

# Lexicografía de idiomas poco documentados

Lev Michael y Christine Beier  
University of California, Berkeley

Pontificia Universidad Católica del Perú  
11 y 13 de abril, 2023

# Bienvenidos

- Este cursillo es diseñado para lingüistas involucrados en (o interesados en (!)) la documentación léxica de idiomas poco documentados.
- Su meta principal es compartir ideas y metodologías que pueden ayudarnos en producir mejores recursos léxicos, tanto para la lingüística como para las comunidades con que trabajamos.

# Bienvenidos

- Abordaremos los siguientes tópicos:
  - Día 1:
    - La importancia de la lexicografía de idiomas poco documentados
    - La estructura y organización de un diccionario multilingüe
    - Los desafíos para obtener definiciones exactas y precisas
    - Métodos para mejorar la exactitud y precisión de las definiciones
    - Cómo pensar sobre la distinción entre entradas distintas y las acepciones (o sea, los sentidos dentro de una sola entrada)

# Bienvenidos

- Día 2:
  - El papel central del análisis gramatical en la lexicografía de idiomas poco documentados
  - Software de base de datos para lexicografía (FLEx)
  - Recomendaciones de flujo de trabajo para el trabajo léxico en idiomas poco documentados
  - Observaciones y recomendaciones sobre metas realistas en la lexicografía de idiomas poco documentados

# La importancia de la lexicografía de los idiomas poco documentados

- Importancia para la comunidad de hablantes (y/o sus herederos)
- Importancia para la lingüística
- Importancia para los campos afines

# Importancia para las comunidades de hablantes y/o heredores

- Los recursos lexicográficos:
  - Funcionan como un registro importante de conocimientos culturales, históricos y ecológicos.
  - Pueden jugar un papel importante en la educación bilingüe y/o en la revitalización lingüística.
  - Pueden funcionar como un símbolo de que el idioma indígena es un idioma real (función antirracista).
  - Pueden funcionar como un símbolo de identidad indígena (función de visibilidad indígena).

## Importancia para la lingüística

- En el contexto (tanto histórico como metodológico) de la documentación y la descripción lingüística, hay una relación muy estrecha entre un diccionario, un corpus de textos y una descripción gramatical.
  - La tríada boasiana: cada parte de la tríada depende en las otras.
- Diccionarios son de suma importancia para los trabajos comparativos-históricos.
  - El estado relativamente subdesarrollado de los estudios histórico-comparativos de las familias lingüísticas sudamericanas se debe en gran parte a una documentación léxica inadecuada.

## El estado actual de la lexicografía de idiomas amazónicas

- La cantidad y la calidad de recursos léxicos (RLs) de idiomas amazónicas siguen incrementado.
- Sin embargo, hay muchos aspectos de la práctica lexicográfica que pueden mejorarse:
  - Muchos RLs no dan definiciones, sino que glosas de una sola palabra.
  - Generalmente, proveen poca información gramatical relevante.
  - Su cobertura de ciertos campos semánticos importantes, como la fauna y la flora, a menudo es insuficiente.
  - A veces se omiten aspectos importantes de la representación fonológica, como la longitud vocálica y el tono.
  - Muchos diccionarios podrían mejorarse con informaciones cultural e histórica local más ricas.





## Impactos en el uso de recursos léxicos por sus comunidades

- Los recursos léxicos con definiciones imprecisas pueden dificultar el trabajo de revitalización del lenguaje.
- Por ejemplo, un estudiante posgrado en UC Berkeley está trabajando con un grupo indígena de California en su programa de revitalización, utilizando recursos léxicos y textuales recopilados en las décadas de 1960 y 1970.
  - Las definiciones en el diccionario tienden a ser muy breves, lo que resulta en una ambigüedad considerable.
  - Cinco verbos distintos se definen como 'romper', sin más detalles.
  - Han tenido que elegir uno de estos verbos como el verbo general que significa 'romper' — a pesar de que los verbos seguramente expresaban significados más precisos.

# Definiciones inadecuadas de la flora y la fauna

- Unos ejemplos *Diccionario Ashéninca–Castellano* (Payne 1980)
  - *amitzi*: especie de ave
  - *chompiniro*: especie de ave negro
  - *macavari*: especie de ave
  - *pachaca*: especie de ave de color blanco
  - *shoniro*: especie de ave
  - *tyointyotzi*: especie de ave nocturno

## Definiciones inadecuadas de la flora y la fauna

- Dado la importancia de la fauna y la flora a las culturas amazónicas, la falta de trabajo serio sobre estos campos semánticos es un vacío importante.
- Imaginamos un hipotético *Diccionario Castellano–Ashéninca* con definiciones como estas:
  - *camión*: tipo de aparato grande
  - *computadora*: tipo de aparato pequeño
  - *televisor*: tipo de aparato

## Propiedades gramaticales relevantes

- Como veremos en el día 2, muchos elementos léxicos tienen propiedades gramaticales que deben identificarse para que su función se documente adecuadamente.
- Por ejemplo:
  - ¿Exhibe el lexema en cuestión irregularidades flexivas?
  - ¿Están restringidos ciertos sentidos del lexema a ciertas categorías TAM?
  - ¿Se requiere que los complementos FN del verbo lleven ciertos tipos de marcas de caso o adposiciones?
  - ¿Un verbo acepta complementos clausulares y FN? Y si lo hacen, ¿son dichas cláusulas finitas o no finitas?
- Muchos diccionarios de idiomas amazónicos no incluyen estos tipos de información.

# Representaciones fonológicas inadecuadas

- A pesar de la importancia de tener representaciones fonológicas correctas, hay una tendencia a omitir rasgos contrastivos supersegmentales.
- Por ejemplo, muchos diccionarios de los idiomas tucanos no incluyen el tono (p.ej., los diccionarios de cubeo (Salser & Salser 1999), desano (Tulio et al. 2000) y koreguaje (Cook and Galow 2001)).

## Representaciones fonológicas inadecuadas

- Sin embargo, el tono es de suma importancia en los idiomas, como vemos en estos triples mínimos del májǰkì:
  - *yéé* 'mojarse', *yèè* 'disparar', *yèè* 'enfermarse'
  - *béó* 'echarse en hamaca', *béò* 'poner en hamaca', *bèò* 'dar nombre'
- De hecho, el diccionario anterior del májǰkì (Velie y Velie 1981) omitió tanto la longitud vocálica como el tono.

# Nuestras experiencias con lexicografía

- Dictionarios de los idiomas:
  - iquito (2022, 2019, 2006)
  - májìkì (en proceso, 2011-13)
  - záparo (2014, 2011)
  - muniche (2009)
- Trabajo de campo también con los idiomas:
  - nanti
  - matsigenka
  - omagua
  - andoa (katsakati)
  - chamicuro



# Nuestras experiencias con lexicografía

- En cada proyecto:
  - Priorizamos el trabajo en equipo para acelerar la fase de investigaciones a fin de difundir recursos lexicográficos y didácticos tan pronto como posible.
  - Organizamos el trabajo lexicográfico del equipo mediante el uso del programa FLEx, instalado en un PC central compartido.
  - Exportamos los diccionarios desde la base de datos de FLEx y luego los formateamos con otras herramientas digitales (LexiquePro; Python, LaTeX).
  - Repartimos los diccionarios en forma impresa en las comunidades mismas (gratis para ellos).
  - Difundimos los diccionarios en formato digital por el internet.





## Trabajo lexicográfico con el iquito



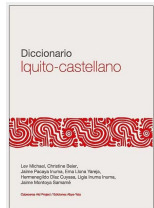
Miembros del equipo de investigadores en 2006

# Trabajo lexicográfico con el iquito

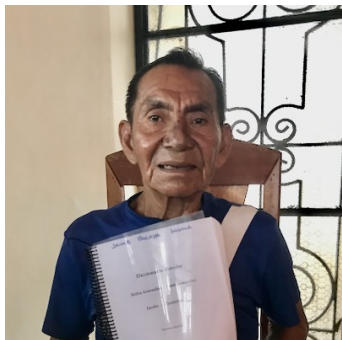
- Características del proyecto:
  - Trabajos de campo: 2001 al 2006, 2008, 2010, 2014 hasta el presente.
  - Un proyecto de mayor escala; equipos de 4 a 7 investigadores visitantes en cada año desde 2002 hasta 2006 (8 semanas, junio-julio).
  - Desde 2006, la gran mayoría del trabajo de campo ha sido hecho por CMB y LDM.
  - La fluidez y el entusiasmo de los especialistas nos ha permitido recurrir en gran medida a los textos en nuestro trabajo léxico.
  - Muchos años de investigación nos permitieron incluir información gramatical sustancial en los diccionarios publicados.

# Trabajo lexicográfico con el iquito

- Productos principales del proyecto:
  - *Diccionario iquito–castellano* (2022, Abya-Yala)
  - *Diccionario escolar ikíitu kuwasíini–tawi kuwasíini, iquito–castellano* (2019, Cabeceras)
  - *Iquito–English Dictionary* (2019, Abya-Yala)
  - *Diccionario bilingüe iquito–castellano y castellano–iquito* (2006, Cabeceras)
  - Varios materiales didácticos (Cabeceras)
  - Diccionario audio-visual (en proceso, *livingdictionaries.app*)
  - Varias tesis de licenciatura (Perú) y de maestría y doctorado (EE.UU.)



## Trabajo lexicográfico con el iquito



El Sr. Jaime Pacaya Inuma con el diccionario escolar en 2019



La Sra. Ema Llona Yareja recibiendo un diccionario en 2023







# Trabajo lexicográfico con el muniche

- Características del proyecto:
  - Trabajos campo: 2008-2009, 2022.
  - Trabajos preliminares, realizados por Sullón en 2008, lo hizo posible sentar las bases de un proyecto de documentación más amplia.
  - LDM obtuvo una subvención de investigación del NSF para un proyecto de documentación rápida en equipo, llamado “Muniche Language Documentation Project” (MLDP), que se llevó a cabo en 2009.
  - Los cinco lingüistas del proyecto trabajaron con siete recordadores del idioma.
  - Cada lingüista se centró en áreas específicas de la documentación para maximizar la productividad en el campo.



# Trabajo lexicográfico con el máíj̀̀k̀̀i

- Contexto del proyecto:
  - En 2006, preocupados por la falta de información y datos disponibles sobre esta lengua, salimos en busca de hablantes.
  - Llegando a Nueva Vida del Yanayacu, la comunidad nos recibió calurosamente y acordamos en colaborar en trabajo lingüístico tan pronto como posible.
  - En 2009, hicimos un acuerdo formal con la federación máíj̀̀nà, el cual incluyó un componente lexicográfico.
  - En 2010, empezamos trabajo de campo en equipo en Nueva Vida. En ese entonces, habían 90-100 materno-hablantes del idioma.
  - En 2011, LDM obtuvo una subvención de investigación del NSF para un proyecto de documentación mayor, llamado el “Máíj̀̀k̀̀i Project” (MP), que se llevó a cabo desde 2011 hasta 2014.



# Trabajo lexicográfico con el májìkì



Miembros del equipo de investigadores en 2012







## Trabajo lexicográfico con el záparo

- Contexto del proyecto:
  - La antropóloga Dra. Brenda Bowser nos invitó a colaborar con ella en trabajo lexicográfico con 5 recordadores del idioma que ella conocía de su propio trabajo de campo.
  - Aunque existían algunos recursos de documentación y descripción limitadas, en ese entonces, esas fuentes incluyen inconsistencias y contradicciones y tuvimos mucho interés en obtener datos adicionales.
  - CMB y Bowser obtuvieron una subvención de investigación del NSF para un proyecto de documentación rápida en equipo, llamado “Záparo RAPID Documentation Project” (ZRDP), que se llevó a cabo en 2010 y 2011.



# Trabajo lexicográfico con el záparo

- Características del proyecto:
  - Trabajos de campo: 2010-2011.
  - El trabajo consistía principalmente en la elicitación, a fin de conseguir datos léxicos y gramaticales.
  - Diseñamos el proyecto sin expectativas de hacer más trabajo de investigación en años posteriores.
  - Hicimos un uso extensivo de los materiales existentes para guiar nuestro trabajo de documentación y descripción.
  - Limitaciones de fluidez entre los recordadores descartó el trabajo basado en textos.
  - Pudimos realizar grabaciones que son un valioso aporte a la documentación del idioma.







# Decisiones de presentación: “yuca”

## **asúrakajina** *sustantivo loc.*

- yucal. *No hay forma plural*

## **asúraaja** *sustantivo*

- yuca en general. *Plural:*  
asúraajaka

## **asúraaja imíini** *sustantivo. Variante de: takitíini*

## **asúraaja niriyaákuuja** *sustantivo*

- tipo de piripiri usado para hacer que los huayos de la yuca crezcan rápido. *Plural:* asúraaja niriyaákuujaka

**asúràaja** (sus.) yuca, un cultivo cuyos *huayos* (raíces tuberosas) largos y ricos en almidón son la base de carbohidratos de la dieta iquito tradicional. La planta, que crece hasta una altura de 1.5-3 m, dependiendo de la variedad, produce un conjunto de *huayos* sobre los cuales se alza un único tallo de muchos nudos y de 3-5 cm de diámetro; del tallo emergen ramas con hojas palmeadas. Además de comerse hervidos, cocidos al vapor y asados, los *huayos* constituyen el ingrediente principal del *itínijja* (*masato*), y también se usan para preparar *waarínà* (*fariña*), una harina tostada y duradera a base de yuca. Si bien la yuca todavía es importante en la dieta de los residentes actuales de San Antonio, está siendo reemplazada por el arroz, los fideos y otros comestibles adquiridos. ▶ NOMBRE CIENTÍFICO *Manihot esculenta*

diccionario escolar  
iquito–castellano (2019)

diccionario iquito–castellano  
(2022)

## La estructura de un diccionario

- Empezamos por exponer la estructura de un diccionario para una lengua poco documentada.
- Primero, describimos las partes principales de un diccionario de este tipo.
- Luego, nos centramos en la estructura de un artículo del diccionario.



## Las partes principales de un diccionario

- La mayoría de los diccionarios de idiomas poco documentados (IPD) son bilingües o multilingües: incluyen el IPD y un idioma global (IG) (aunque hay excepciones, como el diccionario Enhlet (Stahl 1998)).
- Esto implica que el diccionario tiene como mínimo dos partes principales:
  1. Una parte que define los encabezados del IPD en términos del IG.
  2. Otra parte que permite al usuario encontrar encabezados del IPD a base de palabras en el IG.

# Las partes principales de un diccionario

- Además, a menudo hay:
  - Informaciones sobre la ortografía utilizada.
  - Una orientación sobre cómo usar el diccionario.
- Otras partes que pueden ser útiles son:
  - Una descripción de los aspectos de la morfología relevantes para la formación de palabras.
    - En lenguas morfológicamente complejas, el conocimiento de la forma de citación (o la raíz) por sí sola puede dejar al usuario muy lejos de poder utilizar el diccionario con fines comunicativos.
  - Glosarios de terminología especializada, por ejemplo:
    - terminología lingüística
    - terminología regional en el IG
- Ejemplos de partes como estas se encuentran en el *Diccionario iquito-castellano* (2022).

## Los artículos: entrada y acepción

- El contenido principal de la parte IPD→IG del diccionario está organizado en **artículos**.
- Cada artículo comienza con una **entrada**, y tiene una o más **acepciones**, que indican diferentes formas en que se interpreta o entiende la palabra indicada por la entrada.
- La cuestión de cómo distinguir entre acepciones de una sola entrada y diferentes entradas es un tema sutil al que volvemos a continuación.
- Veamos algunos artículos del diccionario iquito-castellano para ver cómo un artículo puede tener una o más acepciones.

**iika** (sus.) **1.** diente, muela o colmillo. *Pos.pref.*  
**2.** pico de ave. *Pos.pref.* **3.** probóscide de insectos que pican, como los zancudos y los tábanos. *Pos.pref.* **4.** punta aguda o borde afilado de una herramienta punzante o cortante, p. ej., la punta de una lanza o el filo de un cuchillo o machete. *Pos.pref.* || ▶ FORMA POSEÍDA IRREG. **îikà** ▶ FORMA POSEÍDA DE TERCERA PERSONA IRREG. **nîikà** ◆ PALABRA RELACIONADA **riwa ííkà** (adj.) de dientes torcidos

Artículo para **iika** con sus acepciones



## Las partes de un artículo

- Entrada (forma de citación)
- Forma(s) básica(s), p.ej., la raíz (forma lexémica)
- Categoría gramatical
- Definición
- Ejemplo de uso
- (Nombre científico)
- (Información grammatical)
  - Por ejemplo, plurales irregulares, alomorfos de la raíz
- (Información cultural)
- (Información sociolingüístico)
- (Formas relacionadas)
  - Por ejemplo, alternancias de valencia, formas derivadas









## El reto básico de las definiciones léxicas

- La tarea científica de desarrollar definiciones exactas y precisas para las formas en un idioma enfrenta dos desafíos básicos:
  1. Cualquier instancia de uso de una forma, o cualquier explicación metalingüística por un hablante, ilumina solo algunas facetas del significado o función de esa forma.
  2. Dado que los lingüistas suelen confiar en un idioma de contacto en el trabajo de campo lingüístico y debido a que el IPD y el idioma de contacto exhiben una organización semántica diferente, las traducciones suelen ser imprecisas.

## Explicaciones metalingüísticas

- Las explicaciones metalingüísticas de los significados de las palabras proporcionadas por los hablantes son imprescindibles.
- Sin embargo, cuando pedimos a un hablante que defina una palabra, típicamente obtenemos una definición que se superpone (o traslapa) parcialmente con el significado núcleo de la palabra; o con un cierto aspecto de su función pragmática, etc.
  - No podemos suponer que se superpone perfectamente: a menudo no es así.
  - El hablante generalmente tiene en mente ciertos usos o ciertos aspectos del significado de una palabra, pero generalmente esto no abarca todo el espectro de funciones o significados.



- Otro ejemplo: la palabra *anásàkà* tiene varios sentidos:
  1. dolor
  2. cólera, rabia
  3. veneno (p.ej., de una víbora)
- Pero cuando preguntamos por el significado de esta palabra, los hablantes por lo general solamente dan el primer sentido.
- Se encuentran los otros sentidos principalmente en los textos.

## Significados y funciones revelados por los textos

- De manera similar, cuando encontramos palabras en un texto traducido, estos contextos de uso nos brindan una comprensión parcial, pero no completa, de las propiedades semánticas, pragmáticas y gramaticales de una forma.
- Además, los textos generalmente tienen traducciones al IG, lo cual presenta el problema de la falta de congruencia semántica entre los dos idiomas.







# Menos distinciones léxicas

- *suwáàni* (adj.)
  1. bien; poseído de cualidades dignas o admirables.
  2. hermoso o guapo, hablando de personas.
  3. moraleja; que posee buenas características sociales y morales, por ejemplo, generoso, educado, obediente, honesto.





- Métodos de trabajo que pueden ayudar a resolver los problemas mencionados:
  1. Trabajar con varios hablantes.
  2. Usar tanto textos como elicitación como fuentes de información.
  3. Hacer uso del concepto de contraste semántico para refinar las definiciones.



- Todos hablantes de un idioma comprenden los significados de la gran mayoría de las palabras del idioma, pero:
- Diferentes personas tienen diferentes habilidades para poder **articular** las diferencias de significado entre las palabras.
  - Esta es una habilidad que se puede desarrollar, especialmente si el trabajo de documentación léxica despierta el interés de los hablantes.
  - Diferentes personas pueden tener más conocimientos sobre campos léxicos específicos (por ejemplo, plantas, actividades de caza o cocina) que otras.
- Si un idioma es dialectalmente diverso internamente, puede ser necesario trabajar con personas familiarizadas con cada uno de los dialectos.

- Pero a veces no es posible trabajar con muchos hablantes.
- En tales casos, se puede obtener un efecto muy similar por trabajar con un hablante sobre el significado de una palabra después de un intervalo de muchos meses o incluso años.
  - El tiempo intermedio a menudo hace que la perspectiva del hablante sobre la palabra se “reinicie”, de modo que proporcione nuevos conocimientos sobre el significado y la función de la palabra.

## Uso de textos

- El desarrollo de un corpus de textos, es decir, una colección de grabaciones de discursos orales que se transcriben y traducen, es una parte central de la mayoría de los proyectos de documentación de idiomas.
- Dichos textos son invaluable para el desarrollo de recursos léxicos por dos razones:
  - Los textos atestiguan formas que se encuentran hacia el lado funcional del espectro léxico-funcional, los que pueden no ocurrir fácilmente a los hablantes como elementos léxicos.
  - Los textos también pueden ejemplificar aspectos menos comunes de los significados de las formas, que tienden a no ocurrirles a los hablantes en contextos de elicitación.



- Uno puede explorar el rango de significados de una forma dada en un corpus usando la función de **concordancia** de FLEx (la aplicación de base de datos léxica y de corpus que describiremos con mayor detalle en el día 2 de este cursillo).
- La función de concordancia proporciona un resumen de todas las instancias de una forma dada en el corpus, con un fragmento del contexto circundante y la opción de ir al texto para examinar el significado y la función de la forma con mayor detalle.

Iquito2021 - FieldWorks Language Explorer

File Send/Receive Edit View Data Insert Format Tools Parser Window Help

Specify Concordance Criteria

Concorded on Entry iwĩĩĩĩ

Specify concordance criteria... Help

Concordance Results

Ref	Occurrence
115.1	Show All
2.AN1	i2 kina = áràtà , náàjì jìitàà kìnà = fíkì2 k = isákùùjì .
146.1	
2.AN1	"liti k = fíkì2 kina = áràtà , iiti kii = mijìtáà2 kii
146.1	
2.AN1	na = ná nàkì , iina máánà 4 ìkìáárikì tìj , waaaraa kaayaaka iijíikùjìna
174.1	
2.AN1	kúràà = na lina yaawāni 4 p = fíkì2 .

Lexicon

Texts & Words

Grammar

Notebook

Lists

Info Baseline Gloss Analyze Tagging Print View Text Chart

146 Word " liti

Morphemes iiti

Lex. Entries iiti

Lex. Gloss here

k = fíkì2

k = fíkì -1 -2

kiíLL = fíkìL -i-1 -2

1SG reside IPFV NON.PST

iiti kii = mijìtáà2

iiti kii = mijìtá -à -2

iiti kiíLL = mijìtá<sub>1</sub> -aa<sub>1</sub> -2

here 1SG live.well.established IPFV NON.PST

kina = áràtà , náàjì jìitàà

kina = áràtà náàjì jìitàà

kina =<sub>1</sub> áràtà<sub>2</sub> náàjì jìitààLL<sub>1</sub>

2PL like thus like

kìnà = fíkì2 k = isákùùjì

Sorted by Price 39/753

Concordancia de texto sobre la raíz **íikì**

Iquito2021 - FieldWorks Language Explorer

File Send/Receive Edit View Data Insert Format Tools Parser Window Help

lquito

**Texts & Words**

Info Baseline Gloss Analyze Tagging Print View Text Chart

Interlinear Texts

- Concordance
- Complex Concordance
- Word List Concordance
- Word Analyses
- Bulk Edit Wordforms
- Statistics

**Lexicon**

Texts & Words

Grammar

Notebook

Lists

Queue: (-/-) No Parser Loaded

Screened by Prof 38/753

146 Word " **iiti** k = iikil2 kina = áràtà , iiti

**Morphemes** k = iikl -i -2 kina = áràtà iiti

**Lex. Entries** kiLL = iikil -ii -2 kina =<sub>1</sub> áràtà<sub>2</sub> iiti

**Lex. Gloss** 1SG reside IPFV NON.PST 2PL like here

kii = mijitàà2 kina = áràtà , náàji jiiàà kinà = iikil2

kii = mijità -à -2 kina = áràtà náàji jiiàà kinà = iikl -i -2

kiLL = mijità<sub>1</sub> -aa<sub>1</sub> -2 kina =<sub>1</sub> áràtà<sub>2</sub> náàji jiiààLL<sub>1</sub> kinà =<sub>1</sub> iikil -ii -2

1SG live.well.established IPFV NON.PST 2PL like thus like 2PL reside IPFV NON.PST

k = isákùuji

k = isákùuji

kiLL = isákùuji

1SG upriver.of

Free Eng "Here I live like you and here I live well, like you, just as you live upriver of me.

Spn "Aquí vivo como ustedes y aquí vivo bien, así como ustedes viven río arriba de mí.

147 Word " Naa = kw = anitíi1kiàà = na kináàjà .

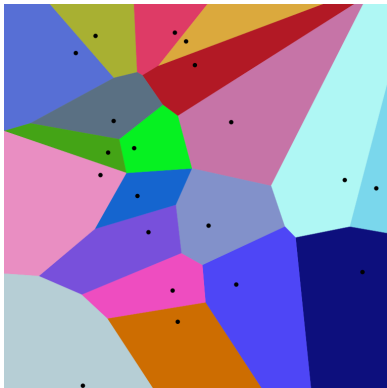
**Morphemes** Naa = kw = anitíi -1 -kiàà = na kináàjà

Un ejemplo en el corpus la raíz **iikil**



## Método de contraste semántico

- Uno de los mejores métodos para mejorar la calidad de los recursos léxicos es utilizar el concepto de **contraste semántico** para refinar las definiciones y descripciones de las funciones.
- Este método toma dos formas:
  - Comparar formas con glosas o definiciones idénticas o similares para aclarar la naturaleza de la diferencia entre ellos
    - palabras → significados/funciones
  - Tomar un campo semántico y explorar qué formas expresan conceptos dentro de ese campo.
    - significados/funciones → palabras



Contrastes semánticos en un campo semántico

# Identificar contrastes semánticos entre palabras

- **Paso 0.**

- En las etapas tempranas de la documentación léxica, con frecuencia se recolectan formas de manera oportunista, obteniendo definiciones aproximadas.

- **Paso 1.** Identifique palabras con glosas iguales o similares en su borrador de recursos léxicos.

- Si usa FLEx, ordene alfabéticamente los campos de glosa y/o definición.
- En los primeros años de desarrollo de un recurso léxico, dicha búsqueda conduce a muchos conjuntos de formas que tienen glosas y/o definiciones sospechosamente similares.
- A medida que la calidad del recurso léxico mejora con el tiempo y el esfuerzo, este tipo de búsquedas arrojan cada vez menos casos que requieren mayor investigación.

- **Paso 2.** Consulte con los hablantes:
  - Pregunte a los hablantes si las formas seleccionadas tienen significados diferentes.
    - En caso afirmativo, pídeles que reflexionen sobre la diferencia de significado y pídeles que traten de explicar la diferencia.
  - Si dicen que no hay diferencia en el significado, discuta si las formas tienen diferencias de algún otro tipo (por ejemplo, estilísticas, dialectales o sociolingüísticas).
- **Paso 3.** Repita los pasos anteriores con varios hablantes.
  - Distintos hablantes pueden tener diferentes puntos de vista sobre el significado de las formas.
  - Los hablantes pueden diferir en la facilidad con la que pueden articular diferencias en los significados de las formas.





- Lo que surgió de la discusión posterior fue:
  - En su sentido más general, *ííkùkù* es un término de parte de cuerpo obligatoriamente poseído que se refiere a la superficie de algo.
  - Si el poseedor es una persona, entonces se entiende que *ííkùkù* se refiere a la superficie de su cuerpo.
    - Esta es la forma que se usaría cuando, por ejemplo, se habla de alguien que se pinta el cuerpo.
  - Pero la superficie del cuerpo es una metonimia de todo el cuerpo.

## Ejemplo 2: 'jalar' en iquito

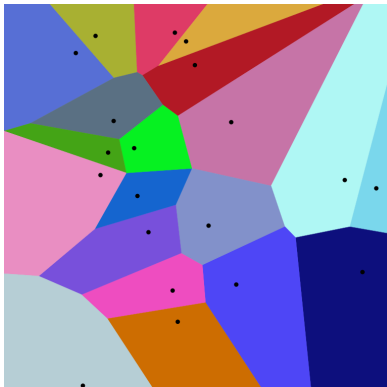
- Si uno consulta el borrador de 2004 del diccionario de Iquito, encuentra dos verbos glosados como 'jalar' (PULL): *j#* y *at#*.
- Sin embargo, cuando presentamos estas dos formas a los hablantes y les pedimos que las contrastaran, inmediatamente nos explicaron que significan cosas bastante diferentes:
  - *j#*: mover algo tirando de él
  - *at#*: enderezar algo largo y flexible en toda su longitud, por ejemplo, una cuerda, tirando de uno o ambos extremos

## Contrastes no semánticos

- Tenga en cuenta que los contrastes entre formas con glosas/definiciones similares no son necesariamente de naturaleza semántica.
- Pueden ser en cambio ser:
  - dialectal
  - estilístico (por ejemplo, juguetón, arcaico)
  - génerolectal
  - eufemístico

## Explorar micro-campos semánticos

- Para aclarar el significado de una forma, a menudo queremos determinar no solo cómo contrasta semánticamente con las formas conocidas, sino también con otras formas que aún tenemos que descubrir.
- Por ejemplo, en el caso de iquito *jalar* (PULL), aclaramos el contraste semántico entre las dos formas comenzamos con *ji* y *ati*, pero hay muchos otros conceptos en JALAR microdominio semántico
  - tirar de algo, pero no moverlo, por ejemplo, porque es demasiado pesado
  - tirar de algo y separarlo de otra cosa
  - tirar de algo para probar su fuerza
  - sacar algo de un contenedor tirando de él
  - ...



Un campo semántico

## Algunas formas del iquito para conceptos en el campo JALAR

- *riikii*: ejercer fuerza sin mover; levantar, empujar o jalar algo sin moverlo
- *atiniiw#ita*: tirar repetidamente de ambos extremos de algo; p.ej., para probar la fuerza de una cuerda, o cuando un perro tira de un trozo de carne con los dientes
- *at#ita*: tensar un objeto parecido a una cuerda que está atado por un extremo tirando de él
- *j#itata*: arrancar o sacar algo unido a un objeto más grande en un punto de contacto relativamente pequeño
- *kajata*: arrancar pelos
- *kir#it#t#*: quitar algo unido a un objeto más grande con un área de contacto relativamente grande
- ...

# Explorar microdominios semánticos

- Hay varios métodos que se puede utilizar para explorar campos semánticos:
  - Uso de recursos existentes diseñados para explorar campos semánticos
  - Creando estímulos nuevos
  - Listado libre ('freelisting')



## Estímulos existentes

- Para algunos dominios existen recursos que proporcionan una **cuadrícula ética** (etic grid) preliminar para el dominio en cuestión:
  - Videos de eventos de cortar y romper (Bohnmeyer et al. 2001)
  - Términos de color: fichas de carta de colores de Munsell
  - Relaciones espaciales topológicas: serie de imágenes de relaciones topológicas (Bowerman y Pederson 1992)
  - Demostrativos (Wilkins 1999)
  - Partes del cuerpo (Enfield 2006)
- Muchos de estos recursos se pueden encontrar en el sitio de TulQuest y en la página de materiales de estímulo y manuales de campo de lenguaje y cognición de MPI.
- Dichos recursos sirven como un punto de partida útil para explorar dominios semánticos, pero pocos de ellos capturan la riqueza semántica completa de un idioma en particular.

## Crear tus propios estímulos

- Creación improvisada de estímulos o cuestionarios:
  - Sea imaginativo en enumerando los posibles conceptos en un dominio (ejemplo de JALAR).
  - Use objetos en el entorno y una variedad de acciones en un dominio, p.ej., utiliza una soga y una variedad de otros objetos para explorar el dominio semántico AMARRAR.
- Aprovechar clasificaciones o tipologías en otras disciplinas:
  - Estudios etnográficos de cultura material, prácticas de subsistencia, prácticas artísticas, etc.
  - Guías de campo para especies mamíferos, aves, herpes, etc., inclusive de mapas de distribución, para crear listas de especies encontradas en el área.
    - (La identificación de especies es una actividad desafiante que requiere sus propias metodologías).







## Campos semánticos especiales

- Tenga en cuenta que hay ciertos campos semánticos que se requerirán de herramientas y metodologías de elicitación especiales.
  - La etnobiología: uso de grabaciones de audio, guías de campo apropiadas y metodologías diseñadas para la identificación de especies/géneros.
  - El parentesco: uso de diagramas de parentesco y metodologías apropiadas.
  - Los colores: uso de fichas de color.