisiones analíticas o terminológicas.

Pero, ¿de dónde vienen esos conceptos de la lingüística general? Desde los años 50, después de la revolución chomskiana en la lingüística, hubo una tendencia a entender categorías en las lenguas del mundo como variantes de conceptos universales. Y con ello venía la tendencia de cosificar conceptos gramaticales que provienen de tradiciones de descripción y de pedagogía de las lenguas europeas. La perspectiva más moderna de la descripción y documentación toma más seriamente los problemas de mapear conceptos de la lingüística general sobre las categorías de una lengua en particular; sobre todo, se toma en serio la posibilidad que estos conceptos tienen una predisposición de mapear de una manera particularmente incongruente a lenguas no europeas (véase Epps, 2010).

La descripción, clasificación y anotación lingüística tienen que ser entendidas como aproximaciones, pero el investigador debe tener la libertad de rechazar conceptos que provienen de la lingüística general para proponer una mejor organización de conceptos intercalados que reflejen la estructura de una determinada lengua. Lo que ha llegado a conocerse como “tipología y descripción multivariante” ha derribado cada vez más los muros escolásticos y platónicos impuestos por la lingüística tradicional (Bickel, 2010).

En vez de reificar categorías lingüísticas tradicionales, se intenta dismantelar categorías lingüísticas hasta su nivel más fino de observación, hasta sus diagnósticos de justificación. Esos diagnósticos se vuelven “variables tipológicas” y explicaciones de correlaciones entre variables dentro y fuera de las lenguas es el dominio de la teoría lingüística (Bond, 2010; Bickel, 2011).

Por ejemplo, tradicionalmente la morfología hace una diferencia entre la derivación y la flexión (Anderson, 1985; Bybee, 1985). Algunas propiedades están resumidas en la **Tabla 4**. Muchos presuponen que la derivación y la flexión son parte de la formación de palabras, que son unidades sintácticas que no pueden ser manipuladas por procesos sintácticos.

Tabla 4*Propiedades de flexión y derivación*

Propiedad	Flexión	Derivación
Distribución	Tiene distribución más fija	Tiene distribución más libre
Opcionalidad	Obligatorio	Opcional
Paradigmaticidad	Se organiza en paradigmas	No se organiza en paradigmas, sino se aplica incrementalmente
Relación con la sintaxis	Indica una dependencia sintáctica	Es interna a la palabra, no tiene relación con la sintaxis.

Normalmente la flexión ocurre más cerca de la raíz. Por ejemplo, en la palabra castellana *habl-ador-es*, el adjetivador *-ador* viene antes del marcador de plural *-es*. Cuando se adopta una visión tipológica más amplia, parece que hay categorías gramaticales que no se clasifican claramente como derivacionales o flexivas. No se trata simplemente de reconocer unas pocas construcciones liminales que traspasan la frontera entre derivación y flexión por razones diacrónicas.

Podemos considerar la categoría de **movimiento asociado** que se ha argumentado como una propiedad territorial de la Amazonía (Guillaume, 2016; Guillaume & Koch, en prensa). El movimiento asociado es una categoría que asocia un movimiento con un verbo principal. Guillaume (2006) y Zariquiey (2018) describen la misma categoría de movimiento asociado como ‘flexión’ y como ‘derivación’, respectivamente. Entonces, ¿cuál es? La verdad es más compleja porque, cuando aplicamos los diagnósticos, no todo converge sobre una respuesta. Podemos mostrar esto examinando los diagnósticos que distinguen entre flexión y derivación en las categorías relevantes. Podemos demostrar fácilmente que los diagnósticos no apuntan a una respuesta clara, son en realidad muy ambiguos, lo que permite a los investigadores seleccionar las versiones de las pruebas que se alinean con cualquier decisión que deseen.

Primero, debemos considerar **los criterios de distribución**. Según este criterio, los elementos flexivos son más fijos que los elementos derivacionales. Sin embargo, una evaluación de este criterio depende de qué elemento se selecciona como pivote con el que evaluar la fijación. Los morfemas de movimiento pueden variar su orden con algunos morfemas y no con otros. Por ejemplo, $\text{=}\beta\text{oná}$ ‘yendo’ puede variar con el morfema =kas ‘desiderativo’, como se puede ver en (17) y (18).

- (17) <tsayabonacásquë>
 tsaya= β ona=kás=ki
 ver=yendo.TR/PL=querer=DEC.PAS
 ‘Él quiso sembrar.’

- (18) <tsayacasbonáquë>
 tsaya=kas=βona=ki
 ver=querer=yendo=DEC:PAS
 ‘Él quiso sembrar.’

Pero el marcador de movimiento asociado no puede variar con el causativo como se ve en (19) y (20). Tiene que ocurrir a la derecha de ese morfema.

- (19) <tsayamabonáquë>
 tsaya=ma=βona=ki
 ver=hacer=yendo=DEC:PAS
 ‘Él le hizo sembrar.’

- (20) *<tsayabonamaquë>
 tsaya=βona=má=ki
 ver=yendo=hacer=DEC:PAS
 ‘**Él le hizo sembrar.’

A continuación, podemos considerar el **diagnóstico de opcionalidad**. Según este diagnóstico, la flexión es obligatoria y la derivación es opcional. Nuevamente, en la medida en que seamos investigadores honestos, notaremos que el diagnóstico es ambiguo. ¿En cuál contexto es obligatorio y según cuál definición? En chácobo no son necesarios para construir una oración ni para construcciones verbales. Pero son obligatoriamente expresados de una manera redundante en el contexto donde el participante está de ida (Tallman, 2018, en prensa).

- (21) <nějo nějo tsëquë jihui niatimaxëni quiha bona mër**abonahiquiha** pib**onahiquiha** noho yochi yoi mër**abonaquia**>

niho	niho	tsiki	hiwi	nia-timaşini	kiá	
este_lado		este_lado	salir	árbol	echar-ADVLZ.NEG	REP

βona	mira=βona=ʔikiá	pi=βona=ʔikiá	noʔó	
hormiga	buscar=YENDO=REP	comer=YENDO=REP	1SG:GEN	
yotʃi	yoi	mira=βona=ki=a		
ají	precioso	buscar=YENDO=DEC:NOPAS=1SG		

‘Se fue moviendo lado a lado por el árbol, nunca echando nada, buscando (**yendo**) hormigas comiéndolas (**yendo**) de paso. “Estoy buscando mi querido ají (**de ida**)”’.

En coordinación asindética son obligatorios sobre el primer verbo cuando el segundo verbo es un verbo de movimiento.

- (22) <abayá tsi kiá jabaníquē>
 a(k)=[**bayá**]_{v1} tsi kiá [**haβa**]_{v2}=ní=ki
 matar=HACER EIR:TR/PL ENL REP correr=REM=DEC: PAS
 ‘Él lo mató, se fue corriendo (se dice).’ TXT 0025:0022

Entonces, muchos de los criterios son ambiguos o tienen que ser fracturados en más de un criterio. Unos criterios convergen en identificarla como derivacional y otros la identifican como flexión. La **Tabla 5** da un resumen de los criterios de derivación/flexión aplicados al movimiento asociado en chácobo.

Tabla 5

Resultados de los criterios aplicados al movimiento asociado en chácobo

Criterio		
i.	Fijado / Libre = <i>má'</i>	Flexión
ii.	Fijado / Libre = <i>kás</i>	Derivación
iii.	Obligatorio en construcciones verbales	Derivación
iv.	Obligatorio en coordinación asindética	Flexión
v.	Marcador de dependencia	Flexión
vi.	Estructura paradigmática	Flexión
vii.	Ocurre antes de ‘la flexión’	Derivación

El criterio vii hace referencia a otros marcadores de flexión. En las lenguas pano normalmente la flexión es entendida como un sufijo o combinación de sufijos que indican tiempo, aspecto y tipo de cláusula (Zariquiey y Valenzuela, en prensa). Todos los elementos entre la raíz y este marcador de inflexión son candidatos para el estado de derivación. En el siguiente ejemplo, -quē se consideraría el elemento de flexión, mientras que tēquēn se consideraría un marcador de derivación.

- (23) <joni siri ahui pistiá tsayatēquēquē>
 honi síri awi pistia=’ tsaya=tikí(n)=ki
 hombre viejo esposa pequeña=ERG mirar=otra_vez=DEC: PAS
 ‘La esposa miró al hombre viejo otra vez’.

Por lo tanto, debido a que aparecen marcadores de movimiento asociado entre la raíz del verbo y el ‘sufijo’ inflexivo final, el criterio vii los identificaría como derivacionales, siguiendo la descripción de Zariquiey (2018) de movimiento asociado en kakataibo.

Un aspecto interesante de muchos de los denominados sufijos derivados de las lenguas Pano es que en realidad no es necesario que aparezcan entre la raíz y el sufijo flexional final. Consideremos la oración en (24) a continuación, donde tēquēn se ubica en la segunda posición.

- (24) <joni siri tēquē ahui pistiá tsayaquē>
 honi siri=tikí(n) awi pistia= tsaya=ki
 hombre viejo=otra_vez esposa pequeña=ERG mirar =DEC:PAS
 ‘La esposa miró al hombre viejo otra vez.’

Si bien los morfemas de movimiento asociado no muestran esta flexibilidad en el ordenamiento, el morfema =*tēquēn* parece mostrar que el criterio vii es, de hecho, ambiguo y debería fragmentarse en más de un diagnóstico. Pero los problemas analíticos con los diagnósticos de flexión / derivación no pueden resolverse tan fácilmente.

En primer lugar, los casos en los que los morfemas sintácticamente flexibles pueden interrumpir la raíz y su aparente marcador de inflexión deberían, siguiendo el criterio de la palabra, hacernos cuestionar si el tramo de estructura que contiene la raíz del verbo y el marcador de flexión es una palabra morfosintáctica. De hecho, en chácobo, el “sufijo” de flexión puede ser interrumpido por un sintagma nominal completo.

- (25) <Caxáquē rabi>
 kaşaki raβi
 jugar=decl:pasado Rabi
 ‘Rabi ha jugado.’

- (26) <Caxá rábiquē>
 kaşa raβi ki
 jugar Rabi decl:pasado
 ‘Rabi jugó (e.j. fútbol)’

¿Nuestro sufijo flexivo candidato es un elemento morfológico o no? Resulta que intentar determinar si un morfema es flexivo o derivativo presupone una distinción entre morfología y sintaxis. Si bien podríamos manifestar que el problema puede resolverse evocando un tipo de morfología de ‘nivel de frase’ (Anderson, 1990, 2005), esto solo introduce otro problema de cómo distinguir entre morfología de frase (‘sintaxis especial’) y sintaxis normal (Tallman, 2018b). Si terminamos rechazando la distinción entre morfología y sintaxis, rompiendo las nociones estructurales en sus diagnósticos y criterios y mostrando que no se alinean a ellos (ver Tallman, 2020a, 2020b; Tallman & Epps, 2020; Tallman & Auderset, en preparación), necesariamente asignamos un significado diferente a los conceptos de ‘flexión’ y ‘derivación’.

Deben describirse los resultados de los criterios que supuestamente nos ayudan a describir la distinción entre flexión y derivación. Pero este ejercicio muestra que las

categorías de la lingüística tradicional no siempre pueden reunirse para ayudarnos a describir con precisión las categorías de una lengua para las que no fueron diseñadas. Como dice Epps:

Ese es el problema básico de tipología: necesitamos comparar las estructuras detrás de las lenguas, pero las estructuras mismas varían de lengua a lengua. Asimismo, eso es un problema para la descripción lingüística: no podemos tener confianza en nuestra metalengua – no podemos razonablemente llamar algo un “verbo” o un “adjetivo” sin aceptar que esas categorías tienen alguna validez tipológicamente. (*mi traducción de Epps 2010, p.6*)

Podemos categorizar la lingüística tradicional y sus diagnósticos como un andamiaje para desarrollar preguntas de investigación. Pero cualquier tipo de identificación de tales categorías en la lengua de interés debe considerarse preliminar e intrínsecamente imprecisa. El fenómeno de interés para el tipologista multivariante es cómo los propios diagnósticos (las variables tipológicas) se correlacionan entre sí y cuáles son las explicaciones para esas correlaciones.

La pregunta no es si podemos racionalizar la existencia de conceptos de la gramática tradicional mapeándolos en lenguas basadas en una selección arbitraria de diagnósticos. El arte de la descripción no consiste solo en aplicar los diagnósticos, sino en desentrañar su ambigüedad en la aplicación. Las elecciones terminológicas deben entenderse como andamios bastante arbitrarios, contruidos alrededor de los observables lingüísticos con fines expositivos (véase Bickel, 2011; Haspelmath, 2010; 2020).

1.4. Conclusión

El estudio de la variación lingüística y las teorías que explican las coincidencias entre los comportamientos observables lingüísticos constituyen una fuerza productiva en el desarrollo de métodos y herramientas de documentación. Por ejemplo, los guiones gráficos y los vídeos del Instituto de Max Planck fueron hechos para comprobar unas teorías lingüísticas y después para investigar la variación lingüística (véase Epps, 2012). Son útiles porque nos hacen pensar en preguntas que no se nos habían ocurrido antes.

Para concluir, la documentación misma tiene una teoría, y muchas metodologías siguen desarrollándose mediante la interacción con la lingüística general. Pero, para avanzar, tiene que estar reconocida como un campo de investigación por sí mismo, que tiene cierto grado de autonomía, no solo por razones sociales, políticas y humanísticas, sino también por razones científicas. Una documentación impulsada por cualquier teoría lingüística va a tener la tendencia de codificar las categorías en la lengua en cuestión, en vez de investigarlas. Además, la lingüística sin la documentación va a tener la tendencia de prejuzgar la veracidad en sus teorías, en lugar de comprobarlas a través de las lenguas del mundo.

No hay nada más anticientífico que hablar despectivamente sobre el trabajo documental y descriptivo como “*meramente* descriptivo”, como si la verificabilidad y la oportunidad de pensar críticamente a través de nuevos datos no fueran aspectos cruciales de la práctica científica.

Referencias

- Anderson, S. (1985). Inflectional morphology. In T. Shopen (Ed.), *Language typology and syntactic description (Volume III): Grammatical categories and the lexicon* (pp. 151-201). Cambridge: Cambridge University Press.
- Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Anderson, S. R. (2005). *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Bickel, B. (2007). Typology in the 21st century: major current developments. *Linguistic Typology*, *II*, 239-251.
- Bickel, B. (2010). Capturing particulars and universals in clause linkage: A multivariate analysis. In I. Bril (Ed.), *Clause Linking and Clause Hierarchy* (pp. 51-104). Amsterdam: John Benjamins Publishing Company.
- Bickel, B. (2011). Multivariate typology and field linguistics: a case study on detransitivization in Kiranti (Sino-Tibetan). (A. K. Peter, B. Oliver, N. David, & M. Lutz, Eds.) *Proceedings of Conference on Language Documentation and Linguistic Theory* 3, 3, 3-13.
- Bickel, B. (2015). Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 901-923). Oxford: Oxford University Press.
- Bickel, B., & Nichols, J. (2002). Autotypologizing Databases and their Use in Fieldwork. In P. Austin, D. Helen, & P. Wittenburg (Eds.), *Proceedings of the International LREC Workshop on Resources and Tols in Field Linguistics*. Nijmegen: ISLE and DOBES.
- Bickel, B., & Zuñiga, F. (2017). The ‘word’ in polysynthetic languages: phonological and syntactic challenges. In M. Fortascue, M. Mithun, & N. Evans (Eds.), *The Oxford Handbook of Polysynthesis* (pp. 158-186). Oxford: Oxford University Press.

- Bochnak, R. M., & Matthewson, L. (Eds.). (2015). *Methodologies in Semantic Fieldwork*. Oxford: Oxford University Press.
- Bohnenmeyer, J. (2015). A Practical Epistemology for Semantic Elicitation in the Field and Elsewhere. In R. M. Bochnak, & L. Matthewson (Eds.), *Methodologies in Semantic Fieldwork* (pp. 13-46). Oxford: Oxford University Press.
- Bohnenmeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips. In S. C. Levinson, & N. Enfield (Eds.), *Manual for the field season 2001* (pp. 90-96). Nijmegen: Max Planck Institute for Psycholinguistics.
- Bond, O. (2019). Canonical Typology. In J. Audring, & F. Masini (Eds.), *The Oxford Handbook of Morphological Theory*.
- Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In S. C. Levinson (Ed.), *Space stimuli kit 1.2*. Nijmegen: Max Planck Institute for Psycholinguistics.
- Bowern, C. (2008). *Linguistic Fieldwork: A Practical Guide*. London: Palgrave Macmillan UK.
- Brody, J. (1986). Repetition as a Rhetorical and Conversational Device in Tojolabal (Mayan). *International Journal of American Linguistics*, 52(3), 255-274.
- Burton, S., & Matthewson, L. (2015). Targeted Construction Storyboards in Semantic Fieldwork. In B. R. M., & L. Matthewson (Eds.), *Methodologies in Semantic Fieldwork* (pp. 135-156). Oxford: Oxford University Press.
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Company.
- Bybee, J. P. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the world*. Chicago: University of Chicago Press.
- Cable, S. (2013). Beyond the past, present, and future: towards the semantics of 'graded tense' in Gĩkũyũ. *Natural Language Semantics*, 21, 219-276.
- Cameron, D. (1995). *Verbal hygiene: The politics of language*. London: Routledge.
- Candea, M. (2018). *Comparison in Anthropology: The impossible method*. Cambridge: Cambridge University Press.
- Carlucci, A. (2015). *Gramsci and Languages: Unification, Diversity, Hegemony*. Chicago: Haymarket.
- Chelliah, S. L., & De Reuse, W. (2011). *Handbook of Descriptive Linguistic Fieldwork*. London: Springer.

- Comrie, B., Haspelmath, M., & Bickel, B. (2008). *Leipzig glossing rules: Conventions for interlinear morpheme-by-morpheme glosses*. Retrieved from Max Planck Institute for Evolutionary Anthropology Department of Linguistics: <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>
- Eagleton, T. (2008). *Literary Theory: An introduction*. Minneapolis: University of Minnesota.
- Eckert, P. (2018). *meaning and Linguistic Variation: The Third Wave in Sociolinguistics*. Cambridge: Cambridge University Press.
- Epps, P. (2010). Linguistic Typology and Language Documentation. In J. J. Song (Ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press.
- Epps, P. L., Webster, A. K., & Woodbury, A. C. (2017). A Holistic Humanities of Speaking: Franz Boas and the Continuing Centrality of Texts. *International Journal of American Linguistics*, 83(1), 41-78.
- Evans, N. J. (Ed.). (2002). *Ethnosyntax: Explorations in Grammar and Culture*. Cambridge: Cambridge University Press.
- Guillaume, A. (2006). La catégorie du mouvement associe en cavineña: Apport à une typologie de l'encodage du mouvement et de la trajectoire. *Bulletin de la Société Linguistique de Paris*.
- Guillaume, A. (2016). Associated motion in South America: Typological and areal perspectives. *Linguistic Typology*, 20(1), 81-177.
- Guillaume, A. (forthcoming). Takanan languages. In P. Epps, & L. Michael (Eds.), *Amazonian Languages. An International Handbook*. De Gruyter Mouton.
- Guillaume, A., & Koch, H. (This volume, chapter 1). *Associated Motion as a grammatical category in linguistic typology*.
- Hanks, W., & Severi, C. (2014). Translating worlds: The epistemological space of translation. *Hua: Journal of Ethnographic Theory*, 4(2), 1-16.
- Haspelmath, M. (2010). Framework-free grammatical theory. In B. Heine, & H. Narrog (Eds.), *The Oxford handbook of grammatical analysis* (pp. 341-365). Oxford: Oxford University Press.
- Haspelmath, M. (2020). The structural uniqueness of languages and the value of comparison for language description. *Asian Languages and Linguistics*.

- Hellwig, B. V., Hulsbosch, M., Somasundaram, A., Tacchetti, M., & Geerts, J. (2020). *ELAN - Linguistic Annotator, version 5.9*. Nijmegen: The Language Archive, MPI for Psycholinguistics. Retrieved from <https://www.mpi.nl/corpus/manuals/manual-elan.pdf>
- Henke, R. E., & Berez-Kroeker, A. L. (2016). A Brief History of Archiving in Language Documentation with an Annotated Bibliography. *Language Documentation and Conservation*, 10, 411-457.
- Himmelman, N. P. (1998). Documentary and descriptive linguistics. *Linguistics*, 6, 161-195.
- Himmelman, N. P. (2006). Language documentation: What it is and what is it good for? In J. Gippert, N. P. Himmelman, & U. Mosel (Eds.), *Essentials of language documentation* (pp. 1-30). Berlin: Mouton de Gruyter.
- Himmelman, N. P. (2012). Linguistic Data Types and the Interface between Language Documentation and Description. *Language Documentation and Conservation*, 6, 187-207.
- Lahaussois, A., & Vuillermet, M. (2019). Methodological Tools for Linguistic Description and Typology. *Language Documentation & Conservation*, 16.
- Lehmann, C. (2001). Language documentation: a program. In W. Bisang (Ed.), *Aspects of typology and universals* (pp. 83-97). Berlin: Akademie Verlag.
- Margetts, A., & Margetts, A. (2012). Audio and Video Recording Techniques for Linguistic Research. In N. Thieberger (Ed.), *The Oxford Handbook of Linguistic Fieldwork* (pp. 13-54). Oxford: Oxford University Press.
- Martin, L. (2000). Parallelism and the Spontaneous Ritualization of ordinary talk: Three Mocho Friends Discuss a Volcano. In K. Sammons, & J. Sherzer (Eds.), *Translating native Latin American verbal art: ethnopoeitics and ethnography of speaking* (pp. 104-124). Washington DC: Smithsonian Institution Press.
- Matthewson, L. (2004). On the Methodology of Semantic Fieldwork. *International Journal of American Linguistics*, 70(4).
- Mithun, M. (2001). Who shapes the record: the speaker and the linguist. In P. Newman, & M. Ratliff (Eds.), *Linguistic Fieldwork: Essays on the Practice Empirical Linguistic Research* (pp. 34-54). Cambridge: Cambridge University Press.
- Prost, G. (1965). Chácobo. In *Gramáticas estructurales de lenguas bolivianas II* (pp. 1-130). Riberalta: ILV.

- Prost, G. (1967). Chacobo. In E. Mattseon (Ed.), *Bolivian Indian Grammars 1* (Summer Institute of Linguistics International Publications in Linguistics 16). Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Seyfeddinipur, M. (2011). Reasons for Documenting Gestures and Suggestions for How to Go About It. In N. Thieberger (Ed.), *The Oxford Handbook of Linguistic Fieldwork*. Oxford: Oxford University Press.
- Tallman, A. J. (2018a). *A Grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon*. University of Texas at Austin.
- Tallman, A. J. (2018b). There are no special clitics in Chácobo (Pano). In N. Webber (Ed.), *Workshop on the Structure and Constituency in Languages of the Americas 21* (pp. 194-209). Vancouver: University of British Columbia Working Papers in Linguistics 26.
- Tallman, A. J. (2020). An ethnographically based linguistic documentation of Araona: A Takanan language of Bolivia. Retrieved from <https://elar.soas.ac.uk/Collection/MPI1207738>
- Tallman, A. J. (2020). Associated motion and posture in Araona (Takana). *Atelier Typologie Semantique*. Lyon: Laboratoire Dynamique du Langage (24th of January 2020).
- Tallman, A. J. (2020). Beyond grammatical and phonological words. *Language and Linguistics Compass*, e12364. Retrieved from <https://doi.org/10.1111/lnc3.12364>
- Tallman, A. J. (2020). Constituency and coincidence in Chácobo (Pano). *Studies in Language*.
- Tallman, A. J. (forthcoming). Associated motion in Chácobo (Pano) in typological perspective. In A. Guillaume, & H. Koch (Eds.), *Associated Motion. Empirical Approaches to Linguistic Typology (EALT)*. Berlin: Mouton de Gruyter.
- Tallman, A. J., & Elías-Ulloa, J. (2020). The acoustic correlates of tone and stress in Chácobo. *Journal of the Acoustical Society of America*, 147, 3028-2042.
- Tallman, A. J., & Stout, T. (2016). The perfect in Chácobo in cross-linguistic perspective. In T. Bui, & R.-R. Ivan (Eds.), *SULA 9: Proceedings of the Ninth Conference on the Semantics of Under-Represented Languages in the Americas*. University of California, Santa Cruz.

- Tallman, A. J., & Stout, T. (2018). Tense and temporal remoteness in Chácobo (Pano). In M. Keough, N. Weber, A. Anghelescu, S. Chen, E. Guntly, K. Johnson, . . . O. Tkachman (Ed.), *Proceedings of the Workshop on the Structure and Constituency in the Languages of the Americas 21* (pp. 210-224). Montreal: University of British Columbia Working Papers in Linguistics 46.
- Tedlock, D. (1977). Toward an Oral Poetics. *New Literary History*, 8(3), 507-519.
- Tedlock, D. (1983). *The spoken word and the work of interpretation*. Philadelphia: University of Pennsylvania Press.
- Uriagereka, J. (1998). *Rhyme and Reason*. Cambridge: MIT Press.
- Valle, D. (2017). *Grammar and Information Structure of Kakataibo*. University of Texas at Austin.
- Vuillermet, M. (2012). *A Grammar of Ese Ejja, A Takanan language of the Bolivian Amazon*. Lyon: PhD thesis, l'Université Lumière Lyon 2.
- Vuillermet, M. (2013). *Associated Motion Elicitation Kit (Version 7)*. Linguistic Department, UC Berkeley.
- Vuillermet, M. a. (2013 ms). *A hunting story ~ Yendo a cazar: A visual stimulus for eliciting constructions that associate motion with other events*. Department of Linguistics, UC Berkeley.
- Webster, A. K. (2008). Running Again, Roasting Again, Touching Again: On Repetition, Heightened Affective Expressivity, and the Utility of the Notion of Linguaculture in Navajo and Beyond. *The Journal of American Folklore*, 121(482), 441-472.
- Wimsatt, W. C. (2007). *Re-Engineering Philosophy for Limited Beings*. Cambridge: Harvard University Press.
- Woodbury, A. C. (2003). Defining Documentary Linguistics. *Language Documentation and Description*, 1, 35-51.
- Woodbury, A. C. (2007). On thick translation in linguistic documentation. *Language Documentation and Description*, 4, 120-135.
- Zariquiey, R. (2018). *A Grammar of Kakataibo*. Berlin: Mouton de Gruyter.

Cómo citar: Gallagher, G. (2023). Fonología. En P. Alandía Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 49-61), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11106978>

CAPÍTULO 2

FONOLOGÍA

Gillian Gallagher*

0. Introducción

La fonología es el estudio de los patrones de sonidos en las lenguas. En contraste con la fonética, que es el estudio de las propiedades físicas de sonidos, la fonología se enfoca en la estructura de contrastes entre sonidos y las combinaciones de sonidos en morfemas y palabras. Cada lengua tiene un conjunto particular de sonidos contrastivos: por ejemplo, el quechua boliviano tiene el sonido [k'], una oclusiva velar glotalizada, que no se encuentra en castellano o guaraní o muchas otras lenguas. De esto se trata la fonología. Cada lengua tiene sílabas con una estructura particular: por ejemplo, en el quechua hay sílabas consonante-vocal (*wa.si* 'casa') y consonante-vocal-consonante (*war.mi* 'mujer'), pero, en contraste con el castellano, no hay sílabas con dos consonantes consecutivas (como en *flaca*). De esto se trata la fonología. Cada lengua tiene sus propias restricciones sobre cómo los sonidos pueden combinarse y cómo se pronuncian en contexto. En quechua, no hay palabras con dos consonantes glotalizadas (*k'achi* 'sal' y *sach'a* 'árbol' existen, pero no *k'ach'a*). También en quechua, la pronunciación de las vocales cambia en el contexto de una consonante uvular ([takini] 'yo canto', pero [takerqa] 'él cantó'). De esto se trata la fonología.

En este capítulo, se revisa las generalizaciones fonológicas como se ven en el sistema de sonidos contrastivos (§2.1), la fonotáctica y la estructura de morfemas (§2.2), la fonotáctica y la morfo-fonología (§2.3), y la morfo-fonología distinta de la fonotáctica (§2.4). La §2.5 está dedicada al préstamo de las palabras entre lenguas.

* Es docente de lingüística en la Universidad de Nueva York, EE.UU. Obtuvo su doctorado en el Instituto de Tecnología de Massachusetts. Sus intereses principales radican en la fonética y la fonología, principalmente en el quechua boliviano.

2.1. El sistema de sonidos contrastivos

Cada lengua tiene unos sonidos **contrastivos**, lo que quiere decir que la diferencia en el sonido corresponde a una diferencia en el sentido. Por ejemplo, en quechua sabemos que [m] y [w] son sonidos contrastivos porque las palabras [masi] ‘amigo, compadre’ y [wasi] ‘casa’ tienen sentidos diferentes y la única diferencia entre [m] y [w] es esa. Los pares de palabras como [masi] y [wasi], que se distinguen por un solo sonido, se llaman **pares mínimos**.

El conjunto de sonidos contrastivos en una lengua tiene una estructura que se ve en las dimensiones fonéticas que se usan. Abajo en las **Tablas 1-3** están los conjuntos de oclusivas en tres lenguas que se hablan en Bolivia: quechua, guaraní y castellano. En estas lenguas, para cada lugar de articulación, hay un contraste de dos o tres consonantes oclusivas con la misma distinción en el modo: aspiración y glotalización en quechua, prenasalización en guaraní y sonorización en castellano.

Tabla 1

Oclusivas en guaraní

	labial	alveolar	velar	velar-labializada
sorda	P	t	k	k ^w
prenasalizada	^m b	ⁿ d	^ŋ g	^ŋ g ^w

Tabla 2

Oclusivas en quechua

	labial	dental	palato-alveolar	velar	uvular
sorda	p	t	tʃ	k	q
aspirada	p ^h	t ^h	tʃ ^h	k ^h	q ^h
glotalizada	pʼ	tʼ	tʃʷ	kʼ	qʼ

Tabla 3

Oclusivas en castellano

	labial	dental	velar
sorda	p	t	k
sonora	b	d	g

En contraste, la **Tabla 4** muestra un conjunto aleatorio de oclusivas, en que cada lugar de articulación tiene un contraste diferente en el modo. Este conjunto es inventado; conjuntos así son poco comunes o inexistentes en las lenguas reales del mundo.¹

Tabla 4

Oclusivas aleatorias (inventado)

	Labial	dental	velar
sorda	P	t	
glotalizada	p'		
sonora		d	g
prenasal			ᵑg

Un aspecto de la fonología de una lengua es el sistema de contrastes entre los sonidos, como se ve en las tablas de arriba. Este sistema es una parte de lo que forma un ‘acento nativo’ en un idioma porque los hablantes nativos pueden percibir y producir bien los sonidos y las dimensiones fonéticas que se usan en su lengua, pero tienen dificultades con los sonidos no nativos. Por ejemplo, en inglés las oclusivas contrastan en aspiración ([p^h] en *pan* ‘sartén’ vs. [p] en *ban* ‘prohibición’), pero en castellano contrastan en sonorización ([p] en *parque* vs. [b] en *barco*). Cuando un anglófono habla castellano puede usar sus oclusivas nativas –dice [p^harke] y [parko]– y así suena como si tuviera un acento extranjero.

2.2. La fonotáctica y la forma de morfemas

Palabras y morfemas se forman de secuencias de sonidos contrastivos y cada lengua tiene sus reglas sobre cómo los sonidos se pueden combinar. Las reglas fonotácticas son responsables de la estructura de las sílabas y de las combinaciones entre sonidos adyacentes y no adyacentes (entre otras generalizaciones). En quechua, por ejemplo, las sílabas tienen la estructura CV (consonante-vocal) o CVC (consonante-vocal-consonante). Por eso, combinaciones de dos consonantes adyacentes solo se encuentran en la mitad de una palabra (1a), nunca al principio o al final de una palabra (1b). En contraste, en castellano es posible tener dos consonantes al principio de la palabra (*flaca*, *pronto*, *clima*), y en inglés combinaciones de consonantes se encuentran también en posición final (*ask* ‘preguntar’, *help* ‘ayuda’).

1 El castellano tiene una africada sorda [tʃ], pero no la africada sonora correspondiente [dʒ]; igualmente, el guaraní tiene una prenasal palatal [ɲj], pero no la sorda [c]. ‘Huecos’ de ese tipo, en un sistema de sonidos contrastivos, son comunes; lo que es raro es que el contraste sea fonéticamente diferente en cada lugar de articulación.

En estos ejemplos, se usa el punto para indicar la frontera entre sílabas, y la estrella para indicar que una estructura es inexistente e imposible.

(1) a. Quechua: palabras con sílabas CV o CVC

ru.mi	CV.CV	‘piedra’
wa.si	CV.CV	‘casa’
mis.k’i	CVC.CV	‘rico’
t’an.ta	CVC.CV	‘pan’
si.pas	CV.CVC	‘mujer joven’
su.maχ	CV.CVC	‘bueno’

b. Quechua: no hay palabras con CCV o VCC

*sk’i.mi	CCV.CV
*ru.misk’	CV.CVCC

Además, la consonante final en sílabas CVC es restringida: no puede ser una oclusiva, debe ser una fricativa o una sonante (nasal, líquida o aproximante). Esta distribución se muestra con los ejemplos en (2a).

(2) a. Quechua: sílabas con fricativas y sonantes al final

k’as.pi	‘palo’	Λim.p ^{hu}	‘limpio’	waΛ.pa	‘gallina’
sox.ta	‘seis’	k’aŋ.ka	‘gallo’	tʃaw.pi	‘medio’

b. Quechua: no hay sílabas con oclusivas al final

*k’at.pi	*sop.ta	*mik’.si
----------	---------	----------

Esta restricción contra las oclusivas en posición final es una propiedad del quechua sureño boliviano. En otras variedades del quechua, que se hablan más al norte en Perú y Ecuador, es posible tener oclusivas en esta posición. En la escritura normalizada de Bolivia, hay palabras que se escriben con oclusiva, pero se pronuncian con fricativa. La escritura representa una pronunciación antigua: la pronunciación ha cambiado en el quechua sureño, pero se ha mantenido en otras variedades. Algunos ejemplos se encuentran en el (3).

<u>ortografía</u>	<u>pronunciación</u>	<u>castellano</u>
<i>suqta</i>	soxta	‘seis’
<i>lliklla</i>	ΛixΛa	‘manta tejida’
<i>runap</i>	runax	‘de la persona’
<i>wawa kaptiy</i>	wawa kaxtij	‘cuando era niño’

Otro tipo de regla fonotáctica en quechua es la restricción contra unas combinaciones de consonantes no adyacentes en una palabra. En quechua hay tres tipos de oclusivas: simples, glotalizadas y aspiradas. Las oclusivas glotalizadas y aspiradas siempre son la primera oclusiva en la palabra, no pueden seguir a una oclusiva (sea simple, glotalizada o aspirada) anterior. Los ejemplos en (4) muestran el patrón. Se ven palabras con oclusivas glotalizadas y aspiradas al principio (4a) y al medio (4b) de la palabra. Todas las palabras en (4b) tienen una fricativa o una sonante como consonante inicial. Las palabras inventadas en (4c) son inexistentes e imposibles porque tienen una oclusiva como consonante inicial.

- (4) a. Quechua: oclusivas glotalizadas y aspiradas en posición inicial
- | | | | |
|----------------|----------|---------------------------|------------|
| q' epij | ‘llevar’ | k^hu ʃuj | ‘cortar’ |
| p' isqo | ‘pájaro’ | p^h uti | ‘angustia’ |
- b. Quechua: oclusivas glotalizadas y aspiradas en posición media
- | | | | |
|--------------------------|------------|-----------------------------------|----------|
| <u>h</u> a ʎ p'a | ‘tierra’ | <u>r</u> a k^h u | ‘grueso’ |
| <u>ʃ</u> a ŋ k'aj | ‘trabajar’ | <u>j</u> u t^h u | ‘perdiz’ |
- c. Quechua: no hay oclusivas glotalizadas o aspiradas siguiendo a otra oclusiva
- | | |
|----------------------------|-------------------------------------|
| * <u>t</u> a ʎ p'a | * <u>t</u> a k^h u |
| * <u>p</u> a ŋ k'aj | * <u>k</u> u t^h u |

Las reglas fonotácticas forman una parte del conocimiento tácito (o sea subconsciente) de la lengua que han adquirido los hablantes. Palabras nuevas o inventadas que siguen las reglas suenan naturales y los hablantes pueden percibir las y producirlas bien. Por otro lado, palabras inventadas que no siguen las reglas suenan extrañas y los hablantes podrían cometer errores en la percepción o en la producción. El efecto de la fonotáctica se observa cuando una lengua se presta palabras de otra lengua con una fonotáctica diferente. Por ejemplo, hay muchas palabras en quechua que tienen su origen en el aymara, pero su pronunciación es un poco diferente en las dos lenguas. En aymara, se permiten palabras con dos oclusivas glotalizadas o aspiradas, pero en quechua estas palabras se pronuncian con una oclusiva glotalizada o aspirada y una oclusiva simple como *t'ant'a* (aymara) pero *t'anta* (quechua) ‘pan’: la pronunciación cambia dependiendo de las reglas fonotácticas de la lengua.

El préstamo de palabras es un contexto natural para ver la fonotáctica. Otra manera de investigar la fonotáctica es en un contexto experimental. En uno de mis experimentos, solicité a quechuahablantes que escuchen palabras inventadas y que

las repitan (Gallagher, 2016a). Grabé a un quechuahablante pronunciando palabras inventadas que fueron consistentes con la fonotáctica del quechua, palabras como [kapi], [map'i], [rip^hu]. De estas grabaciones, corté y pegué los sonidos para formar palabras en nuevos audios con combinaciones de oclusivas inexistentes en quechua, palabras como *[pak'a] o *[p'ak'a]. Los participantes en el experimento cometieron errores en la repetición de palabras de ese tipo. Cuando escuchan una palabra con dos oclusivas glotalizadas o aspiradas la repiten con solo una (*[p'ak'a] se repite como [p'aka]), y cuando escuchan una palabra con una combinación simple-glotalizada o simple-aspirada la repiten con la glotalización o aspiración al principio de la palabra (*[pak'a] se repite como [p'aka]). De un experimento así, podríamos concluir que las restricciones en las combinaciones de oclusivas son una parte de la gramática implícita de un quechuahablante.

2.3. La fonotáctica y morfo-fonología

La fonotáctica puede causar variación en la forma fonológica de morfemas cuando se combinan para formar palabras. La adición de un sufijo o prefijo puede cambiar el contexto fonológico de los sonidos en la palabra. Por ejemplo, en quechua hay un sistema de **alofonía** en las tres vocales contrastivas. En la mayoría de los contextos las vocales se pronuncian como [i u a], pero en el contexto de una consonante uvular se pronuncian con retracción como [e o ɑ]. Esa variación se puede ver en la forma de palabras simples en (5).

- (5) a. Vocales [i u a] en contextos sin uvulares
- | | | | |
|--------|----------|-------|-----------|
| kusa | ‘bueno’ | muhu | ‘semilla’ |
| kiʎa | ‘luna’ | miʎaj | ‘feo’ |
| katari | ‘víbora’ | maki | ‘mano’ |
- b. Vocales [e o ɑ] en el contexto de una uvular
- | | | | |
|--------------------|----------|-------|----------|
| qosa | ‘esposo’ | hoq'o | ‘mojado’ |
| q ^h eʎa | ‘flojo’ | erqe | ‘hijo’ |
| qɑʎu | ‘lengua’ | maqaj | ‘pelear’ |

Esta variación se llama **alofonía** porque está por debajo del nivel de conciencia de los hablantes y pasa automáticamente. La retracción de las vocales ocurre en palabras simples con uvulares, como se ve arriba en (5), y también ocurre con la adición de una consonante uvular a una palabra, como se muestra en (6) con ejemplos de verbos complejos. Los verbos en quechua se forman con sufijos que indican tiempo, concordancia y otros sentidos más complejos. En (6), la vocal final

de la raíz verbal cambia cuando va seguida por el sufijo de tiempo pasado *-rqa* que tiene un sonido uvular.

(6)	<i>sin uvular</i>		<i>en contexto uvular</i>
	wajk'uni	‘cocino’	wajk'orqani ‘cociné’
	tarpuni	‘siembro’	tarporqani ‘sembré’
	hap'ini	‘tengo’	hap'erqani ‘tuve’
	takini	‘canto’	takerqani ‘canté’

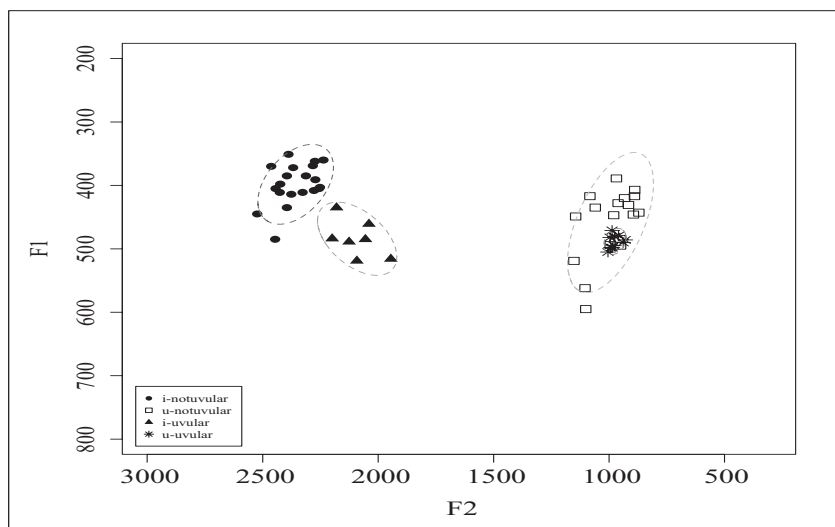
Los pequeños cambios de pronunciación de este tipo son muy comunes en la fonología. Cada lengua tiene sus reglas sobre cómo se pronuncian los sonidos en contexto, y la pronunciación en palabras o en frases se ajusta para seguir estas reglas fonotácticas. En el caso de la retracción de vocales en quechua, hay una motivación fonética para el patrón. Las consonantes uvulares se producen con la lengua muy atrás en la boca. La retracción de la vocal previa o siguiente minimiza el movimiento de la lengua entre vocal y consonante. En otras lenguas con uvulares, hay un patrón similar en que solo una parte de la vocal se retrae.

Ya que la alofonía está por debajo del nivel de conciencia, puede ser útil hacer un análisis acústico para identificar los cambios y las propiedades específicas de los sonidos. Varios investigadores han hecho esto con hablantes de varias variedades de quechua (Pasquale, 2001; Molina Vital, 2011; Gallagher, 2016b; Holliday & Martin, 2017). Sobre todo, la descripción de arriba es consistente con los datos acústicos: en general, las vocales son más bajas y más posteriores en el contexto de una consonante uvular. Pero hay mucha variación. Algunos hablantes retraen las vocales más que otros. A veces una vocal en un contexto sin uvular se pronuncia un poco más baja y posterior, y a veces una vocal en un contexto uvular se pronuncia un poco más alta y anterior. Camacho Rios (2016) ha notado que los monolingües ancianos de Tarabuco no retraen sus vocales tanto como los hablantes más jóvenes y bilingües.

Vale la pena reflexionar un poco sobre la variación y la descripción de la fonotáctica. Las reglas fonotácticas se describen con símbolos discretos ([i] vs. [e]) o propiedades fonéticas (‘alto’ y ‘medio’). Pero el habla real consiste en sonidos continuos y variables que no corresponden directamente a estas categorías discretas. Un ejemplo de la variación se muestra en la **Figura 1**, que es una visualización de las vocales [i], [u], [e], [o] de un hablante de quechua. Cada punto representa el valor de los dos primeros formantes (F1 y F2) de una vocal en una palabra. Los círculos representan las categorías según la descripción de arriba: [i] es **i-notuvular** (no uvular) en puntos negros, [e] es **i-uvular** en triángulos negros, [u] es **u-notuvular** en cuadrados vacíos y [o] es **u-uvular** en estrellas.

Figura 1

Las vocales [i], [e], [u], [o] de un hablante de quechua



Este hablante produce [i] y [e] distintivamente: no hay superposición en los círculos de los puntos y los triángulos. Para las vocales posteriores, la distinción es menos clara. Las vocales en un contexto sin uvular muestran una variación grande entre una vocal más alta como [u] y algo más mediana como [o]. Mientras hay esta variación en el contexto sin uvular, en el contexto uvular la producción de la vocal es más precisa: las estrellas muestran poca variación y todas están en la región media. Lo que muestra esta figura es que la descripción de una lengua en general y las producciones de un hablante particular podrían ser diferentes. En quechua, las vocales son más bajas en general en contextos con consonante uvular que en contextos sin una consonante uvular, pero no es que cada producción de cada vocal de cada hablante sea claramente alta o baja. Otro punto importante es que los símbolos fonéticos como [i e] y [u o] no son perfectamente precisos, la [i] en una lengua (o de un hablante) no debe ser igual a la [i] en otra lengua (o de otro hablante).

El sistema vocálico en quechua se aprecia en el acento de los quechua hablantes que hablan castellano como segunda lengua. Es común, para un hablante de quechua, cambiar las vocales altas [i u] y las vocales medias [e o], por ejemplo, decir [intiru] para *entero*, porque en quechua estas vocales tienen una relación predecible y se producen con mucha variación. En castellano, las vocales [i u e o] son todas contrastivas (o sea, no tienen relación predecible) y se producen con menos variación. En quechua, es común escuchar la misma palabra pronunciada con una vocal alta o media, por ejemplo, *arí* 'sí' se puede oír más como [ari] o más como [are]. Esta diferencia no cambia el sentido en quechua y por eso es menos perceptible. Dependiendo del

sistema fonológico nativo, la misma distinción fonética puede sonar más grande o más pequeña.

2.4. Morfo-fonología

La morfo-fonología se podría ver en las alternancias fonotácticas —como las vocales en quechua que cambian con el sufijo pasado *-rqa*—, pero también hay alternancias morfo-fonológicas que no se relacionan con la fonotáctica. Un ejemplo claro es la diptongación en el sistema verbal del castellano. Algunos verbos muestran diptongación (*e* → *ie*, *o* → *ue*) y este proceso parece fonológico porque su aplicación sigue la posición del acento. La vocal es un monoptongo cuando no es acentuada y es un diptongo cuando tiene acento. Este patrón se ve en (7); el apóstrofo indica la sílaba con estrés.

- (7) a. diptongación *e* → *ie* en posición acentuada
- | | |
|-----------|----------|
| sen'tar | ‘siento |
| sen'tamos | ‘sientan |
- b. diptongación *o* → *ue* en posición acentuada
- | | |
|-----------|---------|
| con'tar | ‘cuento |
| con'tamos | ‘cuenta |

Este cambio en el sonido de la raíz verbal no es a causa de la fonotáctica del castellano: es posible que una vocal con acento sea un monoptongo [e o]. En (8) se encuentran ejemplos de verbos que no son afectados por diptongación, por el contrario, se mantienen como monoptongo en todas las formas.

- (8) a. monoptongo *e* en posición acentuada
- | | |
|-----------|--------|
| ven'ter | ‘vendo |
| ven'temos | ‘vende |
- b. monoptongo *o* en posición acentuada
- | | |
|-----------|--------|
| mon'tar | ‘monto |
| mon'tamos | ‘monta |

Muchas lenguas tienen patrones, así, hay una alternancia en sonidos que se puede describir en términos fonológicos, pero el cambio solo se aplica a algunas palabras.

Otro ejemplo de una alternancia morfo-fonológica, pero no fonotáctica, viene del quechua. En quechua hay varios sufijos en los que la vocal cambia entre [u] y [a]. La variante con [a] aparece antes de otro sufijo con [u]. En (9) hay un ejemplo con el sufijo *-rqu* ‘hacer algo felizmente y ágilmente’. En (9a) este sufijo se pronuncia con [u]. En (9b), cuando el sufijo [-ku] lo sigue, se pronuncia con [a]. Una descripción más completa de esta alternancia se encuentra en Camacho Rios (2019).

- (9) a. wayk'u-rquu-ni ‘Cocino felizmente y ágilmente’
 b. wayk'u-rqaa-ku-ni ‘Cocino felizmente y ágilmente para mí’

La alternancia entre [u] y [a] es regular para este sufijo (y algunos otros como *-yku* ‘perfectivo’ y *-ku* ‘reflexivo’), pero no pasa a causa de la fonotáctica. Como se ve en (9a), secuencias de sílabas con la vocal [u] son permitidas. La [u] en el verbo *wayk'u* ‘cocinar’ siempre es [u], nunca cambia a [a] (**wayk'a-rqu-ni* es imposible).

En resumen, la morfo-fonología de una lengua consiste en las alternancias en la forma fonológica de un morfema. Estas alternancias pueden darse para satisfacer las reglas fonotácticas o podrían ser independientes de la fonotáctica.

2.5. Fonología en el préstamo de palabras

La fonología tiene un gran efecto en el préstamo de palabras de una lengua a otra. En el préstamo, una lengua puede modificar los sonidos individuales según su propio inventario de sonidos y según su propia fonotáctica.

El quechua usa muchas palabras de origen castellano, y su pronunciación depende del nivel de bilingüismo del hablante y de cuándo se prestó la palabra. Los préstamos de hace cientos de años tienen una pronunciación más modificada para confirmar la fonotáctica del quechua. Por ejemplo, [uwiha] para *oveja* o [awila] para *abuelo*. En estas dos palabras, el sonido [β] del castellano (v o b en la ortografía) no tiene correspondencia en quechua y se cambió a [w], que existe en quechua. Las vocales medias se cambiaron a vocales altas. Las palabras prestadas más recientemente pueden mantener una pronunciación muy cercana a la del castellano y los bilingües en particular pueden usar esa pronunciación. Así se escuchan frecuentemente ‘lado’, ‘escuela’ ‘a veces’, etc. en quechua corriente, con vocales medianas y consonantes sonoras.

El guaraní paraguayo es similar al quechua con respecto a que hay palabras que fueron prestadas del castellano en los años 1500, que cambiaron su estructura para seguir las reglas fonotácticas del guaraní. El proceso de prestar palabras y cambiar su pronunciación se llama **nativización**. Normalmente, cuando las palabras se nativizan, los hablantes no saben que son **préstamos**, pasan a ser ‘guaraní’. En el tiempo moderno, hay muchos bilingües que hablan guaraní y castellano, y para ellos es común usar palabras del castellano con su pronunciación castellana cuando hablan

guaraní. Los bilingües no necesitan nativizar o modificar las palabras desde una de sus lenguas a otra.

La tesis de maestría de Pinta (2013) investiga la variación en el préstamo de palabras del castellano. En guaraní, la mayoría de las palabras tiene el acento en la última sílaba. En el préstamo, la posición del acento se cambió en unas palabras, pero se preservó en otras.

(10) a. cambio de acento

<u>castellano</u>	<u>guaraní</u>	
‘baka	va’ka	<i>vaca</i>
a’sukar	asu’ka	<i>azúcar</i>
‘keso	ke’su	<i>queso</i>

b. preservación de acento

a’merika	a’merika	<i>América</i>
‘bino	‘vino	<i>vino</i>
‘bweno	‘weno	<i>bueno</i>

La estructura de sílabas es diferente en castellano y guaraní. La sílaba en guaraní tiene una estructura fija de (C)V (consonante-vocal) y las palabras prestadas en (11a) se forman por adición de una vocal o supresión de una consonante. En (11b) se ve que unas palabras mantienen una sílaba con una consonante final o con dos consonantes al principio, estas palabras no se modifican para satisfacer el patrón de palabras nativas en guaraní.

(11) a. adaptación a CV

<u>castellano</u>	<u>guaraní</u>	
a’ros	a’ro	<i>arroz</i>
‘krus	kuru’su	<i>cruz</i>
‘bolsa	vo’sa	<i>bolsa</i>

b. ‘grasja	gra ’sja	<i>gracia</i>
er’mana	er’mana	<i>hermana</i>
kolek’tiβo	kolek’tivo	<i>colectivo</i>

Los ejemplos en (10) y (11) también muestran la modificación de sonidos individuales que no existen en guaraní. La vibrante múltiple [r] de castellano se presta como [r] y la oclusiva [b] se presta como [v].

Cuando las palabras prestadas se modifican, estos cambios nos muestran la gramática fonológica en acción. Pero el préstamo de palabras también puede introducir nuevas estructuras fonológicas, como se ve en (10b) y (11b), y así puede cambiar la gramática fonológica. Con una población bilingüe, las palabras no modificadas se hacen más comunes y, con el tiempo, las estructuras nuevas pueden hacerse ‘nativas’. Ese proceso puede haber pasado en la historia del quechua. Las variedades de quechua del sur tienen oclusivas glotalizadas y aspiradas que no se encuentran en las variedades de quechua de más al norte. Una hipótesis del origen de estos sonidos es que vienen del contacto con el aymara por palabras prestadas y hablantes bilingües (Stark, 1975; Mannheim, 1991). El aymara tiene oclusivas glotalizadas y aspiradas y estos sonidos se encuentran en quechua sureño en muchas palabras con origen aymara (e.g. *mut'i* ‘mote’ *p'aki*- ‘fragmento, romper’). En el tiempo moderno, los sonidos glotalizados y aspirados son nativizados y se encuentran también en palabras que no tienen origen aymara. Ahora, son sonidos **nativos** en quechua que suenan normalmente para los hablantes.

2.6. Conclusión

Este capítulo ha revisado los tipos de patrones de sonido que componen la fonología. La gramática fonológica se ve en el conjunto de sonidos que una lengua usa y en el sistema de contrastes entre estos sonidos. La fonotáctica es una parte de la fonología que describe qué estructuras son permisibles en las palabras. La fonotáctica se observa en las combinaciones de sonidos y en la estructura de sílabas en palabras simples (con un morfema), y se aprecia también en los cambios en la pronunciación de sonidos cuando los morfemas se combinan para formar palabras complejas. Finalmente, hay generalizaciones morfo-fonológicas que son independientes de la fonotáctica. La gramática fonológica se puede observar en contextos experimentales y también en el préstamo de palabras entre lenguas.

Referencias

- Camacho Rios, G. (2016). Analysis of the acoustic effects of uvular and non-uvular sounds on high vowels: a comparison of two Bolivian Quechua dialects. En Isabelle Leblic & Lameen Souag (eds.), *Du Terrain a la Théorie*. Lacito: Villejuif, France. 275-284.
- Camacho Rios, G. (2019). *Verb morphology in South Bolivian Quechua: A case study of the Uma Piwra rural variety*. [MA thesis, University of Texas, Austin].
- Gallagher, G. (2016a). Asymmetries in the representation of categorical phonotactics. *Language* 92(3): 557-590.
- Gallagher, G. (2016b). Vowel height allophony and dorsal place contrasts in Cochabamba Quechua. *Phonetica* 73.101-119.
- Holliday, N. & Martin, S. (2017). Vowel systems and allophonic lowering among Bolivian Quechua/Spanish bilinguals. *Journal of the International Phonetic Association* 48(2). 199-222.
- Mannheim, B. (1991). *The language of the Inka since the European invasion*. University of Texas Press: Austin, TX.
- Molina-Vital, C. (2011). Reconsidering Cuzco Quechua vowels. Ms. Rice University. Disponible en: http://www.academia.edu/5994608/Reconsidering_Cuzco_Quechua_Vowels
- Pasquale, M. (2001). *Quechua and Spanish language contact: Influence on the Quechua phonological system*. [Doctoral dissertation, Michigan State University, East Lansing].
- Pinta, J. (2013). *Lexical strata in loanword phonology: Spanish loans in Guarani*. [MA thesis, University of North Carolina, Chapel Hill]. Disponible en: <https://cdr.lib.unc.edu/concern/dissertations/bk128c287>.
- Stark, L. (1975). A reconsideration of Proto-Quechua phonology. *XXXIX International Congress of Americanists*, Vol 5, Lima, Perú. 209-219.

CAPÍTULO 3

Cómo citar: Gallinate, G. A., & Lapierre, M. (2023). Fonética. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 62-99), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11106986>

FONÉTICA

Gabriel A. Gallinate & Myriam Lapierre***

0. Introducción

La **fonética**, como subdisciplina de la lingüística, se ocupa del estudio de los **sonidos** del habla para comprender los mecanismos que componen las lenguas en materia del sonido. Los **fonetistas** pueden dedicarse a la tarea de examinar cómo se producen los sonidos en el tracto vocálico de los hablantes de una lengua. También pueden observar los rasgos acústicos y las ondas sonoras que componen un **fono** (sonido) desde una perspectiva física. Del mismo modo, los especialistas en esta área pueden explicar de qué manera y por qué un hablante de una lengua en específico percibe ciertos sonidos. Estos tres campos de acción corresponden a las principales áreas en las que la fonética se divide tradicionalmente como veremos enseguida.

Si un fonetista aborda el estudio de los movimientos que se realizan con los órganos fonadores al momento de producir la vocal [i] en algunas lenguas de la Amazonía boliviana, empleará la **fonética articulatoria** para lograr sus objetivos descriptivos. En ese sentido, explicará qué posiciones adopta la lengua o en qué condiciones se encuentran los demás articuladores como los labios o la mandíbula al pronunciar dicha vocal. Además, tendrá que indicar el estado de las cuerdas vocales al instante de emitir la vocal en cuestión.

Cuando un estudioso de la fonética, en cambio, analiza los componentes físicos del sonido como sus rasgos acústicos, su duración, su frecuencia, su intensidad o su entonación, con evidencia en las ondas sonoras que los constituyen, se servirá de la **fonética acústica** con el propósito de entender cómo se manifiestan los sonidos del habla.

* Es lingüista doctorando en la Universidad de Texas en Austin. Obtuvo su licenciatura en lingüística aplicada en la Universidad Mayor de San Simón, Bolivia. Sus intereses académicos comprenden la descripción de lenguas bolivianas, la fonética, la fonología, la morfofonología, la lexicografía y la documentación de lenguas en peligro de extinción.

** Es docente de fonología en la Universidad de Washington. Obtuvo su doctorado de la Universidad de California, Berkeley. Sus intereses están en la lingüística documental y descriptiva, especialmente en el área de la fonología de los segmentos complejos y nasales.

Por su parte, recurrirá a la **fonética auditiva** o perceptiva en situaciones cuando desee indagar acerca de cómo los hablantes de una lengua reciben, perciben y procesan los sonidos del habla que escuchan (Carr, 2008). De esa manera, se puede explicar por qué los hablantes monolingües de quechua boliviano del sur no distinguen las vocales altas [i, u] de las medias [e, o].

En este capítulo introductorio, solo abordaremos el campo de la fonética articulatoria. Esto será útil para adentrarnos en el mundo de los sonidos de la mano de la ciencia fonética, porque tener una buena comprensión de la fonética articulatoria ayuda a comprender los conceptos de la fonética acústica y de la auditiva. La fonética articulatoria también representa el conocimiento mínimo que un lingüista descriptivo debe tener para lograr la tarea de describir el sistema de sonidos de un idioma. Los estudiantes cuyo interés en la fonética vaya más allá del contenido de este capítulo podrán estudiar para ampliar su conocimiento, abordando en ella las ramas adicionales de la fonética.

3.1. Los órganos fonadores

Cada vez que los hablantes de una lengua quieren producir enunciados oralmente, expulsan el aire que inspiran desde los pulmones hasta fuera de la boca. No obstante, ese acto es simultáneo a una serie de movimientos a lo largo del tracto vocal, que posibilitan que cada sonido articulado tenga esas propiedades que lo diferencian de los demás sonidos.

En este apartado, veremos cuáles son los demás órganos encargados de producir los sonidos, cómo se articulan las consonantes y las vocales. Primeramente, revisaremos el sistema de representación gráfica que usamos para transcribir fonéticamente los sonidos de las lenguas.

3.1.1 El alfabeto fonético internacional

El alfabeto fonético internacional, también conocido como AFI, es un conjunto de símbolos que los lingüistas usan para representar todos los sonidos posibles del habla humana. Este alfabeto puede ser usado para transcribir los sonidos de cualquier idioma. El AFI se subdivide en dos secciones que corresponden con una división importante entre tipos de sonidos: la sección de las consonantes y la sección de las vocales. Las **consonantes**, como por ejemplo [p], [s] y [k], se producen con obstrucción del aire en algún punto del tracto vocal, mientras que las **vocales**, como por ejemplo [i], [a] y [u], se producen sin obstrucción del aire.

Empecemos hablando de la articulación de las consonantes. La **Figura 1** muestra la tabla de las consonantes del alfabeto fonético internacional. La dimensión

horizontal de esta tabla corresponde a **los puntos de articulación** de los sonidos (ver §3.2.1), y la dimensión vertical corresponde a **los modos de articulación** (ver §3.2.2). Estas dos dimensiones representan las características primarias de la fonética articulatoria.

Figura 1

La tabla de las consonantes pulmonares del Alfabeto Fonético Internacional

CONSONANTES (PULMONARES)

© 2020 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retrofleja	Palatal	Velar	Uvular	Faríngea	Glotal
Oclusiva	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Vibrante múltiple	ʙ			r					ʀ		
Vibrante simple		ʋ		ɾ		ɽ					
Fricativa	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Fricativa lateral				ɬ ɮ							
Aproximante		ʋ		ɹ		ɻ	j	ɰ			
Aproximante lateral				l		ɭ	ʎ	ʟ			

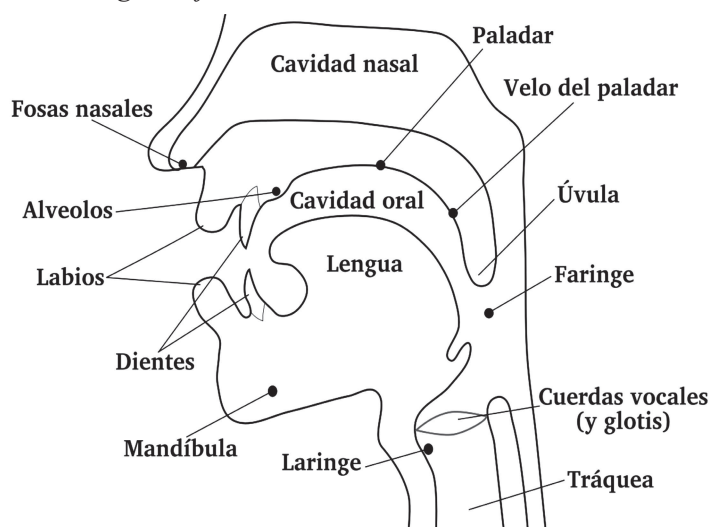
Los símbolos de la derecha de una celda representan sonidos sonoros, y los de la izquierda son sordos.
Las áreas sombreadas indican articulaciones que se consideran imposibles.

El punto de articulación de una consonante es el lugar en el tracto vocálico donde se produce una obstrucción. Por ejemplo, una consonante **bilabial** se produce con una obstrucción del aire entre los labios, y una consonante **dental** se produce con una obstrucción del aire con la punta de la lengua acercándose a los dientes superiores.

La **Figura 2** muestra una vista sagital (diseccionada) de los órganos fonadores. El plano sagital es un corte anatómico que divide al cuerpo en dos mitades; es decir, la mitad derecha y la mitad izquierda. El eje equis u horizontal de la tabla de las consonantes del AFI corresponde a la vista sagital de los órganos fonadores, suponiendo una orientación de los labios hacia la izquierda de la tabla. Imaginemos que la cavidad oral es un tubo que empieza entre los labios y termina en la glotis. En la Figura 2, observamos un ángulo de 90 grados en el tubo, en la coyuntura entre la cavidad oral y la cavidad faríngea. ¿Qué pasaría si pudiéramos enderezar el tubo? Lo que observaríamos es que el tubo del tracto vocal estaría recto, igual al eje equis de la tabla de las consonantes de AFI. En este escenario, la primera columna de la tabla corresponde con el punto de articulación más anterior en el tracto vocálico, es decir, los labios; mientras que la última columna corresponde con el punto de articulación más posterior en el tracto vocálico, es decir, la glotis.

Figura 2

Vista sagital de los órganos fonadores

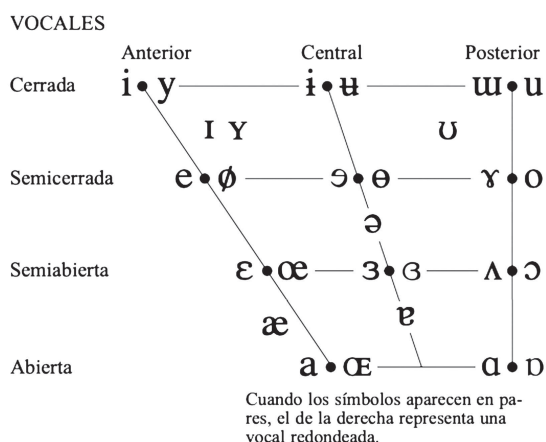


El eje *y* o vertical de la tabla de las consonantes del AFI corresponde con el grado de constricción de la consonante, más o menos equivalente al punto más alto de la lengua al acercarse al paladar. La primera fila de la tabla corresponde al modo de articulación que involucra obstrucción total del aire, es decir, las **oclusivas**, que se producen cuando un articulador toca otro. Las últimas filas de la tabla corresponden al modo de articulación que involucra menos obstrucción del aire en el tracto vocal, es decir, las **aproximantes**, que se producen con una aproximación de un articulador hacia otro.

Seguidamente, hablemos de la articulación de las vocales. La **Figura 3** muestra la tabla de las vocales del AFI. La tabla representa las tres dimensiones articulatorias de las vocales, todas correspondiendo a su punto de articulación. La primera dimensión de la tabla, la dimensión horizontal, corresponde al grado de **posterioridad** de la lengua en la boca durante la producción de cada vocal. La segunda dimensión, la dimensión vertical, corresponde al grado de **altura** de la lengua. Finalmente, la tercera dimensión corresponde al grado de **redondeamiento** de los labios, y se representa por los pares de vocales con el mismo grado de posterioridad y altura, separadas por un punto negro. En cada caso, la vocal a la izquierda representa una vocal no redondeada, y la vocal a la derecha representa una vocal redondeada (ver §3.3 para las características y propiedades de las vocales).

Figura 3

La tabla de las vocales del Alfabeto Fonético Internacional



Muchas veces, los estudiantes que se encuentran iniciando el estudio de la fonética piensan que las vocales anteriores se producen con el **ápice** o la **lámina de la lengua**, mientras que las vocales posteriores se producen con el dorso de la lengua. ¡Pero se equivocan! Todas las vocales se producen con el **dorso de la lengua** (ver la Figura 6). Las vocales anteriores se producen con el **predorso**, como las consonantes palatales, y las vocales posteriores se producen con el **postdorso**, como las consonantes velares. Este hecho ayuda a entender ciertos **cambios de sonidos** frecuentes en los idiomas del mundo. Por ejemplo, las consonantes velares como [k] muchas veces se pronuncian con un punto de articulación palatal [c] antes de las vocales anteriores como [i].

También es importante notar que los lingüistas usan dos series de términos intercambiables para hablar del grado de altura de las vocales. Cuando hablamos de una vocal **cerrada** o **abierta**, estamos hablando del grado de **apertura** de la mandíbula: las vocales abiertas se producen con el mayor grado de apertura de la mandíbula, mientras que las vocales cerradas se producen con el menor grado de apertura. También podemos hablar de esta dimensión articulatoria de las vocales refiriéndose al grado de **altura** de la lengua en vez del grado de **apertura** de la mandíbula. Las vocales **altas** se producen con el dorso de la lengua en un punto alto en la boca, acercándose al paladar, mientras que las vocales **bajas** se producen con el dorso de la lengua en un punto bajo de la boca, acercándose a la base de la boca. Existe una correlación casi exacta entre la apertura de la mandíbula y la altura de la lengua: por eso, podemos decir que todas las vocales altas también son vocales cerradas, y que todas las vocales bajas también son abiertas.

3.1.2. La transcripción fonética

Una transcripción fonética se hace con los símbolos del AFI que representan cada uno de los sonidos de la palabra o frase que queremos transcribir. El AFI utiliza símbolos que provienen principalmente de los alfabetos latino y griego. Existen algunas excepciones, con unos símbolos latinos o griegos modificados, y otros adoptados de otras escrituras. Esto significa que el AFI utiliza muchos más símbolos que los que aparecen en el alfabeto castellano. Para dominar la transcripción fonética, debemos aprender a conocer el sonido que representa cada uno de los símbolos del AFI.

El AFI utiliza símbolos que aparecen en el alfabeto castellano, pero, muchas veces, los sonidos que representan estos símbolos en el AFI son distintos de los sonidos que representan en el alfabeto castellano, particularmente en el caso de las consonantes. Para desambiguar el uso de letras del alfabeto castellano y símbolos del AFI, transcribimos el AFI entre [corchetes] y la escritura castellana entre <comillas angulares>. Por ejemplo, la palabra <alpaca> se transcribe de la forma siguiente en el alfabeto fonético internacional: [alpaka].

La <c> del castellano a veces se pronuncia [s], como en <cena>, y a veces se pronuncia [k], como en <comer>. En el AFI, el símbolo [c] representa el sonido de la <k> antes de una <i>, como en <kilo> [cilo], y el símbolo [k] representa el sonido de esa misma letra <k> antes de <a> u <o>, como en <koala> [koala]. (¡Presta atención a la posición de tu lengua al producir <k> en esas dos palabras!) La secuencia de letras <qu> en castellano también representa el sonido [k], como en la palabra <quechua>, mientras que el símbolo [q] del AFI representa el sonido de la letra <q> en aymara y quechua, como en la palabra quechua <alqo> ‘perro’. La <z> del castellano también se pronuncia [s], como en la palabra <zona>, mientras que la [z] del AFI se pronuncia como la <z> del inglés, como en la palabra inglesa <zoo>. La secuencia de letra <ch> del castellano representa el sonido [tʃ], como en la palabra <chile>.

Cada una de las letras <b, d, g> del castellano, de hecho, representa dos símbolos del AFI. Por ejemplo, en la palabra <dedo>, pronunciada [deðo], la primera <d> se pronuncia como la oclusiva dental [d̪]¹ y la segunda <d>, como la fricativa/aproximante interdental [ð]. En la palabra <griego>, pronunciada [griɐ̯o], la primera <g> se pronuncia como la oclusiva velar [g] y la segunda <g> se pronuncia como la fricativa velar [ɣ]. En la palabra <beber>, pronunciada [beβer], la primera se

1 El símbolo diacrítico [d̪] representa un diente y se coloca debajo de otro símbolo del AFI para indicar que la consonante en cuestión tiene un punto de articulación dental. La modificación del símbolo [d], que normalmente representa una oclusiva alveolar sonora, por el diacrítico dental [d̪] representa una oclusiva dental sonora.

pronuncia como la oclusiva bilabial [b] y la segunda como la fricativa bilabial [β]. Además, las letras y <v>, que se diferencian en el alfabeto castellano, ambas representan los mismos sonidos del AFI, es decir [b] y [β]. En la palabra <vaca>, la <v> representa el sonido [b], y en la palabra <ave>, la <v> representa el sonido [β]. Esta alternancia sistemática entre oclusivas y fricativas del mismo punto de articulación es lo que llamamos un **proceso fonológico**. Dejamos este asunto por ahora (ver Capítulo 2).

La <j> del castellano se pronuncia [x], como en <jaguar>, mientras que la <x> del castellano se pronuncia [ks], como en la palabra <léxico>. La letra <y> representa el sonido de la fricativa palatal [j], a veces también pronunciada como la aproximante palatal [j̞], como en la palabra <yuca>. La secuencia de letras <ll> representa el sonido de la aproximante lateral [ʎ], también a veces pronunciada como la aproximante palatal [j̞], como en la palabra <calle>. La secuencia de letra <rr> representa el sonido [z] en castellano de algunas regiones de Bolivia, como por ejemplo en la palabra <perro>. La letra <r> del castellano también a veces se usa para representar ese sonido, como en la palabra <rojo>. En los demás casos, la letra <r> del castellano representa el sonido [r] del AFI, como por ejemplo en la palabra <pero>. Por otro lado, el símbolo [r] del AFI representa el sonido de la vibrante alveolar, igual como se pronuncia la secuencia de letras <rr> en castellano mexicano o europeo.

Finalmente, la letra <ñ> del alfabeto castellano no existe en el AFI, y el sonido que esa letra representa se transcribe [ɲ] en el AFI. La letra <h> no representa ningún sonido en el alfabeto castellano, pero sí representa un sonido específico en el AFI, es decir, la fricativa glotal [h], como en la palabra del inglés <hello>.

En general, la transcripción de las vocales es más sencilla que la transcripción de las consonantes. El castellano tiene cinco vocales, [a, e, i, o u]. Conveniente para los estudiantes de fonética, cada uno de esos símbolos se pronuncia de la misma manera en el AFI y en el sistema ortográfico del castellano. Lo que puede causar confusión son los muchos símbolos adicionales usados para representar todas las vocales del AFI. Por ejemplo, la vocal central alta [i̞] es muy común en los idiomas hablados en la parte amazónica de Bolivia, pero esa vocal no existe en castellano. En muchos sistemas ortográficos de idiomas amazónicos, se usa la letra <y> para indicar el sonido [i̞]. ¡Pero ojo! El símbolo [y] en el AFI indica una vocal anterior alta y redondeada, como la que se escribe con la letra <u> en francés. Por ejemplo, la palabra francesa <rue> ‘calle’ se pronuncia [ʁy].

¡Hay que prestar mucha atención a estos símbolos del AFI para no confundirlos con las letras del alfabeto castellano! Es uno de los errores más frecuentes cometidos por los estudiantes que inician con el estudio de la fonética.

Además, cada uno de los símbolos del AFI puede ser modificado por **diacríticos**, es decir, por marcas pequeñas que se superponen por encima o por debajo del símbolo básico para cambiar alguna característica fonética de la pronunciación de ese sonido. Por ejemplo, una **tilde** [˜] se puede superponer al símbolo [a] para indicar que esa vocal baja es nasal y se pronuncia con el velo del paladar bajado, produciendo un flujo de aire nasal. Es decir que el símbolo compuesto [ã] indica una vocal baja nasal. Existen muchos diacríticos en el AFI, y es importante familiarizarse con todos ellos. Uno de los diacríticos más usados en los idiomas de Bolivia incluyen la tilde [˜] para indicar nasalización (ver la §3.3.4), el apóstrofo [ˈ] para indicar eyectivización (ver la §3.4.2.4.1), el superíndice [ᵒʰ] (ver la §3.2.3) para indicar aspiración y el símbolo de alargamiento [ː] para indicar un sonido largo. Regresaremos a estos temas en las secciones relevantes.

3.2. Consonantes

Las consonantes son una clase de sonidos que son pronunciados obstruyendo el aire que sale de los pulmones al momento de hablar. La obstrucción del aire puede hacerse de distintos grados, y en distintos lugares de la boca y la garganta. Además, el flujo de aire se puede modificar de distintas maneras para que este salga en forma de pequeñas explosiones, haciendo fricción o saliendo por la nariz, por ejemplo. Asimismo, una consonante puede ser pronunciada con sonoridad o sin ella, lo cual crea tipos de sonidos distintos en las lenguas.

En esta sección, primero veremos los lugares de la cavidad oral donde las obstrucciones pueden ocurrir al pronunciar una consonante, es decir, los **puntos de articulación**. Seguidamente, hablaremos acerca de las maneras en las que el aire es modificado al salir de la boca, es decir, los **modos de articulación**. Finalmente, hablaremos de cómo se genera la **sonoridad** en estos sonidos. Cada consonante tiene un valor especificado por cada uno de estos tres criterios.

3.2.1 Articuladores y puntos de articulación

Los órganos fonadores, también conocidos como articuladores, son el conjunto de músculos que producen cambios en la forma del tracto vocal para producir sonidos distintos durante el habla. Los órganos fonadores se mostraron en la Figura 2. Al producir el sonido de una consonante, siempre están involucrados dos articuladores: un articulador activo, es decir un órgano móvil encargado de moverse y de generar obstrucciones en alguna región del tracto vocal, y un articulador pasivo, es decir un punto estático al que el articulador móvil se acerque.

Cuando hablamos del punto de articulación de un sonido, nos referimos al par del articulador activo junto a un articulador pasivo. La lista y descripción de cada punto de articulación se da abajo, organizada por el articulador activo.

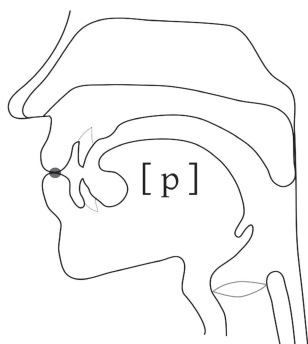
3.2.1.1. Sonidos labiales

Los **labios** son extensiones de tejidos de piel y músculo importantes tanto para pronunciar muchos sonidos como para darles forma a algunos. Están localizados en la entrada de la cavidad oral y sirven de articulador activo durante la producción de sonidos **bilabiales** y **labiodentales**.

Los sonidos **bilabiales** se articulan con los labios (**Figura 4**) que se juntan total o parcialmente, generando una obstrucción total o parcial entre los labios. Ejemplos de estos sonidos son las bilabiales [p], [b] y [m] en palabras del castellano como ['pipa], ['bombo] y ['mimo] respectivamente.

Figura 4

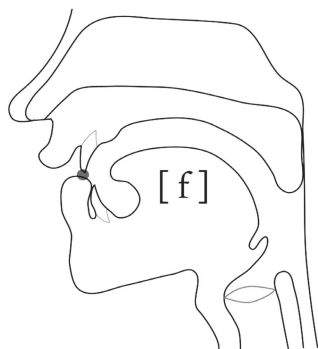
Sonido bilabial [p]



Los sonidos **labiodentales** son articulados con el labio inferior que se eleva para juntarse total o parcialmente contra los dientes superiores (**Figura 5**), como las consonantes labiodentales [f] del castellano o [v] del francés e inglés. Ejemplos de estos sonidos se encuentran en la palabra ['fasil] del castellano y ['vit] 'rápido' del francés.

Figura 5

Sonido labiodental [f]

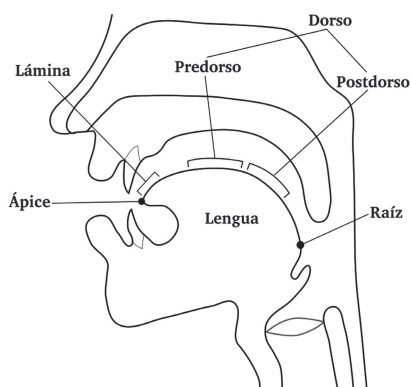


3.2.1.2. Sonidos linguales

La **lengua** es un articulador esencial para la producción de sonidos, puesto que puede adoptar distintas posiciones y formas en la cavidad oral. Las partes de la misma son el **ápice**, la **lámina**, el **dorso** (compuesto por el **predorso** y el **postdorso**) y la **raíz**; cada una es un articulador activo para la producción de ciertos tipos de consonantes. Ello queda ilustrado en la **Figura 6**.

Figura 6

Partes de la lengua



3.2.1.2.1 Sonidos apicales

El **ápice** de la lengua sirve de articulador activo durante la producción de sonidos **alveolares** y **retroflejos**. Los **alvéolos** están localizados entre el paladar duro y los dientes superiores. Son una protuberancia dura cubierta con crestas de piel que se pueden sentir si se tocan con la punta de la lengua detrás de los dientes. El contacto de la lengua con esta región de la cavidad oral ayuda a articular sonidos **alveolares** como [l], [t] o [ɾ].

Los **sonidos alveolares** se producen con el **ápice** de la lengua aproximándose a los alvéolos como en las consonantes [t], [s] y [l] en palabras quechua como ['tata] 'papá', ['sara] 'maíz' y ['lixra] 'ala'. La **Figura 7** ilustra una obstrucción alveolar durante la producción de la oclusiva alveolar [t].

Figura 7
Sonido alveolar [t]

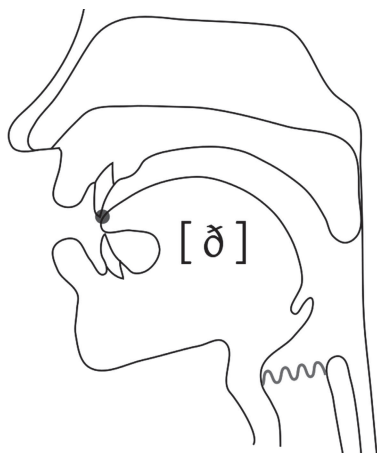


Los **sonidos retroflejos** se pronuncian con el ápice de la lengua flexionado hacia atrás en la región posterior a los alvéolos. Tienden a transcribirse con una colita hacia la derecha debajo de la consonante como en [t̠]. Varios idiomas de Bolivia tienen estos sonidos, entre ellos el maropa [t̠̠], el chácobo [ɬ̠], el tacana [d̠], el baure [ɽ̠] y el castellano hablado en las regiones andinas como La Paz y Cochabamba. En esta variedad de castellano, la consonante <rr> se pronuncia como una fricativa retrofleja sonora [ʒ̠] como en las palabras *perro* ['peʒ̠o] y *roto* ['ʒ̠oto].

3.2.1.2.2. Sonidos laminales

La **lámina de la lengua** sirve de articulador activo durante la producción de sonidos **dentales** y **post-alveolares**. Los **dientes** se sitúan inmediatamente detrás de los labios. Están hechos de material óseo y están ordenados en filas de dientes superiores e inferiores. La mayoría de los sonidos que se articulan en este lugar lo hacen con la lengua tocando o acercándose a los dientes superiores. Dichos sonidos se llaman *dentales*, como [ð], pero también existen sonidos *labiodentales*, como ya vimos en la sección anterior, como [f] o [v].

Los **sonidos dentales** son articulados cuando el ápice o lámina de la lengua entra en contacto con los dientes, generalmente con los dientes superiores como en la **Figura 8**. Sonidos como [ð] y [θ] son dentales. El idioma maropa de la Amazonía tiene este tipo de sonidos en palabras como ['ðite] 'ratón' y ['ðuðe] 'pato'.

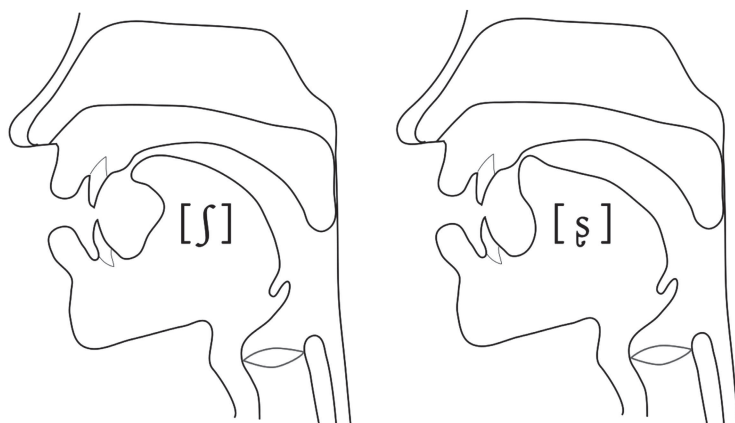
Figura 8*Sonido dental [ð]*

Los sonidos **post-alveolares** se articulan con la lámina de la lengua detrás de los alvéolos. Ejemplos de estos sonidos son las consonantes [ʃ, ʒ, tʃ, dʒ]. En la lengua ese ejja de Bolivia, las consonantes [ʃ] y [tʃ] se encuentran en palabras como ['ʃafa] 'flor' y [tʃitʃi'kʷia] 'paloma cafecita'.

Nótese que las consonantes post-alveolares y las retroflejas se parecen en el punto de articulación donde la obstrucción de aire ocurre. La única diferencia está en que, para una post-alveolar, la lámina de la lengua debe tocar la región posterior de los alvéolos, mientras que, para las retroflejas, es el ápice que hace esto. Véase la **Figura 9** para observar la articulación de ambos.

Figura 9

Puntos de articulación de una fricativa post-alveolar sorda [ʃ] y una fricativa retrofleja sorda [ʂ]



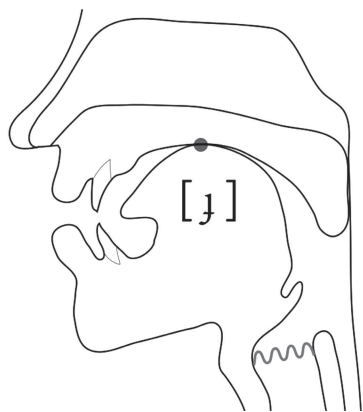
3.2.1.2.3. Sonidos dorsales

El **dorso de la lengua** sirve de articulador activo durante la producción de sonidos **palatales**, **velares** y **uvulares**. El **paladar (duro)**, situado en la parte superior de la cavidad oral, es una superficie ósea cóncava recubierta con piel en el interior de la boca. Los sonidos que involucran el predorso de la lengua acercándose a, o entrando en contacto con el paladar duro, se llaman *palatales*, tal es el caso de [j], [ɕ] o [ɲ]. El **velo del paladar** (o paladar blando) es una superficie muscular situada detrás del paladar (duro). Cuando el postdorso de la lengua se acerca a, o en contacto con, el velo del paladar (o paladar blando), se producen sonidos *velares* como [k, g, x, ɣ]. La úvula, por su parte, es un pequeño apéndice que cuelga en la sección inferior del velo del paladar. Algunos sonidos que se producen acercando el postdorso de la lengua a la úvula se llaman *uvulares*, como por ejemplo [q, ɤ, χ].

Los sonidos **palatales** se articulan con el paladar duro como órgano pasivo. Ejemplos de estos son las consonantes [ɲ, j, ʎ, c, ʝ] en palabras como ['ɲo] 'chancho de tropa', en ese ejja, ['jaku] 'agua' y ['ʎixʎa] 'manto', ambas en quechua boliviano del sur. La **Figura 10** ilustra la articulación de las consonantes palatales. En ese caso, el dorso de la lengua se considera el articulador activo, y el velo del paladar se considera el articulador pasivo.

Figura 10

Sonido palatal sonoro [j]



Nótese que el velo del paladar también puede levantarse y entrar en contacto con el espacio de la faringe, previniendo el escape del aire por la cavidad nasal y restringiendo la salida de este por la cavidad oral. Si el velo del paladar descende, abriendo el pasaje velofaríngeo, obtenemos sonidos nasales como [m] y [n], o nasalizados como [ã] y [ũ]. Cuando hablamos de la articulación oral o nasal de un sonido, consideramos el velo del paladar como el *articulador activo*, y no el

articulador pasivo, porque en estos casos es el órgano móvil (ver §3.2.2.2 y §3.3.4, para más información acerca de las consonantes y vocales nasales, respectivamente).

Los sonidos **velares** se producen con el dorso de la lengua acercándose al velo del paladar, como las consonantes [k, g, x, ŋ], que se encuentran en palabras como ['kikin] 'igual' y ['puχʌj] 'jugar' en quechua del sur. La oclusiva velar sonora [g] está ilustrada en la **Figura 11**.

Figura 11

Sonido velar [g]



Los sonidos **uvulares** se producen con el postdorso de la lengua interactuando con la úvula. El quechua boliviano tiene sonidos de este tipo, como [q] y [ɣ] en las palabras quechuas [qan] 'tú, vos' y ['oɰa] 'oca, camote'. La consonante uvular [q] está ilustrada en la **Figura 12**.

Figura 12

Sonido uvular [q]

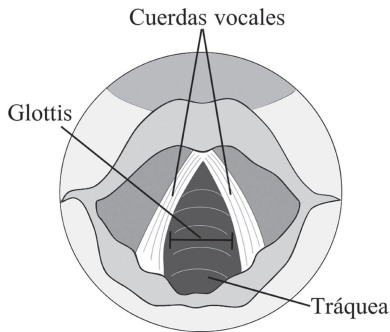


3.2.1.3. Sonidos laríngeos

La **laringe** es una estructura cartilaginosa en forma de tubo que conecta la tráquea con la faringe. A su vez, está compuesta de tres regiones que son el cartílago tiroides, el cartílago cricoides y los anillos traqueales. El aire pasa a través de este pasaje cuando respiramos y cuando producimos sonidos al hablar. Dentro de la laringe, se encuentran dos membranas llamadas cuerdas vocales. Las **cuerdas vocales** (o pliegues vocales) pueden moverse para separarse una de la otra, para acercarse o juntarse estrechamente. El espacio que se genera cuando estas están separadas se denomina **glottis** y es el espacio por donde el aire puede escaparse desde los pulmones. La **Figura 13** muestra las cuerdas vocales y la glottis.

Figura 13

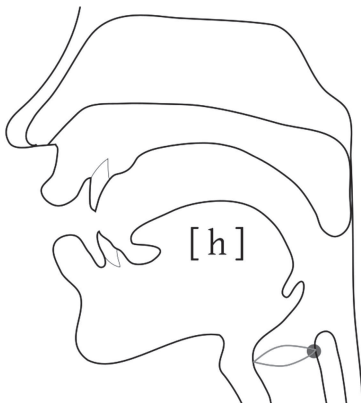
Las cuerdas vocales y la glottis



Los sonidos **laríngeos** o **glotales** se producen gracias a las cuerdas vocales en la laringe, que se encargan de obstruir el aire desde los pulmones hacia la cavidad oral, total o parcialmente, como se aprecia en la **Figura 14**.

Figura 14

Sonido glotal [h]



La oclusiva glotal [ʔ] es común en muchas lenguas de Bolivia como el ese ejja o el movima, y la fricativa glotal [h] está presente en aún más lenguas. Estos sonidos quedan ilustrados en palabras ese ejja como [k^{wa}'ʔiʔi] ‘picaflor, colibrí’ y ['hihiiani] ‘alumbrar, arder el fuego’.

3.2.2. Modos de articulación

Otro criterio para distinguir clases de consonantes es la forma en la que el aire sale del tracto oral o nasal, o el grado de constricción que se realiza, denominado **modo de articulación**. Dado que los grados y tipos de constricción al pronunciar una consonante varían, tenemos distintas maneras de articulación descritas en las siguientes subsecciones.

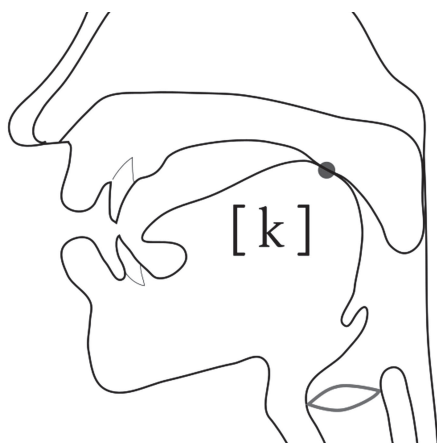
3.2.2.1 Oclusivas

Al pronunciar el sonido [p] como en la palabra *papá* [pa'pa], existe una obstrucción completa de los labios, ya que es un sonido bilabial. Luego de esta obstrucción total, el aire dentro de la boca se comprime de manera que, cuando la constricción se libera, el aire sale de manera abrupta. La liberación del aire a veces genera un poco de fricción parecida a una explosión pequeña. Esto hace que los sonidos producidos de esta manera sean llamados **oclusivos**.

Existen dos tipos de consonantes oclusivas: las **plosivas**, propiamente dichas, y las **africadas**. Las consonantes **oclusivas** son aquellas que producen apenas una pequeña explosión luego de la obstrucción y la compresión del aire. Estos son sonidos como [p, b, t, d, k, g, q, ɢ]. La **Figura 15** ejemplifica la articulación de la oclusiva [k], con el postdorso de la lengua que entra en contacto con el velo del paladar.

Figura 15

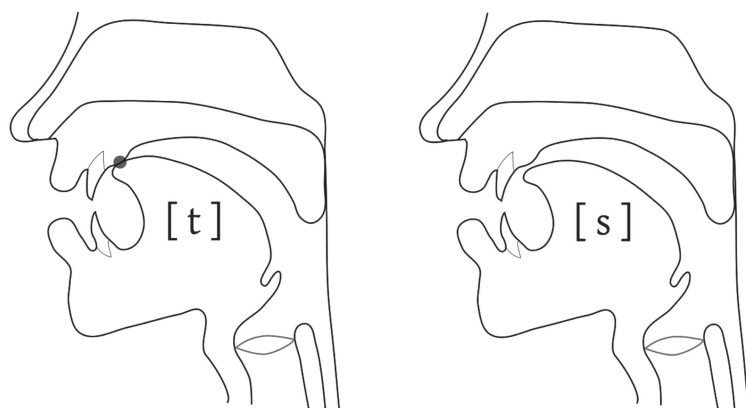
Articulación de la oclusiva velar sorda [k] como en la palabra coca ['koka]



Las consonantes **africadas**, por su parte, generan una fricción o turbulencia igual a la generada por las consonantes fricativas (ver §3.2.2.3) al resaltar la obstrucción en el tracto vocal. Por ejemplo, al pronunciar la primera consonante de la palabra **choza**, notamos que el sonido [tʃ] contiene un momento donde la lámina de la lengua obstruye la región postalveolar y luego el aire se libera generando mucha fricción. Otros sonidos africados comunes son [dʒ, ts, dʒ, tʃ, dʒ] (ver §3.4.1.1 para más información acerca de las consonantes africadas). La **Figura 16** ilustra la pronunciación de la consonante africana alveolar sorda [ts]: al inicio de la producción de ese sonido, el ápice de la lengua crea una obstrucción en los alvéolos, como para producir la oclusiva alveolar sorda [t] (izquierda); al resaltar la obstrucción, el ápice de la lengua se posiciona como para producir la fricativa alveolar sorda [s] (derecha).

Figura 16

Las dos fases de articulación de la africana alveolar sorda [ts]

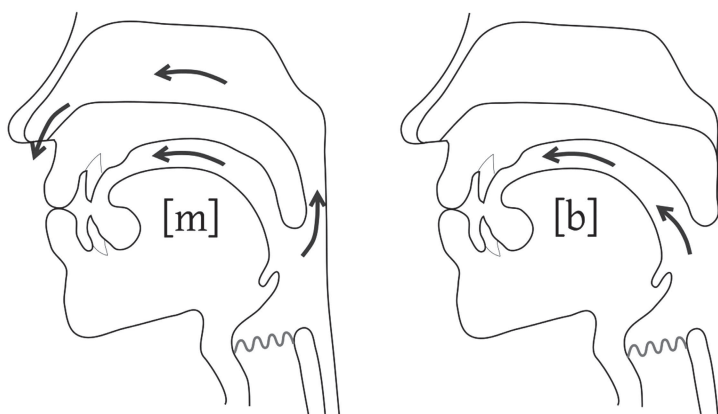


3.2.2.2 Nasales

Las consonantes u oclusivas nasales tienen el mismo proceso de articulación que las oclusivas, pero se producen con el velo del paladar descendido. Existe una oclusión en algún punto del tracto vocal, pero el aire no se comprime tanto porque este sale continuamente por la **cavidad nasal**, en vez de hacerlo por la cavidad oral. Ello puede notarse en ambas consonantes de la palabra *mimo* [mimo]. Esto se logra porque el velo del paladar se relaja y se baja, de manera que el pasaje velofaríngeo, es decir el pasaje que conecta la cavidad nasal a la cavidad oral, queda abierto y el aire sale por este tracto mientras que la obstrucción continúa en la cavidad oral. Ejemplos de consonantes nasales son [m, n, ɲ, ŋ, ɳ]. Véase la **Figura 17** para comparar el velo del paladar al articular la consonante bilabial nasal [m] (izquierda), y su contraparte oral, la bilabial consonante oclusiva sonora [b] (derecha).

Figura 17

Articulación de una consonante nasal bilabial sonora con el velo descendido (izquierda), comparado con una consonante oclusiva bilabial sonora con el velo del paladar levantado (derecha)

**3.2.2.3. Fricativas**

Las consonantes **fricativas** reciben ese nombre porque la constricción que un articulador activo realiza en la boca no es total, sino estrecha. Esto hace que el aire salga continuamente por la boca en forma de turbulencia o **fricción**. Al pronunciar sonidos como [s] o [ʃ] en la palabra movima *sojła* [sojła] ‘dentadura’, podemos notar dicho proceso. Otras consonantes fricativas son [v, β, s, z, ʃ, ʒ, ɣ, ʁ, x, χ, h]. La **Figura 18** ilustra la articulación de la fricativa labiodental sonora [v].

Figura 18

Articulación de la fricativa labiodental sonora [v]



3.2.2.4 Líquidas

Las consonantes líquidas son sonidos para los que hay una obstrucción pequeña en una parte del tracto vocal, pero no suficiente para que el aire no se quede atrapado ni tampoco salga en forma de fricción. Durante la producción de las consonantes líquidas, el aire fluye libremente en el tracto vocal. Hay distintos tipos de consonantes líquidas que comprenden sonidos como [l] o [r].

Las consonantes **laterales** se producen creando una obstrucción total con alguna parte central de la lengua, pero no con sus partes laterales. Es decir que el aire se escapa por uno o ambos lados de la lengua fluyendo por las mejillas internas, como [l, ʎ, ʟ, ʠ]. Por su parte, las consonantes **vibrantes** son aquellas para las que un articulador activo solamente obstruye el aire por periodos de tiempo muy pequeños durante un instante muy breve. La vibrante simple [r] como en *parar* [parar] solo necesita un toque del ápice de la lengua contra los alvéolos, mientras que la vibrante múltiple [ʀ] como en la palabra *perro* [pero] requiere varios toques del ápice de la lengua contra los alvéolos, y estos se realizan muy rápidamente en cuestiones de milisegundos. Otros sonidos vibrantes son [ʀ, ʁ].

En araona, una lengua Tacana de Bolivia, existe un sonido que se pronuncia como una fusión entre la vibrante simple [r] y la lateral alveolar [l], transcrita como [ɭ] que es una vibrante simple lateral alveolar sonora. Ejemplos de este sonido se dan en (1).

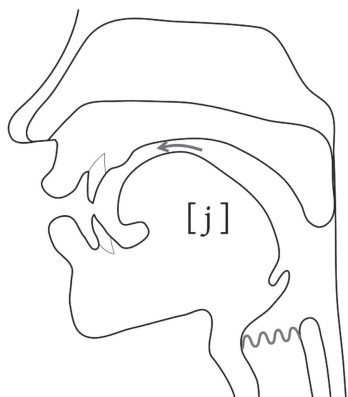
- | | | | | |
|-----|----|------------|----------------------|-----------------|
| (1) | a. | [kʷa'ɭoɭo] | ‘redondo’ | |
| | b. | [no'ɭiɭi] | ‘serpiente pucarara’ | |
| | c. | [ʰoɭi] | ‘felino’ | (Tallman, 2019) |

3.2.2.5. Aproximantes

Las consonantes aproximantes se pronuncian con el articulador activo acercándose (o aproximándose) al articulador pasivo, pero sin generar constricción del aire. De esta manera, el aire sale más libremente sin generar turbulencia, como por ejemplo durante la producción del sonido [w] en la palabra quechua *wasi* ['wasi] ‘casa,’ o en la palabra maropa *wambu* ['wambu] ‘chancho de tropa’. En la **Figura 19**, se ilustra la articulación de la aproximante palatal [j], como en la palabra maropa *yami* ['jami] ‘carne’.

Figura 19

Articulación de la aproximante palatal sonora [j]

**3.2.3. Sonoridad**

Al articular consonantes, un tercer criterio para describirlas es el de **sonoridad**. Este término se refiere al estado de las cuerdas vocales durante la producción de alguna obstrucción en el tracto vocal al pronunciar una consonante. Si las cuerdas vocales están estrechamente cerradas al pronunciar una consonante, el flujo dinámico del aire que pase por ellas provocará que vibren y obtendremos una consonante **sonora**. Notamos esto si tocamos nuestra laringe o garganta con la mano mientras pronunciamos por varios segundos la consonante bilabial nasal [m] como en [mmmmm]. Sentiremos vibraciones en esta parte de la garganta, lo que nos indica que estamos ante vibración de las cuerdas vocales y, por lo tanto, de un sonido sonoro.

Por otra parte, si las cuerdas vocales están separadas al pronunciar un sonido, este es clasificado como **sordo**. Esto se nota más fácilmente si pronunciamos la consonante [s] por varios segundos mientras tocamos nuestra laringe o garganta: [sssss]. Al hacer este ejercicio, no sentimos ninguna vibración en esta parte de la garganta porque las cuerdas vocales no pueden vibrar cuando están separadas.

Como pudo verse en el cuadro de consonantes del AFI en la Figura 1, las consonantes en cada punto y modo de articulación vienen generalmente en pares. Es decir que, en casi cada celda de la tabla, existen dos sonidos que tienen el mismo punto y el mismo modo de articulación. Esto se debe a que puede haber una consonante fricativa velar sorda [k] y otra consonante fricativa velar sonora [g]. Para [k], las cuerdas vocales están separadas, mientras que para [g], las mismas están estrechamente juntas generando vibración. La **Figura 20** muestra la articulación de estas dos consonantes cuya articulación es idéntica, salvo por su sonoridad: la sonoridad es notable en [g] gracias a la línea ondulada.

Figura 20

Distinción de sonoridad entre [k] (izquierda) y [g] (derecha)



En el AFI, las consonantes de la izquierda de una celda son sordas y las de la derecha son sonoras. La manera de agrupar a consonantes de acuerdo con su sonoridad se logra dividiéndolas entre sordas y sonoras. Así, [p, t, k, q, f, θ, s, ʃ, ʂ] son ejemplos de consonantes sordas, y [b, d, g, ɠ, v, ð, z, ʒ, ʐ] son sus pares sonoros.

En la lengua yurakaré, la consonante bilabial sorda [p] es distinta a la consonante bilabial sonora [b], al igual que la alveolar sorda [t] es diferente a la alveolar sonora [d]. Es decir, existe una distinción que solo se encuentra en la sonoridad como se ve en los ejemplos (2) obtenidos de van Gijn (2006, p. 22).

- | | | | |
|-----|----|---------|--------------------------|
| (2) | a. | [abæbæ] | ‘está buscando’ |
| | b. | [apæpæ] | ‘su abuelo’ |
| | c. | [duta] | ‘(se) quema’ |
| | d. | [tuta] | ‘él nos dice a nosotros’ |

Algo que es notable en las lenguas es que hay una inclinación a que las consonantes oclusivas sean básicamente sordas. Esto implica que, si una lengua tiene oclusivas sonoras [b, d, g], también tendrá sus contrapartes sordas [p, t, k]. El yurakaré y el mosetén son idiomas que tienen ambos pares de sonidos, mientras que el quechua solo tiene consonantes oclusivas sordas.

Por su parte, las consonantes fricativas también siguen esa tendencia, pero hay idiomas en los que puede haber **consonantes fricativas sonoras**, pero sin una contraparte sorda como el maropa, que tiene las fricativas [β, ð, s, ɛ, h], de las cuales las dos sonoras [β, ð] no tienen una contraparte sorda [φ, θ]. Muchas veces, estas se parecen más a consonantes aproximantes.

Con respecto a las consonantes nasales, es común que estas sean predominantemente sonoras en los idiomas. En Bolivia, la mayoría de las lenguas tiene consonantes nasales sonoras como [m] o [n]. El ayoreo es una de las pocas lenguas conocidas por tener **consonantes nasales sordas** [ṃ, ṇ, ṅ] además de las consonantes nasales sonoras [m, n, ɲ, ŋ]. Nótese que el ensordecimiento de un sonido sonoro se transcribe con un pequeño círculo debajo o encima de la consonante [Ḷ Ḷ̌].

La división entre consonantes sordas y consonantes sonoras es muy importante en las lenguas del mundo. Los idiomas pueden presentar sonidos sonoros y sordos que los hablantes distinguen. Ese es el caso de los sonidos oclusivos sordos [p, t, k] del castellano que tienen su contraparte sonora [b d g]. Esto se ve en las palabras [‘paso] vs [‘baso] o [te] vs [de].

Al producir una secuencia de una consonante sorda seguida por una vocal sonora, como por ejemplo [te], la vibración de las cuerdas vocales empieza más o menos en el momento que se inicia la producción de la vocal [e], después de que se libere la constricción de la oclusiva [t]. Las **consonantes sordas aspiradas** causan un retraso pequeño en el momento que empieza la vibración de las cuerdas vocales, porque crean un sople breve de aire al liberar la obstrucción en el tracto vocal.

Hay lenguas que presentan un contraste que no involucra consonantes sordas y sonoras, sino que contrastan **consonantes sordas** con **consonantes sordas aspiradas**. El quechua boliviano es un ejemplo en el que sonidos como [p, t, tʃ, k, q] se distinguen de [pʰ, tʰ, tʃʰ, kʰ, qʰ]. Por ejemplo, la palabra [tanta] ‘agrupación, reunión’ es distinta a [tʰanta] ‘usado’.

Más allá aún se encuentran las lenguas que contrastan **consonantes oclusivas sordas** con **consonantes sonoras** y además **consonantes sordas aspiradas**. El idioma quichua de Salasca, hablado en Ecuador, presenta dichos sonidos. Estas consonantes son las oclusivas sordas [p, t, k], las sonoras [b, d, g] y las aspiradas [pʰ, tʰ, kʰ]. Esto queda ilustrado en las palabras [pungu] ‘puerta’, [bun̥ga] ‘abeja’ y [pʰaki] ‘roto’ (Chango y Marlett, 2006).

Vimos que reconocer si una consonante es sonora o sorda puede sentirse al verificar si hay vibración en nuestras cuerdas vocales. Si pensamos en las vocales, sin embargo, veremos que estas son exclusivamente sonoras. Basta con pronunciarlas para darnos cuenta de ello. De hecho, las vocales sordas en las lenguas del mundo no son frecuentes porque es más natural pronunciar estos sonidos si son sonoros. En la siguiente sección, veremos más detalladamente esta clase de unidades.

3.3. Vocales

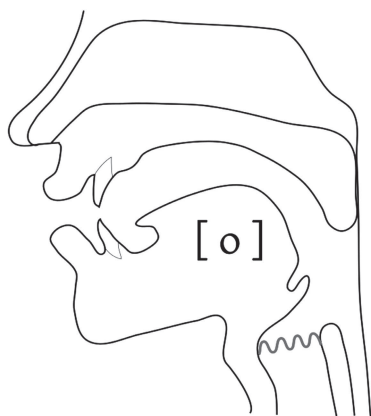
Desde el punto de vista fonético, una **vocal** es un tipo de sonido para el cual **ningún tipo de constricción** u obstrucción se genera en la cavidad oral al pasar

el aire desde los pulmones hasta los labios. Para articular vocales, modificamos la forma y dimensiones del tracto vocal, gracias a las varias posiciones que adoptan la lengua, los labios y la mandíbula. En la producción de las vocales, el dorso de la lengua siempre es el articulador activo, mientras que el paladar es el articulador pasivo. Al producir una vocal, las **cuerdas vocales siempre vibran**.

Por ejemplo, las vocales de la lengua ese ejja son básicamente cuatro [i e a o] (ver §3.1.1 y **Figura 3** para una vista de los demás símbolos de vocales en el AFI). Si hacemos el ejercicio de pronunciar esas vocales una seguida de la otra, notaremos que el dorso de nuestra lengua se posiciona de distintas maneras con cada una de ellas. Esto se debe a que las vocales se distinguen por su grado de **altura** (§3.3.1) y **anterioridad** (§3.3.2), es decir, la posición de nuestra lengua dentro de la boca. Al decir la vocal [a] y luego la vocal [o], notaremos que, para esta última, ambos labios toman ligeramente una forma de círculo. Esta distinción entre la posición de los labios se refiere al grado de **redondeamiento** (§3.3.3). Finalmente, para producir las cuatro vocales del ese ejja, los hablantes de este idioma deben elevar el velo del paladar para que toque la pared de la faringe y cierre el pasaje velofaríngeo (ver la **Figura 21**). Esto se debe al hecho de que las vocales del ese ejja son orales específicamente con respecto a su grado de **nasalización** (§3.3.4).

Figura 21

Vocal posterior media [o]



3.3.1. Altura

El grado de **altura** de una vocal se refiere a la **distancia vertical** entre el dorso de la lengua y el paladar. Hagamos el ejercicio de elevar el dorso de la lengua hasta tocar el paladar y luego bajarlo hasta la posición donde se encontraba al inicio repetidas veces. Notaremos que, al producir este movimiento, el espacio entre ambos

articuladores se reduce o se amplía, según cuán alta esté la lengua o cuán baja esté con respecto al paladar.

Cuando producimos vocales, usamos este mecanismo articulatorio para distinguirlas una de la otra. Por ejemplo, si producimos la **vocal baja** [a], podemos sentir que nuestra mandíbula ha descendido y que nuestra lengua está en una posición baja. En comparación, si producimos la **vocal alta** [i], podemos sentir que nuestra mandíbula ha subido y que nuestra lengua está en una posición alta. Esto implica que el espacio en nuestra boca es más amplio al producir una [a] que una [i]. En ese sentido, según su altura, las vocales pueden clasificarse en tres tipos generales: **bajas**, **medias** y **altas**. Entre estos grados de altura, podemos encontrar puntos intermedios como las vocales medias-bajas y medias-altas, pero nosotros las trataremos como medias por simplificación.

Las **vocales bajas** se articulan con la mandíbula descendida y la lengua baja, de manera que la boca quede abierta y el dorso de la lengua esté distante del paladar. En la lengua yurakaré (aislada), encontramos dos vocales bajas [æ] como en [tæhtæ] ‘pierna’ o [pæpæ] ‘abuelo’ y [a] como en [bakta] ‘plano’ o [ana] ‘esto’ (Gijn, 2006, pp. 22, 25, 28).

Las **vocales altas** son pronunciadas con el dorso de la lengua en la posición más elevada posible, pero, aun así, permite que el aire escape libremente, es decir, sin generar ninguna obstrucción o turbulencia en el flujo del aire. El yurakaré tiene tres vocales altas: [i] como [pii]² ‘hermano mayor de una mujer’, [i] como en [pii] ‘camino’ y [u] como en [bujbu] ‘idioma’ (Gijn, pp. 28, 48).

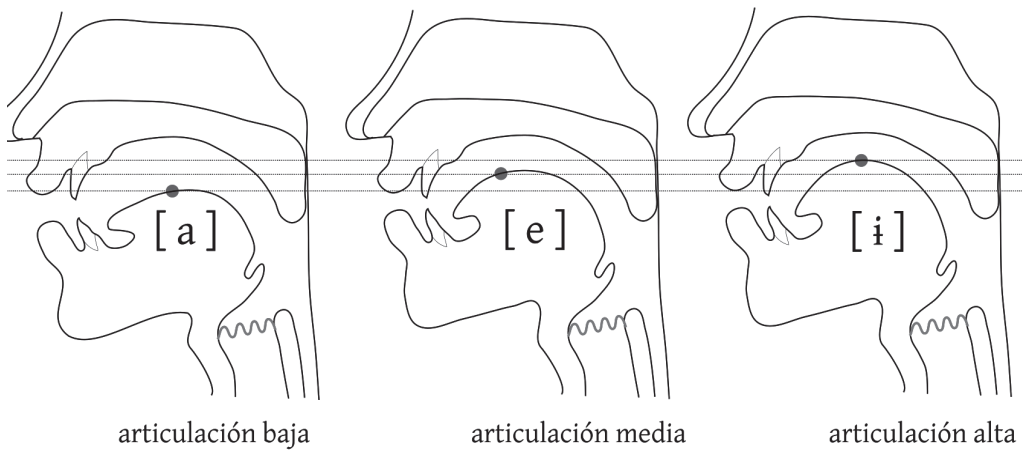
Por su parte, las **vocales medias** se producen con la mandíbula y el dorso de la lengua en una posición intermedia que toma para producir vocales altas y bajas. En yurakaré, también contamos con este tipo de vocales, específicamente las vocales [e] como en [tehte] ‘abuela’ y [o] como en [foŋko] ‘hoyo’ (Gijn, p. 28).

La **Figura 22** ilustra la diferencia en el grado de altura de las vocales [a, e, i]. En esta imagen, podemos comparar la altura del punto más alto de la lengua con las líneas horizontales punteadas. Podemos ver que este punto recae sobre la línea más baja para la vocal baja [a] (izquierda), sobre la línea más alta para la vocal alta [i] y sobre la línea del medio para la vocal media [e]. Como el yurakaré, la mayoría de los idiomas del mundo distinguen estos tres grados de altura, y muchos otros distinguen vocales con grados de altura intermedios entre alto y medio, o entre medio y bajo.

2 La secuencia de dos vocales en estos ejemplos pertenece a distintas sílabas en vez de ser vocales largas.

Figura 22

Los tres grados de altura vocálica para las vocales [a, e, i]



En esta sección, hemos visto que los sonidos vocálicos pueden variar en su calidad gracias a su altura. No obstante, como queda evidenciado en el par de palabras en yurakaré [pii] y [piɪ] que contienen vocales altas, la altura no puede distinguir entre las vocales [i] e [ɪ]. Entonces, ¿cuál es el mecanismo articulatorio que hace que sean distintas?

3.3.2. Anterioridad

Las vocales también pueden distinguirse por su anterioridad. Esta dimensión articulatoria se llama grado de **anterioridad** y se refiere a la **posición horizontal** del dorso de la lengua durante la producción de las vocales. Pronunciemos la secuencia [i, u, i, u, i, u] para notar cómo el movimiento del dorso de la lengua nos orienta con respecto a la anterioridad de estas vocales. Al producir la vocal anterior [i], podemos sentir que nuestra lengua se mueve hacia la zona anterior de la boca, mientras que, si pronunciamos una [u], notaremos que la lengua ha realizado un movimiento hacia atrás. De esta manera, según su anterioridad, las vocales pueden clasificarse en tres tipos generales: **anteriores**, **centrales** y **posteriores**.

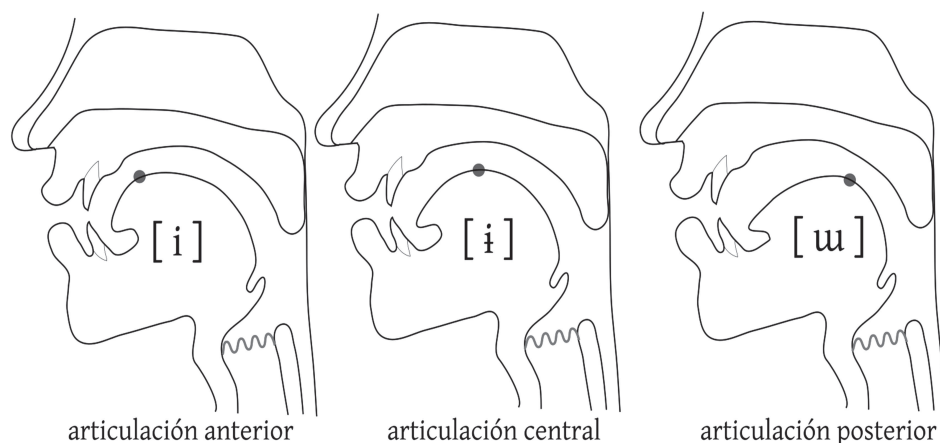
Las **vocales anteriores** se articulan con el predorso de la lengua movido hacia adelante, apuntando a los alvéolos o los dientes. Las vocales [i, e, ɛ] son todas vocales anteriores. En cambio, las **vocales posteriores** se articulan con el postdorso de la lengua movido hacia atrás, acercándose a la dirección del velo del paladar. Las vocales [u, o, ɔ] son posteriores.

Ahora, hablamos de una **vocal central** si la lengua no toma ninguna de las dos direcciones mencionadas arriba y, en su lugar, el dorso de la lengua permanece en un punto intermedio en la cavidad oral. La famosa vocal *schwa* [ə] del inglés es una vocal central.

Volviendo al caso de la lengua yurakaré, vimos que dos vocales [i] e [i] en [pii] ‘hermano mayor de una mujer’ y [pi:] ‘camino’ coincidían en su altura, pero difieren en otro rasgo que es el de la anterioridad. La vocal [i] es una vocal anterior, mientras que [i] es una vocal central. Intentemos pronunciar la vocal [i] durante varios segundos y, poco a poco, recorramos el dorso ligeramente hasta alcanzar la región central de la cavidad oral; si movemos la lengua demasiado hacia atrás, llegaremos a pronunciar, en su lugar, una vocal posterior como [u] o [u]. La diferencia en la anterioridad de las vocales altas (no-redondeadas) [i, i, u] se muestra en la **Figura 23**.

Figura 23

Localización de las vocales anteriores, centrales y posteriores [i, i, u]



Así como el yurakaré, el guaraní ava de Bolivia también presenta vocales que se diferencian en su articulación a través de su grado de anterioridad, es decir [i, i, u]. La vocal [i] es anterior, [i] es central y [u] posterior. Los ejemplos en (3), tomados de Daviet (2016, p. 62) presentan pares de palabras que se distinguen solamente por el grado de anterioridad de una vocal.

- | | | | | |
|-----|-------------|---------|-----------|----------|
| (3) | a. ['nãmbi] | ‘oreja’ | c. ['iwi] | ‘tierra’ |
| | b. ['ãmbi] | ‘moco’ | d. ['uwi] | ‘flecha’ |

Terminamos nuestra discusión de la anterioridad de las vocales mirando datos del weenhayek, una lengua Mataco hablada en Tarija. El weenhayek tiene 6 vocales

que contrastan la anterioridad de manera muy ilustrativa para 6 vocales, como se muestra a continuación (ejemplos de Claesson, 2017, p. 4; Alvarsson & Claesson, 2014, p. 431):

(4)	vocales anteriores		vocales posteriores
	i [tix] ‘lo cava’	u	[tux ^w] ‘lo come’
	e [tex] ‘aunque’	o	[ʔatox ^w] ‘lejos’
	a [ʔaq] ‘tu atadura’	ɑ	[ʔɑq] ‘tu comida’

En esta sección, hemos visto que los sonidos vocálicos pueden variar en su calidad gracias a su anterioridad. Hasta ahora, hemos hablado de dos vocales posteriores altas, [u] y [ʊ]. Claramente, estas dos vocales se distinguen una de la otra, pero todavía no hemos estudiado el mecanismo articulatorio que nos sirva para diferenciarlas. En la próxima sección, hablaremos del grado de **redondeamiento** de las vocales, lo que nos ayudará en esta tarea.

3.3.3. Redondeamiento

Mientras que la altura y la anterioridad se refieren a la posición del dorso de la lengua al producir una vocal, el grado de **redondeamiento** de una vocal se refiere a la forma de los labios. Esta característica articulatoria indica si los labios están **extendidos**, como para pronunciar las vocales [i, e, ɪ, a], o si forman un **círculo** mientras ambos labios sobresalen, como al pronunciar las vocales [u, o]. A las primeras, las denominamos **vocales no redondeadas**, mientras que, a las segundas, las llamamos **vocales redondeadas**. Podemos notar la diferencia en la forma que nuestros labios adquieren al pronunciar las vocales [i, u] repetidas veces.

Para ilustrar este rasgo, intentemos pronunciar la vocal no redondeada [i] durante varios segundos. Mientras lo vamos haciendo, sin soltar el aire, hagamos un esfuerzo en redondear nuestros labios formando un “pico”. Notaremos que el sonido que estamos produciendo es muy distinto a la [i]. Esto se debe a que, al redondear los labios, la calidad de la vocal cambia, y en su lugar tenemos la vocal redondeada [y]. Intentemos pronunciar esos sonidos alternando la forma de los labios de extendidos a redondeados, y podremos percibir mejor el cambio de vocales [i, y, ɪ, y].

En ciertos idiomas, existen palabras que se diferencian simplemente por el grado de redondeamiento de los labios. El wari’, idioma amazónico de Brasil, es uno de estos idiomas. Esta lengua tiene vocales no redondeadas y redondeadas como [i], [y] en [tʃim] ‘noche’ y [ʔyʔym] ‘madrugada’, o [e], [ø] en [tsek] ‘día’ y [kaʔmø] ‘capibara’ (Everett y Kern, 1997).

Otro idioma conocido por tener pares de vocales redondeadas y no redondeadas es el francés. Esta lengua tiene las vocales [i, e, ɛ] y [y, ø, œ], que respectivamente forman contrapartes donde la única diferencia es el grado de redondeamiento de los labios. De este modo, las palabras *lit* [li] ‘cama’ y *lu* [ly] ‘leído’ se distinguen solo por el grado de redondeamiento de la vocal. De la misma manera, las palabras *sait* [se] ‘sabe’ y *ceux* [sø] ‘estos’, y *père* [pɛʁ] ‘padre’ y *peur* [pœʁ] ‘miedo’ son ejemplos de pares mínimos donde la forma de los labios marca la distinción.

3.3.4. Nasalidad

El grado de **nasalización** de una vocal se refiere a la posición del velo del paladar al pronunciar una vocal. Además de la altura y la anterioridad de la lengua, y el redondeamiento de labios, la nasalización es una cuarta dimensión articulatoria que sirve para diferenciar pares de vocales. De hecho, ya hemos hablado de la nasalización con respecto a las nasales [m, n, ɲ, ŋ] en §3.2.2.2. Cuando se trata de las vocales, la nasalización se marca con una **tilde** [~] sobre la vocal, por ejemplo [ã].

Las **vocales nasales** se articulan con el velo del paladar descendido y el pasaje velofaríngeo abierto, permitiendo que el aire salga simultáneamente por la cavidad nasal y la cavidad oral. Por su parte, las **vocales orales** se articulan con el velo del paladar subido y el pasaje velofaríngeo cerrado, dejando que el aire salga solo por la cavidad oral. En la **Figura 24**, se muestra una vocal nasal [ã] (izquierda) y una vocal homorgánica oral [a] (derecha).

Figura 24

Posición del velo del paladar para la vocal nasal [ã] vs. la vocal oral [a]



En Bolivia, existen varios idiomas con vocales nasales. De hecho, la mayoría de los idiomas amazónicos, como todas las lenguas Tupí-Guaraní o el ayoreo, contienen pares de vocales nasales y orales. El yuquí, una lengua Tupí-Guaraní del trópico de Cochabamba, tiene varias vocales nasales, como se aprecia en (5).

(5)	<i>vocales nasales</i>			<i>vocales orales</i>	
a.	[tãĩ]	‘madre’	vs	[tai]	‘casa’
b.	[ˈdʒitiĩ]	‘tipo de abeja’	vs	[ˈdʒiti]	‘hijo, niño’
c.	[ˈisã]	‘su pie’	vs	[ˈisa]	‘su oreja’

En esta sección, hemos estudiado varias maneras en las que las vocales se distinguen desde un punto de vista articulatorio: la altura, la anterioridad, el redondeamiento y la nasalidad. Muchas vocales se producen con un solo valor para cada uno de estos parámetros articulatorios, pero existen vocales que se definen por tener más de una fase articulatoria, cada una con un valor distinto para uno de estos parámetros articulatorios. Por ejemplo, la palabra [tai] ‘casa’ del yuqui en (5a) es una de estas vocales, empezando con la articulación de la vocal [a] y terminando con la articulación de la vocal [i]. A estas vocales complejas llamamos **diptongos**. En la próxima sección, trataremos este asunto, hablando tanto de las consonantes complejas como de las vocales complejas.

3.4. Segmentos complejos

Un aspecto interesante de los sonidos del habla humana es que no todos se articulan con un solo punto o una sola manera de articulación. Existen segmentos que se componen de dos o más sonidos o articulaciones. En esta sección, miraremos dos fenómenos: los segmentos complejos y las articulaciones secundarias.

3.4.1. Sonidos complejos

En muchos idiomas, existen **sonidos complejos**, que tienen dos fases fonéticas, cada una transcrita con un símbolo distinto del AFI. Por ejemplo, la consonante africada del castellano [tʃ] está compuesta de una fase oclusiva alveolar [t], seguida de una fase fricativa post-alveolar [ʃ]. La **ligadura** [͡] es un diacrítico que se pone encima de una secuencia de símbolos del AFI para indicar que esa secuencia forma un solo sonido complejo. Es importante diferenciar una *secuencia de sonidos* de un *sonido complejo*. El sonido complejo, aunque sea compuesto de dos fases, no se puede dividir en dos sonidos. En nuestro ejemplo, el sonido [ʃ] no existe independientemente de la africada [tʃ] en castellano. Por lo tanto, decimos que la consonante africada [tʃ] es un solo sonido, y se percibe como un sonido único por los hablantes nativos del castellano.

3.4.1.1. Africadas

Las africadas son consonantes complejas cuyas dos fases fonéticas se describen como una **fase oclusiva** seguida de una **fase fricativa**. Muchos idiomas bolivianos tienen series complejas de consonantes africadas. Por ejemplo, el chipaya, también

conocido como uru chipaya, tiene nueve africadas, con tres puntos de articulación (alveolar, post-alveolar y retroflejo) y tres modos de articulación (simple, aspirado, eyectivo) distintos (Cerrón-Palomino, 2006). Todas esas consonantes africadas aparecen en la Tabla 1, donde vemos que cada una se distingue de las otras en su articulación.

Tabla 1

Consonantes africadas del chipaya

	alveolar	post-alveolar	retrofleja
simple	[ts]	[tʃ]	[tʂ]
aspirada	[tsʰ]	[tʃʰ]	[tʂʰ]
eyectiva	[tsʼ]	[tʃʼ]	[tʂʼ]

En su inventario de sonidos, el chipaya también incluye las fricativas simples [s, ʃ, ʂ] y las oclusivas alveolares [t, tʰ, tʼ]. Pero es interesante notar que el inventario del chipaya *no incluye* ninguna fricativa aspirada o eyectiva. Por lo tanto, podemos decir que las africadas del chipaya no se pueden dividir en dos sonidos distintos, porque los sonidos [ʃʰ, ʂʰ, ʃʼ, ʂʼ] no existen independientemente de las africadas en las cuales se observan.

3.4.1.2. Consonantes complejas nasales-orales

Muchos idiomas del mundo y muchos idiomas hablados en Bolivia incluyen en su inventario de sonidos consonantes complejas con una fase oral y una fase nasal. Por ejemplo, el sirionó (Tupí-guaraní) tiene una serie de consonantes complejas nasales-orales con tres puntos de articulación (bilabial, alveolar y post-alveolar): [mb̃, nd̃, ɲd̃] (Gasparini, 2012). Durante la articulación de estas consonantes, el velo del paladar está bajado al inicio de la consonante, dejando el aire pasar por la cavidad nasal, y está subido al final de la consonante, solo dejando pasar el aire por la cavidad oral. En sirionó, como en muchos otros idiomas amazónicos, estas consonantes complejas aparecen antes de las vocales orales. En otras palabras, podemos decir que las consonantes nasales [m, n, ɲ] ‘ganan’ una fase oral cuando aparecen inmediatamente antes de una vocal oral. Llamamos a esta transformación articulatoria un proceso de **post-oralización**.

El sirionó también tiene una segunda serie de consonantes nasales-orales [nt̃, ɲd̃, ɲg̃, ɲg̃ʷ], que son las realizaciones fonéticas de las oclusivas plenamente orales con los mismos puntos de articulación, es decir [t, tʃ, k, kʲ] después de una vocal nasal. En otras palabras, esta segunda serie de consonantes nasales-orales resulta de un proceso de **post-nasalización**, porque las consonantes orales ‘ganan’ una fase nasal inicial por ser precedidas de una vocal nasal.

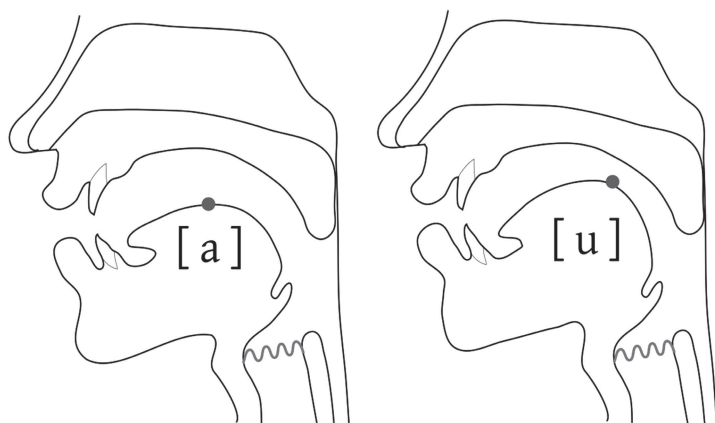
También existen consonantes orales-nasales, ej. [bm̃, dñ], pero estas son mucho más infrecuentes que las consonantes nasales-orales. Por ejemplo, el karitiana (Tupí, Rondônia, Brasil; Storto, 1999) tiene una serie de consonantes nasales [m, n, ŋ] que se *pre* y post-oralizan antes y después de una vocal oral, resultando en tres series de consonantes complejas con una fase nasal y una fase oral: las pre-oralizadas [bm̃, dñ, gŋ̃], las post-oralizadas [mb̃, nd̃, ŋg̃] y la circun-oralizadas [bmb, dnd, gŋg].

Es importante diferenciar una consonante **post-oralizada** de una consonante **pre-nasalizada**, porque ambas se pronuncian de la misma manera: con una fase nasal seguida de una fase oral. La diferencia entre estos dos tipos de consonantes es que la consonante post-oralizada es una consonante nasal que adquiere una fase oral final por un sonido oral en su ambiente, mientras que la consonante pre-nasalizada es una consonante oral que adquiere una fase nasal inicial por un sonido en su ambiente. Esta diferencia, entre **post-oralización** y **pre-nasalización**, es una diferencia **fonológica** (ver Capítulo 2).

3.4.1.3. Diptongos

El concepto de un sonido complejo no se restringe a las consonantes: las vocales también pueden tener dos fases fonéticas distintas. Muchas veces, nombramos a estas vocales complejas con el término de **diptongos**. Por ejemplo, el idioma ese ejja, además de las cuatro vocales simples (o **monoptongos**) [i, e, a, o], también tiene tres diptongos [io, ia, oe] en su inventario de sonidos. La primera fase del diptongo [io] se describe como una vocal anterior alta no redondeada, y su segunda fase se describe como una vocal posterior media redondeada. La primera fase del diptongo [ia] se describe como una vocal anterior alta no redondeada, y su segunda fase se describe como una vocal central baja no redondeada. Finalmente, la primera fase del diptongo [oe] se describe como una vocal posterior media redondeada, y su segunda fase se describe como una vocal anterior media no redondeada.

Las dos fases de un diptongo pueden tener la calidad de cualquier vocal del AFI, pero una de esas dos fases tiende a ser una vocal alta, es decir [i] o [u]. A un diptongo que empieza con una fase media o baja y termina con una fase alta, llamamos **diptongo creciente** (**Figura 27**), y a un diptongo que empieza con una fase alta y termina con una fase media o baja, **diptongo decreciente o descendente** (**Figura 28**). Dicho eso, lo importante para diferenciar un diptongo de una secuencia de dos vocales simples es que la vocal compleja de un diptongo tiene que formar parte de la misma sílaba, y no de dos sílabas distintas.

Figura 25*Fases de articulación de un diptongo creciente [au]***Figura 26***Fases de articulación del diptongo decreciente [ja]*

Muchos idiomas de la familia Tupí-Guaraní tienden a tener diptongos fonéticos, aunque muchas veces los lingüistas que analizan estos idiomas no hablan de esas secuencias de vocales como diptongos. Por ejemplo, el idioma kawaiwete (Mato Grosso, Brasil) tiene en su inventario de sonidos seis monoptongos orales [i, ε, ɨ, a, u, ɔ] y cinco monoptongos nasales [ĩ, ẽ, ỹ, ã, õ]. Además, tiene una serie de once vocales bifásicas [ii, ui, ei, ai, ɔi, eu, au, ãĩ, õĩ, ãĩ, õũ] (Lapierre, 2021). Aunque la mayoría de los autores que describen este idioma no llaman diptongos a estas vocales complejas, desde un punto de vista fonético, son claramente diptongos porque tienen dos fases distintas que forman parte de una sola sílaba.

3.4.2. Articulación secundaria

Existen tipos de sonidos que se componen de la realización de dos puntos de articulación simultáneamente y, aun así, son considerados como una sola unidad fonética. Esto se debe al hecho de que el movimiento de ciertos articuladores del tracto vocálico son independientes uno del otro. Por ejemplo, es posible mover los labios y el dorso de la lengua al mismo tiempo, lo que resulta en un sonido labio-velar. Como en el ejemplo dado, la mayoría de los sonidos con una articulación secundaria son consonantes. Existen distintos tipos de articulaciones secundarias que se desarrollan a continuación.

3.4.2.1. Labialización

Si redondeamos los labios al producir cualquier consonante, obtendremos un sonido **labializado** que transcribimos con el diacrítico [◌^w] siguiendo el símbolo básico. Por ejemplo, si pronunciamos una consonante oclusiva velar sorda [k] y además agregamos los labios como un segundo punto de articulación, lograremos el sonido [k^w], una consonante oclusiva labio-velar sorda. Las lenguas de la familia lingüística Takana son conocidas por tener sonidos labio-velares, como en el maropa en palabras del tipo [k^waβe] ‘yuca, mandioca’. Aunque las consonantes labio-velares tal como [k^w] y [w] son las consonantes labializadas más frecuentemente encontradas en los idiomas del mundo, es posible que otras consonantes sean labializadas también. La lengua mojeño trinitario posee tres sonidos de este tipo, es decir [p^w m^w h^w] (Rose, 2019).

3.4.2.2. Palatalización

Las consonantes con una articulación secundaria palatal se producen con el dorso de la lengua en una posición anterior alta. Estas consonantes, también conocidas como consonantes **palatalizadas**, se transcriben con el diacrítico [◌^j], pues la aproximante [j] se coarticula con la consonante en cuestión. El idioma chiquitano ignaciano de Bolivia posee los dos sonidos palatalizados [p^j k^j], como en las palabras [op^jokoʃ] ‘pescado’ y [k^jaraʃ] ‘zorro’ (Ciucci y Macoñó, 2018).

3.4.2.3. Velarización

Las consonantes **velarizadas** se producen con el dorso de la lengua en una posición posterior alta, además de la articulación básica de la consonante en cuestión. La transcripción de estos sonidos se indica con la consonante principal seguida de un diacrítico que indica la velarización [◌^v] o [◌̠]. De ese modo, la consonante lateral alveolar sonora del inglés pronunciada al final de una sílaba, como en la palabra *pool* ‘piscina’, puede transcribirse de dos maneras distintas: [pul^v] o [pul̠].

3.4.2.4. Laringealización

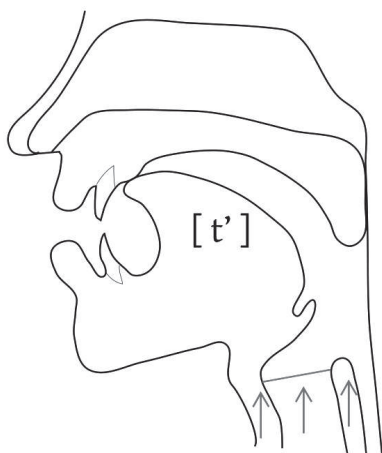
Finalmente, existen consonantes cuya articulación secundaria involucra a la laringe. A estas consonantes las llamamos **laringealizadas**. Aunque la mayoría del tiempo las consonantes laringealizadas son oclusivas, como [b, p, t, d, k, g], cualquier consonante puede ser modificada por el estado de la laringe. Existen dos tipos de consonantes laringealizadas: las consonantes eyectivas y las consonantes implosivas, que veremos a continuación.

3.4.2.4.1. Eyectivas

Si al momento de la obstrucción total de una consonante oclusiva la glotis está cerrada y la laringe elevada (**Figura 25**), esta coarticulación hace que el aire dentro de la boca se comprima mucho, lo cual produce una explosión más fuerte que la de una oclusiva común, característica de una **eyectiva**. Empleamos un apóstrofo siguiendo al símbolo de la consonante principal para transcribir un sonido eyectivo [ɔ']. Los idiomas quechua y aymara de Bolivia se destacan por tener una serie compleja de consonantes eyectivas, es decir [p', t', tʃ', k', q']. Encontramos estas consonantes eyectivas en la palabra quechua [sik'imira] y la palabra aymara [k'isimira], que significan 'hormiga'. La palabra quechua conocida para 'seco, resaca' [tʃ'aki] también contiene una consonante africada eyectiva. Es interesante notar que la mayoría de las consonantes eyectivas encontradas en el mundo son oclusivas sordas.

Figura 27

Oclusiva eyectiva alveolar [t']



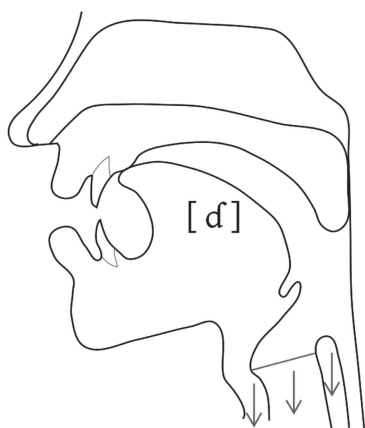
3.4.2.4.2 Implosivas

Si la obstrucción de una consonante está acompañada de una glotis cerrada y de un descenso de la laringe (**Figura 26**), esto producirá que el aire de la boca

no pueda salir al deshacerse la obstrucción, sino que ingresará de manera abrupta. Precisamente porque el aire entra por la boca en lugar de salir como en cualquier otro sonido, estos sonidos se llaman **implosivos**. Las consonantes implosivas se transcriben de la siguiente manera: [ɓ] para una implosiva labial, [ɗ] para una implosiva alveolar y [ɠ] para una implosiva velar. Existen también otros puntos de articulación para las implosivas como palatal y uvular [ʄ ɠʷ].

Figura 28

Implosiva alveolar sonora [ɗ]



El idioma ese ejja hablado en Bolivia se destaca por poseer este tipo de consonante, específicamente implosivas bilabiales y dentales [ɓ, ɗ]. Ejemplos de palabras con implosivas en ese ejja se dan en (6).

- | | | | | |
|-----|----|---------------------|-----------------|-------------------------------|
| (6) | a. | <i>ibabi</i> | [i'ɓabi] | ‘mantarraya’ |
| | b. | <i>dijjidiiji</i> | [ɗiçi'ɗiçi] | ‘mono nocturno’ |
| | c. | <i>dojobesakani</i> | [ɗoho'ɓesakani] | ‘llevar al otro lado del río’ |

Como en el caso del ese ejja y al contrario de las consonantes eyectivas, la mayoría de las consonantes implosivas encontradas en los idiomas del mundo son oclusivas sonoras.

3.5. Conclusión

En este capítulo, hemos hablado de fonética, es decir del estudio de la producción y percepción de los sonidos del habla humana. Empezamos hablando del Alfabeto Fonético Internacional (AFI), un sistema de transcripción que puede representar cualquier sonido de cualquier lengua humana. Luego, nos referimos al

tracto vocal y a los órganos fonadores que usamos para producir los sonidos del habla.

Hemos visto que las consonantes son sonidos que se producen con alguna obstrucción en el tracto vocal, y se diferencian por tres parámetros articulatorios: (1) su punto de articulación, es decir el lugar de la obstrucción en el tracto vocálico; (2) su modo de articulación, es decir el grado de constricción; y (3) su sonoridad, es decir el estatus de vibración o no de las cuerdas vocales. Las vocales, por su parte, son sonidos que se producen sin ninguna obstrucción, dejando que el aire fluya libremente a través del tracto vocal.

A continuación, vimos que las vocales se diferencian por cuatro parámetros articulatorios: (1) su altura, es decir el nivel de altura del dorso de la lengua; (2) su anterioridad, es decir el nivel de anterioridad del dorso de la lengua; (3) su redondeamiento, es decir si su articulación involucra redondeamiento de los labios; y (4) su nasalidad, es decir si su articulación involucra descenso del velo.

Terminamos nuestro estudio de la fonética tratando los segmentos complejos y articulaciones secundarias. Vimos que existen segmentos complejos que se componen de dos o más sonidos o articulaciones como, por ejemplo, las consonantes africadas, las consonantes nasales-orales y los diptongos. Finalmente, vimos que existen consonantes con articulaciones secundarias que se componen de la realización de dos puntos de articulación simultáneamente.

Referencias

- Alvarsson, J.Å. & Claesson, K.. (2014). 'Weenhayek (Mataco). En Mily Crevels & Pieter Muysken (eds.), *Lenguas de Bolivia. Tomo III. Oriente*. La Paz: Plural Editores. 415-465.
- Carr, Philip. (2008). *A Glossary of Phonology*. Edinburgh: Edinburgh University Press Ltd.
- Cerrón-Palomino, R. (2006). *El Chipaya o Lengua de los Hombres del Agua*. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú.
- Chango Masaquiza, F. & Marlett, S. A. (2008). Salasaca Quichua. En *Journal of the International Phonetic Association* 38(2): 223-227.
- Ciucci, L. & Macoñó Tomichá, J. (2018). *Diccionario básico del chiquitano del Municipio de San Ignacio de Velasco*. Santa Cruz de la Sierra: Industria Maderera San Luis S. R. L., Museo de Historia. UAGRM (Editores).
- Claesson, K. (2016). *Estudios de la Gramática del Idioma 'Weenhayek 1*. En línea: <http://noctenes.org/onewebmedia/Gram%C3%A1tica%20'weenhayek%201%2003.2017.pdf>, (acceso el [7 de abril de 2020]).
- Daviet, W. (2016). *Observations sociolinguistiques et analyse de la phonologie du dialecte ava du guaraní bolivien*. [Tesis de maestría 2], Université Lumière Lyon 2. Francia.
- Emkow, C. (2019). *A Grammar of Araona*. (Outstanding grammars from Australia, 19). Munich: LINCOM Academic Publishers.
- Everett, D. L. & Kern, B. (1997). *Wari: the Pacaas Novos Language of Western Brazil*. Routledge.
- Gasparini, N. (2012). *Observations sociolinguistiques et esquisse de la phonologie du siriono: Langue tupi-guarani de Bolivie*. [Tesis de maestría], Université Lumière Lyon 2. Francia.
- Gijn, R. van. (2006). *A Grammar of Yurakaré*. [Tesis doctoral], Radboud Universiteit Nijmegen.
- Ladefoged, P. & Johnson, K. (2011). *A Course in Phonetics*. Sexta Edición. Boston, MA: Wadsworth, Cengage Learning.
- Roach, P. (2009). *English Phonetics and Phonology. A practical course*. Cuarta Edición. Cambridge: Cambridge University Press.

- Rose, F. (2019). Rhythmic syncope and opacity in Mojeño Trinitario. En *Phonological Data and Analysis*, Vol. 1:2, 1-25.
- Storto, L. (1999). *Aspects of Karitiana grammar*. [Tesis doctoral], Massachusetts Institute of Technology (MIT).
- Tallman, A. J.R. (2019). Stress, syllabification and empty morphs in Araona (Takana). *Atelier Typologie phonologique*, Laboratoire Dynamique du Langage, CNRS.
- Vuillermet, M. (2012). *A Grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. [Tesis doctoral], Université Lumière Lyon 2.

CAPÍTULO 4

MORFOLOGÍA

Cómo citar: Nikulin, A. (2023). Morfología. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 100-129), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11111038>

Andrey Nikulin*

0. Introducción

La morfología es la rama de la lingüística que estudia la estructura interna de las palabras, es decir, de qué elementos están compuestas, cómo esos elementos se combinan y qué significado generan. Si tomamos como ejemplo la palabra *subalcaldesas* del castellano, es evidente que esta palabra contiene al menos cuatro elementos, y cada elemento contribuye en algo al significado de la palabra. El elemento **sub-** significa algo como ‘inferior’: al decir *subalcaldesas* y no *alcaldesas*, nos estamos refiriendo a autoridades de un rango inferior (al menos en Bolivia las subalcaldesas están a cargo de los distritos municipales, mientras que las alcaldesas son responsables de los municipios). El elemento **alcalde** denota la máxima autoridad ejecutiva; por tratarse de una raíz morfológica (ver 4.2.1), no podemos remover este elemento de la palabra. El elemento **-esa** señala que las personas a las que nos estamos refiriendo son mujeres (compárese la palabra *subalcaldesas* con *subalcaldes*). Finalmente, el elemento **-s** señala que se trata de más de una entidad (o, en este caso, más de una persona): si dijéramos *subalcaldesa*, estaríamos refiriéndonos a una persona, mientras que con el elemento **-s** dejamos claro que las subalcaldesas son varias.

A diferencia de la fonología, la morfología se interesa únicamente por los elementos constitutivos de las palabras que poseen un significado o al menos una función. Si bien podemos decir que la misma palabra *subalcaldesas* (pronunciada [ˌsuβalkalˈdesas] o [ˌsuβʔalkalˈdesas]) está compuesta de cinco sílabas ([su], [βal], [kal], [de] y [sas]) o incluso de trece sonidos/fonemas, esas divisiones no tienen

* Con sede en Goiânia, Brasil, es profesor en el Núcleo Takinahak de Formação Superior Indígena (Universidade Federal de Goiás). Obtuvo el doctorado en Lingüística de la Universidad de Brasília (2020). Sus intereses se centran en la fonología histórica, así como en la fonología areal. También está muy interesado en la morfología y morfosintaxis de las lenguas chiquitano, maxakali y jê del norte. Su trabajo de investigación actual se centra en la familia Macro-Jê y en la descripción de todos los aspectos de la variedad migueleño del chiquitano. También ha contribuido a una serie de proyectos similares con bases de datos, que abarcan los inventarios fonológicos de Eurasia y la Base de Datos Lexicoestadística Global.

nada que ver con el significado y por ello son, hasta cierto punto, irrelevantes para la morfología. (Dijimos “hasta cierto punto” porque la morfología y la fonología pueden interactuar de formas bastante complejas, como veremos en la sección 4.3.)

Este capítulo tiene como objetivo introducir al lector o a la lectora a la morfología, que se ilustra con datos de lenguas habladas en distintas partes de Bolivia. La sección uno presenta los conceptos fundamentales de la morfología, tales como el morfema, la palabra y la exponencia. Los tipos de morfemas se discuten en la sección dos y los tipos de variantes de morfemas, en la tres. En la sección cuatro, se presentan los conceptos de flexión y derivación. Finalmente, en cinco se discuten las diferencias entre los grados de complejidad morfológica promedio de las lenguas del mundo y de Bolivia. La sección sexta cierra el capítulo.

4.1. Morfología: definición y conceptos básicos

Haspelmath y Sims (2010, p. 2) definen la morfología como “el estudio de covariación sistemática entre la forma y el significado de las palabras”. Es decir, la morfología busca comprender las regularidades que se observan cuando uno compara palabras morfológicamente relacionadas. Observe, por ejemplo, los siguientes pares de palabras de paunaka (una lengua Arawak hablada en el departamento de Santa Cruz, provincia de Ñuflo de Chávez) en (1).

- (1) paunaka (< achane < arahuacas; Terhart, 2022, p. 184)
- | | | | |
|-------------------|------------|------------------|---------------|
| a. <i>kasune</i> | ‘pantalón’ | <i>nikasune</i> | ‘mi pantalón’ |
| b. <i>kuepia</i> | ‘riñón’ | <i>nikuepia</i> | ‘mi riñón’ |
| c. <i>nȳkȳiki</i> | ‘olla’ | <i>ninȳkȳiki</i> | ‘mi olla’ |
| d. <i>pusane</i> | ‘bolsa’ | <i>nipusane</i> | ‘mi bolsa’ |
| e. <i>yumaji</i> | ‘hamaca’ | <i>niyumaji</i> | ‘mi hamaca’ |

Un patrón se repite de forma sistemática en estos datos: los objetos poseídos por la persona que está hablando (o, como se dice en la lingüística, por un poseedor de la primera persona del singular) son denotados por palabras que comienzan con la sílaba *ni-*. El elemento *ni-* está ausente en las palabras que designan los mismos objetos no poseídos. De esta forma, podemos decir que en paunaka las palabras que denotan objetos **covarían** del siguiente modo: cuando el objeto no es de nadie, la palabra ocurre en su forma “básica” (sin ningún elemento adicional), mientras que si la persona que está hablando desea referirse a un objeto que le pertenece a la palabra se le agrega el elemento *ni-* a la izquierda. Este patrón constituye una relación entre la forma y el significado.

En muchos casos, el análisis morfológico de una palabra consiste en la identificación de los elementos que la componen, tan pequeños como sea posible,

donde cada elemento normalmente posee una forma y una función. Esos elementos son conocidos bajo el nombre de **morfemas**. Por ejemplo, las palabras del paunaka como *nikasune* ‘mi pantalón’ y *nikuepia* ‘mi riñón’ (1a–b) contienen dos morfemas cada una, algo que comúnmente se señala mediante guiones en la literatura especializada: *ni-kasune*, *ni-kuepia*. Al proceso de división de una palabra en morfemas se le suele decir **segmentación**, o sea, podemos segmentar las formas *nikasune* y *nikuepia* como *ni-kasune*, *ni-kuepia*. Las palabras que pueden segmentarse en dos o más morfemas son conocidas como **morfológicamente complejas**. En cambio, las palabras *kasune* ‘pantalón’ y *kuepia* ‘riñón’ (1a–b), por su parte, al poseer solo un morfema, no son segmentables ni morfológicamente complejas; ellas se denominan **monomorfémicas**.

Sería equivocado decir que la morfología se ocupa únicamente de la segmentación de palabras en morfemas: las relaciones morfológicas pueden darse incluso entre palabras que no comparten ningún morfema (como en los casos de suplección; ver 4.3.) o entre palabras que no cambian de forma, aunque se modifique su significado (como en los casos de conversión; ver 4.2.4).

4.1.1. ¿Qué es una palabra?

Hasta ahora, al hablar de la morfología y de su objeto de estudio, hemos hecho un amplio uso del término “palabra”. ¿Pero qué es exactamente una palabra? Curiosamente, como se explicará a continuación, el concepto de palabra es uno de los más problemáticos en la lingüística; de hecho, carece de una definición universalmente aceptada.

Antes de que podamos pasar a la discusión de los problemas más serios asociados con el término “palabra”, será necesaria una aclaración respecto a los distintos usos de este término en el lenguaje informal. (Supongamos por ahora, para fines de simplicidad, que sabemos determinar dónde comienzan y dónde terminan las palabras en cada lengua, aunque luego veremos que la situación es bastante más compleja). Tomemos como un ejemplo las palabras *amarillo*, *amarilla*, *amarillos* y *amarillas* del castellano. Desde un punto de vista, son cuatro palabras distintas: se escriben y se pronuncian de maneras distintas, además de presentar diferentes características gramaticales. Pero si consultamos un diccionario, no encontraremos cuatro entradas separadas para *amarillo*, *amarilla*, *amarillos* y *amarillas* –se espera que cada hablante del castellano sepa que todas estas palabras son en realidad variantes de una misma palabra, representada en el diccionario por la entrada AMARILLO. Las palabras en el primer sentido son conocidas como **formas** y en el segundo sentido como **lexemas** (o ítems léxicos). De esta manera, se dice que *amarillo*, *amarilla*, *amarillos* y *amarillas* son cuatro formas distintas de un único lexema AMARILLO. Además de estos dos usos del término “palabra” en el lenguaje informal, existe un

tercer uso del mismo término, cuando queremos referirnos a cada secuencia de letras (o fonemas) concreta en un texto o un discurso; las palabras en este sentido son conocidas como **instancias**. Por ejemplo, la frase ¡Ya yo no porque ya yo ya! (una expresión interjetiva típica del habla tradicional de Santa Cruz) contiene cuatro lexemas y cuatro formas distintas (*ya, yo, no, porque*), pero siete instancias.

La discusión en el párrafo anterior, como ya dijimos, partía del supuesto de que uno sabe determinar los límites de una palabra (así sea de una forma, de un lexema o de una instancia). Efectivamente, en la escuela aprendemos a poner espacios entre palabras y no solemos cuestionar mucho el hecho de que las palabras son, de algún modo, unidades necesarias. Pero desde un punto de vista científico, nuestras convenciones ortográficas no quieren decir gran cosa, ya que son meros acuerdos determinados por instituciones, que no están siempre basados en los hechos de la estructura lingüística e involucran muchas reglas arbitrarias¹. Además, muchas lenguas simplemente no cuentan con una tradición de escritura consolidada (y entre las que la tienen hay algunas que no utilizan ningún tipo de espacios, como, por ejemplo, el tailandés). Por ello, si queremos definir el concepto de “palabra” de un modo que tenga sentido desde un punto de vista lingüístico, necesitamos acudir a algún otro criterio o criterios.

Se han propuesto varios criterios que podrían permitir dividir el flujo de sonidos del habla en palabras, de los que mencionaremos solo algunos. Un criterio es el de **separabilidad**: una palabra prototípica no puede ser interrumpida por un elemento (o por una pausa), mientras que entre palabras distintas normalmente se puede agregar algún elemento. Por ejemplo, no hay cómo poner nada en el medio del sustantivo *salchipapas* del castellano —es absolutamente imposible decir algo como **salchi-con-queso-papas* (el asterisco denota que una determinada expresión está mal formada)— y por eso podemos concluir que *salchipapas* es una palabra según el criterio de separabilidad. En cambio, una expresión como *sopa de maní* sí puede interrumpirse por otros elementos (*sopa espesa de maní, sopa de puro maní*), y diríamos que según el criterio de separabilidad son tres palabras. Otro criterio es la **autonomía**: una palabra prototípica puede ocurrir como una respuesta completa a una pregunta, a diferencia de unidades menores que una palabra (es decir, a una pregunta como *¿Comiste sopa de qué?* uno puede contestar simplemente *maní*, pero sería mucho más raro preguntar algo como *¿Comiste salchi-qué?* y recibir *papas* como respuesta, a menos que sea en un contexto humorístico).

1 Todos los hablantes del castellano en la vida nos hemos preguntado alguna vez si debemos escribir una determinada expresión pegada o separada (por ejemplo, si es “aparte” o “a parte”, “enseguida” o “en seguida”). Dudas semejantes emergen entre hablantes de otros idiomas que utilizan una escritura con espacios. Esto sugiere que la división del discurso en palabras no es necesariamente parte de las intuiciones de los hablantes nativos.

Algunos otros criterios que se han propuesto para las lenguas del mundo tienen que ver con la fonética y la fonología. En las lenguas que poseen acento, muchas veces se dice que la palabra es la unidad que porta un único **acento** (en este sentido *salchipapas* sería una palabra, mientras que *sopa de maní* serían dos –obsérvese que *de* no porta un acento y no constituye una palabra según este criterio)² –. También se ha propuesto que la palabra puede definirse como el **dominio de aplicación de reglas fonológicas**. Por ejemplo, en el dialecto del castellano hablado en Venezuela y en otros países del Caribe *s* y *n* se pronuncian, respectivamente, [h] y [ɲ], no solamente en posición final de sílaba, sino también en posición final de palabra, aunque en el flujo del habla continua esas consonantes aparezcan como ataques silábicos cuando la palabra siguiente empieza por vocal: *los otros* se pronuncia [lo'hotroh] y no *[lo'sotroh] (pero *nosotros* [no'sotroh]), *un ánimo* se pronuncia [u'ɲanimo] y no *[u'nanimo] (pero *unánime* [u'nanime]). Estas evidencias fonológicas sugieren que, según el criterio de transformación de consonantes finales, expresiones como *los otros* y *un ánimo* constituyen dos palabras, aunque según el criterio acentual estas mismas expresiones constituyen una palabra.

Algunos autores han sugerido que sería relevante distinguir entre dos tipos de palabras en las lenguas del mundo –las palabras **morfológicas** (identificadas según criterios como los de separabilidad y autonomía) y las palabras **fonológicas** (identificadas según los dominios de acento y aplicación de otras reglas)–. Sin embargo, ya vimos arriba que criterios fonológicos a veces dan resultados distintos al aplicarse a los mismos datos (o, como se suele decir en la lingüística, que estos criterios a veces **no convergen**). Los distintos criterios morfológicos tampoco convergen en muchas lenguas del mundo. Por ello, el concepto de “palabra” es bastante problemático, y no parece haber una manera objetiva de decidir cómo dividir un discurso en palabras, ni siquiera si uno distingue entre los conceptos de palabra morfológica y fonológica –todo dependerá de cuál criterio será declarado como más importante–.

Un estudio particularmente interesante que sugiere que los conceptos de “palabra”, “palabra morfológica” o “palabra fonológica” no son necesariamente relevantes en todas las lenguas del mundo es el trabajo de Tallman (2021), que analiza en gran detalle los hechos del chácobo, un idioma de la familia Pano–tacana, hablado en el Beni (provincias de Vaca Díez, Yacuma y José Ballivián). Tras definir y aplicar 24 diagnósticos (incluyendo 16 morfológicos y 8 fonológicos) que podrían, en principio, delimitar palabras en chácobo, Tallman concluye que poquísimos de esos criterios convergen, y muestra, usando métodos estadísticos, que la poca convergencia que sí se observa podría darse por casualidad.

2 El acento al que nos estamos refiriendo aquí no es la tilde ortográfica, sino algo que se oye – es lo que caracteriza las sílabas más “fuertes” (o más prominentes) que otras (ver el capítulo Fonética para más detalles).

4.1.2. Exponencia

Las lenguas del mundo pueden expresar la misma información de maneras bastante distintas. Por ejemplo, en castellano se puede emplear el adverbio *también* –que tiene todas las características de una palabra independiente– para expresar que una determinada entidad es de algún modo similar a otra previamente nombrada (como en *ustedes también*). Por el contrario, en quechua boliviano sureño este mismo significado no se expresa por medio de una palabra independiente, sino mediante el elemento ligado *-pis*, como, por ejemplo, en *qankunapis* ‘ustedes también’, donde *qankuna* significa ‘ustedes’ y el elemento *-pis* expresa lo mismo que *también* en castellano. Los patrones de expresión de información lingüística mediante palabras o morfemas se conocen bajo el nombre de **exponencia**.

Una primera distinción importante tiene que ver con la obligatoriedad de expresar un determinado tipo de información. En castellano, por ejemplo, es obligatorio expresar en cada sustantivo si este es singular o plural: si decimos *chompa*, está claro que nos referimos a una única *chompa*, mientras que si decimos *chompas* el interlocutor entiende de inmediato que las *chompas* son varias. Pero en algunas otras lenguas la expresión del número en los sustantivos es opcional. Por ejemplo, *pira* significa tanto ‘pez’ como ‘peces’ en el guaraní boliviano. Si uno desea subrayar que los peces son varios, uno diría *pira reta* (literalmente ‘la muchedumbre del pez’), pero a diferencia del castellano uno puede elegir si expresa o no la pluralidad. De este modo, se puede decir tanto *ouma pira* ‘ya vinieron los peces’ como *ouma pira reta* ‘ya vinieron los peces’ (Gustafson, 2014, p. 335).

La exponencia puede ser **biunívoca** (un morfema por una unidad de información), **redundante** (la misma información se expresa en dos o más morfemas) o **cumulativa** (un morfema expresa más de un tipo de información).

Aunque desde un punto de vista lógico la exponencia biunívoca sería la más apropiada para expresar la información, las lenguas del mundo difícilmente la emplean en la totalidad de sus sistemas gramaticales. Sin embargo, algunas lenguas andinas llegan cerca de ese ideal. Por ejemplo, en la frase *khuchijanakatakiw* ‘es para mis chanchos’ del aymara (variedad de Muylaq’, hablada en el Perú), uno puede identificar cinco morfemas, cada uno con su función: *khuchi* ‘chanchos’, *-ja* ‘poseído por la persona que está hablando, pero no por el interlocutor (= primera persona)’, *-naka* ‘múltiples entidades (= plural)’, *-taki* ‘para (= benefactivo)’, *-w* ‘declarativo’ (Coler, 2014, p. 2211).

Un ejemplo de exponencia redundante es la llamada concordancia (de persona, número y género) del castellano: compárense las expresiones *la bolsa pesada*, *el bulto pesado*, *las bolsas pesadas* y *los bultos pesados*, donde es fácil ver que es obligatorio expresar el género y el número del sustantivo en el artículo definido y en

el adjetivo *pesado*. En baure (un idioma de la familia arawak hablada en el Beni), los sustantivos se dividen en clases, y algunos otros elementos en la frase —como los numerales— reciben un morfema (llamado **clasificador**) que muestra a cuál clase pertenece el sustantivo. Por ejemplo, uno dice *p-oe-sh mokovore* ‘una papaya’, *p-e-sh rekirok* ‘una tutuma’, *p-e-sh mokovis* ‘un zapallo’, *po-se-sh senti* ‘una sandía’. En estos ejemplos se ve que en el medio del numeral *po-...-sh* ‘uno’ se pone un morfema que señala, de manera redundante, si el objeto cuantificado es una fruta dulce (morfema *-i-*; la secuencia *po-i-* se convierte en *poe-*) o no dulce (morfema *-e-*) o si posee una forma oval (morfema *-se-*), entre otras posibilidades (Danielsen, 2007, p.142).

Un ejemplo de exponencia cumulativa es el morfema *-ste* del castellano (como en *almorzaste, dijiste, fuiste*), que expresa al menos tres cosas al mismo tiempo: que la acción denotada por el verbo ocurrió una vez en el pasado (tiempo pretérito), que el sujeto del verbo es el interlocutor (segunda persona) y que el interlocutor actuó solo (singular).³

4.1.3. Plantillas de morfemas

En algunos casos puede ser útil presentar la información acerca de la estructura morfológica de una determinada clase de palabras en una determinada lengua usando **plantillas de morfemas**. Se trata de esquemas que muestran cuáles morfemas pueden (o deben) ocurrir en cuál posición dentro de una palabra. Por ejemplo, la estructura de los sustantivos no poseídos en chiquitano migueleño (familia macro-ye, Santa Cruz, provincia de Velasco) puede representarse de la siguiente manera.

-2	-1	0	+1	+2
(consonante de ligación) <i>r- / n-</i>	(marcador de género*) <i>o- / u-</i> = animado no humano <i>y- / ñ- / ɿ-</i> = masculino <i>Ø-</i> = femenino o no animado	raíz	(diminutivo) <i>-ma'(a) / -ña'(a)</i>	(plural) <i>-kaa / -kyaa</i>
			sufijo <i>-j(i) / -s(i) / -xh(i)</i>	

* = solo en el habla de los hombres

Algunos ejemplos de sustantivos del chiquitano migueleño (todos representativos del habla masculina) se dan en (2).

(2) chiquitano migueleño (< chiquitano < macro-ye)

3 En realidad, el uso del morfema *-ste* contiene mucha más información: señala que el verbo está en el modo indicativo y en la voz activa. En el ámbito de este capítulo no podemos discutir en detalle cómo los términos “modo indicativo” y “voz activa” se relacionan con la manera que expresamos la información.

		-2	-1	0	+1	+2	
a.	(n)okunumasima'ákaa	(n-)	o-	kunumasi	-ma'á	-kaa	'pollitos'
b.	(n)okunumasima'	(n-)	o-	kunumasi	-ma'		'pollito'
c.	(r)okurubasíkaa	(r-)	o-	kurubasí		-kaa	'pollos'
d.	(r)okurubasij	(r-)	o-	kurubasi		-j	'pollo'
e.	ñoñi'íkaa		ñ-	oñi'í		-kaa	'hombres'
f.	ñoñi'ij		ñ-	oñi'í		-j	'hombre'
g.	pa'íkaa			pa'í		-kaa	'mujeres'
h.	pa'ij			pa'í		-j	'mujer'

Las cinco columnas en la plantilla representan las cuatro posiciones morfológicas, numeradas de -2 a +2, relevantes para la descripción de la estructura de los sustantivos no poseídos del chiquitano. Las posiciones señaladas con paréntesis no son obligatorias: por ejemplo, la consonante de ligación (*r-* o *n-*, posición -2), que solo ocurre ante vocales, puede estar presente o ausente en una palabra. La posición -1 (marcador de género) está disponible solo en el habla de los hombres (en el habla de las mujeres chiquitanas el género de los sustantivos no se expresa en la gramática): los sustantivos que denotan entidades animadas no humanas —que para los chiquitanos incluyen varios tipos de árboles, las estrellas, los cometas y la miel— reciben el prefijo *o-* o su variante *u-*; los sustantivos que denotan a hombres reciben el prefijo *y-* o su variante *ñ-* ante vocales (si la raíz empieza con una consonante, esta se palataliza, pero la *y-* no aparece), mientras que los sustantivos que denotan a mujeres o entidades inanimadas no reciben ningún prefijo. La raíz (posición 0) es obligatoria. Las posiciones +1 (diminutivo) y +2 (plural) son ambas opcionales, y cuando ambas están rellenas, el sufijo plural siempre sigue al diminutivo, como en ña'íma'ákaa 'niños' (no se puede decir *ña'í-kaa-ma', con el orden invertido). Sin embargo, si la palabra no contiene ninguno de los sufijos diminutivo o plural, es obligatorio utilizar un sufijo *-j(i)* (o sus variantes *-s(i)* o *-xh(i)*), que no puede coocurrir ni con *-ma'(a)/-ña'(a)* (+1) ni con *kaa/kyaa* (+2) y por ello se analiza como perteneciente a AMBAS posiciones. No es fácil definir la función del sufijo *-j(i)/-s(i)/-xh(i)* con pocas palabras; uno podría decir que es un sufijo del singular no diminutivo de sustantivos no poseídos por un poseedor referencial.

De igual manera, se puede representar los afijos verbales del quechua boliviano sureño mediante una plantilla (los sufijos no productivos de base verbal no están representados).

0	+1	+2	zona +3 (en cualquier orden)		+4	+5
raíz o base	(inceptivo) <i>-ri</i>	(recíproco) <i>-na</i>	(adverbial) <i>-rqu</i> 'persistentemente' <i>-rpa</i> 'de repente'	(asistivo) <i>-ysi</i>	(recíproco) <i>-na</i>	(perfectivo) <i>-yu / -yku</i> 'perfectivo' <i>-yacha</i> 'mov. asociado'

+6	+7	+8	+9	+10	+11	+12	+13
(afectivo) <i>-ri</i>	(valencia) <i>-chi</i> ‘causativo’ <i>-chiku</i> ‘pasivo’	(valencia) <i>-ku</i> ‘reflexivo’ <i>-mu</i> ‘mov. asociado’ <i>-pu</i> ‘aplicativo’ y combinaciones lexicalizadas	(futuro próximo) <i>-naya</i>	(progresivo) <i>-sha</i>	(solo) <i>-lla</i>	(1 ^p) <i>-wa</i>	(cuando) <i>-qti</i>

+14	+15	+16	+17	+18	+19	+20
(tiempo / 2 ^p) <i>-sqa</i> ‘narrativo’ <i>-rqa</i> ‘reportativo’ <i>-na</i> ‘potencial’ <i>-su</i> ‘2’ <i>-sa</i> ‘1 (futuro)’	persona del sujeto <i>-ni</i> ‘1(>3), no futuro’ <i>-nki</i> ‘2(>...)’ <i>-yki</i> ‘1>2SG’ <i>-n</i> (por defecto)	(futuro) <i>-q(a)</i>	(persona) <i>-n</i>	(plural) <i>-chik</i> (1+2, 2PL) <i>-yku</i> (1+3) <i>-ku</i> (3PL)	(solo) <i>-lla</i>	(enclíticos)

El ejemplo (3) muestra que los morfemas tienen que ocurrir en un orden fijo, siguiendo lo expuesto en la plantilla de arriba. Los elementos obligatorios son la raíz (o base, posición 0) y el morfema que expresa la persona del sujeto (posición +15), destacados en negrilla.

(3) quechua boliviano sureño (< quechua II < quechua; Peralta Zurita, 2006, p. 215)

0	+3	+7	+8	+10	+15	+18
<i>Wañu</i>	<i>-rpa</i>	<i>-chi</i>	<i>-mu</i>	<i>-sqa</i>	<i>-n</i>	<i>-ku.</i>
morir	-de_repente	-CAUS	-MOV	-PAS.NAR	-PERS	-3PL

‘Los habían hecho matar (a los perritos) por allá.’

Para más informaciones sobre el orden de los morfemas en el quechua boliviano sureño, ver Camacho Ríos y Tallman (por aparecer).

4.1.4. Glosa

Muchas veces en la literatura lingüística—incluso en este libro—se usan ejemplos de lenguas que los lectores probablemente no hablen. Las **glosas interlineales** se han inventado para que las personas puedan entender con más facilidad cómo es la estructura morfológica de los ejemplos lingüísticos de lenguas desconocidas. La mayoría de las publicaciones contemporáneas se basan en las **reglas de glosado de**

Leipzig,⁴ según las cuales los ejemplos lingüísticos se citan en tres líneas, como en (4) y (5).

- (4) chiquitano migueleño (< chiquitano < macro-ye; Nikulin, por aparecer)

<i>Chi- 'e-zosiú-ka=pi</i>	<i>Ø-xh-amẽ'ẽ.</i>
NEG-1SG-tener_fuerza-F.N3=NEG	NF-1SG _♀ -caminar
‘No puedo caminar.’	

- (5) puquina (lengua aislada; Emlen et al., 2023)

<i>Co-na</i>	<i>afcha-fo</i>	<i>fi-sca-n-qui-nch.</i>
DEM-LOC	estar-PTCP	saber-RFL-PL-1-DECL
‘Sabemos que está aquí.’		

La primera línea contiene el dato relevante en la lengua discutida (muchas veces en letras inclinadas), con la respectiva segmentación morfológica –los morfemas que constituyen una palabra se separan mediante guiones, los clíticos (3.2.2) mediante el símbolo =, las palabras mediante tabulaciones.⁵ Los morfemas cero (3.2.4) pueden señalarse mediante el símbolo Ø–.

La segunda línea –la **glosa**– representa el significado de cada morfema segmentado en la primera. Para los morfemas léxicos se usa simplemente una traducción, aunque aproximada, al **metalenguaje**, es decir, la lengua en que se redacta el texto (en el caso de este libro es el castellano). Para los morfemas gramaticales se utilizan abreviaciones especiales, que se ponen en versalitas (en los ejemplos (4) y (5) son los siguientes: 1 = primera persona, 3 = tercera persona, N3 = cualquier persona que no sea la tercera, F = finito, DECL = declarativo, DEM = demostrativo, NEG = negación, NF = no finito, LOC = locativo, PL = plural, PTCP = participio, RFL = reflexivo, SG = singular, ♀ = habla de las mujeres). Se utilizan las tabulaciones para garantizar que cada palabra en la primera línea aparezca alineada verticalmente con su glosa. Los guiones y el símbolo = se utilizan exactamente como en la primera línea (una glosa correctamente hecha tiene que contener la misma cantidad de guiones y de = en la primera línea y en la segunda). Si un morfema en la primera línea expresa simultáneamente dos o más significados (ver 4.1.2 sobre la exponencia cumulativa), se utiliza un punto en la segunda línea para separar esos significados (por ejemplo, *mayún* y *-ni* en el ejemplo (5) se glosan como “río.ACC” y “1SG.NF”, respectivamente). Por razones estéticas, no se suele emplear el punto para separar la persona del número (es decir, se utilizan las glosas 1SG, 3PL y no 1.SG, 3.PL).

4 Véase <<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>> para la versión completa de las reglas de glosado de Leipzig (en inglés).

5 En algunos casos especiales, las reglas de glosado de Leipzig admiten el uso de otros símbolos, como los corchetes angulares para los infijos (2.3) o ~ para señalar la reduplicación.

Si un morfema en la primera línea se traduce al metalenguaje por una expresión de múltiples palabras, se utiliza el guion bajo en vez del espacio (por ejemplo, *–zosiú–* en el ejemplo 4 se glosa como “tener_fuerza”).

Finalmente, en la tercera línea se da la traducción libre, normalmente entre comillas sencillas. Los ejemplos en los trabajos lingüísticos suelen numerarse y contener la información sobre la lengua a la que pertenecen y sobre su procedencia, a menos que esa información esté clara en el texto acompañante.

4.2. Tipos de morfemas

En esta sección se presentan los distintos tipos de morfemas, que pueden clasificarse según su significado o función (4.2.1.), según su grado de autonomía (4.2.2) o posición respecto a la raíz (4.2.3). En 4.2.4 discutimos la cuestión de los morfemas cero.

4.2.1. Tipos de morfemas según su significado/función

En una primera aproximación, se podría decir que los morfemas se subdividen en los que tienen un significado relativamente concreto, fácil de definir (a esa clase de significado se le suele decir **significado léxico**) y los que tienen un significado abstracto (o **gramatical**). Tomemos como ejemplo la palabra *descubrimientos* del castellano, que claramente contiene, como mínimo, cuatro morfemas: *des–* (como en *desplumar*), *–cubri–* (como en *cubrir*), *–miento–* (como en *yacimiento*) y *–s* (como en *hallazgos*). De estos morfemas, diríamos que *–cubri–* tiene un significado léxico (algo como ‘la acción de tapar, de poner algo encima o delante’), mientras que los demás son más difíciles de definir y por ello son considerados morfemas gramaticales: podríamos decir que *des–* significa ‘acción en que se cancela una acción X efectuada anteriormente’, *–miento* significa algo como ‘proceso o hecho que corresponde a una acción’, mientras que *–s* significaría ‘más de una entidad’. En muchas lenguas del mundo es común que una palabra tenga un único morfema con significado léxico, como *–cubri–* en *descubrimientos*, y entonces a ese morfema se le dice **raíz**. Los demás morfemas son conocidos como **afijos**. De esta manera, diríamos que la palabra *descubrimientos* consiste de una raíz (*–cubri–*) y tres afijos (*des–*, *–miento*, *–s*).

A veces ocurren situaciones en que no es del todo fácil identificar una única raíz en una palabra. En palabras como *hacia*, *nomás* o *al* del castellano, por ejemplo, no parece haber ningún morfema con un significado léxico. Algunos lingüistas dirían que en estos casos el morfema más prominente es la raíz (la raíz de *hacia* sería entonces *hacia*, la de *nomás* sería *–más*, mientras que en *al* no parece tener ningún sentido preguntarse si la raíz sería *a-* o *-l*); otros lingüistas dirían que estas palabras no tienen raíces.

En algunas lenguas, hay morfemas que presentan un comportamiento morfológico asociado con afijos, pero tienen un significado bastante concreto, haciendo difícil la tarea de determinar si se trata de afijos o raíces. Por ejemplo, en chácobo hay una serie de morfemas que pueden ocurrir antes de una raíz verbal léxica, pero cuyo significado también es léxico: estos morfemas denotan partes del cuerpo (Tallman, 2018, pp. 600–1). De esta manera, agregando esos morfemas a un verbo como *ashi* ‘lavar’, uno deriva verbos tales como *bahashi* ‘lavar el brazo o la axila’, *bēhashi* ‘lavar la cara’, *cahashi* ‘lavar la espalda’, *chihashi* ‘lavar las nalgas’, *huihashi* ‘lavar la pierna’, *jahashi* ‘lavar la boca’, *johoshi* ‘lavar el tobillo’, *mahashi* ‘lavar la cabeza’, *mēhashi* ‘lavar la mano’, *nohashi* ‘lavar el estómago’, *pahashi* ‘lavar la oreja’, *pēhashi* ‘lavar el ala’, *pihashi* ‘lavar la costilla’, *quēhashi* ‘lavar los labios’, *rahashi* ‘lavar la rodilla’, *rēhashi* ‘lavar la nariz’, *shihashi* ‘lavar el pecho’, *tahashi* ‘lavar la mejilla’, *tēhashi* ‘lavar el cuello’. Una vez que los morfemas en cuestión solo se usan con una raíz verbal, no pudiendo designar a partes del cuerpo como sustantivos, se los suele analizar como afijos y no como raíces, pero claramente son afijos algo atípicos desde un punto de vista semántico.

Sin embargo, en muchas lenguas es posible que un lexema contenga múltiples raíces que sí se usan por sí solas. Estos lexemas se llaman **compuestos**. Un ejemplo de un sustantivo compuesto es la palabra *popetakixhi* ‘abarca’ del bésiro (un idioma de la familia Macro-ye, hablado más que todo en la provincia de Ñuflo de Chávez en Santa Cruz), compuesta de las raíces *-pope* ‘pie’ y *-taki* ‘cuero, piel, corteza’ (*-xhi* es un afijo del singular no diminutivo de sustantivos no poseídos por un poseedor referencial). En el quechua boliviano sureño, según Condori et al. (2021), hay compuestos de los tipos “sustantivo + sustantivo” (*rumi-sunqu* ‘insensible’ = ‘piedra + corazón’), “sustantivo + verbo” (*runa-mikhu* ‘caníbal’ = ‘gente + comer’), “adjetivo + sustantivo” (*muqu-wasa* ‘jorobado’ = ‘emпинado + espalda’). En ese ejja, un idioma Pano–tacana hablado en el Beni, tanto los sustantivos como los verbos pueden ser compuestos (6).

- (6) ese ejja (< tacana < pano–tacana; Vuillermét, 2012, p. 396)

Jememe-so=ka kekwa-pojo-ka-ani mei=a.

motacú-semilla=CTRS horadar-partir-3^A-PRS piedra=INSTR

‘Las semillas de motacú las abrimos con una piedra.’

Tanto el sustantivo *jememeso* ‘semilla de motacú’ como el verbo *kekwapojo* son compuestos.

Arriba dijimos que los morfemas con un significado léxico son normalmente conocidos como raíces y los con un significado más gramatical o abstracto, como afijos. Pero en algunas lenguas se han identificado morfemas que no parecen tener

ningún significado, ni léxico ni gramatical. Su única función es contribuir a que la palabra esté bien formada. Estos morfemas son a veces llamados **morfoides**. Por ejemplo, en el chiquitano migueleño se han identificado los morfoides *-z-* (ocurre tras afijos de la primera persona del singular o de la primera persona inclusiva, pero solo con algunas raíces que comienzan por una vocal) y *r-/n-* (según si hay algún sonido nasal en la palabra, únicamente si la palabra comienza por una vocal). En el ejemplo (7) los morfoides del chiquitano migueleño se glosan como TH (*-z-*) y L (*r-/n-*).

(7) chiquitano migueleño (< chiquitano < macro-ye)

<i>Au</i>	<i>za'a</i>	<i>pát-j</i>	<i>Ø-mo</i>	<i>na'</i>	<i>enêro-j</i>	<i>ka'a</i>
LOC	DEM	mes-X	3SG-DAT	DEM	enero-X	para_que
<i>zup-aũjoko</i>		<i>iĩi'</i>	<i>Násiya,</i>	<i>r-é-z-ai.</i>		
	1+3-enfermarse	yo	Ignacia	L-1SG-TH-hijo		

‘En el mes de enero nos enfermamos yo, Ignacia y mi hijo.’

En ayoreo (un idioma de la familia Zamuco hablado en el departamento de Santa Cruz y en Paraguay), es muy común que aparezca una vocal entre un afijo de persona y una raíz (a esas vocales se les suele decir “vocales temáticas”). Por ejemplo, los sustantivos *ko* ‘jarra’ y *jnakari* ‘hijo crecido’, al combinarse con el afijo *yok-* ‘nuestro’, resultan en formas como *yok-i-go* ‘nuestras jarras’, *yok-u-jnakari* ‘nuestros hijos crecidos’ (Ciucci & Bertinetto, 2017, p. 294), donde *yok-* es un afijo, *-go* y *-jnakari* son raíces, mientras que *-i/-u-* se pueden clasificar como morfoides.

4.2.2. Tipos de morfemas según su autonomía

Los morfemas pueden clasificarse según su grado de autonomía: algunos pueden formar por sí solos una palabra, mientras que otros jamás ocurren sueltos. A los primeros se les dice **raíces libres** y a los últimos **morfemas ligados** (que se suelen subdividir en **raíces ligadas** y **afijos** según si tienen un significado léxico o más bien gramatical) –nótese que todos los afijos son ligados por definición–. En (8) se dan algunos ejemplos de raíces libres y ligadas y de afijos del ’weenhayek, un idioma de la familia Mataguaya hablado en Tarija (provincia de Gran Chaco).

(8) ’weenhayek (< mataguayas; Alvarsson & Claesson, 2014, pp. 441, 443-4)

raíces libres:	raíces ligadas:	afijos:
<i>'iwee'lah</i> ‘luna’	<i>-paaky'o'</i> ‘pie’	<i>-qaa-</i> ‘posesión indirecta’
<i>hiyaawu'</i> ‘chamán’	<i>-tsut</i> ‘bastón’	<i>-taj</i> ‘aumentativo’
<i>siky'uus</i> ‘sábalo’	<i>-qaajtsuk</i> ‘tambor’	<i>'noo-</i> ‘poseedor indeterminado’
<i>hoosan'</i> ‘hacha’	<i>-wut</i> ‘tronco’	<i>-ek</i> ‘participio’

A diferencia de las raíces libres que constituyen, por sí solas, palabras bien formadas, las raíces ligadas y los afijos necesitan combinarse con morfemas adicionales, por ejemplo: *'noopaky'o'* 'el pie de alguien (poseedor indeterminado)', *la-wut* 'su tronco', *'noo-yhin-ek* 'tejido'.

Lenguas distintas muestran tendencias distintas en cuanto a la clasificación de las raíces en libres y ligadas. En inglés casi todas las raíces son libres. En castellano, así como en muchas otras lenguas del mundo, las raíces nominales suelen ser libres (*mano*, *nube*, *verde*, *chompa*, *motacú*), y las verbales son ligadas (*llor-*, *beb-*, *estornud-*). Como vimos arriba, en *'weenhayek*, así como en muchas lenguas amazónicas, solo una parte de las raíces nominales son libres, mientras que otras son ligadas. Finalmente, hay lenguas como el *bésiro*, en que casi todas las raíces, tanto nominales como verbales, son ligadas (9).

(9) *bésiro* (< chiquitano < macro-ye; Parapaino Castro, 2008)

- a. *tusí-xhi* 'pecho', *ná-tusi* 'tu pecho', pero **túsi*
- b. *siukú-xi* 'olla (de barro)', *siukú-ma* 'ollita', *siukú-ka* 'ollas', *siuku-mán-ka* 'ollitas', pero **síuku*
- c. *a-kusiú* 'tené fuerza', *a-kusiú-ka* 'tenés fuerza', *kusiur-u* 'tiene fuerza', pero **kusíu*
- d. *i-piakí-kia* 'me caí (de arriba)', *a-pakí-kia* 'te caíste', *pákí-o* 'se cayó', pero **páki*

Las formas precedidas por un asterisco en (9) son imposibles en *bésiro*, luego *-tusi-*, *-siuku-*, *-kusiú-*, *-paki-* y casi todas las demás raíces son ligadas en esa lengua, ya que exigen al menos un prefijo o un sufijo.

Algunos morfemas ligados con un significado gramatical, que se conocen bajo el nombre de **clíticos**, presentan características intermedias entre afijos y palabras independientes. A diferencia de los afijos prototípicos, los clíticos son "promiscuos" en cuanto al tipo de elemento al que se agregan y ocupan una posición algo menos rígida dentro de la oración (en varias lenguas la posición de los clíticos es rígida, pero se determina respecto a la oración entera y no a palabras específicas). Al mismo tiempo, los clíticos difieren de palabras independientes en carecer de significado léxico y ser fonológicamente dependientes (un clítico prototípico es incapaz de portar acento, pudiendo ocurrir únicamente al lado de una palabra que sirve de su **hospedero**). Un ejemplo del castellano serían los clíticos "me" y "lo" en expresiones como *alcanzámelo* o *me lo alcanzaste*. Últimamente la validez del concepto de clíticos se ha cuestionado en la literatura, ya que no existe una definición clara que permita distinguir entre clíticos y afijos en cualquier lengua. Sin embargo, nos parece importante mencionar los clíticos porque muchas personas recurren a ese término en sus trabajos descriptivos.

4.2.3. Tipos de afijos/clíticos según su posición

Los afijos que se ubican a la izquierda de la raíz se conocen bajo el nombre de **prefijos**, y el respectivo fenómeno se denomina **prefijación**. Los prefijos se usan en casi todas las lenguas bolivianas, sobre todo en las tierras bajas. En el siguiente ejemplo de la lengua aislada yurakaré, hablada en los departamentos del Beni (Mojos) y Cochabamba (Chapare y Carrasco), aparecen los prefijos de la primera persona del singular *ti-* (‘mis amigos’, ‘mi casa’) y de la tercera persona del plural *ma-* (‘están’).

- (10) yurakaré (lengua aislada; van Gijn, 2006, p. 117)

<i>Lëshie</i>	<i>ti-jumpaño</i>	<i>ero=w</i>	<i>ma-ssë-Ø=w</i>	<i>ti-sibë=y.</i>
dos	1SG-amigo=PL		3PL-estar_parado-3=PL	1SG-casa=LOC

‘Dos amigos míos están en mi casa.’

Los clíticos prefijados se llaman **proclíticos**, y el respectivo fenómeno se conoce como **próclisis**. En (11) se muestra el uso del proclítico posesivo de la segunda persona del singular *mikye=* en ese ejja.

- (11) ese ejja (< tacana < pano–tacana; Vuillermet, 2012, p. 330)

<i>Mikye=bakwa</i>	<i>tii-’ao-naje.</i>
2SG.GEN=hijo	crecer-largo-PAS

‘Tu hijo creció.’

Los afijos que siguen a la raíz se conocen bajo el nombre de **sufijos**, y el respectivo fenómeno se denomina **sufijación**. Los sufijos se encuentran en prácticamente todas las lenguas habladas en Bolivia. En los siguientes ejemplos del aymara en (12), el sufijo *-na* denota una ubicación (“locativo”), el sufijo *-ta* expresa que se trata de la persona con la que se está conversando (“segunda persona”), el sufijo *-xa* señala que se trata de algo que ya se ha estado discutiendo (en este caso sería el zorro; “tópico”), el sufijo *-ru* representa un destino (“ilativo”), mientras que el sufijo *-i* marca la tercera persona.

- (12) aymara (< jaqi; Cerrón-Palomino & Carvajal Carvajal, 2009, p.190)

a.	<i>Pampa-na</i>	<i>ik-ta.</i>	
	pampa-LOC	dormir-2	
	‘Vos dormís en la pampa.’		
b.	<i>Qamaqi-xa</i>	<i>uyu-ru</i>	<i>mant-i.</i>
	zorro-TOP	corral-ILL	entrar-3
	‘El zorro entra al corral.’		

Los sufijos pueden derivar lexemas nuevos, como en las siguientes nominalizaciones de agente en chácobo (13).

- (13) chácobo (< pano < pano–tacana; Tallman, 2018, p.136)
- | | | |
|----|-----------------------------|--|
| a. | <i>xětě</i> ‘oler’ | → <i>xětě-xěně</i> ‘olisqueador, el que huele’ |
| b. | <i>chiquish</i> ‘ser flojo’ | → <i>chiquish-xěně</i> ‘persona o cosa estúpida’ |
| c. | <i>ara</i> ‘llorar’ | → <i>ara-xěně</i> ‘llorón’ |
| d. | <i>pi</i> ‘comer’ | → <i>pīi-xěně</i> ‘comilón’ |
| e. | <i>oxa</i> ‘dormir’ | → <i>oxa-xěně</i> ‘dormilón’ |
| f. | <i>yoma</i> ‘robar’ | → <i>yoma-xěně</i> ‘ladrón’ |

Los clíticos sufijados se llaman **enclíticos**, y el respectivo fenómeno se conoce como éncisis. Por ejemplo, en el chiquitano migueleño se emplean los enclíticos =*ityo* ‘también’, =*zo* ‘¿no ves?’, =(a)*tai* ‘nomás’, =*re*’é ‘pues (habla femenina)’ y =*te*’é ‘pues (habla masculina)’, como en (14).

- (14) chiquitano migueleño (< chiquitano < macro-ye)
- Chauki, chapíe=re’é tyákuta=re’é a-ye-ka-ti Ø-asar-a-ñi’.*
- listo gracias=pues_♀ porque=pues_♀ 2SG-venir-F.N3-CTPT 2SG.NF-mirar-NF-1SG^P
- ‘Listo, gracias pues porque pues viniste a verme.’

Los patrones de afijación que se acaban de describir se conocen bajo el nombre de **morfología concatenativa**: las palabras se forman mediante yuxtaposición de morfemas en un orden claro, uno tras otro. En algunas lenguas, sin embargo, ciertos afijos no se adjuntan al margen izquierdo o derecho de la base, sino que se insertan EN EL MEDIO de otros morfemas. A los afijos de ese tipo se les dice **infijos**, y el respectivo fenómeno se denomina **infijación**. El ejemplo (15) proviene del movima, un idioma aislado hablado en el Beni (provincia de Yacuma). Obsérvese que el infijo *-ka-* del modo *irrealis* (o su alomorfo *-a-* tras consonantes) se inserta dentro de las raíces *aroso* ‘arroz’, *intilakwa* ‘hombre’ y *enferme:ra* ‘enfermera’, ocupando la posición tras la segunda sílaba de la palabra (o tras la primera, si esta es fonológicamente pesada⁶). En los ejemplos glosados, se suele indicar los infijos mediante paréntesis angulares.

- (15) movima (lengua aislada; Haude, 2006, p. 80)
- | | | |
|----|-----------------|--------------------------|
| a. | <i>Kas</i> | <i>aro<ka’>so.</i> |
| | NEG | arroz<IRR> |
| | ‘No hay arroz.’ | |

6 Las sílabas pesadas en movima son las que contienen una vocal larga y las que terminan en una consonante (como *en-* en *enferme:ra*). Las sílabas que terminan en una vocal corta son livianas (por ejemplo, todas las sílabas en *aroso* o *intilakwa* son livianas).

- b. *Kas iti<ka>lakwa-n-chi:ye.*
 NEG hombre<IRR>-NL-EP.niño
 ‘No está el niño.’
- c. *Kas en<a>ferme:ra.*
 NEG enfermera<IRR>
 ‘No está la enfermera.’

Los clíticos infijados se conocen como **endoclíticos** y el respectivo fenómeno como **endóclisis** (o **mesóclisis**). Los endoclíticos son bastante raros en las lenguas del mundo, pero ocurren, por ejemplo, en la norma culta del portugués. Compárense las formas *lo visitaríamos*, del castellano, y *visitá-lo-íamos*, del portugués, con el mismo significado. Aquí el clítico de tercera persona masculina del singular se inserta en el medio de la forma verbal *visitaríamos*, funcionando como un endoclítico.

En algunas pocas lenguas del mundo, los afijos pueden insertarse incluso en **VARIAS** partes dentro de la raíz, de manera discontinua. A ese tipo de afijos se les dice **transfijos** (y el respectivo fenómeno se conoce como **transfijación**). No sabemos de claros casos de transfijación en las lenguas bolivianas, pero encontramos un posible ejemplo en el mojeño trinitario, un idioma Arawak hablado en la provincia de Cercado en el Beni (16).

- (16) mojeño trinitario (< achane < arahuacas; Rose, 2014, p. 230)
yono ‘ir’ + *-a* IRR → *yana* ‘ir.IRR’

En este caso, el afijo del modo irrealis *-a* podría describirse como un transfijo que reemplaza ambas vocales de la raíz, pero hay otras interpretaciones posibles (la autora del trabajo citado, por ejemplo, propone que *-a* sea un sufijo que desencadena un proceso de armonía vocálica).

Finalmente, en algunos idiomas hay afijos, llamados **circunfijos**, que parecieran tener una doble exponencia: un elemento se agrega a la izquierda de la base (como si fuera un prefijo) y otro a su derecha (como si fuera un sufijo). El respectivo fenómeno se conoce bajo el nombre de **circunfijación**. El ejemplo (17) ilustra el circunfijo *a-...-ti* en maropa, un idioma de la familia pano–tacana del Beni (provincia José Ballivián). Ese morfema puede tener una función reflexiva (significa que el paciente de un verbo transitivo es el mismo que el agente) o recíproca (significa que un verbo transitivo tiene múltiples agentes que efectúan la respectiva acción uno hacia el otro), entre otras.

- (17) maropa (< tacana < pano–tacana; Guillaume, 2012, pp. 203, 208)
- | | | | |
|----|----------------------------------|-----------|------------------|
| a. | <i>M-Ø-a-wucha-ti-a=beu</i> | <i>te</i> | <i>kwati=du.</i> |
| | 1 SG-PAS-R/R-calentar-R/R-PAS=PF | SEP | fuego=LOC |
- ‘Me calenté ya en el fuego.’

- ‘Ahora vamos a pelear (lit. pegar uno al otro).’

(18) gwarayu (tupí–guaraní < tupí; Danielsen 2020:20)

- b. *ere-Ø-yuka* → *nd-ere-Ø-yuka-i*
 2SG^A-3P-matar NEG-2SG^A-3P-matar-NEG
 ‘lo mataste’ ‘no lo mataste’

- chiquitano migueleño (< chiquitano < macro-ye)

- b. *a-jiña-ka* → *chi-’a-jiñá-ka-pi*
 2SG-querer-F.N3 NEG-2SG-querer-F.N3-NEG
 ‘querés’ ‘no querés’

- En el último ejemplo, el segundo elemento del circunfijo (*-pi*) entra antes del sufijo de movimiento centrífugo *=ti*, pero después del sufijo de finitud *-ka*.

4.2.4. Morfemas cero y conversión

En las descripciones de distintas lenguas se han identificado los llamados **morfemas cero** (normalmente se trata de **afijos cero**, aunque en algunas lenguas se ha postulado la existencia de **raíces cero**). Estos morfemas son “invisibles”, es decir, no tienen ninguna forma fonológica. Se los postula cuando algún significado gramatical (o –raramente– léxico) no se expresa en la lengua, aunque otros significados similares sí se exponen mediante morfemas. Por ejemplo, en chiquitano migueleño los verbos no finitos en la llamada voz inversa normalmente presentan el sufijo de no finitud –o (o uno de sus alomorfos) seguido de un sufijo que señala la persona y el número del objeto: –ñĩ ‘a mí’, –i ‘a vos’, –zomĩĩ ‘a nosotros (exclusivo)’, –oñĩĩ ‘a nosotros (inclusivo)’, –año ‘a ustedes’, –iño ‘a ellos, a ellas’ (20a–b). Pero cuando el objeto es de la tercera persona del singular (‘a él’, ‘a ella’), no ocurre ninguno de los dos sufijos (20c). Algunos lingüistas dirán que la forma verbal en (20c) contiene dos sufijos cero (señalados mediante el símbolo \emptyset), mientras que otros preferirán decir que la no finitud y el objeto de la tercera persona del singular simplemente no se expresan morfológicamente.

(20) chiquitano migueleño (< chiquitano < macro-ye)

- | | | |
|----|--|---|
| a. | <i>A-jiña-ka</i> | <i>a-i-tyasur-o-ñĩĩ</i> . |
| | 2SG-querer-F.N3 | 2SG-INV-llamar-NF-1SG |
| | ‘Me querés llamar.’ | |
| b. | <i>A-jiña-ka</i> | <i>a-i-tyasur-o-zomĩĩ</i> . |
| | 2SG-querer-F.N3 | 2SG-INV-llamar-NF-1+3 |
| | ‘Nos querés llamar.’ | |
| c. | <i>A-jiña-ka</i> | <i>a-i-tyasu(-\emptyset-\emptyset).</i> |
| | 2SG-querer-F.N3 | 2SG-INV-llamar(-NF-3SG) |
| | ‘Lo querés llamar.’, ‘La querés llamar.’ | |

Todos los lingüistas están de acuerdo que es mejor evitar postular morfemas cero sin necesidad, pero algunas teorías son tan rígidas que no admiten los morfemas cero en ninguna circunstancia y exigen explicaciones alternativas para las situaciones en que uno podría querer identificar un morfema cero.

Los **alomorfos cero** (ver 4.3. para el concepto de alomorfía) son más fáciles de identificar que los morfemas cero. Se trata simplemente de variantes de morfemas que poseen una forma cero, bajo determinadas circunstancias fonológicas, pero que en otros ambientes fonológicos sí tienen una forma “palpable”. Por ejemplo, en chiquitano migueleño uno no observa ningún afijo en las formas de algunos sustantivos flexionados para la segunda persona del singular: *ikiki* ‘tu uña’, *ẽẽ* ‘tu

mano', δ' ópaki 'tu hombro' (compárense las formas *xh-ikiki* 'mi uña', *ixh-ě'ě* 'mi mano', *xh-δ'ópaki* 'mi hombro', todas del habla femenina). ¿Podríamos decir que en ese idioma la segunda persona del singular no es marcada en los sustantivos? Resulta que no: los sustantivos cuya raíz empieza por una consonante sí reciben un prefijo en la misma situación: á-pope 'tu pie', á-tusi 'tu pecho', a-tá'ani 'tu cabeza, tu pelo' (pero í-kyope 'mi pie', í-chusi 'mi pecho', ichá'ani 'mi cabeza, mi pelo'). De este modo, uno se ve obligado a decir que las formas como *ikiki*, *ě'ě*, δ' ópaki sí contienen un prefijo de persona pero que, por razones fonológicas, toma la forma cero ante vocales: \emptyset -ikiki, \emptyset -ě'ě, \emptyset -δ'ópaki (en vez de *a-ikiki, *a-ě'ě, *a-δ'ópaki).

La **conversión** es un fenómeno morfológico que consiste en que una misma base puede usarse como perteneciente a distintas categorías léxicas sin ningún cambio adicional. Por ejemplo, en inglés la conversión forma verbos a partir de sustantivos y viceversa de manera productiva: compárese el verbo *to call* 'llamar' con el sustantivo *call* 'llamada'. La conversión es común también en el quechua boliviano sureño, como en los pares *t'ika* 'flor' y *t'ika-y* 'florecer', *ch'uwa* 'claro, destilado' y *ch'uwa-y* 'escurrir, destilar'; el sufijo -y en *t'ika-y* y *ch'uway* no es un verbalizador sino un marcador de infinitivo y no aparece en los verbos conjugados.

4.3. Alomorfía y suplección

Los **alomorfos** de un morfema son sus variantes que tienen formas distintas, pero significados o funciones idénticas.

La alomorfía puede presentar un **condicionamiento fonológico**: uno puede elegir el alomorfo apropiado analizando únicamente el ambiente fonológico. Por ejemplo, en guarasugwe (un idioma Tupí originario del río Guaporé, zona fronteriza entre Santa Cruz y Brasil), el prefijo de tercera persona activa normalmente tiene la forma *o-*, pero se convierte en *u-* si la sílaba siguiente contiene una *a* o una *ə* (21). De este modo, *o-* y *u-* son alomorfos fonológicamente condicionados en guarasugwe.

(21) guarasugwe (< tupí-guaraní < tupí; Ramirez *et al.*, 2017, p. 427)

- | | | |
|----|-----------------|-------------|
| a. | <i>o-móno</i> | 'se muere' |
| b. | <i>ó-ki</i> | 'llueve' |
| c. | <i>o-muđíwa</i> | 'pinta' |
| d. | <i>u-pápa</i> | 'cuenta' |
| e. | <i>u-káđi</i> | 'se pierde' |

Las raíces también pueden sufrir alomorfía fonológicamente condicionada; en este caso dicen que las raíces presentan **alternancias**. Por ejemplo, en el idioma ayoreo las consonantes iniciales de raíces pueden alternar de varias formas según el contexto (22). En este caso se diría, por ejemplo, que la raíz *poti* 'comida' presenta un alomorfo *boti*

(22) ayoreo (< zamuco; Ciucci & Bertinetto, 2017, p. 297)

- | | | | | | |
|----|----------------|------------------------|---|--------------------|----------------------|
| a. | <i>poti</i> | ‘su comida’ | → | <i>b-a-boti</i> | ‘tu comida’ |
| b. | <i>pâtarai</i> | ‘su diente’ | → | <i>m-a-matarai</i> | ‘tu diente’ |
| c. | <i>dosadi</i> | ‘su costado, su pared’ | → | <i>b-a-rosadi</i> | ‘tu costado’ |
| d. | <i>narane</i> | ‘su homóplato’ | → | <i>b-a-rarane</i> | ‘tu homóplato’ |
| e. | <i>charipi</i> | ‘silla, apero’ | → | <i>b-a-yaripi</i> | ‘tu silla, tu apero’ |

En otros casos, el uso de los alomorfos de un morfema puede estar determinado por factores morfológicos. La alomorfia con un **condicionamiento morfológico** se observa en la interacción de los prefijos de persona con los de voz en chiquitano migueleño: como se muestra en (23), tanto los prefijos de persona como los de voz cambian de forma según la combinación específica. Los alomorfos separados por barras son condicionados fonológicamente (para simplificar la exposición, no se citan los alomorfos de los prefijos de persona que ocurren ante raíces verbales que empiezan por una vocal).

(23) chiquitano migueleño (< chiquitano < macro-ye)

	raíz verbal	prefijo antipasivo <i>ba-/ma-</i>	prefijo inverso <i>iy-/iñ-</i>
	(ante cons.)		(ante vocal) (ante cons.)
1 _♀	<i>i-</i>	<i>xh-a-</i>	<i>Ø-y-/Ø-ñ-</i> <i>Ø-Ø-</i>
1 _♂	<i>i-</i>	<i>y-a-/ñ-a-</i>	<i>Ø-y-/Ø-ñ-</i> <i>Ø-Ø-</i>
2	<i>a-</i>	<i>a-a-</i>	<i>a-iy-/a-iñ-</i> <i>a-i-/a-ĩ-</i>
3	<i>Ø-</i>	<i>Ø-ba-/Ø-ma-</i>	<i>Ø-y-/Ø-ñ-</i> <i>Ø-Ø-</i>
1+2	<i>o-/u-</i>	<i>Ø-ba-/Ø-ma-</i>	<i>ba-iy-/ma-iñ-</i> <i>ba-i-/ma-ĩ-</i>
1+3	<i>zoi-/zoĩ-</i>	<i>zup-a-</i>	<i>zop-iy-/zop-iñ-</i> <i>zop-i-</i>
2+3	<i>au-/auñ-</i>	<i>ap-a-</i>	<i>ap-iy-/ap-iñ-</i> <i>ap-i-</i>
3+3	<i>bo-/mo-</i>	<i>p-a-</i>	<i>op-iy-/op-iñ-</i> <i>op-i-</i>

La **reduplicación** es el tipo de alomorfia que se da cuando un afijo, en vez de tener una forma fija, toma la forma de algún otro elemento, copiándolo total o parcialmente. Por ejemplo, según Crevels y Muysken (2012, p. 362), en cayubaba, una lengua aislada hablada en el Beni (provincia de Yacuma), un prefijo que repite la primera sílaba del verbo señala “estado continuo de estar/ser” (*vere* ‘correr’ → *ve-vere* ‘actividad de correr’, *tëpë* ‘volar’ → *të-tëpë* ‘actividad de volar’), mientras que un sufijo que repite la última sílaba del verbo significa “acción continua” (*tididi* ‘llevar en el hombro’).

La **suplección** es un fenómeno en que un morfema (una raíz o un afijo) posee alomorfos que no presentan ninguna semejanza formal y que no pueden derivarse

uno del otro mediante reglas fonológicas. En (24) se reproducen los paradigmas flexionales (4) de las adposiciones dativas ('para, a') del guarasugwe (Ramírez et al., 2017, pp. 438–9) y del chiquitano migueleño. En ambas lenguas, ocurren alomorfos supletivos: *-o* o *-tsupe* en guarasugwe; *-emo*, *-(i)mo/-ño*, *-me/-ñe* en chiquitano migueleño (la alternancia entre *m* y *ñ* es fonológica en esta última lengua).

(24)	guarasugwe	chiquitano migueleño
1	<i>ts-ó</i>	<i>(h)i-ño</i> ♀, <i>iñ-emo</i> ♂
2	<i>ne-ó</i> [njɔ]	<i>(h)a-emo</i>
3	<i>Ø-tsúpe</i>	<i>Ø-mo</i> ~ <i>Ø-(h)imo</i> (masculino: <i>Ø-(i)mo-ti</i> 'i ♂)
1+2	<i>ðáne-o</i> [ðanjɔ]	<i>(h)o-emo</i>
1+3	<i>óre-o</i> [orjɔ]	<i>zoiñ-emo</i>
2+3	<i>pe-ó</i> [njɔ]	<i>(h)au-me</i>
3+3	—	<i>ñoi-ñe</i> (masculino: <i>Ø-mo-ma</i> ♂)
sin flexión	<i>tsúpe</i>	<i>Ø-mo</i> (masculino: <i>Ø-mo-ti</i> 'i ♂)

Los afijos gramaticales también pueden presentar **alomorfía supletiva**. Es el caso de los marcadores de plural del quechua boliviano sureño (25). El sufijo original *-kuna* se mantiene con sustantivos (o pronombres) que terminan en una consonante (*qam* 'vos' → *qam-kuna* 'ustedes', *llaqtapiq* 'el del pueblo' → *llaqtapiq-kuna* 'los del pueblo'). Con sustantivos que terminan en una vocal, debido al contacto prolongado con el castellano se ha adoptado el sufijo *-s* (*llaqta* 'pueblo' → *llaqta-s* 'pueblos'), pero no se lo puede usar si le sigue un afijo que empiece con una consonante, para no violar las restricciones fonotácticas del quechua (por ello no se dice **wasi-s-nku-man*, sino *wasisninku-man* 'a sus casas', (25b)). De este modo, *-kuna*, *-s* y *-sni* son tres alomorfos supletivos de un único morfema.

(25) quechua boliviano sureño (< quechua II < quechua; Peralta Zurita 2006, pp. 48, 127)

- a. *Jaqay llaqta-s-pi qam-kuna llusqi-nki-chik carru-man-kama.*
 DEM pueblo-PL-LOC VOS-PL salir-PERS-2-2PL carro-ALL-hasta
 'En las ciudades ustedes salen directamente hacia el carro.'
- b. *Kay llaqta-pi-q-kuna ri-n-ku wasi-sni-nku-man.*
 DEM pueblo-LOC-AG-PL ir-PERS-3PL casa-PL-3PL-ALL
 'Los que son del pueblo van a sus casas.'

El término "suplección" no debe utilizarse para pares de lexemas que poseen significados similares pero cuyo uso se determina por alguna propiedad gramatical. Se trata de un error bastante común en la literatura, que debe evitarse. Por ejemplo, en 'weenhayek uno tiene pares de verbos que difieren en cuanto a su transitividad,

como *tujw* ‘comer (transitivo)’ y *t’ek* ‘comer (intransitivo)’. En yurakaré hay pares de verbos que contrastan en número, como *mala* ‘ir (singular)’ y *bali* ‘ir (plural)’, *tuwi* ‘morirse (singular)’ y *shama* ‘morirse (plural)’, *dele* ‘caerse (singular)’ y *ñeta* ‘caerse (plural)’ (van Gijn, 2006, p. 192). En bésiro hay pares de sustantivos poseídos y no poseídos: *xáin-xhi* ‘puchi’ y *n-a’á-xi* ‘su puchi’, *ñanáun-xi* ‘chaco’ y *n-i-yóo-xi* ‘su chaco’ (Parapaino Castro, 2008). En todos estos casos se trata de pares de lexemas distintos que simplemente tienen traducciones idénticas al castellano, y analizarlos como supletivos no tendría más sentido que decirles “supletivos” a tríos de sustantivos del castellano como *piel* / *cuero* / *corteza*, *brazo* / *ala* / *gajo*.

4.4. Flexión y derivación

En 4.1.1 vimos que en la morfología se hace una distinción entre **lexemas** (palabras que son parte del vocabulario de una lengua) y sus **formas**, variantes de un lexema que aparecen en distintos contextos gramaticales. La relación entre una palabra y sus formas se conoce bajo el nombre de **flexión**. Por ejemplo, diríamos que los sustantivos en castellano pueden flexionarse para el número, y un lexema como *salteña* tiene dos formas flexionadas: *salteña* y *salteñas*. La flexión se opone a la **derivación**, término con el que se denotan los mecanismos morfológicos que se emplean para formar palabras nuevas (como en *salteña* → *salteñería*).

Si reunimos todas las formas flexionadas de un lexema, obtenemos lo que se conoce bajo el nombre de **paradigma**. En (25) damos un ejemplo de un paradigma de un verbo transitivo (‘golpear’) del guarasugwe, donde se pueden observar prefijos de persona que corresponden tanto al agente (sujeto) como al paciente (objeto) del verbo.

(26) guarasugwe (< tupí–guaraní < tupí; Ramirez et al., 2017)

	a mí	a nosotros (sin vos)	a nosotros (a mí y a vos)	a vos	a ustedes	a él (a ella, a ellos/as)
yo	—	—	—	<i>oro-núpə</i>	<i>orooro-núpə</i>	<i>a-i-núpə</i>
nosotros (sin vos)	—	—	—			<i>oro-i-núpə</i>
nosotros (yo y vos)	—	—	—	—	—	<i>ða-i-núpə</i>
vos	<i>tse-núpə</i> <i>ape</i>	<i>ore-núpə</i> <i>ape</i>	—	—	—	<i>ere-i-núpə</i>
ustedes	<i>tse-núpə</i> <i>peðópe</i>	<i>ore-núpə</i> <i>peðópe</i>	—	—	—	<i>pe-i-núpə</i>
él (ella, ellos/as)	<i>tse-núpə</i>	<i>ore-núpə</i>	<i>ðane-núpə</i>	<i>ne-núpə</i>	<i>pe-núpə</i>	<i>o-i-núpə</i>

Obsérvese que en algunas combinaciones el verbo no cambia según si el sujeto es de la primera persona del singular o de la primera persona exclusiva (‘nosotros,

pero sin vos’): *oro-núpə* ‘yo te golpeo’ o ‘nosotros te golpeamos’, *orooro-núpə* ‘yo los golpeo a ustedes’ o ‘nosotros los golpeamos a ustedes’. A este fenómeno se le denomina **sincretismo**. Por lo general, no está siempre muy claro si las formas sincréticas representan dos (o más) formas distintas de un lexema que coinciden o si se trata de una forma con varias funciones. Las formas con *ape* y *peðópe* (*tse-núpə* *ape* ‘vos me golpeáis’, *ore-núpə* *ape* ‘vos nos golpeáis’, *tse-núpə* *peðópe* ‘ustedes me golpean’, *ore-núpə* *peðópe* ‘ustedes nos golpean’) probablemente no deban considerarse flexionadas: se trata de expresiones analíticas (ver 4.5) que involucran las formas *tse-núpə* ‘golpearme’ y *ore-núpə* ‘golpearnos a nosotros (sin vos)’, con prefijos *tse-* 1^P, *ore* 1+3^P.

En lenguas distintas, lexemas de clases distintas se flexionan para categorías distintas. Por ejemplo, los sustantivos en castellano se flexionan únicamente para el número, pero en el quechua boliviano sureño los sustantivos pueden recibir flexión de número (*-kuna* /*-s* /*-sni* PL), de persona/número del poseedor (*-y* 1, *-yki* 2, *-n* 3, *-nchik* 1+2, *-yku* 1+3, *-ykichik* 2+3, *-nku* 3+3) y de caso: *-q* ‘de’ (genitivo), *-pi* ‘en’ (locativo), *-man* ‘a’ (alativo), *-manta* ‘desde’ (ablativo), *-paq* ‘para’ (benefactivo, finalidad), *-wan* ‘con’ (instrumental/comitativo), *-ntin* ‘junto a, incluso con’.

quechua boliviano sureño (< quechua II < quechua; Peralta Zurita, 2006, pp.46, 69, 104, 154)

- (27) *Wasi-y-man jamu-ni.*
 casa-1-ALL venir-1.NFUT
 ‘Vine **a mi** casa.’
- (28) *Chanta jatun=chu jallp’a-ykichik?*
 luego grande=INT tierra-2+3
 ‘¿Entonces son grandes **sus** tierras **de** ustedes?’
- (29) *Entero plaza-pi aqha-n-ku á.*
 todo plaza-LOC hacer_chicha-PERS-3PL ah
 ‘**En** toda la plaza hacen chicha.’
- (30) *Kay yaku jamu-n jaqay Karpaqayma-manta.*
 DEM agua venir-PERS DEM Karpaqayma-ABL
 ‘El agua viene de allá **desde** Karpaqayma.’

A veces algunas formas que esperaríamos encontrar en un determinado paradigma no se usan por una u otra razón, lo que genera una laguna. A paradigmas de este tipo se les dice **paradigmas defectivos**. Es lo que ocurre en castellano con verbos como *balbucir* o *colorir*, que carecen de algunas formas: no se suele decir

cosas como *yo *balbuzco* o *yo *coloro*. De igual manera, aunque muchos sustantivos del idioma baure pueden ocurrir tanto en el singular como en el plural (31a–d), algunos sustantivos carecen de una forma singular (31e–f) o plural (31g–h).

(31) baure (< arahuacas; Danielsen, 2007, pp. 128–9)

- | | |
|--|--|
| a. <i>pari</i> ‘casa’ | <i>pari-nev</i> ‘casas’ |
| b. <i>witer</i> ‘murciélago’ | <i>witer-nev</i> ‘murciélagos’ |
| c. <i>wajis</i> ‘estrella’ | <i>wajis(o)-nev</i> ‘estrellas’ |
| d. <i>ni-poiy</i> ‘mi pie’ | <i>ni-poyi-nev</i> ‘mis pies’ |
| e. <i>in</i> ‘agua’ | (no se puede decir * <i>ino-nev</i>) |
| f. <i>ses</i> ‘sol’ | (no se puede decir * <i>ses-nev</i>) |
| g. — (no se puede decir * <i>ahi</i>) | <i>ahi-nev</i> ‘niños’ |
| h. — (no se puede decir * <i>ni-sheche</i>) | <i>ni-sheche-nev</i> ‘mis hijos e hijas’ |

La flexión, que consiste en el cambio de la forma de un lexema en determinados contextos gramaticales, se opone a la **derivación**, que se define como el proceso de creación de lexemas nuevos por medios morfológicos. A veces es difícil establecer si un determinado mecanismo morfológico debe definirse como flexional o derivacional, ya que algunas estrategias de creación de lexemas en ciertas lenguas son tan **productivas** que pueden aplicarse a cualquier lexema de una determinada clase, poniendo en tela de juicio la posibilidad de determinar si las formas resultantes constituyen lexemas nuevos o si son simplemente formas gramaticales de los lexemas originales. Por ejemplo, ¿sería *cafecito* (o *cafecingo*) un lexema distinto de *café*?, ¿o podríamos decir que es más bien una forma diminutiva de *café*? Los mecanismos derivacionales **no productivos**, en cambio, operan en un número reducido de lexemas. En (32) se da un ejemplo de un sufijo no productivo *-wa* ‘el que suele comer tal cosa’ del idioma guarasugwe.

(32) guarasugwe (< tupí–guaraní < tupí; Ramirez et al., 2017)

- a. *kapiʔi* ‘paja’ + *-wa* → *kapiʔ-wa* ‘capiguara’
- b. *mutéwi* ‘garrapata’ + *-wa* → *mutewí-wa* ‘chimachimá (gavilán garrapatero)’
- c. *tapiʔi* ‘anta’ + *-ki* ‘piojo’ + *-wa* → *tapiʔi-kí-wa* ‘caracara negro (gavilán que come piojos de anta)’

El conjunto de lexemas que comparten una única raíz (y, por ende, que fueron derivados a partir de esa raíz) se conoce bajo el nombre de **familia de palabras**. Un ejemplo de una familia de palabras es el conjunto de lexemas del castellano que incluyen la raíz *plum-*: *pluma*, *plumaje*, *desplumar*, *plumoso*. En chiquitano migueleño, a partir de la raíz *cha(b)-* (como en el verbo *chab-o* ‘beber, tomar’), uno

puede derivar palabras como *ma-chám-an-a* ‘invitar a tomar’ (causativo) y *cha-po-j* ‘vaso’ (nominalización de instrumento).

4.5. Lenguas con poca y mucha morfología

No todas las lenguas son igual de complejas en cuanto a su morfología: en algunas lenguas casi todas las palabras contienen exactamente un morfema, mientras que en otras es completamente normal que una palabra contenga múltiples afijos, como en el quechua boliviano sureño (*quedakapu-na-lla-y-paq-puni* ‘decididamente era para que me quede nomás siempre’; Peralta Zurita, 2006, p.110).

Un ejemplo de un idioma con poca morfología es el moré (familia Chapacura), hablado en la provincia de Mamoré en el Beni. Según Angenot-de Lima (2002), ni los verbos ni los sustantivos del moré pueden recibir afijos. En los ejemplos dados en (33), las únicas palabras morfológicamente complejas son la partícula *n-on* (de *na*: IPF + *?on* M) y el nombre propio *sa*:=*?ε*: (literalmente ‘eructar’, un compuesto de *sa*: ‘pudrido’ y *?ε*: ‘defecar’).

(33) moré (chapacura; Angenot-de Lima, 2002, pp. 394, 397)

a. *?iwan?* *na*: *?aŋ* *m^wijak* *?utip*
llegar IPF NEU₁ taitetú Utip
‘Utip llegaba con el taitetú.’

b. *pam* *n-on* *?utip* *pa* *?kɔm* *pɔŋ* *sa*: *sa*:=*?ε*:
sacudir IPF-M Utip PREP agua despertar EXH Saé
‘Saé sacude a Utip para que el agua lo despierte.’

A las lenguas como el moré, en que los afijos son prácticamente inexistentes, se les suele decir **lenguas aislantes**. Algunos ejemplos de lenguas aislantes habladas fuera de Bolivia incluyen el chino (mandarín) y el yoruba. La mayoría de las lenguas bolivianas, sin embargo, no son aislantes, sino **analíticas** (es decir, las que hacen un uso moderado de morfología), **sintéticas** (las que usan mucha morfología) o incluso **polisintéticas** (las que usan una cantidad extraordinaria de morfología). Es importante entender que no se trata de una clasificación tipológica rígida –es decir, no se puede analizar un determinado idioma (como el castellano o el quechua boliviano sureño) como estrictamente analítico, sintético o polisintético–, sino de un continuo. La tabla a continuación muestra el número promedio de morfemas (o el “índice de complejidad morfológica”) en varias lenguas bolivianas, calculado con base en textos representativos.⁷

7 Los cálculos toman en cuenta apenas la morfología concatenativa (eso es, los morfemas separados por guiones en la glosa interlineal) y la palabra es definida ortográficamente. Usamos la segmentación morfológica propuesta en las obras citadas para fines de uniformidad, aunque en algunos casos estamos en desacuerdo con los análisis de los autores. Evidentemente, las preferencias por uno u otro análisis afectarían los valores de los índices de complejidad.

Idioma	Familia	Índice	morfemas/ palabras	Texto
cayubaba	aislada	2,77	202/73	Crevels y Muysken, 2012, pp.370–2
itonama	aislada	2,49	808/324	Crevels, 2012, pp. 283–91 (“La carau”, “Mi marido”)
bésiro	Macro-ye	2,34	795/340	Sans, 2013, pp. 47–57 (“El guajojó”)
aymara	Jaqi	2,19	70/32	Cerrón-Palomino y Carvajal Carvajal, 2009, pp. 211–2 (“Las andanzas del zorro”)
leko	aislada	2,15	626/291	van de Kerke, 2009, pp. 287–331
mosetén de Covendo	Mosetén–chimané	2,08	327/157	Sakel, 2009, pp. 366–9
paunaka	Arahuaca	2,05	1081/527	Terhart, 2022, texto A.1
baure	Arahuaca	1,99	589/296	Danielsen, 2007, pp. 448–53 (“La rana y el peni”)
quechua boliviano	Quechua	1,98	1521/767	cuento “Jiqi Jiqi” (autoría: Santiago Guevara, Asteria Delgadillo)
uru	Uru–chipaya	1,97	57/29	Hannss, 2009, p. 112
guaraní	Tupí	1,88	985/523	Gustafson, 2014, pp. 353–64 (“Yaguariya”)
maropa	Pano–tacana	1,85	280/151	Guillaume, 2012, pp. 224–7 (“Cuando encontré chanchos de tropa”)
’weenhayek	Mataguaya	1,75	180/103	Alvarsson & Claesson, 2014, pp. 459–61
ese ejja	Pano–tacana	1,69	315/183	Vuillermet, 2012, pp. 698–703 (“El palo santo”)
chipaya	Uru–chipaya	1,67	45/27	Cerrón-Palomino, 2009 (“Cuento de los patitos”)
moré	Chapacura	1,66 ¹	720/435	Angenot-de Lima, 2002, pp. 746–53 (“El pájaro Tapam Mamam”)
movima	aislada	1,59	319/201	Haude, 2006, pp. 562–5 (“Tigre y perro”)
chácobo	Pano–tacana	1,53	1299/850	Tallman, 2018, pp.1247–73 (“La madre del viento sur”)
ayoreo	Zamuco	1,42	151/106	Ciucci, 2019, pp.182–3, ejemplo 4

A las lenguas (poli)sintéticas con un alto grado de exponencia biunívoca (ver 4.1.2.), como el quechua boliviano sureño o el aymara, se les suele decir **lenguas aglutinantes**. A las lenguas sintéticas en que predomina la exponencia cumulativa (4.1.2.), como el ayoreo o el chiquitano, a veces se les dice **lenguas fusionales**.

4.6. Conclusión

En este capítulo introductorio se han presentado los principales conceptos de la morfología, acompañados de algunos ejemplos de lenguas originarias bolivianas. Se ha focalizado en la descripción lingüística, y no se ha ni siquiera mencionado las distintas **teorías** morfológicas, como, por ejemplo, la morfología distribuida. A los lectores y las lectoras que deseen profundizar en sus conocimientos de la morfología general, se les recomienda la lectura del manual de Haspelmath y Sims (2010, en inglés), mientras que una excelente introducción a las distintas teorías morfológicas se encuentra en el libro organizado por Audring y Masini (2018, también en inglés).

Referencias

- Alvarsson, J.-Å., Claesson, K. (2014). Weenhayek (mataco). In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo III: Oriente*. La Paz: Plural. 415–65.
- Angenot-de Lima, G. (2002). *Description phonologique, grammaticale et lexicale du Moré, langue amazonienne de Bolivie et du Brésil*. [Tesis doctoral, Universiteit Leiden].
- Audring, J., Masini, F. (eds.). (2018). *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199668984.001.0001
- Camacho-Ríos, G.; Tallman, A. (Por aparecer). Constituency and wordhood in South Bolivian Quechua. In: Tallman, Adam J. R.; Auderset, Sandra; Uchihara, Hiroto (eds.). *Constituency and convergence in the Americas*. Berlín: Language Science Press.
- Cerrón-Palomino, R. (2009). Chipaya. In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural. 29–77.
- Cerrón-Palomino, R., Carvajal Carvajal, J. (2009). Aimara. In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural. 169–213
- Ciucci, L., Bertinetto, P. M. (2017). Possessive inflection in Proto-Zamucoan. *Diachronica* 34(3):283–330. doi:10.1075/dia.34.3.01ciu
- Ciucci, L. (2019). A culture of secrecy: the hidden narratives of the Ayoreo. *International Journal of Language and Culture*, 6(1).175–194. doi:10.1075/ijolc.00021.ciu
- Condori, N., Gabriel, Y., Gallinate, G. A. (2021). La composición adjetival y sus peculiaridades dentro del sistema del quechua del sur de Bolivia. *Página y Signos*, 16. 27–50.
- Crevels, M. (2012). Itonama. In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo II: Amazonía*. La Paz: Plural. 233–94
- Crevels, M., Muysken, P. (2012). Cayubaba. In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo II: Amazonía*. La Paz: Plural. 341–74.
- Danielsen, S. (2007.) *Baure: an Arawak language of Bolivia*. Leiden: CNWS Publications. (Indigenous Languages of Latin America 6.)

- Danielsen, S. (2020). *Esbozo del idioma guarayu de Hoeller (2019 [1929/1932])*. [Gwarayu Ñe'ësa 3]. Urubichá: edición propia. <https://doi.org/10.5281/zenodo.4111047>
- Emlen, N. Q., Mossel, A., van de Kerke, S., Adelaar, W. F. H. (2023). Puquina. In: Urban, M. *Oxford Guide to the Languages of the Central Andes*. Oxford: Oxford University Press.
- Guillaume, A. (2012). Maropa (reyesano). In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo II: Amazonía*. La Paz: Plural. 191–229.
- Gustafson, B. (2014). In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo III: Oriente*. La Paz: Plural. 307–68.
- Hannss, K. (2009). Uchumataqu (Uru). In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural. 79–115.
- Haspelmath, M., Sims, A. (2010). *Understanding morphology*. Londres: Hachette.
- Haude, K. (2006). *A grammar of Movima*. [Tesis doctoral, Radboud Universiteit Nijmegen].
- Nikulin, A. (Por aparecer). Morfología de finitud y estrategias de subordinación en chiquitano migueleño. *Amerindia*.
- Parapaino Castro, P. (2008). *Isiukiché nikorokó Bésiro: guía de escritura del idioma Bésiro*. Santa Cruz de la Sierra: Unión de Artesanos de la Tierra-UNIA RTE.
- Peralta Zurita, E. (2006). *Descripción morfológica de la palabra quechua: un estudio basado en el quechua de Yambata, Norte Potosí*. [Tesis doctoral, Universidad Mayor de San Andrés].
- Ramirez, H., Vegini, V., Vitorino de França, M.C. (2017). O warázu do Guaporé (tupi-guarani): primeira descrição linguística. *LIAMES. Línguas Indígenas Americanas* 17(2). 411–506. doi:10.20396/liames.v17i0.8647468
- Rose, F. (2014). Negation and irrealis in Mojeño Trinitario. In: Michael, Lev; Granadillo, Tania (eds.). *Negation in Arawak Languages* [Brill's Studies in the Indigenous Languages of the Americas, 6]. Leiden: Brill. 216–40doi:org/10.1163/9789004257023_011
- Sakel, J. (2009). Mose tén y Chimane (Tsimane'). In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural. 333–75.
- Sans, P. (2013). *Elementos de la gramática del Bésiro: sociolingüística — fonología — morfología — textos*. Manuscrito, San Antonio de Lomerío.

- Tallman, A. J. R. (2018). *A grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon*. [Tesis doctoral, University of Texas at Austin]. doi:10.26153/tsw/1343
- Tallman, A. J. R. (2021). Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–83. doi:10.1075/sl.19025.tal
- Terhart, L. (2022). *A grammar of Paunaka*. [Tesis doctoral, Europa-Universität Flensburg].
- van de Kerke, S. (2009). Leko. In: M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I: Ámbito andino*. La Paz: Plural. 287–331.
- van Gijn, R. (2006). *A grammar of Yurakaré*. [Tesis doctoral, Radboud Universiteit Nijmegen].
- Vuillermet, M. (2012). *A grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. [Tesis doctoral, Université Lumière Lyon 2].

CAPÍTULO 5

Cómo citar: Myler, N. (2023). Sintaxis. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 130-140), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11111076>

SINTAXIS

Neil Myler *

0. Introducción

La sintaxis es la parte de la lingüística que se enfoca en las reglas que rigen la estructura de las frases y de las oraciones. El orden de los elementos de la frase forma parte de la sintaxis. Por ejemplo, el hecho de que el verbo preceda a su sujeto en movima, mientras que en quechua el orden sea el opuesto, representa una diferencia sintáctica entre los dos idiomas.

(1) Movima

V	S	
Joy-chel	is	kompanyera-sne.
ir-se	las	amigas-suyas

“Sus amigas se fueron.” (adaptado de Haude, 2006, p. 258)

(2) Quechua boliviano

S	V
Masisnin	ripurqanku.
Amigas.sus	se.fueron

“Sus amigas se fueron.”

Sin embargo, la sintaxis involucra mucho más que el orden de las palabras. Cuando estudiamos sintaxis, nos fijamos mayormente en la estructura jerárquica en que se combinan los elementos básicos de la oración, y en las relaciones que existen entre estos. Resulta que las palabras no se combinan como las cuentas de un rosario, una tras otra. En realidad, en cualquier oración, y en cualquier idioma, hay elementos que se relacionan entre sí unos más estrechamente que con otros. En muchos casos, estas relaciones más estrechas son obvias. Por ejemplo, aunque yo no

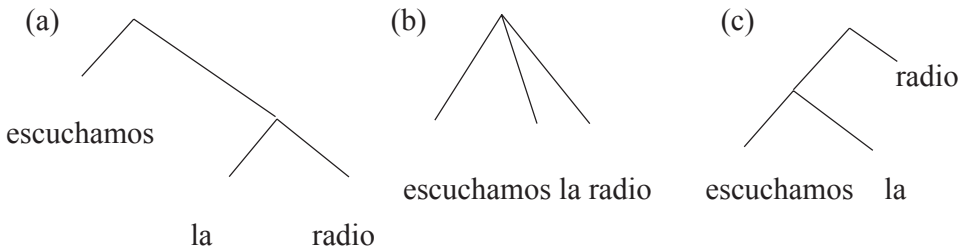
* Es Profesor Asociado en la Universidad de Boston. Se interesa principalmente en la morfología, la sintaxis y las relaciones entre estas disciplinas. Su tesis de doctorado, llevada a cabo en la Universidad de Nueva York, se enfocó en las oraciones posesivas, con mayor énfasis en los idiomas quechuas. Una versión revisada fue publicada como libro en 2016, con el título *Building and Interpreting Possession Sentences*.

sea nativohablante del castellano, tengo una intuición fuertísima que la palabra *la* en la siguiente oración está vinculada más directamente con la palabra *radio* que con la palabra *escuchamos*, y también estoy seguro de que los lectores nativohablantes estarán de acuerdo conmigo en cuanto a este juicio.

(3) Escuchamos la radio.

Es común y útil en los estudios sintácticos representar las hipótesis sobre estas relaciones estructurales de forma diagramática. La intuición que acabamos de observar sugiere que la hipótesis (a) en (4) es superior en este caso que las hipótesis (b) o (c).

(4) Hipótesis para el ejemplo (3)

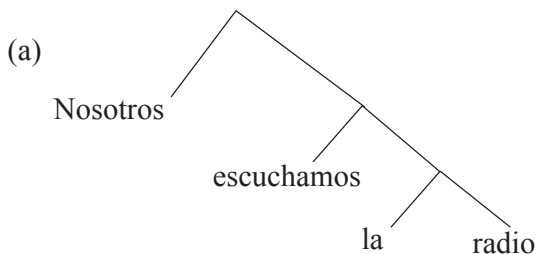


No obstante, las estructuras sintácticas no siempre son tan evidentes. Tomemos la siguiente oración.

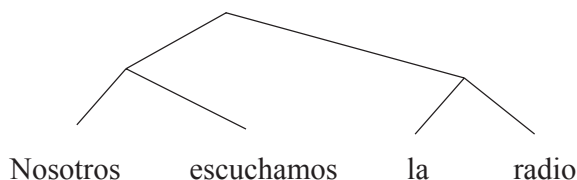
(5) Nosotros escuchamos la radio.

Retengamos nuestra conclusión anterior que *la radio* forma un grupo que excluye al verbo, y preguntémonos cómo es el resto de la estructura de esta oración. Podría ser que el verbo resulte más estrechamente relacionado con el objeto directo *la radio* que con el sujeto *nosotros*; si es así, el diagrama (a) debajo constituiría la mejor hipótesis. En cambio, podría ser el sujeto el que forma una unidad con el verbo que excluye al objeto, como en el diagrama (b). O podría ser que el verbo esté igualmente vinculado con el sujeto y con el objeto, como en el diagrama (c). Si el lector está leyendo este capítulo en el contexto de una clase, debería comparar su opinión sobre esta pregunta con las de sus compañeros de curso.

(6)



(b)



(c)



Queda claro que las intuiciones ingenuas que hemos utilizado arriba no son suficientes en este caso para resolver la cuestión –se necesitan argumentos que conformen principios para establecer la estructura–. Dicho de otra forma, necesitamos pruebas sintácticas de diagnóstico.

Describir bien la sintaxis de un idioma consiste, en gran parte, en el desarrollo y la aplicación de tales pruebas de diagnóstico. Además, entender los principios universales de la sintaxis en la mente humana, lo que es la meta final de la teoría sintáctica, requiere descripciones buenas de muchos idiomas de diversos tipos. El progreso hacia este objetivo se hace mucho más rápido cuando los que llevan a cabo las indagaciones sintácticas son los mismos nativohablantes del idioma estudiado. Que este capítulo sea el primer paso de los lectores en su camino hacia la mejor descripción de la sintaxis de su idioma nativo.

El resto de este capítulo se organiza de la forma siguiente. Empieza con una explicación de los tipos de datos que pueden utilizarse en las investigaciones sintácticas. En la segunda parte, se discute acerca de las pruebas de la estructura de constituyentes. La tercera parte trata de la relación entre la morfología y la sintaxis, y de los problemas que surgen a la hora de distinguir entre estas partes de la gramática. En la cuarta parte, se revisa unos ejemplos de la variación sintáctica en las lenguas bolivianas y se describe ciertos conceptos importantes para entender esta variación. Finalmente, se presenta una breve conclusión.

5.1. Tipos de datos sintácticos

Los tres tipos de datos más importantes para la sintaxis son los **juicios de aceptabilidad** de los nativohablantes, las traducciones de oraciones o frases específicas obtenidas de los nativohablantes y los textos naturales (un término que se refiere aquí a cualquier muestra de discurso natural producida por un nativohablante –que sea una conversación espontánea grabada, un cuento oral o escrito, un artículo

de un periódico o blog, el resultado de una búsqueda en el internet, etc.—). Cada uno es útil en diversos contextos. Los tres tipos de datos son complementarios y, por eso, es muy ventajoso saber cómo obtenerlos y utilizarlos (ver Capítulo 1).

El tipo menos autoexplicativo para un(a) lingüista principiante es el juicio de aceptabilidad. Los hablantes saben mucho más sobre sus idiomas de lo que dicen en la vida cotidiana. Los idiomas son capaces de expresar millones de cosas que la gente nunca diría en un contexto típico. No obstante, no todas las combinaciones lógicamente posibles de las palabras tienen el mismo estatus. Compárense estas dos oraciones, basadas en dos ejemplos extremadamente famosos de Chomsky (1957) (explicaremos pronto los símbolos “#” y “*”).

(7) #Las ideas verdes incoloras duermen furiosamente.

(8) *Furiosamente duermen incoloras verdes ideas las.

Ninguno de los dos tiene ni siquiera una miga de sentido. Sin embargo, el ejemplo (8) es problemático de una manera más profunda que el ejemplo (7). El ejemplo (7) es absurdo desde el punto de vista de su significado, pero conforma los patrones gramaticales del castellano que los hablantes conocen implícitamente: el artículo precede a su sustantivo, el que precede a los adjetivos que lo modifican, etc. Indicamos que esta oración es semánticamente extraña (aunque sintácticamente normal) poniéndole el símbolo “#”. Por el contrario, el ejemplo (8) es aún menos aceptable para los nativohablantes, no porque su significado sea más absurdo que el del (7), sino porque la manera en la que sus palabras están combinadas no cumple con los requisitos estructurales del idioma. Cuando un ejemplo es inaceptable a causa de sus rasgos gramaticales, lo marcamos con el símbolo “*”. Deberíamos subrayar que la causa de un juicio de inaceptabilidad no siempre es evidente, y es importante tener cuidado a la hora de interpretar tales datos. Sobre todo, se debería construir cuidadosamente los ejemplos para los cuales se pretende obtener juicios, para asegurarse de que los juicios de inaceptabilidad se deban a los factores gramaticales de interés, y no a otros factores (por ejemplo, factores pragmáticos o semánticos). También debería notarse que los nativohablantes del mismo idioma pueden variar en sus juicios de aceptabilidad. A veces esta variación se debe a factores sin interés lingüístico (fallas de atención, etc.), pero los casos sistemáticos de variación pueden indicar una verdadera diferencia entre las gramáticas mentales de los hablantes. Tales diferencias son de esperar, incluso si descartamos la variación debida a la existencia de dialectos regionales, porque cada persona construye su propia gramática mental cuando adquiere su idioma en la niñez. Es inevitable, entonces, que surjan algunas diferencias gramaticales incluso entre los miembros de la misma comunidad. La diversidad de este tipo es el enfoque de la **sintaxis microcomparativa**, y tal como la **sintaxis macrocomparativa** (la comparación de idiomas más distantemente

relacionados o sin relación alguna), es de alta importancia para la teoría sintáctica (véase, por ejemplo, Kayne, 1996).

Los juicios de aceptabilidad son indispensables porque nos pueden informar muy rápida y claramente sobre la naturaleza de las reglas sintácticas mentales de los hablantes, mostrándonos lo que esas reglas no permiten. Tales inferencias *negativas* que los juicios de aceptabilidad facilitan son mucho más difíciles de obtener por medio de las traducciones o de los textos naturales. Es por eso que muchos fenómenos sintácticos importantísimos para la teoría lingüística se quedaron sin descubrir hasta mitades del siglo XX, la época en que los juicios de aceptabilidad se hicieron frecuentemente empleados.

No obstante, los otros tipos de datos no pueden descartarse. Las traducciones son un punto de partida indispensable si estamos investigando un idioma que no conocemos bien, aunque son susceptibles de errores causados por los malentendidos (de parte del lingüista, que puede interpretar mal la estructura de la traducción, o de parte del consultor, que puede entender mal la oración cuya traducción el lingüista ha pedido). Los textos naturales también son importantes para la investigación sintáctica, además de su obvia importancia para las investigaciones sociolingüísticas y pragmáticas. El uso de los textos naturales nos enseñará contornos sintácticos y rasgos gramaticales de los cuales no nos daríamos cuenta si nos limitáramos a las traducciones o a los juicios de aceptabilidad basados en ejemplos contruidos. Además, cuando los textos son combinados en un corpus en que se pueden realizar búsquedas, es fácil verificar si un ejemplo contruido se encuentra o no en el corpus. Si el ejemplo existe en el corpus, pero nuestra hipótesis predice que no debería ser posible, esto constituye una prueba contra la hipótesis. Las desventajas de los textos naturales son (i) que muchas cosas que son posibles e interesantes son de uso poco común, y pueden estar ausentes de un corpus a pesar de ser gramaticales, y (ii) a veces los errores del habla pueden sobrevenir.

Esto concluye nuestra examinación de tres fuentes de datos sintácticos complementarias, y las ventajas y desventajas de cada una.

5.2. Pruebas de estructura de constituyentes

Una secuencia de elementos que forman un grupo estructural distinto se llama un **constituyente**. Un fenómeno gramatical que se define en términos de constituyentes se llama una **prueba de estructura de constituyentes** (en inglés, *a constituency test*). Los idiomas de Bolivia (y de todo el mundo) varían según los fenómenos gramaticales que tienen; por eso, las pruebas de estructura de constituyentes son específicas al idioma para el cual han sido desarrolladas. Sin embargo, la lógica de estas pruebas es universal y funciona en cualquier idioma. He aquí cómo funciona el proceso. Observamos los fenómenos sintácticos de un idioma e intentamos, a cada

rato, formular reglas para describir estos fenómenos. Nos damos cuenta de que, en muchos casos, estas reglas se refieren a algunos tipos de secuencias y no a otros. Esto sugiere que necesitamos representar estas secuencias como constituyentes en nuestra descripción gramatical.

Aunque las pruebas de estructura de constituyentes sean lingüísticamente específicas, la mayoría de ellas pueden categorizarse en los siguientes tipos: (i) pruebas de sustitución, (ii) pruebas de desplazamiento, (iii) pruebas de elipsis y (iv) pruebas de coordinación. Ahora veremos algunos ejemplos de cada tipo. De tarea, el lector podría desarrollar más pruebas de cada tipo para su lengua materna, basadas en los ejemplos que proveemos a continuación.

5.2.1. Pruebas de sustitución

Si una secuencia se puede reemplazar con una palabra tal como un pronombre, sin cambiar el significado básico de la oración, es una prueba de que es un constituyente. Por ejemplo, sabemos que *al nuevo bar de la esquina* es un constituyente en el ejemplo (9), porque podemos sustituirlo con *allí* en el ejemplo (10).

(9) Nos vamos al nuevo bar de la esquina.

(10) Nos vamos allí.

Igualmente, se puede demostrar que la secuencia *chay huch'uy wayna* “ese muchacho pequeño” es un constituyente en el ejemplo (11) del quechua boliviano reemplazándola con el pronombre *pay* en el ejemplo (12).

(11) Chay huch'uy wayna tantata munan. (quechua boliviano)
Ese pequeño muchacho pan.ac quiere
“Ese muchacho pequeño quiere pan.”

(12) Pay tantata munan. (quechua boliviano)
Él/ella pan.ac quiere
“Él quiere pan.”

5.2.2. Pruebas de desplazamiento

Si una secuencia puede desplazarse como un grupo, sin cambiar el sentido básico de la oración, se puede concluir que es un constituyente. Abajo se demostró que la secuencia subrayada es un constituyente, desplazándola al inicio de la oración en cada caso.

(13) Te vi en esa plaza ayer.

(14) En esa plaza, te vi ayer.

- (15) Qayna chay plazapi rikhurqayki. (quechua boliviano)
 Ayer esa plaza.en te.ví
 “Ayer te vi en esa plaza.”
- (16) Chay plazapi, qayna rikhurqayki. (quechua boliviano)
 Esa plaza.en ayer te.ví
 “En esa plaza, te vi ayer.”

5.2.3. Pruebas de elipsis

Si es posible omitir una secuencia cuando el contexto permite al oyente entender lo que el hablante quiere decir, es una prueba de que forma un constituyente. En los ejemplos siguientes, se indica el constituyente omitido mediante tachado.

- (17) Quisimos escuchar la radio, pero no pudimos ~~escuchar la radio~~.
- (18) Chay kinsa hatun allqu paypa kanku, kay iskay ~~hatun allqu~~ nuqaq kanku.
 Esos tres grandes perros de él son estos dos grandes perros de mí son
 “Esos tres perros grandes son suyos, estos dos son míos.” (quechua boliviano)

5.2.4. Pruebas de coordinación

Este tipo de prueba es más controvertido, ya que a veces sus resultados son contradictorios, cosa que a algunos lingüistas les ha conducido a concluir que la coordinación no es una prueba de estructura de constituyentes. No obstante, la mencionamos aquí porque es muy utilizada.

Muchos idiomas tienen partículas que combinan secuencias de varios tipos –las palabras *y*, *o*, y *pero* son ejemplos del castellano–. Las secuencias que pueden juntarse de esta manera son constituyentes (suponiendo que confiamos en esta prueba).

- (19) Estas muchachas y esos muchachos están bailando.
- (20) Tukuy warmis tantata mikuchkanku, aqhata ujyachkanku-taq.
 Todas mujeres pan.ac están.comiendo chicha.ac están.tomando-y
 “Todas las mujeres están comiendo pan y tomando chicha”. (quechua boliviano)

5.2.5. Resumen de la parte 2

En esta parte hemos presentado el concepto de una prueba de estructura de constituyentes, y hemos dado algunos ejemplos de los tipos más frecuentemente utilizados. Para practicar, animamos al lector a que vuelva al final de la introducción y que, utilizando los ejemplos de pruebas que hemos provisto, descubra cuál de los diagramas en (6) es el más adecuado para representar la estructura sintáctica de la oración *nosotros escuchamos la radio* en castellano.

5.3. Morfología, sintaxis, morfosintaxis

La gramática teórica tradicional suele distinguir entre la morfología y la sintaxis. Se supone que la diferencia entre las dos es que la morfología se fija en la estructura interna de las palabras y en las formas de estas, mientras que la sintaxis se enfoca en la estructura interna de las frases y de las oraciones.

Aunque esta distinción sea tradicional, y aunque parezca sencilla, a menudo es difícil aplicarla en la práctica. Esta dificultad es muy prevalente en el análisis de las lenguas indígenas de las Américas (véanse Muysken, 1981; Pike, 1945; Tallman, 2020 y Weber, 1983 para diversas perspectivas). Resulta que los constituyentes sintácticos no siempre respetan los límites de las palabras. Tomemos el ejemplo de la frase *kay huch'uy waynawan* “con este muchacho pequeño” en el quechua boliviano, que tiene la descomposición morfológica indicada (este ejemplo está basado en la discusión de Myler, 2016, pp. 38-39):

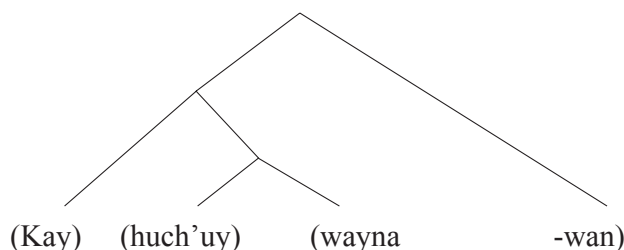
- (21) Kay huch'uy wayna-wan (quechua boliviano)
 Este pequeño muchacho-con
 “Con este muchacho pequeño”

Como la ortografía lo indica, el sufijo *-wan* (que significa “con”) forma parte de una palabra compleja cuya raíz es el sustantivo *wayna*. Sin embargo, la prueba de sustitución demuestra que *kay huch'uy wayna* forma un constituyente que excluye a *-wan*.

- (22) Kay huch'uy wayna-wan → Pay-wan
 “con este muchacho pequeño” → “con él”

El diagrama de abajo muestra la estructura de constituyentes del ejemplo. Los límites morfológicos/fonológicos de las palabras se representan con paréntesis.

(23)



Tales incompatibilidades entre los límites de las palabras y los límites de los constituyentes sintácticos hacen imposible el mantenimiento de una distinción clara entre la morfología y la sintaxis. Para practicar o para discutir en clase, el lector debería preguntarse si existen casos en su lengua materna en que los límites de palabras no estén de acuerdo con la estructura de constituyentes (de ser el lector quechuahablante, debería crear más ejemplos del mismo fenómeno en quechua).

5.4. Variación sintáctica en las lenguas indígenas bolivianas: algunos conceptos y ejemplos

Por falta de espacio, nos concentraremos aquí en dos temas solamente: el ordenamiento básico de las frases al nivel de la oración y el sistema de marcas de caso. Ambos tienen que ver con los recursos gramaticales con los que cuentan las lenguas para señalar las relaciones sintácticas como “sujeto” y “objeto”.

Empecemos comparando los tres ejemplos siguientes de los idiomas movima, ese ejja y quechua.

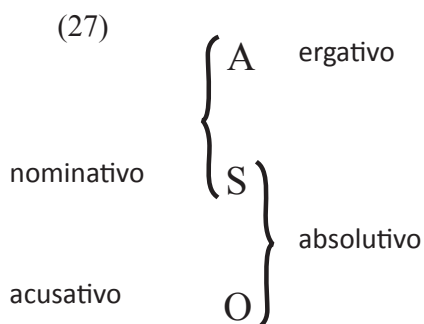
- (24) Che man<a>ye =is pa:ko os rulrul. (movima; Haude, 2006, p. 259)
 Y encontrar.se<DR> art.pl perro art.n.pl jaguar
 “Y los perros se encontraron con el jaguar”.

- (25) E-sho'i=a e-aa ज्या-sowa-ka-ani. (ese ejja; Vuillermet, 2012, p. 285)
 NPF-niño=erg NPF-brazo lanzar-levantar-3ª-pres
 “El niño se levanta el brazo”.

- (26) Wawa maki-n-ta juqhari-n. (quechua)
 Niño mano-3poss-ac levantar-3subj
 “El niño se levanta la mano”.

Debería notarse que estos idiomas permiten otros órdenes, sobre todo quechua y ese ejja (según Haude, el orden de las palabras en movima es algo fijo); no obstante, estos ejemplos reflejan el orden “neutro” de cada idioma, o sea, el orden más común cuando ninguna parte de la oración tiene más énfasis que el resto. Una prueba simple para establecer tal orden es preguntar cuál orden es el más natural como respuesta a una pregunta como “¿qué pasó?”.

De inmediato nos damos cuenta de que el movima se distingue de los otros en que el orden es Verbo-Sujeto-Objeto, mientras que en ese ejja y en quechua encontramos el orden Sujeto-Objeto-Verbo. Además, a pesar de su similitud en el orden lineal, una diferencia importante entre ese ejja y quechua surge cuando examinamos cuidadosamente la morfología del sujeto y del objeto. En ese ejja, el sujeto del verbo transitivo lleva un sufijo casual llamado “ergativo”, y el objeto del verbo transitivo no lleva sufijo alguno —está en una forma llamada “absolutivo”, una forma que se utiliza también para el sujeto de un verbo intransitivo—. En cambio, en quechua es el objeto el que lleva un sufijo casual específico, llamado “acusativo” y el sujeto se queda sin marcar, una forma que también se encuentra con el sujeto de un verbo intransitivo y que se llama “nominativo”. Entonces, ese ejja y quechua son representantes de dos sistemas diferentes de marcas casuales: el ese ejja es un idioma “ergativo-absolutivo” y el quechua es un idioma “nominativo-acusativo”. Dixon (1994, p. 9) provee un diagrama que aclara la diferencia entre los dos sistemas; la etiqueta “A” se refiere al sujeto de un verbo transitivo, “S” se refiere al sujeto de un verbo intransitivo y “O” se refiere al objeto de un verbo transitivo. El sistema nominativo-acusativo se ve a la izquierda, y el ergativo-absolutivo a la derecha del diagrama.



Debería notarse que no todos los sistemas casuales pueden clasificarse simplemente como “nominativo-acusativo” o “ergativo-absolutivo”. Hay muchos casos de idiomas que combinan aspectos de los dos sistemas, y otros en que no hay marcas casuales. También debería notarse que no siempre es fácil identificar un orden de palabras “neutro” en un idioma, ya que algunos permiten muchas opciones, aunque todas las partes de la oración tengan el mismo grado de énfasis. Guillaume (2006, p.16), por ejemplo, afirma que el cavineño tiene un “orden libre de constituyentes”, y que las relaciones gramaticales pueden inferirse solamente gracias a las marcas casuales, que siguen un sistema ergativo-absolutivo. Para practicar o para investigar más, el lector podría indagar cuáles otros órdenes de palabras y sistemas de caso existen en los idiomas bolivianos.

5.5. Conclusión

En este capítulo, hemos estudiado algunos conceptos básicos de la sintaxis: que se trata de la estructura jerárquica de las oraciones, además de su orden lineal, y que tenemos que utilizar pruebas de estructura de constituyentes para comprobar nuestras hipótesis sobre la naturaleza de estas estructuras. También hemos visto que no siempre es posible mantener una distinción clara entre la sintaxis y la morfología, especialmente en las lenguas indígenas americanas. Finalmente, hemos examinado dos rasgos sintácticos que varían mucho entre las lenguas de Bolivia.

Referencias

- Chomsky, N. (1957). *Syntactic Structures*. Mouton de Gruyter.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press.
- Guillaume, A. (2006). *A Grammar of Cavineña*. Mouton de Gruyter.
- Haude, K. (2006). *A Grammar of Movima*. Radboud Universiteit Nijmegen.
- Kayne, R. S. (1996). Microparametric Syntax. Some Introductory Remarks. Reprinted in Richard S. Kayne (2000). *Parameters and Universals* (pp. 3-39) Oxford University Press.
- Muysken, P. (1981). Quechua word structure. In H., Frank (ed.) *Binding and Filtering* (pp 279–329) Longmans.
- Myler, N. (2016). *Building and Interpreting Possession Sentences*. MIT Press.
- Pike, K. (1945). A Problem in Morphology Syntax Division. *Acta Linguistica*, 5, 125-136.
- Tallman, A. (2020). Beyond grammatical and phonological words. *Language and Linguistics Compass* 14:1-14.
- Vuillermet, M. (2012). A Grammar of Ese Ejja, a Takanan language of the Bolivian Amazon. Doctoral Dissertation: Université Lumière Lyon 2.
- Weber, D. (1983). The relationship of morphology and syntax: evidence from Quechua. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*. 161-181.

Cómo citar: Birchall, J. (2023). Lexicografía y elaboración de diccionarios. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 141-159), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11111121>

CAPÍTULO 6

LEXICOGRAFÍA Y ELABORACIÓN DE DICCIONARIOS

Joshua Birchall*

0. Introducción¹

Los diccionarios son uno de los principales componentes en la descripción de una lengua, junto con las colecciones textuales y las gramáticas descriptivas. La elaboración de un diccionario es una tarea monumental que se basa en la documentación del conocimiento léxico de los hablantes nativos de la lengua. Un diccionario puede cumplir una gran variedad de funciones, como herramienta en la educación, base de estudios científicos del léxico de una lengua o como referencia general para la comunidad lingüística. Un diccionario puede tener un alcance restringido y centrarse en una sola área de conocimiento léxico, tal como *La Ventosa Diidxazá Lexico-Botanical Dictionary* (Pérez Báez y Kaufman, 2019), o puede tener un alcance general y tratar de proporcionar una amplia gama de información lingüística en diferentes dominios semánticos, como el primer diccionario de la variedad cusqueña del quechua de González Holguín (1608).

Los diccionarios han existido durante muchos milenios siendo los primeros ejemplos conocidos los diccionarios bilingües de la lengua sumeria durante el Imperio acadio en Asia occidental, alrededor del 2600 a. C. (van Sterkenburg, 2003). Sin embargo, no fue hasta el siglo XX, con la publicación del *Manual de Lexicografía*

* Es profesor en la Universidad de Nuevo México. Tiene un doctorado en lingüística por la Universidad Radboud Nijmegen. Sus intereses se centran en la documentación, la descripción y la comparación de lenguas indígenas de Sud América. Está involucrado en una serie de iniciativas para desarrollar diccionarios multimedia y una enciclopedia cultural con comunidades que hablan lenguas en peligro de extinción en el suroeste de la Amazonía. Otras investigaciones cuyas en curso se relacionan con la lingüística histórica, como los métodos filogenéticos, la reconstrucción comparativa de la gramática y el léxico y el desarrollo de bases de datos léxicas.

1 Agradezco a Ester Torrico por la ayuda en traducir una versión preliminar de este artículo del inglés. También agradezco a Gladys Camacho y a los participantes del I Congreso Nacional de Lexicología y Lexicografía de Bolivia por sus comentarios.

de Zgusta (1971) que la lexicografía pasó a ser considerada generalmente como una subdisciplina dentro de la lingüística. En este sentido, Zgusta define el diccionario:

Un diccionario es una lista ordenada sistemáticamente de formas lingüísticas socializadas y compiladas a partir de los hábitos de habla de una comunidad de habla determinada y comentada por el autor de tal manera que el lector cualificado comprenda el significado... de cada forma por separado y es informado de los hechos pertinentes relativos a la función de esa forma en su comunidad. (Zgusta, 1971, p. 17)

Esta definición deja en claro que un diccionario es mucho más que una simple lista de palabras, es una recolección sistemática de formas lingüísticas combinada con información sobre los usos sociales, semánticos y gramaticales de estas formas. Una serie de informaciones y convenciones se han convertido en práctica habitual entre los lexicógrafos, y los recientes avances en la forma en que se llevan a cabo los estudios lingüísticos, así como el mayor acceso a la tecnología digital, han dado forma a la lexicografía en las últimas décadas.

Este capítulo tiene como objetivo proporcionar una descripción general de cómo se puede llevar a cabo un proyecto de diccionario para que sirva de guía práctica sobre cómo reunir y organizar los datos, y cómo estructurar una entrada léxica y el diccionario en general. La discusión aquí se centra en el diseño de diccionarios para comunidades lingüísticas en peligro de extinción, ya que esta es la situación actual de la mayoría de las lenguas indígenas de Bolivia. Los ejemplos se basan en la experiencia del autor al realizar un proyecto de diccionario multimedia en la lengua moré-kuyubim, en peligro de extinción, hablada en el departamento del Beni por la comunidad Moré, y también en otros trabajos lexicográficos en Bolivia.

6.1. Preparación para un proyecto de diccionario

Para comenzar un proyecto de diccionario, hay dos preguntas importantes que afectan al diseño de todo el proyecto: ¿quién utilizará el diccionario y cómo lo utilizará? Las respuestas a ambas preguntas afectan el diseño general del diccionario.

6.1.1 ¿Quién utilizará el diccionario?

Una de las primeras cosas que se debe tener en cuenta antes de preparar un diccionario es en qué lengua se va a escribir. Si la comunidad de la lengua meta tiene una tradición bien arraigada de escribir en su lengua y si los usuarios previstos del diccionario son hablantes nativos de esta, es habitual elaborar un **diccionario monolingüe** en dicha lengua. En este caso, la lengua que se describe, lo que podemos llamar **lengua meta**, es la misma que se utiliza para describir las formas lingüísticas, lo que podemos llamar **metalenguaje**. Los diccionarios monolingües

son especialmente comunes en los programas de educación nacional, pero no siempre son útiles para el usuario que aún intenta aprender la lengua. Por otro lado, si una comunidad todavía utiliza la lengua como medio de comunicación diario y solo tiene acceso a un **diccionario bilingüe**, un diccionario monolingüe puede ser un aporte bienvenido en la recolección de materiales disponibles sobre la lengua y puede ayudar a aumentar su conocimiento y prestigio.

En situaciones en las que los usuarios esperados de un diccionario no son hablantes nativos y utilizan el diccionario principalmente para aprender más sobre el léxico de la lengua meta, puede ser apropiado utilizar la variante estándar de comunicación diaria en la comunidad como el metalenguaje. En Bolivia, especialmente en las tierras bajas, muchas comunidades indígenas suelen utilizar una variedad local del castellano para la comunicación diaria y especialmente para la comunicación escrita, siendo la lengua indígena la más utilizada de forma hablada y/o por la generación de adultos mayores de la comunidad. Por ejemplo, en la comunidad Moré del departamento de Beni, esta generación es la que utiliza activamente la lengua indígena mientras que la nueva generación tiene un conocimiento pasivo de la lengua porque se comunica principalmente en castellano. Además, existe una variedad de la lengua que se habla tradicionalmente al otro lado del río Itenez/Guaporé por el grupo indígena Kuyubim. Para ayudarlos en el proceso de recuperación y revitalización de su lengua dentro de la comunidad, el diccionario de la lengua moré-kuyubim incluye tanto el castellano y el portugués como los metalenguajes.

Como se señala en Crevels (2012, p. 173), los datos del censo entre los grupos indígenas en Bolivia (y en otros lugares) a menudo no distinguen entre las personas que realmente utilizan la lengua en su vida cotidiana y las que desean aprenderla o se consideran a sí mismas como hablantes de la lengua sobre la base a la identidad étnica. Si el compilador del diccionario fuese o no miembro de la comunidad de habla, esto pone de manifiesto la necesidad de que el trabajo del diccionario se lleve a cabo con una estrecha colaboración de la propia comunidad para comprender mejor y atender a su situación sociolingüística específica y a sus necesidades educativas. Es importante señalar que los diccionarios pueden ser de utilidad no solo para los hablantes de la lengua y los activistas lingüísticos, sino que también pueden ser útiles para los educadores como base de materiales pedagógicos o para los miembros de la comunidad que deseen materiales de referencia sobre su lengua de herencia (Kroskrity, 2015, p. 140). El público al que va dirigido el diccionario y su uso previsto también afectan a todas las etapas del proceso de elaboración del diccionario, desde la decisión de qué documentar hasta la forma de organizar y presentar las entradas léxicas y el propio diccionario, como se expone en las siguientes secciones.

6.1.2. ¿Cómo utilizará la gente el diccionario?

Un componente adicional del proceso de elaboración del diccionario, que debe tenerse en cuenta al momento de planificar el proyecto del diccionario, es la forma en la que se utilizará y distribuirá principalmente entre la comunidad de usuarios. Los medios impresos tienen costos adicionales de producción y distribución, pero muchas comunidades de hablantes suelen atribuirles un valor adicional a los materiales impresos en comparación con los digitales. Los materiales impresos pueden distribuirse fácilmente en copias digitales y en forma de documentos electrónicos, como archivos PDF, lo que constituye una forma especialmente útil de recibir comentarios sobre la versión preliminar del diccionario. Sin embargo, una vez que se produce una versión final en forma impresa, suele ser difícil hacer revisiones a los materiales debido a los altos costos de reimpresión. Muchos editores comerciales no permiten la distribución gratuita de copias digitales o impresas de la versión final del diccionario una vez publicado, pero actualmente existen iniciativas en la publicación con acceso abierto que permiten la descarga gratuita de copias digitales de los diccionarios y una opción de impresión bajo demanda de pago para quienes prefieran una copia impresa. Las instituciones públicas como los museos y las universidades pueden permitir la publicación de materiales impresos a través de opciones de financiamiento no comercial para su distribución gratuita, especialmente para materiales destinados a la comunidad de usuarios, pero generalmente con un número más limitado de ejemplares.

Los diccionarios completamente digitales se están convirtiendo en una opción más frecuente entre los lexicógrafos y los usuarios de diccionarios. Especialmente en los casos en los que no hay una gran comunidad de hablantes que utilice la lengua a diario, la capacidad de incorporar la pronunciación de los lexemas, así como los diferentes ejemplos del uso de estos lexemas como los archivos multimedia pueden ser indispensables para los estudiantes de lenguas. Un beneficio adicional de los diccionarios digitales, a diferencia de los medios impresos, es que es fácil producir versiones revisadas, ya sea para ampliar el material original publicado o para revisar los registros. Sin embargo, es posible que no todos los miembros de la comunidad tengan acceso a la tecnología adecuada para utilizar el diccionario con regularidad. Los materiales con alto contenido multimedia almacenados en un sitio web o una aplicación a menudo requieren conexiones a internet rápidas y estables para su uso óptimo. Ya sea que estén establecidos en la web o en aplicaciones, los recursos completamente digitales tienen algunos inconvenientes adicionales, como la necesidad de actualizar la versión publicada para garantizar la compatibilidad con los estándares tecnológicos cambiantes y los diferentes sistemas operativos de las computadoras personales y de los dispositivos móviles. Mientras que un libro impreso puede seguir cumpliendo su función durante décadas, sin una actualización

periódica, un producto digital puede volverse funcionalmente obsoleto. Los lexicógrafos pueden explorar diferentes opciones de publicación que mejor se adapten a las necesidades de los usuarios. Como se expone en la siguiente sección, una metodología basada en la elaboración de un diccionario a partir de una base de datos léxica bien diseñada permite una mayor flexibilidad en las opciones de publicación y garantiza una mayor longevidad y utilidad de los datos lingüísticos que constituyen la base del propio diccionario.

6.2. Documentación léxica

Al comenzar la fase de documentación léxica de un proyecto de diccionario, se debe tener en cuenta una serie de preguntas: ¿qué materiales ya están disponibles?, ¿cómo se pueden recopilar los datos primarios? y ¿cómo organizar los datos una vez recopilados? La comprensión clara de estas preguntas permite a un lexicógrafo planificar y ejecutar adecuadamente la fase de documentación léxica de un proyecto de diccionario.

6.2.1. ¿Qué materiales ya están disponibles?

Cuando se planea desarrollar la documentación del léxico de una lengua, es importante estudiar los materiales existentes de esta. Dicha documentación existente puede abarcar desde notas de campo escritas a mano hasta gramáticas descriptivas publicadas y colecciones de documentación multimedia de acceso público, almacenadas en archivos profesionales como el Archivo de Lenguas en Peligro de Extinción (ELAR) o el Archivo de Lenguas Indígenas de América Latina (AILLA). Tener un buen conocimiento de los materiales disponibles puede ayudar a que la fase de documentación léxica de un proyecto de diccionario sea más eficiente al proporcionar una línea de base que ayuda a orientar las prioridades de documentación.

Una vez conocida la documentación disponible, la lexicografía debe considerar cómo se puede utilizar la documentación anterior para mejorar la calidad y la cobertura del propio diccionario, así como para facilitar el proceso de documentación léxica. En algunos casos, como en los materiales audiovisuales existentes, puede ser posible incorporar directamente al diccionario en desarrollo partes del trabajo de documentación anterior, dependiendo de los derechos intelectuales adscritos a estos materiales y de sus condiciones de uso, y, como siempre, citando y atribuyendo la autoría adecuadamente. En otro caso, puede ser más útil usar los materiales anteriores para orientar las nuevas sesiones de documentación. A continuación, se examinarán dos ejemplos diferentes de cómo los materiales lingüísticos existentes pueden guiar el proceso de documentación léxica.

Por ejemplo, cuando el autor comenzó el trabajo de campo con los Moré en 2017 para elaborar un diccionario multimedia, estaba familiarizado con aspectos generales de la cultura material y de la estructura social de los grupos vecinos de

Chapacura en Brasil, como los Wari' y los Oro Win, pero tenía poco conocimiento de los elementos culturales específicos de los Moré. Sin embargo, a través de una consulta cuidadosa a una serie de trabajos etnográficos, como en Ryden (1942) y Leigue Castedo (1957), y gracias a que este último incluye también una modesta lista de palabras, fue posible identificar aspectos específicos de la cultura y de la historia de los Moré que, aunque no se ven ni se discuten con frecuencia en el contexto moderno de la sociedad Moré, son importantes para el pueblo, para su historia y su forma de vida tradicional. Después de discutir sobre los materiales con varios colaboradores de Moré que estaban vivos durante el período de contacto con la sociedad boliviana en general, a principios y mediados del siglo XX, fue posible realizar entrevistas de elicitación para registrar estas palabras y los ejemplos de su uso (véase Birchall et al., 2022).

Otro ejemplo reciente del uso de materiales existentes para la lexicografía moderna puede verse en Danielsen et al. (2019) con la lengua guarayo. El equipo de investigación adaptó el diccionario publicado en Hoeller (1932) a la ortografía comunitaria actual y luego trabajó con hablantes jóvenes de la lengua para corregir cualquier error y grabar el audio de las diferentes entradas. El proyecto se publicó con notas históricas, culturales y lingüísticas adicionales en la revista en línea de acceso abierto *Dictionaria*.

6.2.2. ¿Cómo se puede recolectar los datos primarios?

Como se señala en Frawley et al. (2002, p. 1), el proceso de documentación léxica y la elaboración de diccionarios avanza “expandiéndose desde una modesta lista de palabras y glosas a algo parecido a una enciclopedia cultural”. Incluso el estudio más preliminar de una lengua suele implicar la recolección de elementos de vocabulario básicos, y es normalmente ahí dónde y cómo comienza el proceso de documentación léxica. A medida que el proceso de documentación avanza, es esencial que el lexicógrafo atienda a la comunidad de futuros usuarios construyendo una colección de entradas léxicas que refleje sus necesidades comunicativas. Para ello, es necesario comprender no solo el contexto físico en el que se utiliza la lengua, como la flora, la fauna y los fenómenos naturales típicos relevantes para la vida cotidiana de la comunidad, sino también el contexto social en el que se habla la lengua. Especialmente cuando el lexicógrafo no es miembro de la comunidad de habla, el conocimiento cultural e histórico de fondo es necesario para alcanzar una profunda y adecuada cobertura léxica que debe adquirirse a través de la observación participante, las entrevistas con miembros de la comunidad bien informados y el estudio cuidadoso del registro etnolingüístico existente.

La documentación del léxico puede utilizar diversas metodologías. Cuando el proyecto del diccionario se lleva a cabo junto con un proyecto de documentación

lingüística más amplio, o como una fase posterior de este, es posible que una parte sustancial del conocimiento léxico ya se haya documentado en un corpus transcrito de habla natural. Los casos de uso del lenguaje dentro de un corpus documental multimedia pueden utilizarse como ejemplos naturalistas del uso de elementos léxicos específicos. Sin embargo, la elaboración de un corpus de este tipo suele ser un proceso largo y prolongado que puede durar muchos años y, por lo tanto, puede no ser factible para proyectos cuyo objetivo principal sea la lexicografía. A menudo, es necesario emplear otros métodos para garantizar una cobertura adecuada del léxico y, a menudo, se utilizan diversas variaciones de las técnicas estándar de **elicitación**.

La **elicitación lingüística** es “la recolección de respuestas a estímulos lingüísticos o no lingüísticos diseñados para estudiar la competencia lingüística de los encuestados y/o sus prácticas de uso de la lengua” (Bohnenmeyer, 2014, p. 21). La práctica prototípica de elicitación es aquella en la que un investigador le pide a un hablante que proporcione una traducción cercana de una palabra o frase específica en la lengua meta que ha sido proporcionada por el investigador en una metalengua. Para fines lexicográficos, las entrevistas de elicitación organizadas en torno a dominios semánticos específicos a menudo producen los mejores resultados, en lugar de tratar de obtener palabras por orden alfabético o por parte del discurso. Si bien a menudo es una sociedad fructífera, especialmente para la documentación de lexemas específicos de forma aislada, basarse únicamente en los equivalentes de traducción puede limitar el alcance de la cobertura léxica, ya que el lexicógrafo suele limitarse a los conceptos presentes en el metalenguaje. También se debe tomar en cuenta los efectos que la estructura del metalenguaje ejerce sobre la lengua meta a través de la **imposición estructural**, el **calco semántico** y otros efectos de **contacto**, especialmente en la producción de enunciados de más de una sola palabra (véase en Lüpke, 2009 y Payne, 1997, pp. 366-371 para debates útiles sobre el papel de la elicitación y la colección de textos en el trabajo de campo lingüístico). La definición anterior de elicitación lingüística también incluye el método cada vez más común de utilizar estímulos visuales como imágenes o vídeos para obtener una respuesta lingüística de un hablante de la lengua. Como señala Majid (2012), el uso de estímulos visuales ayuda a limitar la influencia del metalenguaje en la respuesta lingüística obtenida. Una técnica adicional que es, en cierto modo, una hibridación, tanto de la elicitación por equivalente de traducción como de la elicitación por estímulos visuales, es la práctica de llevar a cabo la documentación lingüística a través de lo que Lüpke (2009) denomina “eventos comunicativos escenificados”, donde se pide a los hablantes que manifiesten su comportamiento lingüístico esperado como participantes en una situación específica. Todas estas diferentes técnicas pueden contribuir a enriquecer la documentación léxica disponible de una lengua y pueden complementar los conocimientos lingüísticos adquiridos mediante la colección de textos y la observación de los participantes.

Por razones prácticas, la documentación de una lengua que no ha sido bien documentada o descrita comienza inicialmente con una mayor dependencia de las técnicas tradicionales de elicitación y luego pasa a la recopilación de textos y a los métodos de observación participante, especialmente cuando el investigador no es miembro de la comunidad de habla. A medida que el investigador construye gradualmente un corpus de conocimiento léxico, es esencial que se sigan protocolos estándar en la documentación lingüística, como la recopilación oportuna de metadatos (o, mejor dicho, información sobre cómo se recolectó esa misma información lingüística), así como la transcripción de los datos a una forma ortográfica que permita una fácil búsqueda de los contenidos. Los metadatos a menudo se catalogan como archivos de texto u otros archivos de formato digital que acompañan a los archivos multimedia dentro de un corpus de documentación. Las transcripciones se suelen realizar con programas informáticos que permiten la alineación temporal entre los medios y el texto, como ELAN o SayMore. Una vez recogidos los datos y sometidos al tratamiento necesario para garantizar su utilidad, estos pueden organizarse en una base de datos léxica que puede utilizarse específicamente para la elaboración de un diccionario.

6.2.3. ¿Cómo organizar los datos una vez recolectados?

A medida que un lexicógrafo recopila un corpus cada vez más amplio de información léxica y ejemplos de uso, se recomienda organizar estos datos en algún tipo de base de datos. En los últimos años se han introducido una serie de buenas prácticas en el desarrollo de bases de datos (ver Capítulo 1) procedentes de disciplinas ajenas a la lingüística, como el desarrollo de software, las estadísticas y la administración de empresas, que pueden aplicarse fácilmente a las bases de datos lingüísticos. En este sentido, Wickham (2014) resume varias de estas prácticas bajo la denominación de principios de “datos ordenados”, es decir: (i) cada variable forma una columna, (ii) cada observación forma una fila y (iii) cada tipo de unidad de observación forma una tabla. En su forma más simple, podemos imaginar una base de datos léxica como una hoja de cálculo y podemos interpretar estos principios en el sentido de que cada fila en la tabla de la hoja de cálculo corresponde a un lexema concreto (una “observación”) y que cada columna de la tabla corresponde a una propiedad específica de ese lexema (una “variable”), como parte del discurso, la clase semántica, la transcripción fonética, la transcripción ortográfica, etc. (véase la sección 6.3., más adelante, para examinar la información que suele asociarse a una entrada léxica). Es importante que cada observación de una tabla tenga su propio número de identificación para que pueda ser referenciada en otras secciones de la base de datos. Para incluir ejemplos de uso vinculados a lexemas específicos, es posible hacer referencia a su número de identificación en una tabla separada donde se recopilan y seleccionan los datos textuales, o bien reproducir los propios ejemplos

dentro de la tabla de lexemas. Cuando se elabora un diccionario digital en lugar de uno impreso, casi siempre se utiliza una base de datos para elaborar el producto final. Bergenholtz y Nielsen (2013) ofrecen comentarios adicionales sobre el diseño de bases de datos lexicográficas que pueden dar cuenta de relaciones más complejas entre formas y sus significados, especialmente los casos de sinonimia (múltiples formas con el mismo significado) y polisemia (una única forma con múltiples significados).

Existen diversos programas para lingüistas de campo que permiten la elaboración de bases de datos de información textual y léxica. Un método común es el uso de software de trabajo de campo producido por el Instituto Lingüístico de Verano (ILV/SIL), originalmente Shoebox, que luego fue reemplazado por Toolbox, y ambos han sido sustituidos por Fieldworks Language Explorer, comúnmente conocido como FLEx. Estos programas cuentan con una serie de herramientas que permiten la recopilación de bases de datos de información léxica derivada directamente de los datos textuales y, además, cuentan con una serie de funciones que proporcionan glosas semiautomáticas de textos recién agregados en función de la información suministrada con anterioridad. Especialmente en el caso de FLEx, algunos investigadores han tenido dificultades para exportar datos léxicos del programa de forma que puedan utilizarse fácilmente para otras aplicaciones.

A la hora de diseñar una base de datos léxica, es importante sopesar las ventajas adicionales de utilizar un software de terceros frente a las desventajas de no disponer de los datos en un formato textual fácilmente transportable y abierto, como los valores separados por comas (.csv) o los valores separados por tabulaciones (.tsv) que se utilizan habitualmente en las disciplinas digitales, aunque algunos software de terceros permiten exportar los datos en dicho formato. Otra opción para almacenar sus datos léxicos es utilizar el Cross-linguistic Data Formats (Forkel et al., 2018), un formato para guardar diferentes tipos de datos lingüísticos diseñados para garantizar la compatibilidad entre herramientas, plataformas y programas existentes. El CLDF tiene una serie de principios de diseño que respaldan esta multifuncionalidad: todos los datos deben almacenarse como archivos de texto con codificación UTF-8 (unicode), los conjuntos de datos deben ser editables a mano en un software de hoja de cálculo comúnmente disponible y las entidades, a las que se hace referencia, dentro de un archivo de datos deben estar vinculadas a catálogos de referencia como Glottolog (Hammarström et al., 2021) para lenguas/dialectos o como Concepticon (List et al., 2021) para conceptos semánticos. Este recurso requiere de una conservación manual adicional de los datos en comparación con el software disponible en el mercado, pero tiene la ventaja adicional de poder modificar fácilmente los datos y utilizarlos para otros fines.

Pasar directamente de los datos primarios recogidos a un formato de diccionario publicable sin una base de datos intermedia puede parecer un paso que ahorra tiempo

en un principio, pero el uso de una base de datos tiene una serie de ventajas. Una de las principales ventajas de un enfoque de base de datos para la lexicografía es una cuestión de revisión y organización. Disponer de todas las entradas léxicas en el mismo formato de búsqueda permite identificar fácilmente las entradas reduplicadas o, en los casos en los que no existe un consenso comunitario generalizado sobre las convenciones ortográficas, las múltiples instancias del mismo lexema transcritas de diferentes maneras y/o con definiciones o traducciones similares, pero no idénticas. La aplicación de diferentes filtros a los datos tabulares en el software de hoja de cálculo habituales, permite una rápida inspección visual de los datos. Una ventaja adicional de un enfoque de base de datos para la lexicografía es la posibilidad de realizar modificaciones globales en la propia base de datos o en cualquier producto derivado de la misma. Independientemente de la estructura de la base de datos utilizada, es de suma importancia que la información almacenada en ella se conserve en un formato de archivo que garantice la longevidad de los datos.

Un último aspecto sobre la organización de datos primarios para un proyecto de lexicografía es qué hacer con los archivos multimedia que acompañarán a un diccionario digital. Lo más importante es que los archivos multimedia, ya sean de audio o video, tengan una convención de nomenclatura uniforme que se les aplique. Una convención de nomenclatura debe permitir la identificación inmediata del contenido de los archivos, siendo cada nombre de archivo único e independiente de la estructura de carpetas más amplia en la que se almacenan los archivos. Tener un nombre único para cada archivo permite al lexicógrafo referirse a cada archivo específico dentro de la base de datos lexicográfica más amplia. En los casos en los que un archivo multimedia se extrae de un archivo multimedia mayor, como una ficha única de una entrevista de elicitación más amplia, se puede incluir información en el nombre del archivo extraído que permita identificar el archivo de origen o mantener una tabla adicional con los metadatos de cada archivo extraído que incluya su procedencia.

6.3. Estructura de una entrada léxica

Una vez que se ha documentado la información léxica que se incluirá en un diccionario y se han organizado los datos, la siguiente tarea es organizar las entradas reales del diccionario. Esto lleva a la pregunta obvia: ¿qué incluir en una entrada de diccionario?

El contenido de una entrada de un diccionario puede variar en función de los objetivos generales del proyecto y del público al que va dirigido. Garrett (2018, pp. 197-198), cuando menciona los objetivos de un diccionario para apoyar los esfuerzos de revitalización de la lengua, observa que “un diccionario tiene que ayudar a la gente a usar la lengua en peligro de extinción (EL)”. Por ejemplo, cuando varias

palabras de la lengua en peligro de extinción (EL) tienen significados similares o traducciones idénticas en la lengua materna, un diccionario para la revitalización de la lengua debe ayudar a los usuarios a comprender las implicaciones de sus elecciones de palabras. ¿Qué palabra es apropiada en distintos contextos de uso? ¿En qué se diferencian desde el punto de vista semántico, pragmático y sintáctico? Los estudiantes que se basan en un diccionario que carece de detalles sobre estos parámetros de uso pueden tender a combinar el vocabulario de la lengua en peligro de extinción (EL) con patrones sintácticos y semánticos de la lengua materna, dando lugar a una lengua materna “relexificada”.

Aunque un lexicógrafo puede tomar diferentes decisiones sobre cómo organizar y presentar los datos lingüísticos dependiendo de si los usuarios previstos normalmente utilizarán el diccionario para el aprendizaje de idiomas en lugar de material de referencia general, existen componentes que son estándar en la lexicografía moderna que se encontrarán en la mayoría de las entradas del diccionario, como el **lema** identificado, una **pronunciación**, una **definición**, la **parte de la oración**, **información gramatical** y **ejemplos de uso**. Muchos diccionarios también optan por incluir **información cultural o etimológica** adicional cuando está disponible.

El **lema** de una entrada de diccionario, a menudo llamado también **encabezado o término**, es la forma canónica del lexema a la que se asignarán todas las demás variantes morfológicas de la palabra. Otra forma de describir un lema es como “una abstracción que subsume otras formas relacionadas tanto morfológica como semánticamente” (Frawley et al., 2002, p. 3). Por ejemplo, en la entrada de la palabra en inglés *break*, se debe incluir también la forma de pasado *broke* y la forma del participio *broken*. Bajo esta misma entrada cubierta por el lema *break*, también se pueden incluir formas relacionadas adicionales, como desglosar y romper. Estas formas relacionadas ayudan a los usuarios del diccionario a encontrar fácilmente la información que buscan.

La estructura morfológica de la lengua que se describe y las convenciones comúnmente usadas por la comunidad de hablantes deberían ayudar a determinar la selección de un lema para una entrada. Independientemente de las elecciones que se realicen, que se puede considerar la “ideología del lenguaje” del diccionario, lo más importante para los usuarios del diccionario es que estas opciones se apliquen de forma coherente en todas las entradas, y que estas opciones se muestren explícitamente en la portada del diccionario (Rhodes et al., 2018). Es habitual que las diferentes formas flexionadas de un mismo lexema se organicen bajo el mismo lema, mientras que los lexicógrafos varían en cuanto a si las formas derivadas deben organizarse bajo el mismo lema. Otro tema por considerar es si el propio lema debe llevar alguna inflexión en su forma presentada al inicio de una entrada. Por un lado, una **inflexión** en la forma del prefijo haría que todas las entradas de la misma parte de la oración comenzaran con la misma letra. Por otro lado, la ausencia total de inflexión en un

lema podría dar lugar a una forma que no se da de manera natural en la lengua, lo que dificultaría que el usuario localice la información que busca.

Debajo del encabezado de un lema, la entrada debe incluir su pronunciación e información gramatical adicional, como parte de la oración. La pronunciación se presenta generalmente como una aproximación fonética estandarizada utilizando el Alfabeto Fonético Internacional. Algunas tradiciones lexicográficas de lenguas sin ortografía fonémica como el inglés pueden adoptar convenciones diferentes, y las lenguas con una ortografía fonémica transparente, pero con una considerable variación dialectal en la pronunciación, pueden optar por no incluir ninguna pronunciación, como en el *Diccionario de la lengua española* (Real Academia Española, 2001). Las categorías gramaticales específicas de la lengua suelen incluirse en una entrada del diccionario. Las categorías generales de partes de la oración, como sustantivo, verbo o adjetivo, son las más comunes, pero también se incluyen categorías específicas de la lengua, como gerundios o pronombres interrogativos, así como otra información gramatical relevante para la entrada concreta de la lengua, como la clase/género de los sustantivos o la clase flexiva de los verbos. Algunos diccionarios también incluyen los paradigmas de las formas flexionadas, especialmente cuando existen patrones múltiples o irregulares en la lengua.

Las entradas del diccionario siempre incluyen una definición del lema y sus significados asociados. Es una práctica habitual definir una palabra sin utilizar la propia palabra para evitar la redundancia. Como señala Geeraerts (2003), una palabra puede tener múltiples significados o sentidos más allá de su denotación, es decir, el tipo de elementos en el entorno a los que se refiere la palabra, como su significado emotivo, su significado gramatical y su significado pragmático. Estos diferentes significados de una palabra se basan en gran medida en el uso que hacen los hablantes de la misma y es la labor del lexicógrafo describir estos diferentes usos. En los diccionarios bilingües, a veces se recurre a la traducción directa cuando existen conceptos correspondientes entre la lengua meta y el metalenguaje. Sin embargo, puede ser problemático suponer que, incluso en los casos en que existe una traducción directa entre las dos lenguas, el uso de los términos es idéntico y el lexicógrafo debe presentar los diferentes sentidos en los que se utiliza la palabra. Tanto en los diccionarios monolingües como en los bilingües, también se pueden incluir palabras relacionadas, como **sinónimos** y **antónimos**, o **expresiones idiomáticas** específicas de la cultura que hacen uso de la palabra.

Varios diccionarios incluirán ejemplos de uso del lema descrito en sus diferentes acepciones. Como se destaca en la cita de Garrett (2018) al comienzo de esta sección, el objetivo de los ejemplos es ayudar a guiar al usuario del diccionario en el uso del término. Puede ser útil para el usuario incluir ejemplos que resalten otras palabras asociadas con la entrada. Por ejemplo, en una entrada del diccionario

sobre la palabra ‘miel’, sería útil que el usuario tuviera un ejemplo que incluyera el verbo apropiado para el consumo de miel, como por ejemplo si la miel se “come”, se “toma”, se “lame”, se “bebe” o de otra manera (todos estos son verbos correspondientes al consumo de miel que el autor ha encontrado en su trabajo de campo con diferentes lenguas en la Amazonía). Además de los ejemplos de uso de un elemento léxico, muchos lexicógrafos incluyen imágenes del referente cuando es posible. Si el diccionario es multimedia, también se pueden incluir videos o audios que proporcionen un ejemplo del uso del término o un ejemplo del uso del referente del término en su contexto apropiado.

Un último dato que suele incluirse en una entrada de diccionario es lo que se sabe sobre la etimología de un término en particular. En las lenguas con una rica tradición filológica, la historia de una palabra puede ser bien entendida y atestiguada a través de documentos históricos de diferentes épocas, así como en reconstrucciones propuestas para el término en protolenguas anteriores. Sin embargo, en el caso de las lenguas que carecen de tal tradición, es posible incluir simplemente términos de otras lenguas que puedan identificarse como afines.

6.4. Estructura del diccionario

Ya hemos mencionado sobre la planificación de un proyecto de diccionario, de la recopilación de datos y la organización de esa información en entradas, pero ahora pasamos a dos preguntas sobre la estructura general del diccionario: ¿cómo se debe representar ortográficamente la lengua meta? y ¿cómo puede organizarse el diccionario?

La pregunta sobre qué ortografía utilizar es a menudo objeto de un acalorado debate en el contexto de la lexicografía de las lenguas en peligro de extinción. En comunidades lingüísticas con una larga tradición de convenciones de escritura, la tarea es bastante sencilla siempre que la ortografía represente suficientemente los sonidos contrastantes relevantes de la lengua. Sin embargo, un diccionario es a veces la primera obra escrita importante en algunas lenguas y puede tener un considerable efecto normalizador sobre la forma escrita de la lengua en su uso futuro. Cuando una lengua no tiene una ortografía establecida y ampliamente adoptada, lo más probable es que los usuarios del diccionario ya estén escribiendo en otra lengua. El diseño de la ortografía plantea dos cuestiones: ¿qué símbolos utilizar? y ¿qué símbolos representan qué sonidos en la lengua?

La mayoría de los lingüistas prefieren utilizar símbolos fonéticos cuando trabajan con una lengua no escrita, sobre todo porque estos símbolos tienen la capacidad de representar aspectos de la lengua que no suelen recogerse en las ortografías basadas en caracteres latinos, como la longitud, la aspiración y el tono.

Sin embargo, desde el punto de vista del usuario, el uso de símbolos puramente fonéticos como la ortografía tiene una serie de inconvenientes, como la falta general de familiaridad en la forma de pronunciar algunos de estos símbolos, así como la dificultad general para escribir esta lengua en una computadora o teléfono sin teclado especial. La mayoría de los lexicógrafos adoptan una ortografía práctica que utiliza el mismo conjunto de símbolos empleados en el idioma oficial, y muchos optan por complementarla con una fonética o fonémica más detallada como parte de la entrada estándar (véase § 6.4.).

Suponiendo que se adopte el mismo conjunto de símbolos que se utilizan en el idioma oficial como base para una ortografía práctica utilizada como medio principal para presentar formas escritas de los lexemas, se plantea entonces la cuestión de cómo representar los sonidos de la lengua meta que no están presentes en el idioma oficial. Los lexicógrafos han adoptado una serie de convenciones, como la representación de las oclusiones glotales o las eyectivas con un apóstrofo y el uso de acentos para representar los contrastes tonales. Independientemente de las elecciones que se realicen, es importante que un lexicógrafo, ya sea de la propia comunidad de habla o no, trabaje en estrecha colaboración con otros miembros de la comunidad que tengan un interés personal en la adopción de la ortografía, como educadores y otros líderes de la comunidad. Las implicaciones de cada elección deben presentarse y debatirse con el objetivo de alcanzar un consenso. En Hinton (2014), se discuten una serie de cuestiones relacionadas con el desarrollo de ortografías comunitarias para lenguas no escritas, donde se destaca que las ortografías pueden tener considerables implicaciones políticas, y es una de las implementaciones más obvias de la ideología lingüística dentro de un diccionario. Hinton identifica una serie de factores importantes que deben considerarse cuidadosamente, el más importante es que la ortografía sea considerada aceptable por la comunidad lingüística. Otros factores incluyen la capacidad de representar con precisión todos los sonidos distintivos de la lengua, la facilidad con la que se puede aprender y la sencillez con la que se puede reproducir en los medios.

La última pregunta por considerar es cómo organizar las entradas de un diccionario. El enfoque más común es presentar las entradas en orden alfabético. Sin embargo, como se discute en Frawley et al. (2002, p. 9), se deben tomar ciertas decisiones a la hora de alfabetizar las entradas dependiendo de los símbolos adoptados en la ortografía, como por ejemplo si las entradas que comienzan con el símbolo *p* irían antes o después de las que tienen *p*, o si un dígrafo como *ch* que representa un solo sonido debería ir antes de las entradas que comienzan con *ci* o bien deberían formar su propia sección del diccionario. Otro enfoque consiste en organizar el diccionario temáticamente en torno a diferentes nociones semánticas

o gramaticales, como los animales o los verbos, con las entradas posteriores organizadas alfabéticamente dentro de cada sección. Esta organización puede ser especialmente útil para los usuarios del diccionario que se dedican principalmente al aprendizaje de lenguas, y es más fácil de aplicar cuando cada lema tiende a asociarse con un solo significado.

Una consideración adicional para la organización de las entradas es si estas deben organizarse según la lengua meta o, en el caso de los diccionarios bilingües, según el metalenguaje. En el caso de los diccionarios multimedia, es habitual permitir una opción para ambas organizaciones en función de la configuración seleccionada por el usuario, pero, en el caso de los diccionarios publicados por escrito, esto puede resultar más complicado. Los usuarios de diccionarios que se preocupan principalmente por aprender el idioma pueden encontrar beneficioso buscar palabras utilizando el metalenguaje. Para apoyar a este tipo de usuarios, una opción razonable puede ser tener la estructura principal del diccionario organizada por la lengua meta con un glosario adicional al final, que está organizado en el metalenguaje. Como ocurre con la mayoría de las opciones en lexicografía, son las necesidades de los usuarios las que deben impulsar el diseño y la implementación del diccionario.

6.5. Lexicografía en el contexto boliviano

Hay una larga tradición de lexicografía con las lenguas indígenas de Bolivia incluso antes de la fundación del Estado Plurinacional de Bolivia. Por ejemplo, uno de los primeros estudios lexicográficos con una lengua de América del Sur fue realizado por el padre jesuítico italiano Ludovico Bertonio con hablantes de la lengua aimara que vivían al sur del lago Titicaca, a principios del siglo XVII (Bertonio, 1612). La tradición de lexicografía aimara continúa hasta hoy en universidades bolivianas.

En las tierras bajas bolivianas, también hay una larga tradición lexicográfica desde el periodo colonial, como, por ejemplo, el *Arte de la lengua Moxa con su vocabulario y cathecismo*, publicada originalmente por el jesuita Pedro Marban a principios del siglo XVIII (Marban, 1894 [1701]). Según Adelaar (2012), es posible que muchos otros trabajos lexicográficos y gramaticales hayan sido producidos durante el periodo colonial, como las “artes” y “tesoros” de los jesuitas, pero que se han perdido por solo ser publicados como manuscritos o que aún están por descubrir en los archivos de la Iglesia católica en América Latina o Europa.

En el siglo XX se retomó el interés sobre la descripción de las lenguas indígenas de Bolivia y sus vocabularios, primeramente por investigadores relacionados a diferentes institutos religiosos, como el diccionario de Guayaros

por el franciscano Hoeller (1932), y otros del Instituto Lingüístico del Verano, como el trabajo de Van Wynen y Van Wynen (1962) sobre la lengua tacana.

En las últimas décadas, varios investigadores académicos han avanzado en la descripción lexicográfica de las lenguas bolivianas a partir de la producción de diccionarios y vocabularios. Muchos de estos trabajos sirven como ejemplos de la implementación de lexicografía moderna mediante el uso de datos multimedia para producir diccionarios digitales. El diccionario de Guarayu de Danielsen et al. (2019) mencionado arriba es un buen ejemplo de cómo se puede utilizar materiales lexicográficos antiguos para producir diccionarios modernos. Otro elemento importante de lexicografía moderna es que los investigadores indígenas han tomado un papel central en la documentación léxica y la elaboración de diccionarios de sus lenguas. Por ejemplo, el diccionario *Ese Ejja Ewowa Ja'a Pokiajji* fue producido recientemente por el lingüista Alejandro Machuqui Dominguez, también conocido como Bawapoji, miembro de la comunidad Ese Ejja (Machuqui Dominguez, 2021). Otro ejemplo es el diccionario *Moré-Kuyubim* producido por el autor de este texto, miembros de las comunidades Moré y Kuyubim y otros colaboradores lingüistas y antropólogos (Birchall et al., 2022). Estos dos últimos diccionarios también tienen versiones multimedia que están disponibles como aplicativos Android. Así se puede ver que Bolivia es un líder de lexicografía moderna en América Latina, pero, dado el estado de riesgo de sus lenguas indígenas, aún queda mucho trabajo lexicográfico por hacer.

6.6. Conclusión

Como se mencionó a lo largo de este capítulo, es importante que se tome una decisión en la elaboración de un diccionario, ya sea en la selección de la forma morfológica de un lema, la ortografía adoptada o en su organización general, y que esta elección esté motivada por las necesidades de los usuarios y que se aplique de forma coherente en todo el trabajo. El lugar apropiado para presentar estas opciones es la introducción del diccionario. Aquí, el lexicógrafo puede proporcionar información adicional a los usuarios para guiarlos en el uso del diccionario, con la presentación de las motivaciones por las que se adoptaron dichas elecciones. En otras palabras, es una oportunidad para que el lexicógrafo presente la ideología lingüística que subyace en la producción del diccionario (véase Kroskrity, 2004 y Rhodes et al., 2018 para discusiones útiles sobre el papel de la ideología lingüística en el desarrollo del diccionario). La introducción también es una oportunidad para proporcionar más información sobre la lengua, como sus sonidos y su gramática, o información sobre la historia y la cultura de las personas que la hablan, así como para reconocer a todos los colaboradores que han contribuido con el diccionario.

Además, en este capítulo, se han abordado una serie de cuestiones relacionadas con la lexicografía, especialmente en lo que respecta a la elaboración de diccionarios para lenguas en peligro de extinción en el contexto boliviano. Existe una necesidad urgente de aumentar el estudio de los léxicos de todas las lenguas, especialmente las de Bolivia. Una de las mejores maneras de hacerlo es mediante la capacitación de los miembros de las comunidades indígenas en los métodos y la teoría necesarios para documentar y describir sus propias lenguas. Un diccionario es una poderosa herramienta para tomar los conocimientos adquiridos a través de la documentación lingüística y la lexicografía y ponerlos en manos de sus usuarios.

Referencias

- Adelaar, W. F. H. (2012). Historical overview: Descriptive and comparative research on South American Indian languages. En L. Campbell y V. Grondona (Eds.), *The indigenous language of South America: a comprehensive guide* (pp. 1-57). De Gruyter Mouton.
- Bergenholtz, H. & Nielsen, J. S. (2013). What is a lexicographical database? *Lexikos*, 23, 77-87.
- Bertonio, L. (1612). *Vocabulario de la lengua aymara*. Francisco del Canto.
- Birchall, J., Chyc, P., Leigue Sae, J. B., Cujubim, M., Costa, C. N. d. (2022). *Dicionário Multimídia Moré-Kuyubim. Versión 1.0*. Museu do Índio. <http://japiim.museudoindio.gov.br/dic/morekuyubim/>.
- Bohnemeyer, J. (2014). A practical epistemology for semantic elicitation in the field and elsewhere. En R. Bochnak y L. Matthewson (Eds.), *Methodologies in semantic fieldwork* (pp. 13-46). Oxford University Press.
- Crevels, M. (2012). Language endangerment in South America: The clock is ticking. En L. Campbell y V. Grondona (Eds.), *The indigenous language of South America: a comprehensive guide* (pp. 167-233). De Gruyter Mouton.
- Danielsen, S., Sell, L. y Terhart, L. (2019). Guarayu. A revised dictionary by Alfred Hoeller. *Dictionaria*, 7, 1-3590. <https://dictionaria.clld.org/contributions/guarayu>.
- Forkel, R., List, J.M., Greenhill, S., Rzymski, C., Bank, S., Cysouw, M., Hammarström, H., Haspelmath, M., Kaiping, G. A. y Gray, R. D. (2018). Cross-Linguistic Data Formats, advancing data sharing and re-use in comparative linguistics. *Scientific Data*, 5, 180205 (2018). <https://doi.org/10.1038/sdata.2018.205>

- Frawley, W., Hill, K. C., y Munro, P. (2002). Making a dictionary: Ten issues. En W. Frawley, K. C. Hill y P. Munro (Eds.), *Making dictionaries: preserving indigenous languages of the Americas* (pp. 1–22). University of California Press.
- Garrett, A. (2018). Online dictionaries for language revitalization. En L. Hinton, L. Huss, y G. Roche (Eds.), *The Routledge handbook of language revitalization* (pp. 197–206). Routledge.
- Geeraerts, D. (2003). Meaning and definition. En P. van Sterkenburg (Ed.), *A practical guide to lexicography* (pp. 83–93). John Benjamins Publishing Company.
- Gonzalez Holguín, D. (1608). *Vocabulario dela lengua general de todo el Peru llamada lengua Quichua, o del Inca: corregido y renovado conforme ala propiedad cortesana del Cuzco*. Francisco del Canto.
- Hammarström, H., Forkel, R., Haspelmath, M. y Bank, S. (2021). *Glottolog 4.4*. Max Planck Institute for Evolutionary Anthropology. <https://glottolog.org>.
- Hinton, L. (2014). Orthography wars. En M. Cahill y K. Rice (Eds.), *Developing orthographies for unwritten languages* (pp. 139–168). SIL International.
- Hoeller, A. (1932). *Guarayo-Deutsches Wörterbuch*. Verlag der Missionsprokura der P. P. Franziskaner.
- Kroskrity, P. V. (2004). Language ideologies. En A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 496–517). Blackwell Publishing.
- Kroskrity, P. V. (2015). Designing a dictionary for an endangered language community: lexicographical deliberations, language ideological clarifications. *Language Documentation & Conservation*, 15, 140–157.
- Leigue Castedo, L. (1957). *El Itenez Salvaje*. Ministerio de Educación.
- List, J. M., Rzymiski, C., Greenhill, S., Schweikhard, N., & Panykh, K., Tjuka, A., Hundt, C. y Forkel, R. (Eds.). (2021). CLLD Concepticon 2.5.0 [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.4911605>
- Lüpke, F. (2009). Data Collection Methods for Field-Based Language Documentation. *Language Documentation and Description*, 6, 53–100.
- Machuqui Dominguez, A. (2021). *Ese Ejjaja Esowijo Ewowi Ja'a Pokiajji: Diccionario Ese Ejja – Español, Español – Ese Ejja*. Foundation for Endangered Languages.

- Majid, A. (2012). A guide to stimulus-based elicitation for semantic categories. En N. Thieberger (Ed.), *The Oxford handbook of linguistic fieldwork* (pp. 54-71). Oxford University Press.
- Marban, P. (1894 [1701]). *Arte de la lengua Moxa con su vocabulario y cathecismo*. B. G. Teubner.
- Payne, T. (1997). *Describing morphosyntax: a guide for field linguists*. Cambridge University Press.
- Peréz Báez, G. y Kaufman, T. (2019). La Ventosa Diidxazá Lexico-Botanical Dictionary. *Dictionaria*, 9, 1-952. <https://dictionaria.clld.org/contributions/diidxaza>.
- Real Academia Española. (2001). *Diccionario de la lengua española*. 22a edición. Editorial Espasa Calpe.
- Reiman, D. W. (2010). Basic oral language documentation. *Language Documentation & Conservation*, 4, 254-268.
- Rice, K. (2018). Reflections on documenting the lexicon. En McDonnell, B., Berez-Kroeker, A. L. y Holton, G. (Eds.), *Reflections on language documentation 20 years after Himmelmann 1998* (pp. 122-131). University of Hawai'i Press.
- van Strekenburg, P. (2003). 'The' dictionary: definition and history. En P. van Strekenburg (Ed.), *A practical guide to lexicography* (pp. 3-17). John Benjamins Publishing Company.

CAPÍTULO 7

Cómo citar: Stewart, J. (2023). Contacto de lenguas y variación. En P. Alandia Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 160-179), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11111143>

CONTACTO DE LENGUAS Y VARIACIÓN

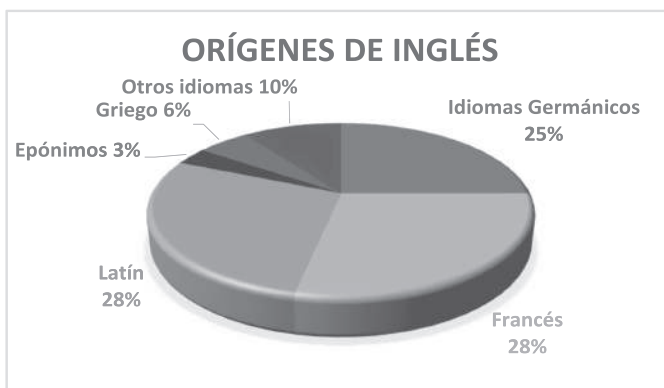
Jesse Stewart*

0. Introducción

Uno de los conceptos fundamentales de la lingüística es la **paridad** (*parity*), la cual explica que no hay un idioma superior a otro y que no existe ningún “idioma puro”. Todos los idiomas documentados en el mundo tienen influencias de otros. Por ejemplo, la mayoría del vocabulario del inglés proviene de cuatro familias de idiomas

Gráfico 1

Orígenes léxicos del inglés



(Gráfico 1); sin embargo, tiene **préstamos** (*loanwords*) de muchos otros, incluso de idiomas hablados en Bolivia; por ejemplo, del quechua provienen las siguientes palabras en inglés: *condor*, *quinoa*, *coca*, *jerky*, *llama*, *pisco*, *puma*, *guano*, *Inca*, *ayahuasca*; del aymara: *alpaca*; y del guaraní: *jaguar*, *cashew*, *macaw*, *piranha*.

La mayoría de estas palabras llegaron al inglés a través del castellano, que también, junto con el portugués, tiene préstamos innumerables de cientos de idiomas indígenas de las Américas. Además, el castellano tiene más de 4000 palabras con origen árabe, que se incluyeron como préstamos durante la conquista árabe-

* Es Profesor Asociado del Departamento de Lingüística de la Universidad de Saskatchewan, Canadá. Tiene doctorado en Lingüística por la Universidad de Manitoba, Winnipeg, MB, Canadá. Ha colaborado con los hablantes de la comunidad de Pijal en la documentación de su lengua, compilando el primer diccionario lingüístico multimedia, entre otros documentos. Sus áreas principales de investigación son fonética, fonología, lenguas mixtas y lenguas en contacto.

musulmana en la Península Ibérica durante más de novecientos años. Algunas de las palabras de origen árabe comunes en castellano son: *almuerzo*, *almohada*, *ajedrez*, *atún*, *bagre*, *café*, *chivo*, *espinaca*, *fideo*, *limón*, *loco*, *máscara*, *quintal*, *talco*, *taza*, *zanahoria*, entre muchas más. Con base en estos ejemplos, se puede definir un **préstamo** como una palabra o parte de una palabra que se comparte de un idioma a otro, y la única manera en la que esto es posible es a través del contacto con hablantes de otros idiomas, ya sea en persona o por documentación (escrita, audio, video, etc.).

Con esas consideraciones, en este capítulo veremos diferentes fenómenos que resultan del contacto de lenguas, desde los préstamos, con sus diferentes efectos en los sistemas en contacto, hasta la emergencia de nuevas lenguas, cuya clasificación presentamos.

7.1. Resultados del contacto de lenguas

El lenguaje es una herramienta dinámica que siempre está evolucionando para encajar con las necesidades comunicativas de sus usuarios. Una de las maneras más comunes para describir nuevos conceptos, tecnologías y cosmovisiones introducidas por otra cultura es simplemente adaptar su vocabulario; sin embargo, los puristas frecuentemente ven los préstamos como degradación o corrupción del idioma.

En realidad, la adaptación de palabras extranjeras es un fenómeno natural utilizado en todos los idiomas del mundo y puede ser usado como marcador histórico de las interacciones interculturales entre humanos, que pueden revelar el tipo y la intensidad del contacto; por ejemplo, cuando los sirvientes ingleses entregaban la comida a sus patrones franceses durante la conquista normanda de Inglaterra, mataban el *pig*, pero servían el *pork* (del francés *porc*). Es por eso que el inglés tiene nombres nativos para los animales y nombres derivados del francés para su comida, respectivamente: *sheep-mutton* (del francés *mouton*); *cow-veal/beef* (del francés *veau/beouf*). Sin embargo, con el tiempo, los préstamos pueden ser difíciles de identificar; normalmente la forma fonológica de la palabra prestada se asimila a la fonología del idioma prestatario y, sin un estudio **etimológico** (*etymological*), esta pasa inadvertida. Un hablante de castellano no pronuncia *arroz* como [ˈʔarz] en árabe رَزْ, en su lugar dice [aˈros], basado en la fonotáctica castellana. Un hablante de japonés no dice [ˈfʷʊʔʔbaɪ] o [ˈfutbal], dice [ɸʊʔʔtɔboːɾʊʔ] para el *fútbol*.

- 1) En los siguientes ejemplos de idiomas hablados en Bolivia, ¿cuántos préstamos del castellano se pueden identificar?
- 2) ¿Se puede identificar alguna tendencia en los cambios en cada idioma?

Tabla 1*Ejemplos de préstamos del castellano en las lenguas bolivianas*

Idioma	Frase	Referencia
Chácobo	naráha poró-wa bari=´ wa=ki	Tallman, 2018, p. 672
Tapieté	kamĩ'õ ou o-a he(se) siyora	González, 2005, p.308
Tapieté	hana she-ru papere kwa-'ä kwa-ä mbi-para	González, 2005, p.381
Baure	kowapa teç rotir howoki.	Danielsen, 2007, p.86
Yuracaré	poybolo=chi bati tẽ-ta-ø poropesor	Van Gijn, 2006, p.327
Bésiro	Ané chaama, xepénte chaama siro kítu naróx, o noseóx.	Sans, 2013, p.63

Además, es necesario señalar que la adaptación del vocabulario extranjero no es la única manera en que las personas acomodan nuevos conceptos en sus idiomas; dos ejemplos comunes incluyen **neologismos** (*neologisms*) y **calcos** (*calques*). Un neologismo es la creación de una palabra nueva o una nueva definición para una palabra existente; una estrategia común para formar neologismos es a través de palabras compuestas; así, en algunos idiomas quechuas, los puristas tratan de reemplazar préstamos del castellano con neologismos formados de palabras quechuas. Por ejemplo, en algunos diccionarios se encuentran palabras como *antapyyu* (caballo de metal: *anta* = metal + *apyyu* = caballo) para reemplazar *pisiklita* (bicicleta) (Chimbo Aguinda et al., 2008, p. 48). Los **calcos** (*calques*) son traducciones literales de otro idioma, por ejemplo, muchos hablantes del castellano andino usan *pues*, o las variaciones *pue*, *pe* con la función de *-mi/-n* que en idiomas quechuas transmiten información directa o de primera mano (M. S. Manley, 2007, p. 202).

(1) *Él camina rápido a la casa, pue.*

Pay wasinmanmi utqhayta purin.

(Quiroz Villarroel, 2000, p. 159)

Cuando se produce el contacto, los factores extralingüísticos colocan a un idioma en una posición socialmente más prestigiosa sobre el otro. En la mayoría de los casos, el idioma “extranjero” toma esta posición y tiene influencia unidireccional en el idioma de los “contactados” (Fought, 2010; Hickey, 2010). Ejemplos de idiomas con un alto grado de cambio inducido por contacto incluyen:

El aymara a través del quechua (2) y el castellano (3):

(2) *Pampa* “planicie” ← *pampa* (q.); *warmi* ← *warmi* (q.) “mujer, esposa” (Hardman, 2013)

(3) *P'usp'uru* ← *fósforo* (es.); *karu* ← *caro* (es.) (Hardman, 2013)

El francés canadiense a través del inglés canadiense:

(4) *fun* ← *fun* (ing.) “divertido”; *J'ai les shakes* “tengo miedo” ← *shake* (ing.) “temblar”

El guaraní (jopará) y quechua a través del castellano:

- (5) *Ndajúi porque chereasy.* (g.) ← *No vine porque estaba enfermo* (es.) (Zajícová, 2009)
 (6) *pero, (p)lasa, estaka* (q.) ← *pero, mercado (plaza), estaca* (es.) (Parker & Ibañez, 1964)

El castellano de EE.UU. a través del inglés de EE.UU.:

- (7) *Vamos a shopear* (es. ee.uu.) ← *Let's go shopping* (ing. ee.uu.); “*vamos de compras*” (es. estándar)

Y el inglés influenciado por idiomas nativos, por ejemplo, el inglés (criollo) hablado en Belice:

- (8) *Wi laik fufu eena kuknat milk.* (ing. bel.) ← *We like ripe mashed plantains in our coconut milk.* (ing. estándar) “A nosotros nos gustan los plátanos maduros majados en la leche de coco”. ← *fufu* (de akan, hablado en Ghana) “una masa hecha de yuca y harina de plátano verde” (Herrera et al., 2007)

Más allá de los idiomas hablados, el contacto lingüístico es muy común en la cultura sorda que casi siempre tiene interacción con una cultura oral que es dominante. A causa de este tipo de contacto, hay muchos conceptos erróneos sobre las lenguas de señas que es importante aclarar: (1) las lenguas de señas son idiomas completos y tienen la misma capacidad comunicativa que los idiomas hablados; (2) no son idiomas artificiales y se desarrollan naturalmente como los idiomas hablados; (3) no son gestos imprevistos similares a la imitación, sino que tienen vocabularios extensivos, sintaxis y fonologías complejas, al igual que los idiomas hablados; (4) no son universales, tienen dialectos y pueden ser tan diferentes como el castellano y el japonés. Sin embargo, a causa del contacto con el idioma hablado, hay bastante influencia, especialmente en el vocabulario y en el uso del alfabeto manual, lo cual puede ser erróneamente interpretado para los que no tienen entrenamiento en la lingüística y lo ven como un intento de transmitir el idioma hablado con la modalidad gestual. Sin embargo, el núcleo gramatical del idioma de señas y del idioma hablado no tienen relación.

Casi todos los conceptos tratados en este capítulo pueden ser aplicados a los idiomas de señas. Es importante destacar que la Lengua de Señas Boliviana (LSB) es reconocida como idioma oficial de Bolivia a través del Decreto Supremo N.º 0328.

Con respecto a las influencias lingüísticas, existe una gran cantidad de literatura que describe lo que sucede cuando dos o más idiomas o dialectos entran en contacto. Como grupo, los idiomas de contacto generalmente exhiben cambios similares; sin embargo, el nivel de cambio puede variar considerablemente.

Los préstamos no son los únicos fenómenos resultantes del contacto de lenguas; en las condiciones adecuadas, todos los elementos lingüísticos pueden estar sujetos teóricamente a transferencia (Thomason, 2001). Sin embargo, los préstamos de

clase abierta (o de palabras de contenido: sustantivos, verbos, adjetivos, adverbios, etc.) son el elemento más común copiado en el idioma “contactado” (Hickey, 2010; Winford, 2010). La justificación de la mayor prevalencia de préstamos de la clase abierta sobre préstamos de la clase cerrada (funcionales/ gramaticales: adposiciones, pronombres, partículas etc.) proviene del hecho de que la integración a la estructura morfosintáctica del lenguaje no es esencial, una tarea a menudo difícil para hablantes monolingües y adultos.

Translingüísticamente, los sustantivos son generalmente prestados con más frecuencia que los verbos y las transferencias de morfología derivativa (cambios lexicales) son más fáciles de realizar que la morfología flexiva (cambios gramaticales) (Thomason, 2010).

(9) Morfología derivada compartida en quechua (izquierda) y aymara (jaqi) (derecha):

- a. Topicalización: -(q|k|x)a vs. -x(a) (Adelaar, 2017)
- b. Información indirecta: -(s|sh)i vs. -chi (Quartararo, 2017)
- c. Semblativo -(sh|s|h|Ø)ina vs. (h|Ø)ina etc. (Muysken, 2015)

Sin embargo, mientras que la situación de contacto se intensifica y el aprendizaje es más «guiado», elementos más resistentes también pueden transferirse y, en situaciones de contacto muy intensas como el castellano y el quechua en los Andes, los elementos del idioma “contactado” se prestan con frecuencia al idioma “extranjero” (Thomason, 2010).

(10) Castellano andino influenciado por los idiomas quechuas:

- a. Fonología: el uso de /z/, y /ʒ/ en el norte de Ecuador – *perro* ['pezo], *lluvia* ['ʒuβia] (Stewart, 2020a)
- b. Fonología: la integración del fonema /ʃ/ (<sh>) en préstamos de los idiomas quechuas y en apodos – *Sebastián* → *Shaba* ['ʃaba]; *Macedonia* → *Mashi* ['maʃi]; *Ancash* ['ankaʃ] (un departamento en Perú) (Hardman-De-Bautista, 1982); *shungo* ['ʃungo] “corazón”; *shunsho* ['ʃunʃo] “tonto” (las dos palabras anteriores son comunes en el español de Ecuador)
- c. Lexical: *choclo* “maíz”; *cuy* “cuy (cobayo, conejillo de indias)”; *guagua* “bebé”; ñaña/ñaño “hermana/ hermano” (Escobar, 2011)
- d. Calcos: *nomás* ← *-lla* (q.) “solo, justo”; *pue* ← *-mi* “información de primera mano” (Quiroz Villarroel, 2000). El uso extensivo de diminutivos y en casi todas las categorías gramaticales (sustantivos, adjetivos, numerales, pronombres, preposiciones, adverbios, gerundios) (Escobar, 2011) – *vaquita*, *azulito*, *unito*, *ellita*, *dentrío*, *lejitos*, *andandito* ← *huagrawa*, *azulgu*, *shucgu*, *paygu*, *ucugu*, *carugu*, *purishpagu* “huevo, azul, uno, ella, dentro, lejos, andando”
- e. Gramática: *dar* + gerundio - *Dame pasando* ← *Yalishpa cuhuai* “por favor, pásamelo” (común en Ecuador) (Bruil, 2008)
- f. Gramática: el uso de flexiones verbales en el futuro como un imperativo - *¡Vendrás mañana!* ← *shamungui!* “¡Ven!” (común en Ecuador) (Haboud, 1998)

- g. Gramática: el uso de un verbo de movimiento + un gerundio subordinado para transmitir el perfecto – *Vengo comiendo*. ← *mikushpa shamuni* “Habiendo comido, vengo”. (común en Ecuador) (Haboud, 1998)

Debido a la relación social entre los hablantes de ambos idiomas, a menudo se desarrolla una actitud lingüística negativa entre los hablantes del idioma “contactado” y el aprendizaje del idioma “extranjero” se convierte en una prioridad. Una vez logrado el aprendizaje, comúnmente es el único idioma transferido a los hijos; sin embargo, el idioma “extranjero” es frecuentemente **marcado** (*marked*), ya que los padres lo adquieren típicamente como L2 y bajo condiciones “no guiadas” (Hickey, 2010a). El aprendizaje o la adquisición no guiada de la lengua (7.2.3) a menudo deja a los hablantes con la necesidad de llenar “vacíos” (conceptos o estructuras en L1 que no existen en L2) en el idioma recién adquirido. En estas condiciones, los préstamos sistémicos de la lengua de sustrato pueden ingresar al nuevo dialecto (Hickey, 2010b).

7.2. Resultados por cambio inducido de contacto de lenguas

7.2.1. Ultracorrección

La **ultracorrección** o **hipercorrección** (*hypercorrection*) ocurre cuando una persona que intenta hablar otro idioma (o dialecto) sobregeneraliza ciertas reglas. El resultado es un sistema marcado o acentuado, que frecuentemente los hablantes nativos pueden detectar. La ultracorrección puede ocurrir en las estructuras gramaticales (11), en la formación de las palabras (morfología) (12), en el significado de las palabras (semántica) (13), en la pronunciación de las palabras (fonológica y fonética) (14) y hasta en la ortografía (15). La ultracorrección no solo es un fenómeno de contacto, se encuentra también en la lengua materna (16).

- (11) Al hablar idiomas como el quechua, el aymara, el yuracaré o el guaraní, muchos hablantes nativos del inglés o francés omiten los sujetos en estructuras donde su uso es deseable. Por ejemplo, un hablante de inglés podría decir: *Na ricurcachu*. (q.) *No vio*. (es.) en una situación donde el sujeto es ambiguo y sería mejor decir *Ñuca panica na ricurcachu* (q.) *Mi hermana no vio*. En este caso, el hablante de inglés o francés omite el sujeto pensando que se escucha más formal o culto, aplicando una regla que no tiene en su idioma, ya que el uso de los sujetos es casi siempre obligatorio en inglés y en francés.
- (12) Al hablar castellano, algunos hablantes nativos de inglés se equivocan en el género de las palabras femeninas que terminan en *-ma* (*forma, pijama*), ya que han aprendido la regla de que las palabras originarias del griego terminadas en *-ma* son masculinas (*programa, sistema, teorema, tema e idioma*).

- (13) Frecuentemente, los hablantes de inglés piensan que el castellano tiene muchas “palabras sofisticadas” (*big words*), ya que el mismo vocabulario que se usa a diario en castellano normalmente es reservado para registros formales o técnicos en inglés; esto se debe a que estos registros usan muchas palabras derivadas del latín, debido a que era el idioma de la ciencia durante mucho tiempo. Entonces, un angloparlante diría: *yo utilicé el agua de la llave*, en lugar de *yo usé el agua de la llave*, en un intento de sonar más culto o usar la palabra apropiada; ya que la palabra “*usé*” en inglés se utiliza a diario y la palabra “*utilicé*” se encuentra en contextos más formales o técnicos.
- (14) Al hablar inglés, muchos hablantes nativos de castellano advierten que las palabras que empiezan con <es-> en castellano frecuentemente corresponden a <s-> en inglés; por ejemplo, *estadio* → *stadium*; *estaca* → *stake*, entonces puede haber un uso excesivo de esta regla en palabras que empiezan con <es-> en inglés; por ejemplo: *escape* [ɪˈskeɪp] → *scape** [skeɪp] “escapar” [eskaˈpaɾ]. Al hablar castellano, muchos hablantes nativos de quechua ecuatoriano ponen /e/ y /o/ en palabras que tienen /i/ o /u/ respectivamente, ya que el quechua de Ecuador no tiene las vocales medias en la misma distribución como en el castellano: por ejemplo, *lluvia* [ˈʎu.βja] → *llobia* [ˈʒo.βja]; *mesa* [ˈmesa] → *misa* [ˈmisa] “mesa”.
- (15) Al escribir en castellano, algunos hablantes de quechua ponen la <h> al comienzo de las palabras que empiezan con vocal, ya que el quechua no usa la <h> muda; por ejemplo, *se **hiban** a bautizar a un niño*. Lo mismo pasa, frecuentemente, con los hablantes de italiano al escribir inglés por motivos similares.
- (16) Muchas palabras prestadas del latín en inglés conservan la forma singular y la forma plural de la declinación originaria, por ejemplo: *bacterium-bacteria* “bacteria-bacterias”, *corpus-corpora*, “corpus-corpus”, *radius-radial*, “radio-radios”, *index-indices*, “índice-índices”, *formula-formulae*, “fórmula-fórmulas”, *focus-foci*, “enfoque-enfoques”. En un intento por sonar más cultos, algunos hablantes de inglés aplican estas reglas a palabras irregulares o que tienen origen en otra declinación o en otro idioma: *octopus* “pulpo” → *octopi** (la forma correcta es *octopuses* u *octopodes*, aunque algunas fuentes aceptan *octopi* a través de la regularización –véase 7.2.2–); *virus* “virus” → *virii** o *virii** (la forma correcta es *viruses*).

7.2.2. Regularización

La **regularización** (*regularization*) es un fenómeno a menudo mencionado en la literatura de la adquisición del lenguaje, en la dialectología y en el cambio

lingüístico. Sin embargo, es uno de los fenómenos responsables del cambio inducido por contacto, especialmente en los pidgins, criollos y lenguas francas (véase 7.3). Se define como una tendencia de reemplazar formas irregulares o excepciones con formas que reflejan patrones generales en el idioma; por ejemplo, en Colombia existe un idioma llamado palenquero que es un criollo basado en el castellano y el kikongo (es hablado en varios países del centro de África). En este idioma, todos los adjetivos regularizan a la forma masculina: *Ese nata é susio*, “Esa nata está sucia”; no hay género gramatical ni conjugaciones verbales, lo cual ha hecho obligatorios los pronombres de sujeto y los tiempos se entienden por las partículas preverbiales (Mackenzie, 1999).

Observemos la **Tabla 2**. Todas las partículas preverbiales tienen sus orígenes en el castellano con la excepción del futuro.

¿Puede identificar sus **cognados** (*cognates*) en el castellano?

¿Cómo fue establecida la base verbal?

¿Cómo podría decirse las siguientes oraciones en palenquero?:

Él va a bailar. Ella contaba. Ustedes saben limpiar.

Tabla 2

Pronombres y tiempos verbales en palenquero

Pronombres	Presente	Habitual	Futuro	Pasado	Imperfecto
Í	ta ablá	asé ablá	tan kumé	á enfermá	taba bibí
Yo					
Bo					
Ele					
Suto					
Utere					
Enú, ané					
	hablar	hablar	comer	enfermar	vivir

Nota. Basado en Mackenzie, 1999

7.2.3 El aprendizaje o la adquisición no guiada

El **aprendizaje** o la **adquisición no guiada** (*unguided learning or acquisition*) son dos fenómenos resultantes del cambio inducido por el contacto. El aprendizaje no guiado sucede a menudo cuando una persona aprende un idioma, pasada la pubertad y sin una pedagogía estructurada. Sucede cuando los niños adquieren el idioma aprendido de sus padres bajo condiciones justamente no guiadas, lo cual resulta en un nuevo dialecto nativo influenciado por el idioma nativo de los papás. Si los niños tienen contacto con compañeros que hablan el idioma estándar o con una

pedagogía estructurada, frecuentemente se convierten en **bidialectales** (*bidialectal*). Como hemos visto, el aprendizaje no guiado a menudo deja a los hablantes con la necesidad de llenar “vacíos” en el idioma recién adquirido para expresar conceptos o usar estructuras de su idioma nativo que no existen en su nuevo idioma. Por ejemplo, al hablar inglés, los hablantes nativos de castellano frecuentemente producen todas las vocales de una manera acentuada o “completa” cuando los hablantes nativos de inglés producen la vocal neutral (schwa [ə]) en posiciones no-acentuadas, como es el caso de [mə'ʃɪn] → [ma'ʃɪn] para decir *machine* “máquina”. Frecuentemente, los hablantes nativos del castellano también regularizan el uso de preposiciones en el inglés, por ejemplo: *in the table, in the TV stand, in the box* etc., en lugar de decir *on the table, on top the TV stand, inside the box* “en la mesa, en el mueble para el televisor, en la caja”. Estos nuevos elementos llegan a ser **nativizados** (*nativized*) en el nuevo dialecto una vez que se pasan a los hijos.

Se encuentra una multitud de ejemplos en el castellano rural andino, usado por hablantes nativos de las lenguas quechuas y aymaras. Sin embargo, nos focalizaremos en la transferencia de la **evidencialidad** (*evidentiality*), la cual se define como un medio que indica la fuente de información comunicada en el discurso. Los idiomas quechuas usan **evidenciales** (*evidentials*) en la forma de sufijos, por ejemplo, para comunicar información directa de primera mano se usa *-m(i)* y para comunicar información conjeturada e inferida se usa *-chr(a)/-chari*. Ya que el castellano no usa evidenciales gramaticales para comunicar este tipo de información, los hablantes nativos de las lenguas quechuas y aymaras han implementado un sistema parecido en el castellano rural andino (17).

(17) Evidencialidad en el castellano rural andino:

- a. El pretérito perfecto simple se usa para comunicar información directa.
 - i. *Hace un ratito dej-é mis llaves sobre la mesa, pero ahora no aparecen.* (Palacios Alcaine, 2005)
- b. El pretérito pluscuamperfecto para comunicar información conjeturada e inferida.
 - i. *Entré en mi casa y olía bien rico, y pensé “¡Qué rico! Alguien hab-ía hecho una torta.”* (Palacios Alcaine, 2005)

7.2.4. Alternancia de código

De acuerdo con Thomason (2001), el mecanismo con mayor número de estudios es la **alternancia de código** (*code-switching*), que sucede cuando un hablante alterna entre dos o más idiomas o dialectos durante una misma conversación, ya que la alternancia de código se encuentra entre personas que dominan o por lo menos

entienden dos o más idiomas. Frecuentemente, el enunciado alternado conserva la mayor parte de la sintaxis y la fonología del idioma respectivo; el ejemplo (18) representa la alternancia de código **intrasectorial** en el quechua ecuatoriano (kichwa) y el castellano; las partes subrayadas son de origen castellano.

- (18) *Nobiapa parte gashtota saquishpami apashpa rin uno pierna de ganado o cabeza de ganado. Shinallata apashpa rina a unos diez o doce costales de naranja, unos veinte cabezas de plátano, unos quince costales de pan. Chay tandaca de sal o de dulce.*

“La familia de la novia contribuye con el gasto que consiste en una pierna de ganado, cabeza de ganado y es común llevar otros presentes como: diez o doce costales de naranja, veinte cabezas de plátano y quince costales de pan; el pan puede ser de sal o de dulce” (basado en Stewart, 2013).

El (19) representa un ejemplo de la alternancia de código **intersectorial** en el castellano e inglés:

- (19) A: *¡Ayer estaba caminando por la calle y encontré veinte dólares!*
 B: *¡Wow, veinte dólares! What did you do with it?*
 A: *I thought I was going to spend it on a new pair of shoes, pero de repente me desperté y me di cuenta de que estaba soñando.*
 B: *¡¿Ay, ques pues?! You're crazy.*

Hay una variedad de desenlaces que pueden suceder, desde préstamos comunes como sustantivos y marcadores del discurso, hasta la formación de un nuevo idioma, por ejemplo, el gurindji kriol hablado en el norte de Australia (McConvell & Meakins, 2005).

Un resultado común de la alternancia de código, y del bilingüismo en general, se llama **diglosia** (*diglossia*), que se define como el uso de uno de los idiomas en ambientes formales; por ejemplo, en el trabajo, trámites oficiales, etc., mientras que el otro idioma se usa en ambientes informales, por ejemplo, en la comunidad o en la casa. La diglosia es muy común en Bolivia entre los idiomas indígenas y el castellano.

A veces los conceptos se cruzan entre los préstamos y la alternancia de código, especialmente en frases cortas o palabras individuales que asimilan total o parcialmente a la fonología del otro idioma. Sin embargo, hay tres características que ayudan a identificar esta diferencia: (1) los préstamos son más recurrentes y aparecen en el vocabulario de los monolingües, (2) la alternancia de código es más espontánea y (3) los préstamos se ajustan a la gramática del idioma prestatario o sus funciones gramaticales se pierden y esto resulta en una **palabra/ forma congelada** (*frozen form*).

7.2.5. Lexificación

El término **lexificación**, en sí, es usado rara vez en la lingüística y no es fácil de definir, aunque se entiende como la transmisión del léxico de una manera general desde un idioma a otro. Sin embargo, añadiendo prefijos al término, se crean conceptos más precisos y útiles para describir fenómenos lingüísticos usados en la investigación del contacto de lenguas.

La **relexificación** (*relexification*) se define como el proceso de “reencadenamiento” o “reetiquetado” de entradas léxicas de un idioma a otro (Lefebvre, 2006, 2005; Lefebvre & Therrien, 2007; Muysken, 1981) y frecuentemente afecta el vocabulario a gran escala. El idioma que presta el vocabulario se denomina **lexificador** (*lexifier*).

La relexificación es una subclasificación de préstamo donde solo la forma fonológica de una palabra es adoptada, es decir que el significado de la palabra y sus funciones gramaticales en el idioma originario no se transfieren a propósito. Por ejemplo, en la “media lengua”, idioma mixto hablado en el Ecuador, basado en léxico del castellano y la gramática quichua, se reemplaza el pronombre *can* “tú/ vos/ usted” por el pronombre *bos* derivado del castellano. Sin embargo, el sentido y la morfología de la palabra sigue siendo igual a la del quichua; es decir, el sistema pronominal no requiere una forma formal como “usted”, ya que no existe en el quichua¹, y para formar el plural (*ustedes*) se añade el pluralizador del quichua *-cuna*. En media lengua, *boscuna*; en quichua, *cancuna* “ustedes”. Por lo tanto, solo el “sonido” de la palabra del castellano se transfirió.

En otros casos, otras propiedades pueden ser transferidas; este proceso se llama **translexificación** (*translexification*) (Muysken, 1981). Por ejemplo, el quichua tiene el verbo *ricuna* “ver” y se puede cambiar el sentido, añadiendo el morfema *-chi*, *ricuchina* “mostrar”; *-ri*, *ricurina* “asomar, parecer, se ve”; *-ra*, *ricurana* “espiar, mirar fijamente” y si la relexificación fuera el único proceso, el idioma mixto tendría *bina* “ver”, *bichina* “mostrar”, *birina* “asomar, parecer, se ve”, y *birana* “espiar, mirar fijamente”. De hecho, estas palabras existen, pero a través de la translexificación existen otras también: *mostrana* “mostrar”, *asomana* “asomar” y *parisiana* “parecer”, las cuales han transferido su fonología junto con el significado del castellano (Muysken, 1981, p. 59).

También existe el término **adlexificación** (*adlexification*), el cual se refiere a la adición del vocabulario cultural, que no ha estado presente previamente en el idioma prestatario. La palabra o variaciones de la palabra *computadora*, del inglés *computer*, es un ejemplo de adlexificación en muchos idiomas; además, el término

¹ *Quiquin* (*kikin*) es una adición reciente para reflejar el uso formal de *usted*.

supralexificación (*supralexification*) se refiere a la expansión o complicación de clases semánticas preexistentes a través de préstamos (Epps & Law, 2019; Grant, 2015). Un ejemplo es la adición de términos para colores de otro idioma para expandir la clase semántica de colores; por ejemplo, muchos idiomas quechuas han adoptado las palabras *bioleta*, *berdi*, y *plumu* “violeta, verde, plomo”.

7.2.6. Metatipia

La **metatipia** (*metatypy*) es el proceso contrario de la lexificación; en lugar de modificar el léxico, la metatipia implica la reestructuración de la morfosintaxis (gramática) y la semántica de un idioma basado en otro, mientras se mantienen las formas lexicales (M. Ross, 2007), y esto se logra principalmente a través de los calcos (traducciones literales de otro idioma). El siguiente ejemplo (20) proviene de Ross (1999) e indica cómo la estructura gramatical en el *takia* (el idioma modificado) es igual a la del *waskia* (el idioma modificador).

- (20) *yai* *Tamol* *An* *ida* Takia
 ane *Kadi* *mu* *ili* Waskia
 Yo hombre DET con.el
 “el hombre y yo”

De acuerdo con Ross (1999), el idioma modificado es emblemático de la identidad de sus hablantes, mientras que el idioma modificador es intercomunitario, una tendencia común en muchos idiomas mixtos (7.3.3).

7.3. Clasificaciones de idiomas de contacto

El lenguaje simplemente es una herramienta social que usamos para comunicar la experiencia humana; mientras que nuestras experiencias van cambiando, nuestro lenguaje tiene que mantenerse actualizado.

A menudo, este proceso de cambio es gradual, ya que el desarrollo de tecnologías y de nuestras culturas normalmente han ocurrido despacio y de forma disonante durante la mayor parte de nuestra historia como especie. Sin embargo, en el mundo globalizado, hemos sido testigos de muchos cambios en un breve período de tiempo y esto se refleja en nuestros idiomas; por ejemplo, durante la pandemia del COVID-19, observamos la creación de un nuevo vocabulario, expresiones y el resurgimiento de palabras poco usadas en la vida diaria (*distanciamiento social*, *epidemia* vs. *pandemia*, *asintomático*, *autoaislamiento*, *cuarentena*, *pródromo*, *aplanar la curva*, *quédate en casa*, *EPP* (*equipo protector personal*), etc.); pero en el lenguaje no hay nada más poderoso para crear cambios drásticos como el contacto de idiomas entre personas desconocidas, con diferentes culturas y cosmovisiones. Ese choque puede ser una fuerza externa tan fuerte que cambia la experiencia humana, la cual puede reflejarse en su idioma; imaginemos ser víctimas de la esclavitud, al estar

controlados por personas que no hablan nuestro idioma o ser víctimas de una guerra, conquista o colonización. Estos eventos van a tener un impacto profundo en cómo entendemos el mundo y en cómo nos comunicaríamos con las personas que están en el poder y, al mismo tiempo, con las otras víctimas que quizás no hablen nuestro idioma. En casos menos traumáticos, imaginemos que deseamos comunicarnos a largo plazo con otros grupos de personas, para el trueque o la compra y venta de bienes. Finalmente, imaginemos un grupo de jóvenes bilingües atrapados entre dos culturas que desarrollan una nueva identidad; todos estos eventos marcados en este párrafo estimulan la creación de un nuevo idioma. A propósito de todo ello, las siguientes subsecciones describen tres lenguas que resultan del contacto.

7.3.1. Pidgins

Los **pidgins** (*pidgins*) son un grupo de idiomas de contacto creados a partir de la necesidad de comunicarse con otros cuando no existe una lengua común. Durante la conquista de Europa por casi todo el mundo, había un gran auge de pidgins, en especial durante la trata de esclavos en los países que ahora forman parte de las islas caribeñas. Los pidgins también se forman por el trueque y para la compra y la venta de bienes con grupos que tampoco tienen una lengua común. Varios de los rasgos más evidentes e identificables de un pidgin son las simplificaciones extensivas de la gramática, el vocabulario, el uso restringido del idioma, la falta de hablantes nativos y la falta de uniformidad estructural. Los pidgins que se basan en las lenguas romances, como el castellano, el francés y el portugués, a menudo muestran una pérdida o reducción en el género funcional, simplificación completa de las flexiones verbales y una reducción en los tiempos.

Dependiendo de las circunstancias sociales, los pidgins pueden tomar varias formas, desde los que tienen vocabularios mixtos de todas las **lenguas fuentes** (*source languages*), los cuales son frecuentemente usados para el comercio, hasta los que tienen influencia primaria de una sola lengua y son frecuentemente usados para comunicarse con el grupo de poder (ya sea económico, social, regional o de opresión). Los ejemplos de pidgins que han existido en América Latina incluyen: *língua geral amazónica* (basado en las lenguas tupis), hablado en Brasil (véase Moore et al., 1994); ndyuka-tiriyó (basado en el ndyuka), hablado en Surinam, y el pidgin amazónico (basado en el castellano y el shuar), hablado en Ecuador (véase Gnerre, 1975; Muysken, 1997; Simson, 1886).

Cuando un pidgin se establece como un idioma de uso común para el comercio o para cuestiones políticas entre varios grupos que no comparten el mismo idioma, se lo conoce como **lengua franca** (*lingua franca*); si su uso se expande, el pidgin puede pasar a la siguiente generación como idioma nativo. En este caso, ya no es conocido como un pidgin, sino como un idioma criollo.

7.3.2. Criollos

Los idiomas **criollos** (*creoles*) reciben este nombre ya que, durante el período de la colonización europea, estos fueron hablados por los criollos, personas de descendencias mixtas o por los esclavos nacidos en las colonias. Frecuentemente, los idiomas criollos toman el nombre del país originario, por ejemplo, el “criollo haitiano”, y, cuando se refieren a los idiomas criollos en general, se usa el nombre del idioma lexificador, por ejemplo, “los criollos de base francesa”.

En las Américas, los criollos de base francesa se hablan en Haití, Luisiana (EE. UU.), en muchos países y colonias de las Antillas (Martinica, Guadalupe e Islas de los Santos), en la Guayana Francesa y en el estado de Amapá en Brasil.

Existe una gran concentración de idiomas criollos en el Caribe a causa de la trata de esclavos y frecuentemente el idioma lexificador viene de Europa; por ejemplo, el Papiamentu, hablado en Aruba y en Curazao, tiene bases de portugués, castellano y muchas palabras de holandés; el Berbice (ya extinto) fue hablado en Guyana y tiene su base en el holandés; el patois jamaiquino hablado en Jamaica tiene su base en el inglés.

Al mismo tiempo, hay criollos de base castellana en todo el mundo; el chabacano de las Filipinas; el palanquero de Colombia (véase Tabla 2); el castellano bozal (ya extinto) fue hablado en Cuba por esclavos africanos, y algunos investigadores categorizan el jopará hablado en Asunción como un criollo, aunque esto es controversial.

Dato curioso:

Existe evidencia de que el inglés moderno pasó por la criollización, ya que, comparado con otras lenguas germánicas, muchas complejidades gramaticales han sido simplificadas o se han perdido por completo; por ejemplo, el género gramatical, el sistema de declinaciones nominales, las conjugaciones verbales y una gran cantidad de préstamos del francés. Esta hipótesis se llama la “hipótesis criolla del inglés medio”, y los que están a favor (véase Bailey & Maroldt, 1977) piensan que la criollización sucedió durante la conquista normanda de Inglaterra en el siglo XI. También hay que resaltar que esta hipótesis es controversial (véase Rothwell, 1998).

En la literatura criolla, frecuentemente se encuentran los términos **sustrato** (*substrate*) y **superstrato** (*superstrate*), los cuales se refieren respectivamente a sus posiciones sociales; el idioma nativo (idioma subordinado) y el idioma lexificador (idioma dominante).

En las primeras generaciones que hablan el nuevo idioma, los idiomas criollos frecuentemente aún mantienen una estructura simplificada y vocabulario reducido, ya que han pasado por la etapa de pidginización; sin embargo, como cualquier idioma natural, se intensifican en complejidad con el tiempo y, si es que los idiomas criollos continúan en contacto con el idioma lexificador, pueden llegar a una etapa de **descriollización** (*decreolization*), donde **convergen** (*converge*) de nuevo con el idioma progenitor. A lo largo de esta travesía, encontramos un continuum, desde un **basilecto** (*basilect*) hasta un **acrolecto** (*acrolect*); el primero es la variedad del criollo con menos prestigio social, mientras que el acrolecto es la variedad del criollo con más prestigio, ya que tiene características más comunes con el idioma lexificador.

7.3.3. Idiomas mixtos

Los **idiomas mixtos** (*mixed languages*), también conocidos como **idiomas mixtos bilingües** (*bilingual mixed languages*), son una categoría especial de contacto del idioma, dado que los creadores son bilingües competentes en ambas lenguas fuentes. Por ende, no se forman por necesidad comunicativa, sino por razones expresivas, para mostrar una nueva identidad, mantener una identidad antigua o simplemente para facilitar la comunicación entre dos cosmovisiones.

Dada su génesis, los idiomas mixtos no pueden ser clasificados genéticamente usando métodos tradicionales de categorización (véase el modelo de descendencia de Stammbaum; Meakins & Stewart, accepted). A diferencia de los pidgins, de los idiomas criollos y de las lenguas francas, los idiomas mixtos son frecuentemente reservados para el uso interno, dentro de una misma **comunidad de hablantes** (*speech community*). Ya que los creadores fueron bilingües competentes, no se encuentra simplificación en sus estructuras gramaticales y no existe una reducción en el vocabulario.

Al existir múltiples factores sociales y lingüísticos que tienen que suceder en el momento y lugar correcto para crear una lengua con tales características, de los 7117 idiomas documentados en el mundo (Eberhard et al., 2020), los idiomas mixtos solo constituyen una pequeña fracción (<0.5%), lo que los convierte en raros.

Más allá de estas descripciones generales, los idiomas mixtos son difíciles de categorizar, dadas sus diferentes situaciones sociohistóricas y estructuras lingüísticas; sin embargo, existen tres clasificaciones que se basan en las divisiones sistemáticas de cada lengua fuente.

En la categoría prototípica, primer grupo, se encuentran divisiones entre el vocabulario y la gramática (así como la media lengua hablada en Ecuador); en

En América Latina, existen dos lenguas reconocidas como mixtas: la media lengua de Ecuador y el kallawaya de Bolivia.

El ejemplo (21) ilustra una oración típica de media lengua y los elementos en cursiva derivan del castellano.

- “Ese hombre tiene una buena voz para cantar”. (Stewart, 2020b)

(22) Cchana-*chi-rqa-iqui* isna-*pu-na-iqui-paq* kallawayá
call-CAU-PAST-1-2 go-MOV-NOM-2-BEN *quechua*

- “Pedí que te llamen para que puedas ir”.

7.4. Conclusión

En este capítulo se ha considerado varios temas que forman parte de una extensa literatura que investiga lo que sucede cuando grupos de personas entran en contacto sin tener un mismo idioma en común. Los rasgos que marcan un idioma por el contacto toman varias formas, desde préstamos léxicos hasta la creación de una nueva lengua.

El grado de influencia se relaciona con el tiempo de contacto, el tipo de tecnologías intercambiadas, la dinámica social entre los grupos, las estructuras lingüísticas de los idiomas, entre otros más. Aunque el contacto de lenguas puede ser un tema complejo, los rasgos preservados en un idioma a causa del contacto aportan evidencia histórica que se ha encontrado en generaciones pasadas; al mismo tiempo, permite comprender los factores sociológicos, psicológicos, lingüísticos y cognitivos que permiten a los seres humanos tomar dos o más idiomas no relacionados y dividirlos para crear un idioma nuevo completamente funcional, basado en diferentes componentes lingüísticos, que incluya una nueva cosmovisión.

Referencias

- Adelaar, W. (2017). A typological overview of Aymaran and Quechuan language structure. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Part III - Typological Profiles of Linguistic Areas and Language Families* (pp. 651–682). Cambridge University Press. <https://doi.org/10.1017/9781316135716.021>
- Bailey, C. J., & Maroldt, K. (1977). The French lineage of English. In J. Meisel (Ed.), *Langues en contact-pidgins-creoles-Languages in contact* (pp. 21–53). Narr Francke Attempto.
- Bruil, M. (2008). *Innovations in the Ecuadorian Converb Systems: Grammatical change in language contact situations*. [MA Thesis]. University of Leiden.
- Chimbo Aguinda, J. J., Ullauri Velasco, M. A., & Shiguango Andi, E. E. (2008). *Shimiyukkamu-Diccionario: Kichwa-Español, Español-Kichwa* (Segunda). Benjamín Carrion.
- Eberhard, D. M., Gary Simons, & Charles Fennig. (2020). *Ethnologue: Languages of the World* (23rd ed.). SIL International. <http://ethnologue.com>
- Epps, P., & Law, D. (2019). *Language Contact* (J. Darquennes, J. C. Salmons, & W. Vandenbussche, Eds.; Vol. 1, p. 866). Walter de Gruyter GmbH & Co KG.

- Escobar, A. M. (2011). Spanish in Contact with Quechua. In M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 323–352). Blackwell Publishing Ltd.
- Fought, C. (2010). Linguistic contact. In R. Hickey (Ed.), *The handbook of language contact* (pp. 229–319). Wiley-Blackwell.
- Gnerre, M. (1975). *A Spanish pidgin of the Shuar* [Manuscript]. University of Rome.
- Grant, A. P. (2015). *The Oxford Handbook of the Word* (J. R. Taylor, Ed.; p. 960). OUP Oxford.
- Haboud, M. (1998). *Quichua y Castellano en los Andes Ecuatorianos. Los Efectos de un Contacto Prolongado*. Abya-Yala.
- Hardman, M. (2013). *Dictionary of the Jaqi languages: Aymara, Jaqaru and Kawki*. Ethnos Project. <http://test.aymara.ufl.edu/dictionary.html>
- Hardman-De-Bautista, M. J. (1982). The mutual influence of Spanish and the Andean languages. *Word*, 33(1–2), 143–157. <https://doi.org/10.1080/00437956.1982.143572>
- Herrera, Y., Manzanares, M., Woods, S., Crosbie, C., & Decker, K. (2007). *Kriol – English Dikshineri; English – Kriol Dictionary* (P. Crosbie, Ed.; 1st ed.). SIL International.
- Hickey, R. (2010). Language contact: Reconsideration and reassessment. In R. Hickey (Ed.), *The handbook of language contact* (pp. 1–29). Wiley-Blackwell.
- Krippner, S. (2020). Gottschalk-Batschkus. In C. E. Gottschalk-Batschkus & J. C. Green (Eds.), *Handbook of ethnotherapies*. BoD.
- Lefebvre, C. (2006). *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole* (Vol. 88). Cambridge University Press.
- Lefebvre, C. (2005). *Relexification: A process available to human cognition*. 27, 125–139.
- Lefebvre, C., & Therrien, I. (2007). On Papiamentu ku. In J. Siegel, J. Lynch, & D. Eades (Eds.), *Language description, history and development: Linguistic indulgence in memory of Terry Crowley* (Vol. 30, pp. 169–182). John Benjamins Publishing.
- Mackenzie, I. (1999). The Linguistics of Spanish [University Personal Webpage]. *Palenquero*. <https://www.staff.ncl.ac.uk/i.e.mackenzie/palenque.htm>

- Manley, M. S. (2007). *Spanish in Contact: Policy, Social and Linguistic Inquiries* (K. Potowski & R. Cameron, Eds.). John Benjamins Publishing.
- McConvell, P., & Meakins, F. (2005). Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics*, 25, 9–30.
- Meakins, F., & Stewart, J. (accepted). Mixed Languages. In S. Mufwene & A. M. Escobar (Eds.), *Cambridge Handbook of Language Contact*. Cambridge University Press.
- Moore, D., Facundes, S., & Pires, N. (1994). *Nheengatu (Língua Geral Amazônica), its History, and the Effects of Language Contact*. Survey Of California And Other Indian Languages.
- Muysken, P. (1981). Halfway between Quechua and Spanish: The case for relexification. In A. R. Highfield (Ed.), *Historicity and variation in Creole studies* (Vols 57–78). Karoma Publishers.
- Muysken, P. (1997). Media Lengua. In S. Thomason (Ed.), *Contact languages: A Wider Perspective* (pp. 365–426). J. Benjamins Pub. Co.
- Muysken, P. (2015). From nominal predicate to dictic clausal highlighter: The development of hina ‘like’. In M. Manley & A. Muntendam (Eds.), *Quechua Expressions of Stance and Deixis*. Brill.
- Palacios Alcaine, A. (2005). La influencia del Quichua en el español andino Ecuatoriano. In *Homenaje al Dr. Luis Jaime Cisneros* (Vol. 1, pp. 389–408). Pontificia Universidad Católica del Perú.
- Parker, G., & Ibañez, A. (1964). *English-Quechua Dictionary—Cuzco, Ayacucho, Cochabamba*. US Department of Health, Education, and Welfare.
- Quartararo, G. (2017). *Evidencialidad indirecta en aimara y en el español de la Paz: Un estudio semántico-pragmático de textos orales* [PhD Dissertation, Stockholm University]. <https://su.diva-portal.org/smash/get/diva2:1091898/FULLTEXT01.pdf>
- Quiroz Villarroel, A. (2000). *Gramática Quechua Boliviano Normalizado* (1st ed.).
- Ross, M. (2007). Calquing and metatypy. *Journal of Language Contact*, 1(1), 116–143. <https://doi.org/doi.org/10.1163/000000007792548341>
- Ross, M. D. (1999). *Exploring metatypy: How does contact-induced typological change come about?* [Keynote talk]. Australian Linguistic Society’s annual meeting, Perth.

- Rothwell, W. (1998). Arrivals and departures: The adoption of French terminology into middle English. *English Studies*, 79(2), 144–165. <https://doi.org/10.1080/00138389808599121>
- Simson, A. J. (1886). *Travels in the wilds of Ecuador*. Sampson Low, Marston, Searle, and Rivington.
- Stewart, J. (2013). *Stories and traditions from Pijal: Told in Media Lengua* (1st ed.). CreateSpace.
- Stewart, J. (2020a). A preliminary, descriptive survey of rhotic and approximant fricativization in Northern Ecuadorian Andean Spanish varieties, Quichua, and Media Lengua. In R. Rao (Ed.), *Spanish Phonetics and Phonology in Contact: Studies from Africa, the Americas, and Spain*. John Benjamins. <https://doi.org/10.1121/10.0000991>
- Stewart, J. (2020b). Media Lengua Dictionary. *Dictionaria*, 13, 1–3191.
- Thomason, S. (2001). *Language contact: An introduction*. Edinburgh University Press.
- Thomason, S. (2010). Contact explanations in linguistics. In R. Hickey (Ed.), *The handbook of language contact* (pp. 29–47). Wiley-Blackwell.
- Winford, D. (2010). Contact and borrowing. In R. Hickey (Ed.), *The handbook of language contact* (pp. 1–29). Wiley-Blackwell.
- Zajícová, L. (2009). *El Bilingüismo Paraguayo. Usos y Actitudes hacia el Guaraní y el Castellano*. Iberoamericana y Vervuert.

CAPÍTULO 8

LINGÜÍSTICA HISTÓRICA

Cómo citar: Ciucci, L. (2023). Lingüística histórica. En P. Alandía Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 180-203), Página y Signos/Funproceib Andes. <https://doi.org/10.5281/zenodo.11111169>

Luca Ciucci*

0. Introducción

Cada idioma, de hecho, por su propia naturaleza, está sujeto a cambios con el tiempo. La lingüística histórica estudia el cambio de las lenguas desde una perspectiva diacrónica, es decir, a través del tiempo, por lo que se denomina también lingüística diacrónica. La diacronía (del griego *dia* ‘a través’ y *chrónos* ‘tiempo’) se opone a la sincronía (del griego *syn* ‘con, junto’ y *chrónos* ‘tiempo’), que se refiere al estudio de la lengua en un momento preciso en el tiempo, no necesariamente contemporáneo. Por ejemplo, tanto el estudio del aymara hablado en La Paz en el siglo XVI (analizando la traducción de los textos religiosos editados por José de Acosta [1584]) como del aymara paceño actual (Cerrón-Palomino & Carvajal Carvajal, 2009) son estudios sincrónicos; la perspectiva diacrónica entra en juego, por ejemplo, si se comparan los datos del siglo XVI con los de hoy para describir el cambio lingüístico que se ha producido en el ínterin.

El presente capítulo tiene como objetivo proporcionar algunos elementos esenciales de lingüística histórica.

8.1. La lingüística histórica y la clasificación genética de las lenguas¹

Una de las tareas de la lingüística histórica es clasificar las lenguas según su afinidad genética. A partir de un idioma, con el tiempo, se desarrollan variaciones

* Luca Ciucci es investigador en la Universidad James Cook de Cairns, Australia. Obtuvo su doctorado en Lingüística en Pisa, Italia en el 2013. Sus intereses principales son la lingüística histórica, la morfología y la tipología lingüística. Se dedica a la investigación de las lenguas zamuco y del chiquitano/bésiro.

¹ Muchos ejemplos en este capítulo provienen de nuestros estudios sobre las lenguas zamuco y el chiquitano (o bésiro). Para evitar repeticiones, mencionamos solo aquí las fuentes utilizadas para estas lenguas, que son: Ciucci (2016) para la morfología de las lenguas zamuco; Ciucci & Bertinetto (2015, 2017) y Ciucci (2021) para la reconstrucción morfológica del proto-zamuco; Ciucci & Pia (2019) por la sociolingüística del ayoreo; Ciucci (2020) para el contacto entre lenguas zamuco y otras lenguas del Chaco. Para el zamuco antiguo, véase Chomé (1958) y Ciucci (2018). Para el antiguo chiquitano (hablado en el siglo XVIII), los datos provienen de Adam & Henry (1880) y de nuestra investigación en varios archivos, para el chiquitano ignaciano, véase Ciucci & Macoñó Tomichá (2018), para el chiquitano migueleño Nikulin (2018/2019); los datos para el chiquitano lomeriano vienen de nuestro trabajo de campo. Por razones de simplicidad, hemos omitido algunas características de la morfología nominal de las lenguas zamuco (ver Ciucci, 2016; Bertinetto et al., 2019). Hemos transcrito los datos de las lenguas zamuco y del chiquitano utilizando caracteres del AFI, mientras que en los otros casos hemos utilizado, en la medida de lo posible, la transcripción de los respectivos autores.

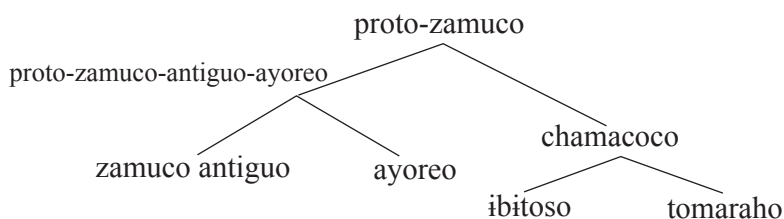
Agradecimientos. Queremos expresar nuestra gratitud a Alexandra Y. Aikhenvald, Pier Marco Bertinetto y Biera Cubilla Zadovsky.

dialectales, y poco a poco estos dialectos dejan de ser mutuamente inteligibles, convirtiéndose así en idiomas diferentes. Estas lenguas hermanas, que descienden de la misma proto-lengua, forman una familia lingüística. De estos nuevos idiomas pueden originarse otros, creando así subgrupos dentro de la familia. Las relaciones entre las lenguas que tienen un ancestro común se representan mediante un árbol genealógico.

Mostramos aquí un árbol genealógico muy simple, el de la familia Zamuco (Figura 1). Las lenguas zamuco descienden del proto-zamuco, que no está documentado históricamente. La reconstrucción del proto-zamuco u otras proto-lenguas es una de las tareas de la lingüística histórica. Dos ramas se separan del proto-zamuco: el primer grupo está constituido por el †zamuco antiguo y el ayoreo, mientras que el segundo, por el chamacoco. El zamuco antiguo se hablaba en el siglo XVIII en el Chaco septentrional y fue descrito en las Misiones de Chiquitos, donde los hablantes abandonaron gradualmente el zamuco antiguo en favor del chiquitano (o *bésiro*).² El ayoreo también se habla en el Chaco septentrional y aunque es similar al zamuco antiguo no debe ser considerado como la continuación directa de esta lengua. Ambos idiomas, sin embargo, tienen muchas similitudes, lo que indica que se separaron en tiempos relativamente recientes (poco antes del siglo XVIII) y que provienen del mismo idioma común, que es descendiente inmediato del proto-zamuco, y que podemos llamar proto-zamuco-antiguo-ayoreo.

Figura 1

Árbol genealógico de la familia Zamuco



A la derecha del árbol está el chamacoco, que forma por sí solo una rama de la familia. El chamacoco comparte poco léxico con el ayoreo y el zamuco antiguo, lo que indica que se separó hace mucho tiempo de la otra rama de la familia; aunque el chamacoco ahora se habla principalmente en el norte de Paraguay, cerca de la frontera con Bolivia, estudios históricos indican que grupos de hablantes pertenecientes a esta rama de la familia vivían en las Misiones de Chiquitos (Combès, 2009). El chamacoco se divide hoy en dos dialectos: el tomarahho y el ibitoso, que

2 Este idioma tiene ahora el nombre oficial de *bésiro*, mientras que *chiquitano* es el nombre tradicional; dado que *bésiro* aún no se emplea en todas las comunidades de hablantes, para mayor claridad me referiré al idioma usando el nombre tradicional.

son dos variedades mutuamente inteligibles y muy similares. Estas variedades se han separado en tiempos recientes, probablemente en la época de la fundación de Puerto Pacheco (ahora Bahía Negra), por los bolivianos en 1885 (Combès, 2009), en territorio chamacoco entonces disputado entre Bolivia y Paraguay.

El hecho de que los idiomas estén organizados en familias es un concepto reciente. La lingüística histórica comienza convencionalmente en 1786, cuando William Jones notó que el sánscrito, el griego antiguo y el latín derivaban de un antepasado común perdido; este, que ha sido reconstruido, era el proto-indoeuropeo. Hoy sabemos que el castellano, el portugués, el francés y el italiano descienden del latín que, a su vez, pertenece a la rama itálica de las lenguas indoeuropeas. Es menos conocido el hecho de que, poco antes de Jones, en 1783, el padre jesuita Filippo Salvatore Gilij se había dado cuenta de que el mojeño (hablado en Bolivia) y el maipure (una lengua extinta que se hablaba en el Alto Orinoco, en Venezuela) tenían una afinidad genética, descubrimiento con el que sentó las bases para la clasificación de las lenguas arawak. La familia Arawak es la que tiene más idiomas y la más extendida de toda Sudamérica: incluye unos 40 idiomas que se hablan actualmente y decenas de otros extintos (Aikhenvald, 2012); estos idiomas se extendían desde Guatemala y las islas del Caribe hasta Argentina y Paraguay. En la sección siguiente hablaremos del subgrupo de las lenguas arawak habladas en Bolivia para ilustrar el método comparativo. Otra gran familia es la de las lenguas Tupí, que incluye diez ramas. La rama más numerosa es la de las lenguas tupí-guaraní (derivadas de un proto-tupí-guaraní), que a su vez incluye ocho subgrupos (Rodrigues y Cabral, 2012); entre las lenguas Tupí-guaraní habladas en Bolivia, el sirionó, el yuki, el guarasu y el guarayu forman una sola rama, mientras que el chiriguano y el tapieté son parte de otra rama, junto con otras lenguas habladas fuera de Bolivia, entre ellas el antiguo guaraní y el guaraní paraguayo.

Existen también idiomas aislados, para los cuales no se puede identificar ningún idioma genéticamente relacionado. En Bolivia hay varias lenguas aisladas como el canichana, el cayubaba, el itonama, el leko, el movima y el yurakaré. El chiquitano es considerado por muchos como una lengua aislada, mientras que estudios recientes lo consideran como una lengua Macro-yê (Adelaar, 2008; Nikulin, 2020).

8.2. El método comparativo

En la sección anterior, vimos que la lingüística histórica clasifica las lenguas en función de su relación genética. Esto es posible gracias al método comparativo, que permite verificar el parentesco entre las lenguas, y reconstruir su antepasado común. Ilustraremos el método comparativo a través de algunas lenguas Arawak. Entre los subgrupos de esta familia, se encuentra el de las lenguas arawak habladas en Bolivia: baure, mojeño, paunaka y †paikoneka. El idioma más cercano a estos

es el *terêna*, hablado en Brasil, pero que, en el pasado, tenía una variedad llamada *chané*, que se hablaba en Bolivia (Carvalho, 2016). Podemos asumir que estos idiomas tuvieron un antepasado común. Como Carvalho llama ‘Bolivia-Paraná’ a las lenguas Arawak de Bolivia y al *terêna*, podemos llamar a la respectiva proto-lengua ‘proto-Bolivia-Paraná’ (abreviado: proto-BP). La comparación de los ancestros comunes que se pueden reconstruir para cada subgrupo de lenguas Arawak permite reconstruir el proto-arawak. Por razones de simplicidad, nos limitaremos aquí a comparar el *paunaka*, el *mojeño* y el *terêna*. Excluimos al *baure*, que parece ser el más innovador, y al *paikoneka*, que está extinto. El *mojeño* tiene dos variedades que se hablan actualmente: el *ignaciano* y el *trinitario*. Estas variedades se han separado recientemente y esto debe tenerse en cuenta en la comparación; se puede decir que ambas proceden de un proto-*mojeño* (para cuya reconstrucción, véase Carvalho & Rose, 2018); el proto-*mojeño* no debe confundirse con el antiguo *mojeño*, comparado por Gilij (1783) con el *maipure*; el antiguo *mojeño* es una variedad paralela al *ignaciano* y al *trinitario*, y en este ejemplo lo omitiremos por razones de simplicidad.

El método comparativo consiste, como su propio nombre indica, en la comparación de palabras que se consideran relacionadas, como las que se reportan en la **Tabla 1**. El objetivo es identificar una serie de correspondencias fonéticas regulares, observar en qué contexto se producen e identificar los cambios fonéticos recurrentes. Posteriormente, será posible reconstruir el inventario de sonidos en el proto-BP y reconstruir palabras individuales. Esta comparación ilustrativa está tomada de un estudio mucho más amplio de Carvalho (2018); por razones de espacio, muchos detalles han sido obviamente omitidos.

Dado que el léxico se puede prestar, debe excluirse de la comparación los préstamos de otros idiomas; por esta razón, a menudo la comparación se refiere al léxico básico (como en la Tabla 1), que se considera más resistente al préstamo.

Tabla 1

Léxico compartido entre paunaka, ignaciano, trinitario y terêna

	Paunaka	Mojeño		Terêna
		Ignaciano	Trinitario	
LUNA	<i>kuhe</i>	<i>kahe</i>	<i>kohe</i>	<i>kohêe</i>
NOCHE	<i>juti</i>	<i>jati</i>	<i>joti</i>	<i>jóti</i>
HORMIGA	<i>kusii</i>	<i>kafîru</i>	<i>kfîru</i>	<i>kosîu</i>
MANO	<i>-bui</i>	<i>-waʔu</i>	<i>-woʔu</i>	<i>-wôʔu</i>
JAGUAR	<i>Isini</i>	<i>ifîni</i>	<i>ʔfîni</i>	<i>sîni</i>
CABEZA	<i>-ʃiti</i>	<i>-ʃuti</i>	<i>-ʃuti</i>	<i>-tûti</i>
SOL	<i>safê</i>	<i>safê</i>	<i>safê</i>	<i>káfe</i>

SABER	-ifu	-eŋa	-eŋo	-êŋo
SUEGRO	-muŋiku	-imaŋuka	-imŋuko	-imóŋuko
PIEDRA	mai	mari	mari	marípa
COSTILLA	-himunepa	hirumane	-hi:monepa	-
LAVAR	-kipu	-sipa-ka	-sip-ko	-kípo
TAPIR	samu	sama	samo	kámo

El caso más simple es cuando un sonido se conserva en todas las lenguas hijas, por ejemplo *m*. Esto nos permite decir que **m* del proto-BP se mantuvo inalterado. El símbolo *** se utiliza para las formas no documentadas, pero reconstruidas.

La vocal *u* del paunaka corresponde a *o* en terêna y trinitario, y a *a* en ignaciano (véase ‘luna’, ‘noche’, ‘saber’, ‘suegro’). Aquí las dos variedades de mojeño difieren: dado que el trinitario tiene *o* como el terêna, se puede suponer que aquí el ignaciano ha innovado y del proto-mojeño al ignaciano **o > a*. El símbolo *>* indica la derivación de *a* de **o*. Este es un ejemplo de **bajamiento vocálico**. Nótese que la vocal *o* nunca aparece en las palabras del ignaciano, de hecho, el cambio fonético también ha producido un cambio fonológico, es decir, la pérdida de **/o/* (que se convierte en */a/*). Hay palabras (por ejemplo, ‘tapir’) en que */a/* está en todos los idiomas y por lo tanto el proto-BP tenía **a*. Como en ignaciano se ha perdido la diferencia fonológica entre **/o/* y **/a/*, se habla de **desfonologización**. En paunaka en cambio **o > u*, lo que es un caso de **elevación vocálica**.

La vocal *u* en el ignaciano, trinitario y terêna siempre corresponde a *i* en paunaka (por ejemplo, en ‘hormiga’, ‘mano’, ‘cabeza’), esto indica que la protolengua tenía **u*, que se anteriorizó convirtiéndose en *i* en paunaka.

Analicemos ahora las consonantes. ‘Piedra’ tiene la consonante *r* en ignaciano, trinitario y terêna, pero no en paunaka, por lo que podemos suponer que en paunaka ocurrió la **elisión** de **r*. La elisión es la pérdida de un sonido en una palabra. En ‘costilla’, la *r* del ignaciano corresponde a Ø (‘cero’), sin embargo, la presencia de *i:* en trinitario (correspondiente a *i* en ignaciano) indica que ha habido **elisión** de *r*; esto es también un ejemplo de **alargamiento compensatorio**, es decir, el alargamiento de un sonido para compensar la pérdida de un segmento contiguo.

Todas las lenguas arriba tienen *k* si la siguiente vocal viene de **o* (‘suegro’, ‘hormiga’). *k* puede sufrir un cambio si la siguiente vocal es *i* (‘lavar’), que también se puede reconstruir en proto-BP. En este contexto, *k* en terêna y paunaka corresponde a *s* en ignaciano y trinitario, por lo tanto, podemos decir que en proto-mojeño **k > s* / __ **i*; el símbolo / se utiliza para introducir el contexto en que se produce el cambio y el símbolo __ indica la posición del sonido en cuestión. El cambio **k > s* / __ **i*

se debe a la asimilación ejercida por **i*. La **asimilación** es un cambio fonético por el cual un sonido se vuelve más similar a otro. En este caso, como **i* es una vocal anterior, la consonante precedente anterioriza su punto de articulación (**k > s*); este tipo de asimilación es regresiva, ya que es ejercida por el segmento que sigue.

En ocasiones, es necesario ampliar la comparación a otros idiomas emparentados que, aunque no formen parte del mismo subgrupo, pueden preservar una situación más arcaica: por ejemplo, delante de *a*, *k* del terêna corresponde a *s* en las otras lenguas. Aquí podemos decir que la consonante original era **k*, si sabemos que ‘tapir’ es *kema* en otras lenguas arawak como bahuana y apurinã, y que en proto-campa se puede reconstruir como **kemari*. En otras palabras, **k* se conserva mejor en otras ramas de la familia Arawak.

La africada *tʃ* del paunaka, ignaciano y trinitario corresponde a *ʃ* en terêna en el contexto *___ *o, *e* (‘sol’, ‘saber’) y a *ʃ* o *t* en terêna en el contexto *___ *u* (‘cabeza’, ‘suegro’). En ambos casos podemos reconstruir **tʃ* para el proto-BP. En consecuencia, en terêna **tʃ > ʃ / ___ *o, *e, *u* (‘sol’, ‘saber’, ‘suegro’). Este es un ejemplo de **lenición**, es decir, de un cambio fonético en que el sonido sufre un debilitamiento articulatorio: aquí una africada se convierte en fricativa. El cambio de un sonido oclusivo (o africado) a fricativo es un tipo de lenición más específico que se denomina **fricativización**. El otro cambio fonético que ocurre en terêna es **tʃ > t / ___ *u* (‘cabeza’), este es un ejemplo de **desafricación**, es decir, el sonido fricativo pierde uno de sus dos componentes (también **tʃ > ʃ* puede considerarse una desafricación).

El sonido del proto-BP no siempre se conserva en al menos uno de sus descendientes. En ‘jaguar’ y ‘hormiga’, *s* del paunaka y del terêna corresponde a *tʃ* del ignaciano y trinitario. Solo con estos datos, sería natural asumir que el proto-BP tenía **s*. En realidad, Carvalho reconstruye **ts* para el proto-BP. De hecho, *s* está documentado en terêna a principios del siglo XX: la palabra para ‘jaguar’, por ejemplo, fue transcrita *tsiini*. Esto significa que el cambio fonético en terêna se ha producido en tiempos relativamente recientes. Hay otros datos (que sería demasiado largo mencionar aquí, véase Carvalho 2018, pp. 411-413) para poder reconstruir **ts*. Aquí hay que destacar que la reconstrucción de **ts* permite explicar fácilmente el cambio ocurrido en los otros idiomas: en efecto, en paunaka y terêna **ts > s / ___ *i* (se trata de una fricativización común), mientras que en ignaciano y trinitario **ts > tʃ / ___ *i*. En este último caso, el contexto es relevante porque en ignaciano y trinitario *i* provoca un tipo de asimilación denominada **palatalización**, es decir, el sonido adquiere una articulación palatal; este fenómeno es común cuando una consonante precede o sigue a una vocal o a una semivocal anterior.

Estos que acabamos de ver son solo algunos ejemplos en los que se establecen las correspondencias fonéticas entre idiomas y se reconstruyen los sonidos del proto-BP. A partir de estas correspondencias fonéticas analizadas por Carvalho (2018), podemos reconstruir algunas palabras del proto-BP. Los únicos sonidos que no se han considerado son los compartidos por todos los idiomas, que en sus respectivos contextos se pueden reconstruir fácilmente en el proto-BT, como **j* en ‘noche’.

Tabla 2

Posible reconstrucción de algunas palabras del proto-BP

	paunaka	Mojeño		terêna	proto-BP
		ignaciano	trinitario		
LUNA	<i>kuhe</i>	<i>kahe</i>	<i>kohe</i>	<i>kohêe</i>	<i>*kohe</i>
NOCHE	<i>juti</i>	<i>jati</i>	<i>joti</i>	<i>jóti</i>	<i>*joti</i>
CABEZA	<i>-ʃʉti</i>	<i>-ʃʉti</i>	<i>-ʃʉti</i>	<i>-tûti</i>	<i>*-ʃʉti</i>
SOL	<i>safɛ</i>	<i>safɛ</i>	<i>safɛ</i>	<i>káɛ</i>	<i>*kaɛ</i>
PIEDRA	<i>mai</i>	<i>mari</i>	<i>mari</i>	<i>marípa</i>	<i>*mari</i>
LAVAR	<i>-kipu</i>	<i>-sipa-ka</i>	<i>-sip-ko</i>	<i>-kípo</i>	<i>*-kipu</i>
TAPIR	<i>samu</i>	<i>sama</i>	<i>samo</i>	<i>kámo</i>	<i>*kamo</i>

Por supuesto, esta es solo una reconstrucción para ilustrar el método comparativo. Una verdadera reconstrucción requeriría un trabajo mucho más extenso: la comparación debería abarcar un número mayor de palabras, también debería incluir otros idiomas del grupo Bolivia-Paraná, como baure y paikoneka. Además, hay que considerar también los datos lingüísticos recopilados en los siglos pasados, por ejemplo: el antiguo mojeño y el baure documentado por los jesuitas y, cuando la reconstrucción es dudosa, se deben tener en cuenta los datos de otras lenguas arawak externas a este grupo.

8.3. El cambio fonético y fonológico

El ejemplo anterior nos ha permitido mostrar algunos cambios fonético-fonológicos. Hemos visto, por ejemplo, la asimilación: muchos cambios son tipos de asimilación, como la palatalización, que es muy común. La **armonía vocálica** es un tipo de asimilación a distancia: en chamacoco *ahakôr* (2SG) ‘hacer, construir’ viene de **ahokârú*: dejando de lado otros cambios, la vocal inicial *a* ha asimilado la vocal siguiente.

El **ensordecimiento** es la pérdida de sonoridad. Al principio y al final de una palabra se puede considerar un tipo de asimilación, ya que sigue o precede un silencio: el ayoreo *nue* (3) ‘ser superior, vencer’ viene de **nue*, con ensondecimiento de *n* inicial.

La **disimilación** es lo opuesto de la asimilación, pero menos frecuente. La disimilación ocurre cuando un sonido cambia para diferenciarse de otro sonido: en yuki (tupí-guaraní) *ririĩ*: ‘pequeño’, las dos róticas se han diferenciado dando **diniĩ*: que luego, debido a la elisión de las vocales, se ha convertido en *-dn*, que es un sufijo diminutivo (Villafañe, 2003, pp. 20, 73).

Hay **fusión** (o **coalescencia**) cuando dos sonidos se unen en uno solo que posee algunas características de ambos; por ejemplo, en ayoreo **baise* > *bese* (2SG) ‘encontrar, alcanzar’.

Arriba hemos visto algunos ejemplos de lenición, su opuesto es la **fortición**, que es el fortalecimiento articulatorio de un sonido; por ejemplo, el pronombre de la segunda persona plural *olak* del chamacoco corresponde a *uwak* en zamuco antiguo y ayoreo. *Olak* deriva de una forma **owak* o **uwak* del proto-zamuco, con la fortición de **w* en *l*.

La **metátesis** es el cambio de posición de los sonidos dentro de la palabra. En guarayu, desde los años 30 hasta la actualidad, se ha producido metátesis entre *r* [r] y *i* (Crowhurst & Trechter, 2014): *andavir* ‘cuervo’ > *andavri*; *potr* ‘flor’ > *potri*; *yukir* ‘sal’ > *yukri* (datos de Danielsen et al., 2019).

Otros cambios pueden estar relacionados con la adición o pérdida de sonidos. Ya hemos mencionado la elisión. Un tipo específico de elisión es la **síncopa**, que consiste en la eliminación de una vocal dentro de la palabra. Si en la Tabla 1 comparamos el trinitario *kʃiru* ‘hormiga’ y *-imʃuko* ‘suegro’ con los otros idiomas, vemos que en trinitario se ha producido una síncopa (Carvalho & Rose, 2018, pp. 19-20). La **apócope** es la caída de una vocal final, por ejemplo, en el siglo XVIII en baure se decía *ehiro* ‘hombre’, *kahapa* ‘mandioca’ y *nitipo* ‘uña’; estas palabras con el tiempo se convirtieron en *hir*, *kahap* y *nitip* (Danielsen, 2007, pp. 51-52). En *ehiro* > *hir* también podemos observar la **aféresis**, es decir, la pérdida de un sonido inicial (generalmente de una vocal).

La **epéntesis** es la inserción (o adición) de un sonido no etimológico en una palabra. Por ejemplo, en chamacoco los sustantivos masculinos cuya raíz termina en *u* toman el sufijo plural *-wo*: *oju-wo* (M.PL) ‘casas’, *du-wo* (M.PL) ‘pueblos’. El sufijo originario, que se puede reconstruir en proto-zamuco, es **-o*; la *w* fue insertada originalmente para evitar una secuencia de *u* y *o*, no permitida por la fonotáctica de la lengua. ‘Epéntesis’ tiene a menudo un significado genérico. Un tipo de epéntesis es la **prótesis**, es decir, la adición de una vocal al comienzo de la palabra. Por ejemplo, el pronombre independiente de la tercera persona singular del chamacoco, *ire*, viene de un pronombre reflexivo de tercera persona, **=re* (conservado en zamuco antiguo y ayoreo). La prótesis de *i* permitió al clítico convertirse en una palabra fonológica en chamacoco.

Hemos enumerado algunos de los cambios fonético-fonológicos más comunes. Como es de esperarse, este inventario no puede ser exhaustivo (para más detalles, véase Trask, 1996; Campbell & Mizco, 2007). El cambio de sonidos solo puede afectar la fonética o tener un impacto en la fonología de la lengua. Desde el punto de vista estrictamente fonológico, hemos visto en la sección anterior la **desfonologización** del contraste entre */o/ y */a/ en mojeño ignaciano. El fenómeno opuesto es la **fonologización**, que es el proceso mediante el cual un alófono de un fonema se convierte en un fonema independiente. Se encuentra un ejemplo de fonologización en la variedad ignaciana del chiquitano, donde se ha creado una oposición entre /p/ y /pi/: un par mínimo es *p'ooɕ* (3SG.NM) ‘su casa’ vs. *poɕ* (FG.SG) ‘casa (en general)’; *p'ooɕ* viene de *ipooɕ* (3SG.NM), una forma documentada en el siglo XVIII; después de /i/, /p/ era palatalizada y realizada como [pʲ]. Debido a la aféresis de /i/ inicial, la elección entre [pʲ] y [p] ya no depende del contexto, sino que determina ella misma la diferencia entre *poɕ* y *p'ooɕ*, y por lo tanto debemos considerar /p/ y /pi/ como dos fonemas distintos.

8.4. La comparación morfológica

La detección de una serie de correspondencias fonéticas sistemáticas entre dos o más idiomas puede ser a menudo suficiente para establecer o no una relación de parentesco. Sin embargo, a veces es necesario el estudio morfológico para determinar dichas relaciones. De hecho, el léxico se puede prestar con relativa facilidad, lo que implica un problema, ya que esto puede acarrear en muchos casos la alteración del resultado de la comparación. Es bien conocido el caso del inglés, que proviene de la rama germánica de las lenguas indoeuropeas, sin embargo, solo el 26% del léxico inglés tiene origen germánico, mientras que la mayor parte proviene, directa o indirectamente, del latín (Williams, 1975). Por tanto, la mera comparación del léxico puede inducir al error. Como bien muestra la historia del inglés, esta lengua no descende del latín. Sin embargo, comparar solamente el léxico de dos o más lenguas, podría inducirnos al error, en el caso de idiomas hablados por poblaciones de las que solo conocemos la historia reciente. Por el contrario, la morfología, y en particular la morfología flexiva, se considera la parte de la lengua más resistente al cambio lingüístico. La comparación morfológica es, pues, un elemento adicional, a veces decisivo, que permite establecer si hay afinidad genética, incluso cuando hay relativamente poco léxico común entre lenguas relacionadas.

La familia lingüística Zamuco es un ejemplo que muestra que la comparación morfológica es relevante para establecer un origen común. Por su parte, el zamuco antiguo y el ayoreo comparten la mayor parte de su léxico, mientras que el ayoreo y el chamacoco tienen solo alrededor del 30% del vocabulario en común. Sobre esta base es posible establecer correspondencias fonéticas, pero es necesario el estudio de la morfología flexiva para establecer las relaciones entre estas lenguas.

Veamos algunas formas verbales en las tres lenguas zamuco. La mayoría de los verbos del zamuco antiguo y ayoreo tienen el prefijo de tercera persona *te-* (1-2). Este prefijo es *te-* o *ts-* en chamacoco (1-2). El estudio de las correspondencias fonéticas nos dice que *ts* es una innovación del chamacoco, y en algunos verbos del chamacoco *te-* y *ts-* se alternan libremente. El zamuco antiguo y el chamacoco distinguen entre tercera persona realis e irrealis: los verbos con *te-/ts-* en la tercera persona realis tienen el prefijo *d-* para la tercera persona irrealis. Como el zamuco antiguo y el chamacoco pertenecen a dos ramas distintas de la familia lingüística, esto indica que el ayoreo ha perdido la distinción entre realis en tercera persona e irrealis y, por lo tanto, podemos reconstruir los prefijos **te-* (3.REAL) y **d-* (3.IRR) para el proto-zamuco. Un grupo más pequeño de verbos tiene el prefijo *t-* en la tercera persona, realis e irrealis, en las tres lenguas (3), por lo que podemos reconstruir **t-* para la tercera persona realis e irrealis de estos verbos en proto-zamuco.

El prefijo de persona precede un tema verbal que comienza por vocal. Cuando esta vocal es alta (1-2), se convierte en *a* o (solo en chamacoco) en *e* en la segunda persona. Este cambio se produce en todos los idiomas y, por lo tanto, ya ocurría en el proto-zamuco, en donde la vocal a reconstruir en la segunda persona es **a*, que se conserva en zamuco antiguo y ayoreo, y que generalmente corresponde a *a* o *e* del chamacoco. El zamuco antiguo y el ayoreo distinguen entre realis e irrealis en la segunda persona (1-3); el chamacoco tiene una única segunda persona singular con prefijo cero, que corresponde a la segunda persona irrealis de las otras dos lenguas. Esto quiere decir que en chamacoco la segunda persona irrealis ha reemplazado a la segunda persona realis. A veces las excepciones morfológicas ofrecen valiosas indicaciones de que una palabra proviene de la misma protolengua. En ‘sentarse’ (2), la primera persona del zamuco antiguo y del chamacoco muestran una inserción *-ij-* que mantiene una irregularidad del proto-zamuco.

- (1) ZA REAL: *a-iraha* (1SG), *d-araha* (2SG), *te-iraha* (3); IRR: *Ø-araha* (2SG), *d-iraha* (3) ‘saber, aprender’
 AY *j-iraha* (1SG), *b-araha* (2SG.REAL), *Ø-araha* (2SG.IRR), *te-iraha* (3), ‘saber, entender’
 CH *t-iraha* (1SG), *Ø-eraha* (2SG), *te-iraha* (3.REAL), *d-i-iraha* (3.IRR) ‘saber, entender’
- (2) ZA REAL: *a-ij-akare* (1SG), *d-akare* (2SG), *te-iakare* (3); IRR: *Ø-akare* (2SG), *d-iakare* (3); ‘sentarse, quedarse’
 AY: *j-akare* (1SG), *b-akare* (2SG), *te-akare* (3); IRR: *Ø-akare* (2SG) ‘sentarse, quedarse’
 CH: *t-ij-akir* (1SG), *Ø-akir* (2SG), *ts-a-kir* (3.REAL), *d-a-kihir* (3.IRR) ‘sentarse, descansar’

- (3) ZA REAL: *a-gu* (1SG), *d-agu* (2SG), *t-agu* (3); IRR: *agu* (2SG) *t-agu* (3) ‘comer algo’
 AY *j-agu* (1SG), *b-agu* (2SG), *t-agu* (3); IRR: *Ø-agu* (2SG) ‘comer, morder’
 CH *t-aw* (1SG), *Ø-ew* (2SG), *t-ew* (3.REAL), *t-ew* (3.IRR) ‘comer’

De estos ejemplos emergen así correspondencias morfológicas sistemáticas que permiten vislumbrar un origen común entre las tres lenguas, y emergerían más si se hubiera reproducido todo el paradigma (para un análisis completo, véanse los trabajos citados en la nota 1).

Cuando surge nueva información sobre un idioma, a menudo también la reconstrucción debe actualizarse; por ejemplo, en un diccionario inédito del zamuco antiguo surgieron algunos verbos con el prefijo *s-* de tercera persona realis, como *s-okâru* (3.REAL) ‘hacer algo’, a compararse con el chamacoco *ε-ijokôr* (3.REAL) ‘hacer’; dado que en chamacoco *ε* deriva de la palatalización de *s*, se puede reconstruir un prefijo **s-* (3.REAL) para un pequeño número de verbos del proto-zamuco.

8.5. El cambio morfológico

Como hemos visto en los ejemplos anteriores, también la morfología está sujeta a cambios; mostraremos aquí algunos tipos de cambio morfológico. Por **analogía** entendemos un cambio en que una forma se vuelve similar a otra para seguir un patrón más regular o, en todo caso, más común. Por ejemplo, el zamuco antiguo y el chamacoco tienen un pequeño grupo de verbos con tercera persona realis con prefijo cero y el tema que comienza en vocal alta; la tercera persona irrealis de dichos verbos tiene el prefijo *d-* con cambio de la vocal inicial de la raíz, como en el verbo ‘dar’ (4). Este grupo de verbos se puede reconstruir para el proto-zamuco, pero son poco frecuentes, ya que la mayoría de los verbos en las lenguas zamuco tiene un prefijo que deriva de **te-* (1-2). La única forma de tercera persona del verbo ‘dar’ en ayoreo se ha ‘regularizado’ por analogía con el principal grupo de verbos, y ha adoptado su prefijo de tercera persona, *te-*. La analogía, por lo tanto, constituye una excepción a las leyes fonéticas, ya que no existe correspondencia fonética regular entre *te* del ayoreo y *Ø* del zamuco antiguo y del chamacoco.

- (4) ‘dar’ ZA *Ø-isi* (3.REAL), *d-osi* (3.IRR); AY *te-isi* (3); CH *Ø-iei* (3.REAL), *d-oei* (3.IRR)

Otro cambio morfológico consiste en la reinterpretación de los límites de un morfema; este es un tipo de **reanálisis** (los procesos de reanálisis también se encuentran en la sintaxis, véase la sección 8.6.). Por ejemplo, el verbo en (5) tiene prefijo cero en la tercera persona del zamuco antiguo y del ayoreo; en chamacoco la consonante inicial de la tercera persona fue reinterpretada como prefijo de tercera persona, dada la existencia de un grupo de verbos con el prefijo *t-* (véase arriba); en consecuencia, la segmentación del verbo ha cambiado también en el resto del paradigma (en el ejemplo se puede comparar la tercera con la primera persona).

(5) ZA: *a-itiboha* (1SG.REAL), *Ø-tiboha* (3.REAL/IRR) ‘bostezar, abrir la boca’

AY: *j-itiboha* (1SG), *Ø-tiboha* (3) ‘bostezar’

CH: *tik-ibuhu* (1SG), *t-ibuhu* (3.REAL/IRR) ‘bostezar, lamer’

Arriba hemos visto que en el chiquitano ignaciano ‘su casa’ es *p'ooŋ* (3SG.NM). En el chiquitano documentado en el siglo XVIII (o chiquitano antiguo), ‘su casa’ era *i-pooŋ* (3SG.NM), con el prefijo *i-* de tercera persona, y por razones eufónicas se podía añadir una *n* inicial a los sustantivos que comenzaban en vocal, lo que permitía evitar una secuencia de dos consonantes dentro de una oración. Por lo tanto, de esta forma, se tenía a menudo *n-i-pooŋ* ‘su casa’. En el chiquitano hablado en Lomerío, ‘su casa’ es *ni-poŋ* (3SG.NM); aquí la *n* inicial fue reinterpretada como parte del prefijo de tercera persona, que se convirtió en *ni-*.

La **gramaticalización** también puede considerarse un tipo de reanálisis: las palabras autónomas dotadas de un significado léxico se vacían semánticamente y se convierten en morfemas vinculados con significado gramatical. Por ejemplo, en muchas lenguas del mundo el diminutivo a menudo proviene de la gramaticalización de la palabra ‘niño’: en zamuco antiguo y ayoreo el sufijo del diminutivo es *-ap*, este es el resultado de la gramaticalización de la raíz de ‘niño/a’, *ap*, que también se puede reconstruir para el proto-zamuco. El chiriguano tiene un sufijo diminutivo *-rai* que viene de la gramaticalización de *tái* ‘hijo’ (Dietrich, 1986, p. 177). En todas las variedades de mojeño (mojeño antiguo, trinitario e ignaciano) el sufijo diminutivo *-ŋiŋa* deriva de ‘pequeño’ (*aŋiŋa* en mojeño antiguo) o más probablemente de *ŋiŋa* ‘hijo/a’ (documentado en las tres variedades). El mojeño trinitario creó un nuevo diminutivo *-çira*, que resulta de la gramaticalización de *çira* ‘semilla’ (Rose, 2018). Arriba mencionamos el diminutivo del yuki, que proviene de ‘pequeño’.

Las partes del cuerpo pueden convertirse en adposiciones espaciales: en ese ejja (*e-*)*wasi* ‘pie’ fue gramaticalizado en *=wasiije* ‘hacia’ y (*e-*)*jakajja* ‘hombro’ en *=jakajjajje* ‘atrás’ (Vuillermet, 2012, p. 90). En tapiete, el aspecto habitual se expresa mediante el sufijo *-pi* que deriva del adverbio *yepi* ‘siempre’, abreviado durante el proceso de gramaticalización (González, 2005, p.158).

8.6. El cambio sintáctico

A menudo, la sintaxis se considera como la parte de la gramática que cambia con mayor facilidad. En el caso de lenguas poco documentadas, es difícil investigar el cambio sintáctico, ya que a menudo no se dispone de textos recopilados en los siglos pasados para compararlos con textos actuales. La morfología puede ser el resultado de un cambio sintáctico, y por eso, como escribió Givón (1971, p. 413), la morfología de hoy es la sintaxis de ayer.

A continuación, mostramos un ejemplo de cambio sintáctico que tuvo consecuencias a nivel morfológico. El wichí (mataguayo) tiene dos grupos dialectales: el pilcomayeño y el bermejeño (Nercesían, 2019). En el wichí pilcomayeño los verbos derivados con un aplicativo expresan el objeto con un sufijo (6a), mientras que los otros verbos tienen un prefijo para el objeto (6b) (excepto en la primera persona).

(6) Wichí pilcomayeño (Nercesían, 2019)

- | | |
|---|---|
| a. <i>o-yämlhi-'am-ho</i>
1SUJ-hablar-2OBJ-APL.ASOC
'Te hablo a vos.' | b. <i>n-a-iw'en</i>
1SUJ-2OBJ-ver
'Te veo.' |
|---|---|

Para analizar el origen del sufijo del objeto, Nercesían (2019) compara el wichí con las otras lenguas mataguayo y analiza las atestaciones de la lengua a partir de finales del siglo XIX. Nercesían reconstruye las siguientes fases:

(i) En wichí el sufijo del objeto se desarrolló a partir de construcciones de verbos seriales asimétricas en que un pronombre libre era el objeto del verbo 1 y el sujeto del verbo 2, que era un verbo intransitivo direccional o locativo: [[SUJETO-VERBO_1] [PRONOMBRE_LIBRE VERBO_2]].

(ii) Posteriormente, tuvo lugar un proceso de **reanálisis**: el verbo 2 fue reinterpretado como posposición aglutinada al pronombre libre, que pasó a ser el objeto del único verbo que quedó: [[SUJETO-VERBO] [OBJETO-POSPOSICIÓN]].

(iii) El objeto y la postposición se aglutinaron al verbo y la posición fue reanalizada como aplicativo: [SUJETO-VERBO-OBJETO-APLICATIVO]. De esta manera, un pronombre libre se convirtió en sufijo.

Un cambio se puede difundir posteriormente más allá del contexto original. Por ejemplo, el wichí bermejeño (7) extendió el sufijo del objeto a todos los verbos: véase (6b) y (7b).

(7) Wichí bermejeño (Nercesían, 2019)

- | | |
|--|---|
| a. <i>n'-yomlhi-'am-hu</i>
1SUJ-hablar-2OBJ-APL.ASOC
'Te hablo a vos.' | b. <i>n'-w'en-'am</i>
1SUJ-ver-2OBJ
'Te veo.' |
|--|---|

La sintaxis puede cambiar debido al contacto con otra lengua, y estos cambios a veces son fáciles de reconocer. Por ejemplo, en el chiquitano lomeriano, los hablantes suelen usar algunas conjunciones del castellano; las oraciones se coordinan con *y*, mientras que *pero* y *porque* a menudo reemplazan a *tapi* 'pero' e *it'opiki* 'porque'.

8.7. El cambio léxico

El léxico puede ser sustituido por uno nuevo, muchas veces prestado de otro idioma: en ayoreo ‘perro’ es *tamokoj*; esta palabra no se encuentra en ninguna otra lengua zamuco y fue tomada del chiquitano *tamokoʃ* ‘perro’. El contacto no se limita al ayoreo y al chiquitano, ya que la palabra *tamuku* ‘perro’ está documentada en terêna, mojeño antiguo y mojeño ignaciano (Combès & Hirtzel, 2007), mientras que el yurakaré tiene *chajmu* ‘perro’ (Van Gijn, 2006).

Un cambio semántico ocurre cuando un lexema cambia de significado con el tiempo. Un cambio de significado puede ocurrir por contigüidad. Por ejemplo, en chiquitano antiguo *aruʃ* (3SG.NM) significaba ‘su labio’ y este significado se mantiene en el chiquitano ignaciano y migueleño; en la variedad de Lomerío, en cambio, *aruʃ* significa ‘boca’: dado que el labio rodea la boca, *aruʃ* designa una parte del cuerpo contigua a su significado original.

El cambio semántico también puede producirse por la similitud entre dos conceptos. En este caso hablamos de transferencia metafórica, que a menudo da lugar a casos de polisemia. Por ejemplo, la transferencia metafórica de ‘ojo’ a ‘hueco’ está documentada en muchas lenguas del mundo; entre ellas se encuentran las lenguas zamuco y el chiquitano, donde hay polisemia entre ‘ojo’ y ‘hueco’. El zamuco antiguo *edo* ‘ojo/hueco’ y el chiquitano antiguo *sutoʃ* ‘ojo/hueco’ han ampliado aún más su significado a ‘ventana’.

Cada población tiene tabúes lingüísticos, es decir, palabras, expresiones o temas que se tienden a evitar. En muchas poblaciones, entre ellas, por ejemplo, los ayoreos, se prefiere evitar pronunciar el nombre de los muertos. Un tipo de tabú muy común se refiere a entidades que se perciben como peligrosas o pertenecen a la esfera de lo sagrado. Esta interdicción lingüística a menudo produce un cambio léxico, porque la entidad tabú se menciona indirectamente. Por ejemplo, en el chiquitano de Lomerío, el dueño de los animales, llamado *nifi jiritʉʃ* ‘jichi del cerro’ o *nifi niuʃ* ‘jichi del monte’, protege a los animales silvestres y puede ser muy peligroso si un cazador mata a demasiados animales. El dueño de los animales también se llama *aʃti pisiʃ* ‘el negro’.³ El uso de *pisiʃ* ‘negro’ para referirse al dueño sirve para evitar designarlo directamente, y esto tiene que ver con un tabú lingüístico, que produce una extensión del significado de *pisiʃ* ‘negro’, y luego, con el tiempo, este proceso puede llevar al reemplazo de las palabras originales. Un caso similar es el de ‘jaguar’ en ayoreo, cuya palabra original es *putugutoj*; en el zamuco antiguo se tiene *putuguto* ‘jaguar’; el ayoreo ha introducido un sinónimo para el mismo animal: *karataj*, que es al mismo tiempo el adjetivo ‘rojo’ y se ha utilizado para referirse al jaguar de forma indirecta, hasta convertirse en la forma ahora más utilizada para su denominación.

3 *Aʃti* ‘el’ es propio del habla del hombre; en el habla de las mujeres se usa *tʉne* ‘el’.

8.8. Cómo y por qué ocurre el cambio lingüístico

Son muchos los factores que determinan el cambio lingüístico. Aquí solo es posible enumerar algunos de ellos. Aunque el carácter de un pueblo, las características anatómicas de una determinada población, el clima y la geografía se han mencionado con frecuencia en el pasado, no hay evidencia de que estos factores ejerzan una influencia sobre el cambio lingüístico.

La lengua es un producto de la sociedad y su cambio depende de factores sociales. El cambio lingüístico es a menudo el producto consciente o inconsciente de un pequeño grupo de personas y luego se extiende a otros hablantes. El lenguaje tiene también entre sus funciones la de expresar la identidad de una comunidad o de un grupo de personas unidas, tal como por la edad o pertenencia social; estas características pueden afectar elementos poco visibles de la lengua. Por ejemplo, en Lomerío y Concepción se habla la misma variedad de chiquitano, sin embargo, hay pequeñas diferencias: en Lomerío se dice *nikoroş* ‘letra’, *noifoşoş* ‘víbora’ y *pakuş* ‘plátano’, mientras que en Concepción estas palabras se pronuncian *nikuruş*, *nuişuşuş* y *pakuşuş*. Quien documenta un idioma en peligro nota a menudo diferencias considerables según la edad de los hablantes; esto no depende solo del hecho de que la lengua en cuestión esté influenciada por otra lengua (como el castellano), sino también de que estos rasgos lingüísticos innovadores expresen, en un momento de cambio social, la pertenencia de los hablantes a un determinado grupo etario.

El cambio lingüístico suele coincidir con un cambio repentino en la sociedad debido, por ejemplo, a una migración, una epidemia o una invasión. Estos acontecimientos producen cambios en las relaciones sociales y pueden, por ejemplo, cambiar la cohesión social o producir la fusión de dos o más poblaciones. Una densa red de relaciones resulta en una mayor cohesión social y resistencia al cambio, como ocurre en pequeñas comunidades de hablantes. Por el contrario, en las grandes comunidades, donde los vínculos son menos estrechos y hay una apertura al contacto externo, el cambio lingüístico es más fácil, ya que la lengua expresa pertenencia o no a una determinada comunidad o grupo de personas.

El contacto con otras lenguas contribuye notablemente al cambio lingüístico. A menudo, dos lenguas permanecen en contacto durante mucho tiempo influyéndose mutuamente, mientras que a veces una termina sustituyendo a la otra. Cuando dos o más poblaciones se unen en un momento de cambio social, frecuentemente adquieren el idioma de la población mayoritaria o dominante; la lengua hablada previamente por cada grupo minoritario se denomina sustrato; aunque esta lengua desaparece con el tiempo, ella interfiere con la lengua aprendida, modificándola. En las Misiones de Chiquitos, la lengua más hablada era la que hoy llamamos chiquitano antiguo, por lo que fue elegido como idioma ‘oficial’ por los jesuitas. En las misiones se hablaban

también muchas otras lenguas no emparentadas con el chiquitano, así como el baure, el zamuco antiguo y el caypotorade, que era una variedad zamuco probablemente cercana al chamacoco actual. Estos sustratos han ejercido una influencia sobre el chiquitano y son uno de los factores responsables de las diferentes variedades de chiquitano que se hablan actualmente en el territorio de cada misión. Asimismo, el zamuco antiguo, antes de desaparecer, sufrió también cambios debido a la influencia del chiquitano.

El cambio lingüístico puede ser también la consecuencia de la separación de dos comunidades, que primero desarrollan dos dialectos diferentes del mismo idioma y luego dos idiomas distintos. Esto es probablemente lo que sucedió hace mucho tiempo cuando el proto-zamuco se dividió en dos ramas (ver sección 8.1.).

Otro posible factor de cambio es el aprendizaje del idioma por parte de los niños. En efecto, se cree que el niño aprende la lengua hablada por los padres, pero nunca la imita a la perfección, produciendo pequeños cambios en ella.

La frecuencia de una construcción o de una palabra también influye en el cambio lingüístico: las expresiones muy frecuentes se acortan con el tiempo. Un caso común es el de los saludos. En el caso del chiquitano de Lomerío, la fórmula de saludo utilizada por los hombres, *ʃamuʃaume*, viene de la abreviatura de *ʃama* (mucho) *uʃia* (bien) *aume* (a vosotros). En el siglo XVIII se documentó la expresión de saludo *ʃama oʃima aemo* (mucho bien a_tí), aún no abreviada.⁴ También la palabra *uʃia* ‘bien, bueno’, que se usa muy a menudo, es más corta que la forma *oʃima*, utilizada en el mismo siglo. De hecho, si una expresión o una palabra tiene un uso frecuente, es más fácil de entender y esto permite simplificarla. Además, con una alta frecuencia de uso se tiende a automatizar el movimiento articulatorio; se hace así un menor esfuerzo articulatorio, lo que resulta en una simplificación fonética. A nivel gramatical, en cambio, la repetición continua favorece la memorización por parte de los hablantes y, por lo tanto, el mantenimiento de irregularidades. Por eso los verbos más utilizados suelen ser irregulares (piénsese en el verbo ‘ser’ del castellano); por el contrario, las formas poco utilizadas tienden a regularizarse con el tiempo.

8.9. Cómo se produce el cambio lingüístico

Un cambio fonético o morfológico se manifiesta principalmente de manera gradual, se extiende inicialmente a algunas palabras, luego puede hacerlo más rápidamente, pero no necesariamente se extiende a todo el léxico. Hemos visto anteriormente que el chamacoco mantiene una diferencia entre la tercera persona realis e irrealis en la mayoría de los verbos: por ejemplo, *no* (3.REAL) ~ *do* (3.IRR) ‘ir’. Actualmente, entre los hablantes de mediana edad o ancianos, se observa una

4 Ya en el siglo XVIII, *aemo* ‘a tí’ es la forma de segunda persona singular, *aume* la de segunda persona plural.

tendencia a la pérdida de esta oposición en algunos verbos, por lo que a veces en lugar de *do* (3.IRR) encontramos *no* (3.REAL). La pérdida de la oposición entre la tercera persona realis e irrealis se está generalizando entre los hablantes más jóvenes, y este es un rasgo criticado por los hablantes más ancianos. De hecho, el cambio lingüístico suele ser visto con recelo, y mientras un grupo de hablantes (en este caso los jóvenes) modifica la estructura de la lengua, otros grupos de hablantes tienen una actitud purista hacia el cambio, que es vista como una forma ‘incorrecta’ de hablar. La lingüística hace tiempo abandonó la actitud prescriptiva, es decir, su única tarea es observar lo que sucede sin emitir un juicio.

En muchos idiomas existe una aversión hacia la introducción de nuevos elementos que sean claramente identificables como provenientes de otra lengua. Por ejemplo, al comienzo del contacto entre el ayoreo y el castellano, los hablantes intentaron evitar los préstamos del castellano, utilizando expresiones como *naiṇane igiṇaj* (chamanes casa) para ‘hospital’, *uguteade pi* (cosas medio_de transporte) para ‘mochila’. En la actualidad, tanto la expresión ayoreo como la palabra castellana se encuentran en competencia de uso, así como *no* (3.REAL) y *do* (3.IRR) ‘ir’ compiten por la tercera persona irrealis; si *no* se impusiera definitivamente, ‘ir’ perdería la oposición entre la tercera persona realis e irrealis, como sucedió con todos los verbos del ayoreo. En otras palabras, el cambio lingüístico del futuro, aunque no sea siempre predecible, está en la variación lingüística de hoy.

Si los hablantes a lo largo del tiempo han elegido entre dos o más formas en competencia, esta elección ha sido dictada por las dinámicas sociales que operaron dentro del pasado de una determinada población, que no siempre son conocidas y que son el objeto de estudio de la sociolingüística. En algunas ocasiones, se elige una forma que proviene de las clases altas porque es prestigiosa; en otras, por diferentes razones, se elige una forma utilizada por las clases bajas o no considerada estándar (en este caso se habla de ‘prestigio encubierto’). La variación lingüística se basa en factores sociales, tales como la clase social, la edad y el sexo. Los primeros estudios en este ámbito se deben al lingüista norteamericano Labov, quien desde los años sesenta ha investigado la relación entre la sociedad y el cambio lingüístico (véase Labov, 1994-2010).

El cambio lingüístico se asocia a menudo a una simplificación a nivel estructural (como lo fonético o morfológico), pero el cambio puede producir una mayor complejidad. Por ejemplo, en el zamuco antiguo y en el ayoreo se distingue una primera persona singular y plural. El chamacoco creó primero una distinción entre primera persona plural exclusiva e inclusiva, y luego una forma de ‘plural mayor’ de la primera persona inclusiva.

- (8) ZA REAL: *a-iraha* (1SG), *a-iraha-go* (1PL) ‘saber, aprender’
 AY *j-i-raha* (1SG), *j-iraha-go* (1PL) ‘saber, entender’
 CH *t-iraha* (1SG), *oj-iraha* (1PL.EXCL), *j-iraha* (1PL.INCL), *j-iraha-lo* (1PLM.INCL)
 ‘saber, entender’

El plural mayor es una característica inusual en las lenguas del mundo y expresa un número muy grande de referentes y/o la totalidad de los posibles referentes. Ciertamente, estos cambios han aumentado la complejidad del sistema verbal del chamacoco y se deben al contacto lingüístico con otras lenguas vecinas.

8.10. Los límites del método comparativo

Mediante el método comparativo, podemos retroceder en el tiempo unos milenios y reconstruir la proto-lengua de la que se originó una familia lingüística, sin embargo, hay límites temporales más allá de los cuales no es posible ir.

A veces se pueden encontrar formas similares en lenguas muy distantes y genéticamente distintas: muchas lenguas sudamericanas tienen una raíz para ‘mano’ que comienza con $(V)mV$. En Bolivia, por ejemplo, el puquina (aislado) tiene *muqa* ‘mano’, el itonama (aislado) *-ma* ‘para’ ‘mano’; **maki* ‘mano’ fue reconstruido para el proto-quechua y **e-me* ‘mano’ para el proto-takana (datos de Zamponi, 2020). No hay datos suficientes para reconstruir la forma exacta de ‘mano’ en proto-zamuco, pero podemos decir que comenzaba con la sílaba **ma* o **ma*. Greenberg (1987) ya señaló similitudes léxicas de este tipo con el objetivo de demostrar que todas las lenguas indígenas americanas, con excepción de las lenguas na-dené y esquimo-aleutianas, según él, tendrían una relación genética distante. Esta hipótesis no es aceptada porque Greenberg cometió serios errores metodológicos y muchos de los datos que reportó son incorrectos. Sin embargo, del trabajo de Greenberg a veces emergen similitudes notables entre lenguas distantes y no relacionadas, como la que concierne a ‘mano’. Esta semejanza se puede explicar en parte por casualidad, dado que en una lengua el número de sonidos y de sus combinaciones es limitado; de hecho, incluso *mano* del castellano comienza con $(V)mV$, pero no tiene relación ni genética ni de contacto con las palabras mencionadas anteriormente.

A veces dos palabras pueden ser similares porque imitan un sonido (en este caso se habla de onomatopeya). En ayoreo el sonido que se produce golpeando es *tok tok*, en chamacoco *tox tox*; aunque estos idiomas están relacionados, la similitud entre *tok tok* y *tox tox* puede deberse simplemente al hecho de que reproducen un sonido externo, por lo que este dato es poco relevante para la comparación histórica.

Con relación al caso citado anteriormente de ‘mano’, Zamponi (2020) señala que, si consideramos solo las lenguas aisladas y las familias lingüísticas, las formas para ‘mano’ que comienzan con $(V)mV$ son mucho más frecuentes en Sudamérica que

en el resto del mundo, y esta diferencia es estadísticamente significativa. Esto puede deberse al hecho de que se ha prestado ‘mano’ de un idioma a otro o al hecho de que algunos de los idiomas en cuestión pueden tener una relación genética distante. En casos como este, en que el contacto se produjo hace mucho tiempo, o en que las lenguas se separaron en tiempos remotos, es prácticamente imposible distinguir entre préstamo y rastros de un origen común. Esto, por supuesto, suponiendo que podamos saber cuándo la presencia de dos formas similares para ‘mano’ se debe a la casualidad, lo que de hecho es imposible para el estado de nuestro conocimiento actual.

La morfología también se puede tomar prestada, aunque con mayor dificultad que el léxico y en circunstancias de contacto areal muy intenso. En algunos casos, utilizar el léxico de otra lengua es tabú para los hablantes, por lo que se pueden prestar elementos gramaticales, pero poco léxico (Aikhenvald, 2012), hecho que sucedió en Bolivia en las áreas de Guaporé-Mamoré y del Chaco (Epps, 2020). Por esto, la lingüística histórica debe tener en cuenta el contexto histórico-cultural y sociolingüístico de las lenguas que estudia. Las lenguas zamuco comparten algunos morfemas con las lenguas circundantes, probablemente debido al préstamo morfológico y no a un origen común; solo cuando el morfema no puede reconstruirse en las proto-lenguas de dos familias distintas, el contacto puede ser demostrado sin duda.

Sin embargo, puede haber casos en que sea difícil distinguir entre casualidad, préstamo morfológico o herencia genética. Por ejemplo, en varios idiomas del Chaco y de la meseta brasileña se puede encontrar la presencia de *i/j* asociado con la primera persona y *a* asociado con la segunda. Greenberg (1987) muestra como ejemplo ‘casa’ en chiquitano antiguo: *i-poo* (1SG) ‘mi casa’, *a-poo* (2SG) ‘tu casa’, que se compara con el chorote (mataguayo) *i-jet* (1SG) ‘mi casa’, *a-wet* (2SG) ‘tu casa’; en zamuco antiguo tenemos *j-igeda* (1SG) ‘mi casa’, *a-igeda* (2SG) ‘tu casa’. El mismo patrón, para el sujeto o el poseedor, se puede encontrar en wichí y en las lenguas mataguayo no habladas en Bolivia, en toba y en otras lenguas guaycurú, en macro-yê (el chiquitano es una posible lengua macro-yê, véase arriba) y en otros idiomas hablados en Brasil y Paraguay (véase datos en Zamponi, 2017, pp. 50-51). Asimismo, en este contexto no puede atribuirse la distribución geográfica de estas características únicamente al azar, ya que el contacto lingüístico y/o la herencia genética pudieron haber jugado un rol en su difusión; en ocasiones, puede ser difícil determinar qué factor intervino y para qué idiomas. Esta y muchas otras similitudes entre las lenguas mataguayo y guaycurú se han atribuido a un origen común (Viegas Barros, 2013) o al contacto lingüístico (Comrie et al., 2010).

8.11. Una breve nota sobre la toponimia

El estudio de los topónimos a veces puede brindarnos información valiosa sobre un idioma determinado. Por ejemplo, en Cerdeña (Italia), el estudio sistemático de los topónimos nos permite reconstruir algunos elementos del paleo-sardo, una lengua sobre la que no existe documentación. El estudio de los topónimos permite explorar el sustrato lingüístico de una región. Aunque el chiquitano tiene actualmente unos miles de hablantes, el idioma se ha mantenido en muchos topónimos en el departamento de Santa Cruz, que muestran la presencia del sustrato chiquitano. La difusión de los topónimos aymaras nos muestra que esta lengua en el pasado tuvo una distribución geográfica diferente (Cerrón-Palomino & Carvajal Carvajal, 2009, p.172). El puquina es una lengua extinta sobre la que existe una documentación muy limitada, muchos topónimos en *-baya* y *-laque* provienen del puquina (como *Quiabaya*, *Italaque*). El estudio de los topónimos de origen puquina aún no se ha abordado, pero puede contribuir a ampliar nuestro conocimiento sobre esta lengua (Adelaar & van der Kerke, 2009, p.127).

8.12. Conclusiones y perspectivas para la lingüística histórica en Bolivia

En este capítulo hemos ilustrado algunos fundamentos de la lingüística histórica, hemos mostrado cómo el método comparativo permite reconstruir proto-lenguas para las que no se dispone de documentación, pero también hemos evidenciado los límites del método comparativo. De hecho, la falta de datos suficientes puede dificultar la comparación, y la reconstrucción con frecuencia puede estar sujeta a cambios, a medida que se disponga de nuevos datos. En los últimos años se han producido muchos trabajos descriptivos sobre las lenguas habladas en Bolivia, que permitirán nuevos estudios desde el punto de vista diacrónico, y que en parte ya lo han permitido, como lo señalan muchos ejemplos de este capítulo. Para la familia Takana, existen las reconstrucciones fonológicas del proto-takana de Girard (1971), Key (1968) y Key et al. (1992), pero no se conocen reconstrucciones morfológicas. En los últimos años han aparecido descripciones nuevas y más detalladas de las lenguas takana, como el cavineño (Guillaume, 2008), el ese ejja (Vuillermet, 2012) y el araona (Emkow, 2019). Estos datos posibilitan nuevos estudios diacrónico sobre la familia en cuestión. Se aplican consideraciones similares a las lenguas Arawak, aunque es una familia que se extiende mucho más allá de Bolivia. En cuanto a la reconstrucción del proto-arawak, el trabajo de referencia es de Payne (1991), pero mientras tanto se han publicado numerosas gramáticas de lenguas Arawak (para Bolivia, se puede mencionar el baure [Danielsen, 2007]). Asimismo, debido a estos nuevos trabajos, Aikhenvald ha estudiado desde el punto de vista diacrónico algunos morfemas que se pueden reconstruir para el proto-arawak (Aikhenvald, 2018; 2021), sin embargo, el avance de estos estudios muestra la necesidad de actualizar el trabajo de Payne (1991).

Las relaciones genéticas entre las lenguas a veces son inciertas, lo que afecta también a muchos idiomas bolivianos. El puquina tiene una clasificación incierta; se ha formulado la hipótesis de una relación genética distante con las lenguas arawak, pero es apenas una hipótesis (Adelaar y van der Kerke, 2009). Para las lenguas quechua se ha propuesto una relación genética distante con la familia Aymara. Hay muchas semejanzas entre las dos familias, pero no se ha podido probar dicha hipótesis, ya que tampoco en este caso es fácil distinguir entre préstamo y parentesco genético (Cerrón-Palomino & Carvajal Carvajal, 2009, p.172). Consideraciones similares se aplican a la familia Takana y a la familia Pano, entre las cuales existe una probable relación genética, sin embargo, sobre la que no hay consenso por parte de los estudiosos (Valenzuela & Guillaume, 2017, pp. 28-29).

En conclusión, aún quedan muchas preguntas abiertas. Trabajos descriptivos, recientes o nuevos, nos permitirán trasladar la reflexión del nivel sincrónico al diacrónico. En este sentido, aunque la lingüística histórica mira hacia el pasado, el futuro sigue siendo rico en perspectivas para los estudios diacrónicos sobre las lenguas indígenas de Bolivia.

Referencias

- Acosta, J. de S. J. (ed.) (1584). *Doctrina Christiana y catecismo para instrucción de los Indios*. Antonio Ricardo.
- Adam, L. & Henry, V. (1880). *Arte y vocabulario de la lengua chiquita*. Maisonneuve y Cia.
- Adelaar, W. (2008). Relações externas do Macro-Jê. O caso do Chiquitano. En Stella Telles & Aldir S. de Paula (eds.), *Topicalizando Macro-Jê*. 9-27. NECTAR.
- Adelaar, W. & van der Kerke, S. (2009). Puquina. En M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I. Ámbito andino*. 125-146.
- Aikhenvald, A. Y. (2012). *The Languages of the Amazon*. Oxford University Press.
- Aikhenvald, A. Y. (2018). Disentangling a versatile prefix: the nature and development of a polysemous marker in Arawak languages. *International Journal of American Linguistics*, 84, 1. 1-49.
- Aikhenvald, A. Y. (2021). Removing the owner: non-specified possessor marking. En Luca Ciucci (ed.), *From fieldwork to reconstruction*, número especial de *Studia Linguistica*.
- Bertinetto, P. M., Ciucci, L. & Farina, M. (2019). Two types of morphologically expressed non-verbal predication. *Studies in Language* 43, 1. 120-194.

- Campbell, L. & Mixco, M. J. (2007). *A glossary of historical linguistics*. Edinburgh University Press
- Carvalho, F. O. de. (2018). Terena, Chané, Guaná and Kinikinau are one and the same language: setting the record straight on southern Arawak linguistic diversity. *LIAMES* 16, 1. 39-57.
- Carvalho, F. O. de. (2018). The historical phonology of Paunaka (Arawakan). *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 13, 2. 405-428.
- Carvalho, F. O. de & Rose, F. (2018). Comparative reconstruction of proto-mojeño and the phonological diversification of mojeño dialects. *LIAMES* 18, 1. 7-48.
- Cerrón-Palomino, R. & Carvajal Carvajal, J. (2009). Aimara. En M. Crevels & P. Muysken 2009. *Lenguas de Bolivia. Tomo I. Ámbito andino*. 169-213.
- Chomé, I. (1958) [ante 1745]. Arte de la lengua Zamuca. Présentation de Suzanne Lussagnet. *Journal de la Société des Américanistes de Paris* 47. 121-178.
- Ciucci, L. (2016). *Inflectional morphology in the Zamucoan languages*. Centro de Estudios Antropológicos de la Universidad Católica.
- Ciucci, L. (2018). Lexicography in the Eighteenth-century Gran Chaco: the Old Zamuco Dictionary by Ignace Chomé. En Jaka Čibej, Vojko Gorjanc, Iztok Kosem & Simon Krek (eds.), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. 439-451. Ljubljana University Press.
- Ciucci, L. (2020). Matter borrowing, pattern borrowing and typological rarities in the Gran Chaco of South America. En Francesco Gardani (ed.), *Borrowing matter and pattern in morphology*, número especial de *Morphology*.
- Ciucci, L. (2021). How historical data complement fieldwork: new diachronic perspectives on Zamucoan verb inflection. En Luca Ciucci (ed.), *From fieldwork to reconstruction*, número especial de *Studia Linguistica*.
- Ciucci, L. & Bertinetto, P. M. (2015). A diachronic view of the Zamucoan verb inflection. *Folia Linguistica Historica* 36, 1. 19-87.
- Ciucci, L. & Bertinetto, P. M. (2017). Possessive inflection in Proto-Zamucoan: a reconstruction. *Diachronica* 34, 3. 283-330.
- Ciucci, L. & Macoñó Tomichá, J. (2018). *Diccionario básico del chiquitano del Municipio de San Ignacio de Velasco*. Ind. Maderera San Luis SRL/Museo de Historia UAGRM.
- Ciucci, L. & Pia, G. E. (2019). Linguistic taboos in Ayoreo. *The Mouth* 4. 31-54.
- Combès, I. (2009). *Zamucos*. Instituto de Misionerología.

- Combès, I. & Hirtzel, V. (2007). Apuntes sobre los tamacocis. *Anuario de estudios bolivianos archivísticos y bibliográficos* 13. 241-267.
- Comrie, B., Golluscio, L. A., González, H. A. & Vidal, A. (2010). *El Chaco como área lingüística*. En Zarina Estrada Fernández & Ramón Arzápalo Marín (eds.), *Estudios de lenguas amerindias 2: contribuciones al estudio de las lenguas originarias de América*, 85-131. Editorial Unison.
- Crevels, M. & Muyken, P. (2009). *Lenguas de Bolivia. Tomo I. Ámbito andino*. Plural Editores.
- Crowhurst, M. & Trechter, S. (2014). Vowel-rhotic metathesis in Guarayu. *International Journal of American Linguistics* 80, 2. 127-173.
- Danielsen, S. (2007). *Baure. An Arawak language of Bolivia*. CNWS Publications.
- Danielsen, S., Sell, L. & Terhart, L. (2019). Guarayu. A revised dictionary by Alfred Hoeller. *Dictionaria* 7. 1-3590.
- Dietrich, W. (1986). *El idioma chiriguano. Gramática, textos, vocabulario*. Instituto de Cooperación Iberoamericana.
- Emkow, C. (2019). *A grammar of Araona*. Lincom Europa.
- Epps, P. L. (2020). Amazonian linguistic diversity and its sociocultural correlates. En Mily Crevels, & Pieter Muysken (eds.), *Language dispersal, diversification, and contact: a global perspective*. Oxford University Press.
- Gilij, F. S. (1783). *Saggio di storia americana*. Tomo III. Luigi Perego.
- Girard, V. (1971). *Proto-Takanan phonology*. University of California Press.
- Givón, T. (1971). Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistic Society*, 7. 394-415.
- González, H. A. (2005). *A grammar of Tapiete (Tupi-Guarani)*. [Tesis doctoral. University of Pittsburgh].
- Greenberg, J. H. (1987). *Language in the Americas*. Stanford University Press.
- Guillaume, A. (2008). *A Grammar of Cavineña*. Mouton de Gruyter.
- Key, M. R. (1968). *Comparative Takanan phonology with Cavineña phonology and notes on Pano-Takanan*. Mouton.
- Key, M. R., Tugwell, R. M. & Wessels, M. (1992). Araona correspondances in Takanan. *International Journal of American Linguistics* 58, 1. 96-117.
- Labov, W. (1994-2010). *Principles of linguistic change*. 3 tomos. Wiley Blackwell.

- Nercesían, V. (2019). Variación dialectal y diacrónica del objeto pronominal en wichí/weenhayek (mataguaya): paradigmas prefijante y sufijante. *Cuadernos de Lingüística de El Colegio de México* 6, 1.
- Nikulin, A. (2018/2019). ¡Manityaka au r-ózura! *Diccionario básico del chiquitano migueleño. El habla de San Miguel de Velasco y de San Juan de Lomerío*. Santa Cruz de la Sierra. Ms.
- Nikulin, A. (2020). *Proto-Macro-Jê: um estudo reconstrutivo*. [Tesis doctoral. Universidade de Brasília].
- Payne, D. L. (1991). A classification of Maipuran (Arawakan) languages based on shared lexical retentions. En Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*. 355-499. Mouton de Gruyter.
- Rodrigues, A. D. & Cabral, A S. A. C. (2012). Tupían. En Lyle Campbell & Verónica Grondona (eds.), *The Indigenous languages of South America. A comprehensive guide*, 495-574. Mouton de Gruyter.
- Rose, F. (2018). The rise and fall of Mojeño diminutives through the centuries. *Studies in Language* 42, 1. 146-181.
- Trask, R. L. (1996). *A dictionary of phonetics and phonology*. Routledge.
- Valenzuela, P. & Guillaume, A. (2017). Estudios sincrónicos y diacrónicos sobre lenguas Pano y Takana: una introducción. *Amerindia* 39, 1. 1-49.
- Van Gijn, R. (2006). *A grammar of Yurakaré*. Tesis doctoral. Radboud Universiteit.
- Viegas Barros, P. J. (2013). La hipótesis de parentesco Guaicurú-Mataguayo: estado actual de la cuestión. *Revista Brasileira de Linguística Antropológica* 5, 2. 293-333.
- Villafañe, L. (2004). *Gramática yuki: lengua tupí-guaraní de Bolivia*. Ediciones del Rectorado, Universidad Nacional de Tucumán.
- Vuillermet, M. (2012). *A grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. Tesis doctoral. Université Lyon 2.
- Williams, J. M. (1975). *Origins of the English language: a social and linguistic history*. Free Press.
- Zamponi, R. (2017). First-person *n* and second-person *m* in Native America: a fresh look. *Italian Journal of Linguistics* 29. 189-230.
- Zamponi, R. (2020). Some precontact widespread lexical forms in the languages of Greater Amazonia. *International Journal of American Linguistics* 86, 4. 527-573.

CAPÍTULO 9

Cómo citar: Gallagher, G. (2023). Lengua y adquisición. En P. Alandía Mercado (Ed.), *Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia* (1ª ed., pp. 204-215), Página y Signos/Funproeib Andes. <https://doi.org/10.5281/zenodo.11111196>

LENGUA Y ADQUISICIÓN

Gillian Gallagher

0. Introducción

Todos hablan un idioma nativo. Mucha gente habla dos, tres o más idiomas con competencia nativa. Lo remarcable es que la adquisición de un(os) idioma(s) nativo(s) en niños es implícita y natural: ocurre sin instrucción formal, ya sea por parte de los padres o de una escuela. En cualquier sociedad, cultura o país, los niños entienden bien su idioma nativo desde los 2 o 3 años y pueden hablar más o menos fluidamente entre los 3 y 5 años. El proceso de adquisición es una parte del desarrollo normal de un niño, como aprender a caminar. En la ausencia de una situación patológica (sea biológica o social), todos adquieren su lengua materna.

La habilidad de adquirir una lengua como lengua materna se encuentra en todos los bebés y disminuye con la edad. Antes de la pubertad es posible lograr competencia nativa en una lengua sin instrucción formal, pero después de la pubertad se hace mucho más difícil. Si un niño de 5 años se muda de un pueblo hispanohablante hasta otro país anglófono, normalmente este niño podría aprender inglés en 1 o 2 años y se escucharía igual a los otros niños angloparlantes. Pero si un adulto de 20 años hace lo mismo, no logrará el mismo nivel de dominio. Hay adultos que viven 40 años o más en un nuevo país, pero siguen hablando con un acento y cayendo en errores gramaticales en su segundo idioma. En la ciencia de la adquisición de lenguas, el periodo antes de la pubertad en el que un niño puede adquirir un idioma nativo se llama '**periodo crítico**'.

Es importante distinguir la competencia en una lengua oral –que es automática y universal– de la competencia en una lengua escrita –la cual es enseñada y particular a una cultura–. Todos hablan, pero no todos leen y escriben. La mayoría de las lenguas en el mundo (entre 3000-7000 dependiendo de cómo se cuenten) son únicamente orales, sin sistema o tradición de escritura o educación formal. Aunque todos aprenden a hablar sin instrucción formal, casi nadie aprende a leer o escribir sin ella. Este capítulo se enfoca en la adquisición de una lengua materna como parte del desarrollo infantil típico que se logra implícitamente, sin instrucción, y casi

igualmente en todos los niños. El uso de la lengua escrita, las diferencias entre la gente en sus habilidades retóricas y artísticas, y la adquisición de segundas lenguas para adultos son todos otros temas que se puede investigar por separado.

A continuación, se revisa los pasos en la adquisición de una lengua materna en general (sección 9.1.) y se presenta varios ejemplos más detallados de estudios sobre la adquisición de fonética y fonología (sección 9.2.), y de palabras (sección 9.3.). La sección 9.4. se enfoca en el bilingüismo y la 9.5 corresponde a la conclusión.

9.1. Pasos en la adquisición de una primera lengua

Los bebés y los niños pueden entender su lengua materna mejor de lo que pueden producirla. En cada paso de la adquisición, sus habilidades de percepción son más sofisticadas que sus habilidades de producción. La producción empieza con palabras aisladas, normalmente sustantivos muy comunes e importantes en su vida –‘leche’, ‘mamá’, ‘perro’, ‘sol’, ‘agua’– y palabras como ‘no’, ‘más’, ‘hola’, que son marcadores de las interacciones interpersonales. Frecuentemente, las primeras producciones de palabras son inteligibles solo para los padres, porque los sistemas de fonética y fonología no han sido desarrollados y no es posible para el bebé producir palabras como un adulto. Fonéticamente, los bebés tienen diferencias físicas que les impiden hablar como adultos. Sus lenguas son mucho más grandes en relación con el tamaño de sus bocas. Fonológicamente, pueden producir diferencias sistemáticas en el proceso de adquisición, por ejemplo, en los niños que hablan inglés, es normal que produzcan una ‘k’ como una ‘t’ sistemáticamente entre los 2 a 3 años. La variabilidad en ese tipo de correspondencia es muy grande –cada niño tiene su propio sistema aunque haya patrones que son más o menos comunes–.

Los niños empiezan a hablar con palabras aisladas y avanzan a frases muy cortas, de 2 o 3 palabras. Normalmente, esas frases son combinaciones de sustantivos, verbos y adjetivos, sin mucha inflexión o complejidad morfológica y con errores en la concordancia (verbal o nominal). Por ejemplo, niños en proceso de aprendizaje del castellano produjeron las frases: ‘los huevo’ (con plural en el artículo *los* pero no en el sustantivo *huevo*), ‘se van avión’ (con verbo plural *van* pero sujeto singular *avión*). Como en la fonología, cada niño tiene su propio sistema de morfología y sintaxis. Puede repetir los mismos errores durante semanas o meses antes de producir una estructura como un adulto.

En la adquisición de lenguas con morfología compleja, los niños tienen que aprender las reglas para las palabras regulares pero también las excepciones a estas reglas. Por ejemplo, en castellano, una niña de 2 años, que se llama María, usó la forma irregular y correcta ‘sé’, pero también usó ‘sabo’, que es incorrecto pero que

sigue el patrón regular (Marcus et al., 1992; Aguado-Orea & Pine, 2001). Errores de ese tipo se llaman ‘regularización’ y son muy comunes en la adquisición de lenguas con ese tipo de morfología.

9.2. Adquisición de fonología y fonética

Antes de producir su primera palabra, el bebé sabe mucho sobre su lengua ambiental. En el primer año, el sistema de percepción se desarrolla hasta que el niño pueda distinguir los sonidos contrastivos de su lengua y segmentar palabras del discurso continuo.

9.2.1. Percepción de contrastes

Hay varios estudios que investigan cómo los bebés perciben los sonidos lingüísticos. Lo que se encuentra es que los bebés más jóvenes (6 meses más o menos) pueden distinguir sonidos diferentes aun si la distinción no se usa en su lengua ambiental. Los bebés más grandes (12 meses más o menos), como los adultos, solo pueden distinguir los sonidos que son contrastivos en su lengua.

Un estudio clásico es el de Werker & Tees (1984). En ese estudio, compararon la percepción de /k/ y /q/ para tres grupos de hablantes: (i) adultos que hablan salish, una lengua en que los dos sonidos son contrastivos; (ii) adultos que hablan inglés, una lengua en que /q/ no existe, y (iii) bebés de 6 meses que crecen en un ambiente angloparlante.

El idioma salish se habla en el noroeste de los Estados Unidos y el sudoeste de Canadá. Como el quechua y el aimara en Bolivia, ese idioma tiene un contraste entre las consonantes velares y uvulares. Los ejemplos en (1) muestran esa distinción en el quechua boliviano con pares mínimos.

- (1) El contraste /q/-/k/ en el quechua boliviano
- | | | | |
|-------|----------|-------|-----------|
| kusa | ‘esposo’ | kusa | ‘bueno’ |
| qiru | ‘jarrón’ | kiru | ‘diente’ |
| waqaj | ‘llorar’ | wakaj | ‘mi vaca’ |

El inglés, como el castellano, no tiene sonidos uvulares. Normalmente, para los adultos, las diferencias no nativas son difíciles de escuchar. Para un hablante monolingüe de inglés o castellano, la distinción entre /k/ y /q/ es difícil de percibir y los dos sonidos se oyen casi igual.

La metodología para cuantificar la percepción en bebés es un procedimiento que se llama “preferencia de giro de cabeza”. El bebé escucha sonidos y aprende que una marioneta aparece cuando el sonido cambia. Aprenden a girar la cabeza para ver a la marioneta al escuchar las sílabas [ba] y [da]. Por ejemplo, escuchan

“[ba]...[ba]...[ba]...[da]” y la marioneta aparece junto con el “[da]”, o escuchan “[ba]...[ba]...[ba]” y no ven nada. Después de aprender la asociación del cambio de sonido con la marioneta, la investigación empieza y los bebés escuchan las sílabas [ki] y [qi]. Si giran la cabeza cuando cambió el sonido, eso quiere decir que escuchan las uvulares y velares como distintas. Un procedimiento similar se usa con los adultos, pero ellos oprimen un botón en lugar de girar la cabeza (y no ven una marioneta).

Werker & Tees han encontrado que los adultos hablantes de salish pueden distinguir /k/ y /q/ en un 100%, pero los adultos angloparlantes solo los distinguen en un 30%. Los bebés anglófonos pueden distinguir en un 80%, lo que es mucho mejor que los adultos angloparlantes que, al igual que los bebés, nunca han escuchado las uvulares en su lengua ambiental. En un segundo experimento, Werker & Tees compararon a bebés de 6 a 8 meses, de 8 a 10 meses y de 10 a 12 meses, y encontraron que los bebés de 10 a 12 meses ya perciben como adultos, pues no pueden distinguir los contrastes no nativos.

De esos datos aprendemos que los bebés empiezan su vida con la habilidad de escuchar muchas distinciones entre sonidos, pero en su primer año aprenden a ignorar las distinciones que no son importantes en su lengua. Otros estudios similares han encontrado que hay contrastes que los bebés tienen que aprender a escuchar si son importantes en su lengua materna. En su estudio, Narayan et al. (2010) investigaron la percepción de un contraste muy poco frecuente en las lenguas del mundo: el contraste entre una nasal alveolar y una velar en posición inicial, [na] vs. [ŋa]. Ese contraste existe en el idioma filipino que se habla en las Filipinas. Encontraron que los bebés que crecen en un ambiente filipino hablante pueden distinguir ese contraste entre los 10 a 12 meses, pero no entre los 6 a 8. Los bebés que crecen en un ambiente angloparlante no pueden distinguir el contraste de los 4 a 5 meses, de los 6 a 7 meses ni de los 10 a 12 meses. De estos datos, Narayan et al. (2010) concluyen que hay dos tipos de contraste. Todos los bebés perciben algunos contrastes al principio de su vida, y su sistema de percepción se desarrolla para ignorar la distinción si no se usa. Otros contrastes no son perceptibles al principio de la vida, y solo los bebés que escuchan una lengua que usa ese contraste pueden aprender a escucharlo.

Los resultados de estos estudios con bebés nos muestran que el sistema de percepción se forma para la lengua ambiental en el primer año. De los 10 a 12 meses, los bebés escuchan bien las distinciones fonéticas que son importantes en su lengua ambiental e ignoran las distinciones que no son contrastivas.

9.2.2. Adquisición de categorías fonéticas

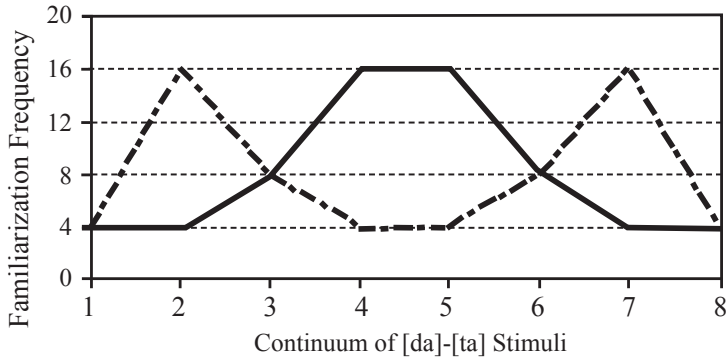
El centro de la adquisición de un sistema fonético es el análisis de la variación. La pronunciación de una palabra es diferente, fonéticamente, cada vez que se pronuncia. Para cada sonido contrastivo en una lengua, hay un rango de variación en sus propiedades fonéticas. Por ejemplo, en quechua el sonido /i/ se puede pronunciar con F1 entre 250 a 550 Hz (el F1 es una propiedad fonética que corresponde a la altura de una vocal). Para un hablante de quechua, todas esas pronunciaciones suenan como /i/, porque se ha aprendido que la variación entre un F1 de 300 y un F1 de 500 no es importante (quiere decir que no es fonémico, no cambia el sentido de una palabra) en su lengua. Para un hablante de castellano, por otra parte, esa diferencia es muy importante. Una vocal con F1 de 300 suena como /i/ y una vocal con F1 de 500 suena como /e/. Eso también es algo que los hablantes de castellano han aprendido, su sistema de percepción se ha desarrollado para escuchar claramente distinciones en F1 que corresponden a /i/ y /e/. El sistema de percepción de los hablantes de quechua se ha desarrollado para ignorar la distinciones en F1 en ese rango.

Maye et al. (2002) investigaron cómo los bebés aprenden a enfocar o ignorar diferencias fonéticas en su primer año. Su hipótesis fue que la importancia reside en el patrón de variabilidad. Por ejemplo, a pesar de que una /i/ de quechua se puede pronunciar con F1 de 250 a 550, la mayoría de las producciones de ese sonido se sitúan al medio de ese rango (entre 300 a 400). La distribución de valores es unimodal, con frecuencia alta en el medio y frecuencia baja en los extremos. Por otra parte, en una lengua en que /i/ y /e/ son contrastivos la distribución de valores sería bimodal, con dos concentraciones de frecuencia alta a los extremos y frecuencia baja en el medio del rango.

En el estudio de Maye et al., los bebés de 6 a 8 meses escucharon sílabas con variación entre [ta] y [da] (pronunciado en inglés). Un grupo de bebés escuchó una distribución de variación unimodal y un grupo escuchó una distribución de variación bimodal, como se muestra en la Figura 1. Cada bebé escuchó sílabas con sonidos de los ocho pasos entre [da] y [ta], pero la frecuencia de cada paso fue diferente para los dos grupos. Para el grupo unimodal, los pasos en la mitad del continuo fueron tan frecuentes como serían en una lengua en que [d] y [t] no son contrastivos. Para el grupo bimodal, hubo dos concentraciones a los extremos del continuo, como en una lengua en que [d] y [t] son dos fonemas diferentes.

Figura 1

El continuo acústico entre [da] y [ta] en ocho pasos y la frecuencia de cada paso en las dos condiciones, bimodal (línea punteada) y unimodal (línea sólida).



La primera fase del experimento fue la familiarización en la que los bebés escucharon las sílabas del continuo [ta]-[da] en distribución unimodal o bimodal. La segunda fase fue el examen en el que los bebés escucharon pares de sílabas: alternancia entre [ta] de punto 1 y [da] de punto 8 o repetición de punto 3 o punto 6. Los investigadores analizaron el tiempo en que los bebés se enfocaron en la dirección del sonido.

Los bebés que escucharon la distribución bimodal respondieron a la diferencia entre [ta] y [da]: miraron más tiempo cuando escucharon una alternancia entre [ta] y [da] que cuando escucharon la misma sílaba. En cambio, los bebés que escucharon la distribución unimodal no mostraron indicios de percibir la diferencia entre [ta] y [da]. Los resultados nos muestran que los bebés pueden usar la información en la distribución de propiedades fonéticas para aprender lo que es contrastivo e importante en su lengua.

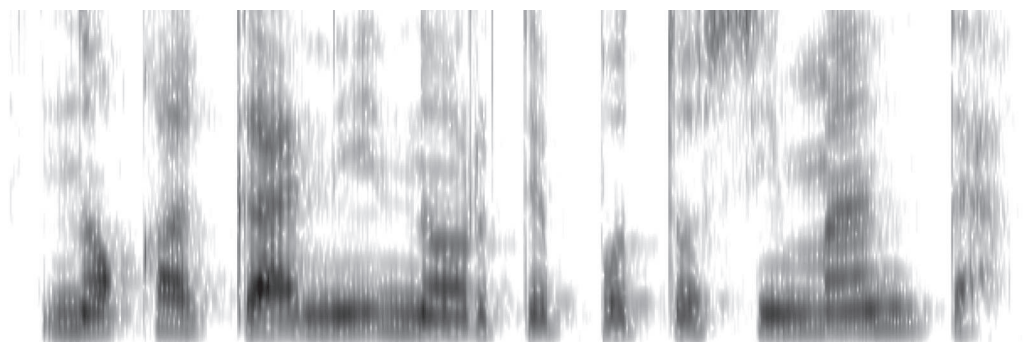
9.3. Adquisición de palabras

Una palabra es una asociación entre sonido y sentido. Para aprender una palabra, un niño tiene que identificar un sonido en las frases que escucha y construir un sentido para ese sonido.

Cuando hablamos, no hacemos pausas entre palabras. Hay un espacio entre ‘palabras’ en la escritura, pero no hay nada en la lengua oral que corresponda a ese espacio. La Figura 2 muestra un espectrograma de una frase hablada en quechua, con los fonemas marcados. (Un espectrograma es una visualización del habla en que el blanco es silencio y el negro es un sonido fuerte). En el espectrograma, se ve que no hay silencio o pausas entre palabras.

Figura 2

Espectrograma de la frase quechua ñuqa kani norte potosimanta 'yo soy de Norte Potosí' con fonemas y palabras marcadas.



ñ u q a | k a n i | n o r t e | p o t o s i m a n t a

En el habla, las pausas indican las fronteras entre frases o ideas, no se usan para indicar palabras. Un paso importante en el desarrollo lingüístico es aprender cómo segmentar las palabras del habla corriente. Una hipótesis es que los niños usan la frecuencia de ocurrencia de sonidos en contexto para identificar palabras. La idea es que una secuencia de sonidos que forma parte de la misma palabra sería más frecuente que una secuencia de sonidos a través de dos palabras. Por ejemplo, la frase de la Figura 2 *ñuqa kani norte potosimanta* en la que las dos primeras palabras se transcriben como /ñuqakani/ sin pausa. La hipótesis es que las secuencias /ñuq/, /oqa/, /kan/ y /ani/ (que son secuencias de tres fonemas en la misma palabra) son más frecuentes en el habla corriente que las secuencias /qak/ o /aka/ (que son secuencias de fonemas de dos palabras).

Esa diferencia en frecuencia existe porque las palabras individuales se pueden usar en frases diferentes y así las secuencias de sonidos a través de las palabras no es siempre la misma. En quechua se puede decir *ñuqa kani* 'yo soy', pero también *ñuqa puñuni* 'yo duermo', *ñuqa mancharikuni* 'yo tengo miedo', *ñuqa purini* 'yo camino', etc. En cada frase aparece la secuencia /uqa/ de *ñuqa*, pero la /qak/ solo aparece cuando se dice *ñuqa kani*, no en las otras frases. Una otra influencia en la frecuencia es la fonotáctica. En quechua, no hay morfemas que tienen consonantes uvulares y velares juntas, por lo que la combinación en /qak/ es de baja frecuencia y siempre indica que hay una frontera (ya sea entre morfemas o palabras).

Un estudio de Saffran et al. (1996) mostró que los bebés de 8 meses usan la frecuencia de secuencias de sonidos para identificar palabras. En la investigación, los bebés escucharon dos minutos del habla continuo sintético con 'palabras' inventadas: por ejemplo *bidaku-padoti-golabu-bidaku*. El único indicio de una palabra fue la

frecuencia de las sílabas. No había ninguna entonación ni pausa para indicar las fronteras entre ‘palabras’, pero la probabilidad de la transición entre sílabas distinguió las sílabas que formaron parte de la misma ‘palabra’. Por ejemplo, la probabilidad de *da* después de *bi* fue 1.0, porque *bidaku* es una ‘palabra’. En cambio, la probabilidad de *kupa* (y todas las sílabas a través de las palabras) fue 0.33 porque solo ocurrió cuando *padoti* siguió a *bidaku*.

Después de escuchar los dos minutos del habla (familiarización), los bebés escucharon ‘palabras’ de tres sílabas (examen), algunas que fueron **palabras familiares** (de la familiarización) y algunas que fueron **palabras nuevas**. Por ejemplo, *bidaku* (palabra) se comparó con *kupado* (no palabra, secuencia de *bidaku* y *padoti*). Las investigadoras midieron el tiempo en el que el bebé miró en la dirección de la voz que pronuncia la palabra. En ese paradigma, los bebés se sientan en las rodillas de su cuidador en un cuarto con dos altavoces en paredes diferentes. Hay luces delante del bebé y en los dos lados, para enfocar la atención del bebé. Empiezan cada ensayo mirando la luz de adelante y giran su cabeza para mirar la luz en la pared con el altavoz de donde viene la palabra. En esta investigación, los bebés miraron más tiempo hacia las ‘palabras nuevas’ que hacia las ‘palabras familiares’, y eso quiere decir que reconocen las ‘palabras familiares’. Después de solo dos minutos de escucha y con solo la frecuencia de transición como indicio, los bebés podían reconocer ‘palabras’ del habla corriente.

Si los bebés usan la frecuencia para identificar el sonido de una palabra, ¿cómo aprenden el sentido de la palabra? Una parte del sentido (claramente de sustantivos simples) se aprende del contexto en que la palabra se usa, pero el problema es que siempre hay muchas posibilidades de qué quiere decir la palabra. Por ejemplo, si estoy con un niño y vemos a un perro paseando y yo digo ‘perro’, es posible que la expresión quiera decir el nombre del tipo de animal que es el perro, pero también es posible que ‘perro’ se refiera a ese perro específicamente (‘perro’ es su nombre), a cualquier animal con cuatro patas, a la cola, al color del perro, a esa raza de perros, a la actividad de pasear, etc. Ese problema fue observado por Quine (1960), un filósofo. A veces los niños cometen errores, pero normalmente aprenden el sentido correcto muy rápido.

Los estudios sobre la adquisición de palabras han notado algunas regularidades sobre cómo los niños construyen un sentido para una palabra nueva. Markman & Hutchinson (1984) encontraron que los niños de entre 2.5 a 4 años asumen que una palabra se refiere a un tipo taxonómico. En su estudio, los niños miraron la foto de un objeto familiar y tuvieron que elegir otro objeto ‘similar’ entre dos opciones. Por ejemplo, si miraron una torta de cumpleaños, tenían que elegir entre otra torta y un regalo. En una condición, las instrucciones incluyeron una palabra nueva, los niños escucharon “¿Ves esto? Esto es un *biv*. ¿Puedes encontrar otro *biv*?”. En la

otra condición, no había una palabra nueva, las instrucciones estuvieron “¿Ves esto? ¿Puedes encontrar otro así?”. En la condición con palabra, los niños eligieron el objeto con relación taxonómica (otra torta en nuestro ejemplo) casi siempre: interpretaron una nueva palabra como nombre del tipo. En la condición sin palabra, sus respuestas estuvieron divididas entre los dos objetos. Así aprendimos que dado que los niños ven la similaridad entre objetos de la misma temática, interpretan las palabras como nombres de tipos.

9.4. Bilingüismo

Muchos niños hablan más de una lengua desde su infancia, son bilingües o multilingües. Normalmente, estos niños viven en comunidades multilingües en las que hay una lengua que se habla en la casa y con la familia, amigos y vecinos, y otra lengua que se usa en las escuelas, en los ámbitos estatales y en actividades formales u oficiales. También es posible tener bilingüismo en la familia si hay dos padres que hablan lenguas diferentes. La otra situación común que resulta en multilingüismo es la migración, donde una familia que habla una lengua vive en una comunidad extranjera en la que la mayoría habla otra lengua. En Bolivia, se encuentran muchos pueblos multilingües y se podría decir que el país en general es una sociedad multilingüe.

Un niño multilingüe aprende sus lenguas sin instrucción formal al igual que un monolingüe, simplemente escucha varias lenguas en su vida. Normalmente las lenguas tienen cada una su contexto cultural y social, y los niños aprenden con quién y en qué contexto se usa cada lengua. Por ejemplo, podría ser que el quechua se usa en la casa con la familia, pero en la escuela se usa el castellano. También es posible que use una lengua con su madre y otra con su padre (o los abuelos, tíos, etc.). El multilingüismo estable se encuentra en sociedades en las que cada lengua tiene su propio contexto de modo que los multilingües usen frecuentemente cada lengua. Por ejemplo, *solo* hablan quechua con sus papás y nunca castellano, pero con sus amigos de la escuela hablan castellano. Cuando los contextos para las lenguas no están bien definidos, es posible que un niño crezca monolingüe en la lengua con más prestigio o con baja competencia en la lengua indígena porque no tiene la exposición suficiente. En el contexto boliviano, Luykx (2004) discute la importancia de preservar el ambiente familiar para las lenguas indígenas porque el uso de castellano cada vez es más común afuera de la casa, en todo el país.

En un contexto de adquisición multilingüe, los niños muestran evidencia de que distinguen sus lenguas desde la infancia, pero también hay influencias entre las lenguas en el desarrollo (y de igual manera en unos adultos multilingües). Mehler et al. (1986) encontraron que bebés de cuatro días podían percibir la diferencia entre dos lenguas y que preferían escuchar la lengua de su mamá. Aunque los niños

monolingües mezclan mucho sus lenguas, distinguen correctamente con quién usar cada lengua. Por ejemplo, si su papá habla castellano y su mamá quechua, el niño usaría más castellano –en forma sintáctica y en palabras– con el papá y más quechua con la mamá, aunque usaría las dos lenguas con los dos padres. Es muy común que los niños multilingües mezclen sus lenguas en los primeros años y con el tiempo aprendan mejor con quién y en qué contextos se usa cada lengua. Cuando mezclan las lenguas, puede que usen una palabra de una lengua en una frase de la otra. Eso es más común cuando su vocabulario es pequeño; podría deberse a que sabe una palabra en solo una lengua. Mezclan también las frases enteras. Por ejemplo, Lindholm & Padilla (1978) dan el ejemplo de un niño bilingüe en inglés y castellano que dijo ‘*I ask him* que yo voy a casa’ en el que usa el inglés ‘*I ask him*’ para el castellano ‘yo le pido’. En ese ejemplo, el niño mezcla las lenguas, pero cada frase tiene una estructura buena y refleja una idea distinta, y de eso concluiríamos que está aprendiendo los dos sistemas (Genessee, 1989).

En la adquisición de dos lenguas, podrían aparecer influencias profundas en las estructuras gramaticales. Sánchez (2006) investigó el habla de los niños bilingües castellano-kechwa en Perú (en la región de San Martín, distrito Wayku). Los participantes en el estudio fueron bilingües castellano-kechwa y monolingües en castellano de 9 a 13 años. Los niños tuvieron que contar una historia basada en una serie de imágenes sobre un sapo, un niño y su perro (Mayer & Mayer, 1992). Los bilingües contaron la historia en kechwa y también en castellano y los monolingües solo en castellano. La comparación interesante es entre el castellano de los bilingües y el de los monolingües. Sánchez encontró que los bilingües usan estructuras gramaticales en castellano con sentidos prestados o influenciados del kechwa. En kechwa, hay un sufijo ‘-yka’ que expresa el tiempo progresivo (2) y un sufijo ‘-naya’ que expresa el deseo (3). Cuando esos dos sufijos se usan juntos, se puede expresar el deseo progresivo o que la acción es inminente (4).

- (2) *kawa-yka-n*
mirar-PROG-3sg
‘Está mirando’.
- (3) *maka-naya-n*
golpear-DES-3sg
‘Quiere golpear’.
- (4) *muku-naya-yka-n*
morder-DES-PROG-3sg
‘Está queriendo morder o está a punto de morder’.

En sus cuentos, los bilingües usaron la estructura ‘estar queriendo <verbo>’ en castellano, para decir que alguien está a punto de hacer algo. Los monolingües no usan esa estructura para ese sentido. Así vemos que la correspondencia entre la estructura morfosintáctica y la forma semántica (el sentido) puede ser diferente para los bilingües y monolingües porque los bilingües pueden mostrar influencia de sus otras lenguas.

9.5. Conclusión

En ese capítulo, se ha revisado algunos resultados importantes sobre el desarrollo lingüístico en los monolingües y multilingües. La adquisición de un(os) idioma(s) nativo(s) es un proceso implícito, con bajo nivel de conciencia, que ocurre sin instrucción formal, en contraste con la adquisición de una segunda lengua como adulto o con las habilidades de alfabetización. Investigaciones han encontrado que los bebés usan la estadística de sonidos para aprender las palabras y los fonemas de sus lenguas. Los errores sistemáticos en la adquisición de la morfología y la sintaxis muestran que los niños construyen reglas que cambian durante el desarrollo.

Referencias

- Aguado Orea, J. & Pine, J. (2001). Overregularizations in Spanish. Presentación en *Child Language Seminar 2001*. Disponible en: https://www.researchgate.net/publication/260082604_Overregularisations_in_Spanish
- Genessee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language* 16(1): 161-179.
- Lindholm, K. & Padilla, A.M. (1978). Language mixing in bilingual children. *Journal of Child Language* 5(2): 327-335.
- Luykx, A. (2004). The future of Quechua and the Quechua of the future: Language ideologies and language planning in Bolivia. *International Journal of the Sociology of Language* 167: 147-158.
- Marcus, G. et al. (1992). Overregularization in language acquisition. *Monographs for the Society for Research in Child Development* 57(4):1-182.
- Markman, E.M. & Hutchinson, J.E. (1984). Children’s sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology* 16(1): 1-27.
- Maye, J.; Werker, J. & Gerken, L.A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101-B111.

- Mayer, M. & Mayer, M. (1992). *One Frog Too Many*. Nueva York: Dial Press.
- Mehler, J., Lambertz, G., Jusczyk, P.W. & Amiel-Tison, C. (1986). Discrimination de la langue maternelle par le nouveau-né. *Académie des Sciences* 3: 637-640.
- Narayan, C., Werker, J. & Speeter Beddor, P. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science* 13(3): 407-420.
- Quine, W.V. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Saffran, J., Aslin, R. & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science* 274(5294): 1926-1928.
- Sánchez, L. (2006). Kechwa and Spanish bilingual grammars: Testing hypotheses on functional interference and convergence. *International Journal of Bilingual Education and Bilingualism*. 9(5). 535-556
- Werker, J. & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7. 49-63.

ABREVIATURAS

. frontera de sílaba	FUT = futuro
- frontera de morfema	GEN = genitivo
1 = primera persona (“yo” o “nosotros/as” en el plural)	INCL= inclusivo
1+2 = primera persona inclusiva (“yo y vos”)	INSTR = instrumental
1+3 = primera persona exclusiva (“yo y otros/as, pero sin vos”)	INT = interrogativo
2 = segunda persona (“vos” o “ustedes” en el plural)	IRR = irrealis
3 = tercera persona (“él, ella” o “ellos, ellas” en el plural)	L = consonante de ligación
A = agente de verbo transitivo	LOC = locativo
ABL = ablativo	M = masculino
ACC = acusativo	MOV = movimiento asociado
ADV = adverbial	N3 = cualquier persona menos la tercera
AG = nominalizador agentivo	NAR = narrativo
ALL = alativo	NEG = negación
APL = aplicativo	NEU1 = neutro 1
ASOC= asociativo	NF = verbo no finito
AY= ayoreo	NFUT = no futuro
BP= Bolivia-Paraná	NL = nasal de ligación
C= consonante	NM= no-masculino
CAUS = causativo	OBJ= objeto
CH= chamacoco	P = paciente (objeto) de verbo transitivo
CONJ = conjetural	PAS = pasado
CTFG = centrífugo (“yéndose”)	PERS = persona por defecto
CTPT = centrípeto (“viniendo”)	PF = perfecto
CTRS = contrastivo	PL = plural
DAT = dativo	PLM= plural mayor
DECL = declarativo	PREP = preposición
DUB = dubitativo	PRS = presente
EP = elemento ligado	PTCP = participio
EXCL= exclusivo	R/R = reflexivo/recíproco
EXH = exhortativo	REAL= realis
FG= forma genérica	RFL = reflexivo
ILL = ilativo	SEP = partícula separadora
IPF = imperfectivo	SUJ= sujeto
F = verbo finito	SG = singular
F1= formante uno	TH = consonante temática
F2= formante dos	V= vocal
	x = singular no diminutivo de sustantivos no poseídos por un poseedor referencial
	♀ = habla de las mujeres
	ZA= zamuco antiguo

Glosario

Alófono: Uno de dos o más segmentos fonéticamente distintos en los que se puede realizar un fonema.

Alomorfo: Una de dos o más formas superficiales que asume un morfema, o, en términos más simples, variante de un morfema.

Área lingüística (lingüística areal): Espacio geográfico en el que coexisten varias lenguas, las cuales comparten un conjunto de rasgos cuya presencia no puede explicarse por relaciones genéticas, restricciones universales sobre las estructuras lingüísticas o desarrollos lingüísticos, ni por la casualidad; esos rasgos no se hallan en lenguas que se encuentran fuera del área, por lo que son el resultado del contacto lingüístico (Heine, 2006, p 90).

Caso: Es una categoría nominal que alude a distintas relaciones gramaticales que pueden contraer los sintagmas nominales de acuerdo con la función que cumplan (nominativo, genitivo, acusativo, dativo, ablativo, vocativo, instrumental y locativo).

Elicitación: Técnica para obtener información de los hablantes sin que los investigadores les pregunten directamente sobre la gramática, el vocabulario o los usos de un idioma. A partir de preguntas, se obtiene los juicios de gramaticalidad de los hablantes sobre determinadas construcciones y sus posibles interpretaciones.

Ergativo: Es el caso morfológico que marca al participante más activo, por ejemplo el sujeto, distinguiéndolo, en una oración transitiva, del objeto. Ergatividad: De acuerdo con Lyons, es el paralelismo sintáctico entre el objeto de un verbo transitivo y el sujeto de un verbo intransitivo.

Finito/no finito: Los verbos marcados por el tiempo, la persona, el número y otras categorías son definidos como finitos. El término “verbo finito” contrasta con “verbo no finito” e infinitivo.

Fonema: En muchas teorías de la fonología, un elemento fundamental, mínimo; es la unidad abstracta de la estructura fonológica, que tiene una función distintiva.

Gramaticalidad/aceptabilidad: En el paradigma generativista, la gramaticalidad es la propiedad de las oraciones que se ajustan a las reglas y principios de una determinada lengua. La aceptabilidad es la adecuación de un enunciado a una determinada situación comunicativa; la adecuación depende más de factores de producción del enunciado que de la gramática.

Isomorfo: Paralelismo entre dos estructuras. Llevado al extremo lingüístico, supone paralelismo entre forma y contenido de dos estructuras.

Par mínimo: Dos palabras de distinto significado que exhiben segmentos diferentes en un punto pero segmentos idénticos en todos los demás puntos.

Paradigma: Conjunto de elementos lingüísticos asociados entre sí, ya sea por su parecido formal, por la función que cumplen o por la idea que expresan. Es cualquier constituyente que pueda reemplazarse entre sí en un determinado contexto.

Pasado hodierno: Pasado remoto.

Realis/irrealis: Desde algunas teorías, son categorías relacionadas con la modalidad. Realis, vinculada a hechos experimentados, constatables, e irrealis, a hechos no experimentados, hipotéticos. “La distinción realis/irrealis refleja juicios de que ciertas ideas surgen de la percepción directa, la memoria, o las expectativas de lo que es normal, mientras que otras tienen su origen en la imaginación” (Chafe, 1995, pp. 363).

La presente edición se terminó
de imprimir el mes de junio de 2023
en Talleres Gráficos “KIPUS”
c. Hamiraya 122 • Telf./Fax.: 591-4-4582716 / 4237448

ISBN: 978-9917-0-2681-5



9 789917 026815