

[ESTUDIOS]

De la elicitación al texto: una nota sobre el trabajo lingüístico de campo*

From elicitation to text: a note on linguistic fieldwork

Igor Vinogradov¹

Instituto de Investigaciones Antropológicas, UNAM, México

vinogradov@iia.unam.mx

Resumen: La actividad principal que nutre la documentación de una lengua es la recopilación de una muestra de discursos (textos) que ocurren de forma natural en una comunidad de hablantes. Los datos textuales a menudo se complementan en las sesiones de elicitación que permiten aclarar el estatus gramatical de frases y palabras fuera del contexto. En el caso de lenguas severamente amenazadas y moribundas, la recolección de discursos naturales se convierte en un desafío, ya que los últimos hablantes suelen utilizar su lengua nativa en oca-

Palabras clave: trabajo de campo, documentación de lenguas, discurso natural, lenguas yumanas, lenguas amenazadas. **Keywords:** fieldwork, language documentation, natural discourse, Yuman languages, endangered languages.

* Este artículo es producto de la investigación realizada en el marco del Programa UNAM-PAPIIT, proyecto IA400624 "Documentación y descripción del kumiay y ku'alh, dos lenguas en peligro de extinción". La primera etapa de la documentación lingüística del ku'alh en 2023 fue financiada por *The Jacobs Research Funds*. Agradezco a Manuel Alejandro Sánchez-Fernández por introducirme a la comunidad indígena de Santa Catarina y por acompañarme en varias ocasiones de trabajo de campo. Este estudio no hubiera sido posible sin el apoyo incondicional de Telma Cañedo, Raúl Sandoval y Daria Mariscal, hablantes nativos de la lengua ku'alh. Asimismo, agradezco a las personas anónimas que dictaminaron la primera versión de este artículo.

¹ Es doctor en Lingüística por la Universidad Estatal de Moscú (2013). Sus investigaciones han contribuido a la documentación y descripción de las lenguas que están en peligro de extinción. Es autor de 2 libros y más de 40 artículos en revistas arbitradas. La mayoría de sus trabajos se dedican a la morfología, sintaxis y discurso de las lenguas mayas en sus aspectos sincrónico y diacrónico, como también a la tipología gramatical de las lenguas de Mesoamérica. Ha realizado trabajo de campo con hablantes de varias lenguas indígenas en México y Guatemala. Tiene cuatro años de experiencia posdoctoral en el Instituto de Investigaciones Antropológicas en la UNAM y en el Departamento de Antropología de las Américas en la Universidad de Bonn, Alemania. Desde septiembre 2022 es investigador asociado en el Instituto de Investigaciones Antropológicas de la UNAM.

siones muy limitadas. Con base en las experiencias de la documentación de la lengua ku'alh (familia yumana, Baja California, México), este artículo plantea la posibilidad de incentivar a los colaboradores nativohablantes a producir discursos naturales en el marco de las sesiones de elicitación con la finalidad de equilibrar la muestra de datos lingüísticos. Los criterios a tener en cuenta en la elicitación se ilustran con ejemplos concretos. Un análisis de sus ventajas y limitaciones permite proponer algunas pautas metodológicas para unir la elicitación tradicional y la obtención de textos orales en el marco de la misma actividad.

Abstract: The main activity that informs language documentation is the collection of a sample of naturally occurring speech (texts) in a community of speakers. Textual data are often supplemented by elicitation sessions that allow to clarify the grammatical status of phrases and words out of context. In the case of severely endangered and moribund languages, the collection of natural speech becomes a challenge, as the last speakers often use their native language on very limited occasions. Based on the experiences of documenting the Ko'alh language (Yuman family, Baja California, Mexico), this paper discusses the possibility of encouraging native-speaking collaborators to produce natural speech during elicitation sessions in order to obtain a more balanced sample of linguistic data. The criteria to be taken into account in elicitation are illustrated with concrete examples. An analysis of its advantages and limitations allows for the proposal of some methodological guidelines for combining traditional elicitation and oral text collection within the same activity.

1. Introducción

El presente escrito presenta una reflexión metodológica que surgió a partir de las experiencias de la documentación lingüística de la lengua ku'alh (familia yumana) en Baja California, México. En las diferentes escalas de lenguas amenazadas (Krauss, 2007; Lewis y Simons, 2010; Lee y van Way, 2016) el ku'alh se clasifica como lengua moribunda. La cantidad de hablantes no excede 20 personas, que

en su mayoría son adultos mayores (Vinogradov y Sánchez-Fernández, 2025: 258), ya que la transmisión de la lengua a nuevas generaciones se ha interrumpido. Los espacios públicos y privados donde se habla la lengua son muy reducidos y para algunos miembros de la comunidad ni siquiera existen. Varias personas, aunque declaran ser hablantes del ku'alh, no utilizan la lengua en su vida diaria.

El ku'alh se habla principalmente en la comunidad indígena de Santa Catarina que se localiza en el municipio de Ensenada, en el estado de Baja California. En esta comunidad los ku'alh conviven con hablantes del pa ipai, otra lengua indígena, quienes constituyen la mayoría étnica. Filogenéticamente, ambas lenguas pertenecen a la familia yumana, pero no tienen parentesco cercano; véase Kroeber (1943) y Joel (1964) para la clasificación de las lenguas yumanas, y Miller (2018) para la perspectiva actualizada. El ku'alh forma parte del subgrupo diegueño de la rama delta-californiano, mientras que el pa ipai pertenece al subgrupo pai-río, junto con varias lenguas (hualapai, yavapai, havasupai, etc.) que se hablan a lo largo del río Colorado en Arizona, Estados Unidos.

Una de las primeras menciones de la existencia en Santa Catarina de una lengua distinta del pa ipai le pertenece a Owen (1960). Sin embargo, después del trabajo pionero de Robles (1964), durante décadas el ku'alh y sus hablantes no atraían la atención por parte de los lingüistas ni antropólogos (Sánchez-Fernández, 2022). Los estudios del ku'alh, de manera más o menos sistemática, empezaron hace quince años. El reporte preliminar del Instituto Nacional de Lenguas Indígenas (INALI, 2010) fue una respuesta a la necesidad de reconocer el ku'alh de manera oficial en los organismos gubernamentales, con el fin de proteger los derechos lingüísticos de sus hablantes. Al mismo tiempo, el ku'alh ha sido incluido en los proyectos de documentación de las lenguas yumanas que se estaban llevando a cabo en los Estados Unidos (Field y Miller, 2018).

De acuerdo con la definición clásica de Himmelmann (1998), la documentación lingüística es una actividad que se dedica al registro de las prácticas y tradiciones lingüísticas de una comunidad de habla; véase también Woodbury (2011). La documentación lingüística busca recolectar, procesar y preservar una muestra de datos empíricos sobre la lengua que sea representativa en cuanto a la estructura de esta lengua y que pueda dar una impresión justa sobre cómo y para qué se utiliza (Lehmann, 2001). Particularmente en su etapa inicial, la documentación se realiza

mediante trabajo de campo, que implica una serie de procedimientos dedicados a la recolección y procesamiento de datos lingüísticos en las comunidades donde residen los hablantes de la lengua, y en estrecha colaboración con ellos (Himmelman, 2006; Austin, 2016).

Las experiencias de la documentación del ku'alh indican que estos procedimientos pueden variar dependiendo del grado de vitalidad de la lengua. Para aumentar la calidad de datos que se recopilan en campo y ampliar la representatividad de la muestra, se sugiere considerar el grado de peligro en que se encuentra la lengua al momento de planificar su documentación. Este artículo plantea la posibilidad de ajustar los procedimientos de trabajo lingüístico de campo relacionados con las técnicas de elicitación, en específico, en el caso de lenguas moribundas.

2. El trabajo lingüístico de campo

El trabajo de campo es una actividad esencial para los lingüistas que se dedican al estudio de lenguas minoritarias y poco descritas. Estas lenguas no cuentan con una cantidad suficiente de recursos documentales, por lo que es imposible utilizar en el estudio lingüístico fuentes secundarias, como gramáticas, diccionarios, materiales de archivo o textos publicados previamente. El objetivo general del trabajo de campo es obtener datos primarios sobre las estructuras lingüísticas. El trabajo lingüístico de campo típicamente incluye dos tipos de actividades: los ejercicios de elicitación y la grabación (en forma de audio o video) de discursos que ocurren en la comunidad de hablantes de manera natural.

La elicitación puede definirse como un conjunto de procedimientos metodológicos que tienen por objetivo obtener datos lingüísticos sobre un fenómeno particular en un contexto controlado. La elicitación en su forma prototípica implica el uso de ciertos estímulos (palabras, frases o expresiones cortas) en la metalengua que el colaborador o la colaboradora nativohablante tiene que traducir a su lengua nativa. En el contexto mexicano, la metalengua por lo general es el español. La colección de estímulos verbales forma un cuestionario lingüístico; por ejemplo, véase el famoso cuestionario de Dahl (1985) para el estudio de categorías verbales

de tiempo, aspecto y modo o el cuestionario de Kassian *et al.* (2010) para la revisión semántica de la lista de ítems léxicos básicos de Swadesh. Para evitar la tarea de traducción y, por consiguiente, para poder colaborar con hablantes monolingües, algunos investigadores optan por técnicas de elicitación monolingüe (Dale, 1978; Everett, 2001).

La mayor crítica metodológica concierne a la naturaleza de los datos elicitados que no representan el uso natural de la lengua, sino se producen en condiciones creadas por el investigador de manera artificial, lo que puede alterar los resultados. Como respuesta a estas críticas, para disminuir la influencia de la metalingua, en varios estudios se ha optado por el uso de estímulos visuales en forma de dibujos, videos o secuencias de imágenes que forman una historia; por ejemplo, véase Kinloch (1971) y Burton y Matthewson (2015). La tarea del colaborador nativohablante en este caso ya no implica traducción o corrección del estímulo verbal, sino consiste en narrar en su lengua nativa lo que observa en los estímulos.

El problema con la confiabilidad de los datos elicitados ya era evidente a principios del siglo pasado, cuando la lingüística antropológica apenas daba sus primeros pasos. Franz Boas fue uno de los primeros investigadores que, partiendo de la hipótesis de que los procesos lingüísticos en mayor medida permanecen inconscientes para los hablantes (Jakobson, 1944), se opuso a la idea de limitar la base empírica a las entrevistas y señaló la necesidad de ampliar la metodología para incluir el análisis de textos o discursos que ocurren naturalmente. La recolección y análisis de textos también tienen sus propios problemas metodológicos, desde la cuestión ética de recopilar conversaciones de otras personas hasta las complicaciones técnicas de realizar una grabación de buena calidad en un ambiente natural, con posibles ruidos al fondo y con varias personas hablando al mismo tiempo. La transcripción y un minucioso análisis lingüístico de las muestras de habla natural también presentan un desafío, ya que requieren mayor grado de conocimiento de la lengua y mucho más tiempo que una sesión de elicitación.

Actualmente, los expertos en la lingüística de campo y en la documentación de lenguas amenazadas y poco estudiadas parecen coincidir que es importante incluir en los materiales de documentación lingüística tanto los datos elicitados como los que provienen del análisis de textos grabados y transcritos (Chelliah, 2001; Matthewson, 2004), con el fin de garantizar la representatividad de datos

primarios, para que un estudio descriptivo que se lleve a cabo con base en este material pueda dar resultados válidos y confiables. No obstante, existen ocasiones cuando es imposible obtener datos de los dos tipos. En el caso de lenguas extintas que ya no tienen hablantes, no es posible realizar las actividades de elicitación. En el caso de lenguas moribundas que están próximas a la extinción y que no han sido documentadas con anterioridad, a menudo es difícil conseguir una muestra de discursos naturales.

Estas lenguas, por definición, cuentan con una cantidad muy pequeña de hablantes fluidos que son bilingües y que (casi) no utilizan su lengua materna. Para las personas que ya no tienen la costumbre de hablar la lengua, las tareas de traducir palabras y componer frases cortas, que se hacen en el marco de una sesión de elicitación, pueden ser más fáciles y más atractivas que producir una narración coherente o interactuar verbalmente con otras personas. En cambio, la petición de contar alguna historia, leyenda o anécdota en la lengua es una tarea demasiado compleja que puede provocar que el colaborador se sienta incómodo o decline la invitación a participar en las actividades de documentación. En este contexto, la elicitación naturalmente se convierte en el método principal de obtención de datos lingüísticos. Sin embargo, el requerimiento de equilibrar las fuentes de datos incluyendo en el análisis textos y discursos producidos de manera natural se mantiene. A continuación, se presenta una posible solución que, al menos en parte, podría ayudar a satisfacer este requerimiento metodológico.

3. De la elicitación al texto

Como se ha mencionado anteriormente, en las entrevistas sociolingüísticas, la mayoría de los hablantes del ku'alh indican que llevan años sin hablar la lengua. Como la costumbre de utilizar el ku'alh en la comunicación diaria se está perdiendo, muchos hablantes tienen dificultades al momento de intentar contar una historia en forma de un monólogo o platicar con otros hablantes. La falta de fluidez en los últimos hablantes de una lengua es bien conocida en la literatura (Evans, 2001). Entre los problemas más comunes están las palabras que se olvidan, los nuevos conceptos que no tienen un equivalente en ku'alh, el uso de préstamos del español y, con menor frecuencia, del pa ipai, así como las alternancias inconscientes de código

entre el ku'alh y el español. Si se logra producir un discurso más o menos coherente, los mismos hablantes a menudo caracterizan el resultado como artificial y forzado.

Una de las posibilidades sería proporcionar a los colaboradores más soporte contextual. Una forma común para hacerlo es mediante estímulos visuales, como películas o series de imágenes (Hayes, 1954; Yegerlehner, 1955; Kinloch, 1971; Bochnak y Matthewson, 2020). Por ejemplo, el cuento de la rana que se hizo famoso en los estudios del discurso narrativo a partir del trabajo de Berman y Slobin (1994), ha sido empleado con éxito en los proyectos lingüísticos enfocados en el pa ipai (Sánchez-Fernández y González, 2023). No obstante, en las experiencias de la documentación del ku'alh los estímulos visuales no han ayudado por varias razones. A causa de diferencias culturales, los objetos y eventos que aparecen en las imágenes, a veces no corresponden a un lexema claro en ku'alh, y esto lleva a que el hablante se detiene en búsqueda de una palabra adecuada y no puede seguir con la narración. Si se trata de una secuencia de imágenes, no siempre está clara la relación lógica o temporal entre ellas. Para personas mayores que no tienen experiencia con álbumes ilustrados, resulta difícil entender las convenciones abstractas que son obvias para los autores de los estímulos, como, por ejemplo, enmarcar en una nubecita lo que el protagonista está pensando, o poner un tache cerca de una imagen para indicar la negación; véase TFS Working Group (2011), por ejemplo.

En cuanto a las conversaciones, el problema principal reside en que todos los hablantes son bilingües, y una plática sobre muchos temas es más fácil y natural en español que en ku'alh. Aunque los colaboradores empiecen con la intención de hablar en ku'alh, en la mayoría de las ocasiones terminan platicando en español o dejan de conversar, sintiendo que el diálogo no sale de manera fluida. Además, existen dificultades logísticas para que dos o más personas de edad avanzada se encuentren en el mismo lugar donde, además, se pueda hacer una grabación de buena calidad (Caldecott y Koch, 2014).

Una posible solución al problema de falta de textos en la documentación del ku'alh se presentó durante una experiencia de elicitación con un hablante particular. Varios trabajos metodológicos han señalado la importancia de llevar a cabo entrevistas de manera flexible, dejando espacio para aportaciones voluntarias de los colaboradores nativohablantes. Clark (1971: 55) discute el problema de mantener el equilibrio entre la recopilación de información adicional que proviene de las

historias que quiera contar el colaborador y el seguimiento de una entrevista estructurada con propósitos claros y bien definidos desde un principio. Más recientemente, Mithun (2001), Matthewson (2004) y Crane y Fleisch (2019) hacen énfasis en que toda la información proporcionada por el colaborador por su propia iniciativa puede ser muy valiosa para la documentación lingüística. Sus opiniones sobre el contexto, las posibles alternativas de expresar lo mismo, los comentarios sobre el estilo, formalidad, el orden de palabras y los matices de significado permiten documentar no solamente el comportamiento lingüístico, sino también ofrecen acceso al conocimiento metalingüístico, una de las partes fundamentales en la documentación de una lengua (Himmelmann, 1998). Respetando este principio metodológico, en varias ocasiones el investigador abrió espacio a las contribuciones propias por parte del colaborador durante la elicitación. Esperando la siguiente pregunta, en una de las sesiones el colaborador empezó a complementar las frases que iba a traducir al ku'alh, como en (1).

(1)

Investigador: *la mujer se fue y nunca regresó* [la frase del estímulo]

a) nich'ak wa ñam ñinimak chwa

nich'ak	w-a	ñam	ñi-nimak	ch-wa
mujer	3-ir	ya	3/1-abandonar	3-estar.sentado

'la mujer se fue y me abandonó'

b) ñinimak chwa

ñi-nimak	ch-wa
3/1-abandonar	3-estar.sentado

'me abandonó'

c) pa l'a – y *nunca regresó*

pa	l'a
llegar	NEG

'no regresó'

d) ñam pa l'a nich'ak

ñam	pa	l'a	nich'ak
ya	llegar	NEG	mujer

'la mujer ya no regresó'

e) ipa nmiy chwa

ipa	nmiy	ch-wa
hombre	estar.enojado	3-estar.sentado

'el hombre está enojado'

f) ñam pa l'a

ñam	pa	l'a
ya	llegar	NEG

'ya no regresó'

g) ñinimak chwa ñam pa l'a

ñi-nimak	ch-wa	ñam	pa	l'a
3/1-abandonar	3-estar.sentado	ya	llegar	NEG

'me abandonó, ya no regresó'

h) yey lich wa – *estoy triste*

yey	lich	wa
corazón	estar.mal	estar.sentado

'estoy triste'

i) yey lich pa'am

yey	lich	pa'am
corazón	estar.mal	quedarse

'me quedé triste'

j) ñay yey lich pa'am

ñay	yey	lich	pa'am
yo	corazón	estar.mal	quedarse

'yo me quedé triste'

k) nich'ak ñam wa ñam

nich'ak	ñam	w-a	ñam
mujer	ya	3-ir	ya

'la mujer ya se fue'

l) ñam pa l'a

ñam	pa	l'a
ya	llegar	NEG

'ya no regresó'

La frase del estímulo dio lugar a varias interpretaciones, inclusive "me abandonó" en (1b), "el hombre está enojado" en (1e) y "me quedé triste" en (1i) que no forman parte del estímulo original en español. La exagerada cantidad de repeticiones se debe a la costumbre que el colaborador desarrolló viendo que el investigador intenta transcribir palabra por palabra cada frase que él pronuncia. Sin embargo, es notable que casi nunca son repeticiones exactas, compárese (1c) y (1d) o (1i) y (1j), por ejemplo. De igual manera, estando acostumbrado a la tarea de traducción, el mismo colaborador inserta en algunos fragmentos los equivalentes en español; véase (1c) y (1h).

Como el colaborador estaba produciendo frases directamente en ku'alh, su aporte ha sido muy bienvenido. A pesar de haber surgido en un contexto de elicitación que fue creado de manera artificial, se puede considerar que estas frases forman un discurso o texto, aunque corto, pero coherente y natural en el sentido de que lo produjo una persona nativohablante de manera espontánea sin tener que traducir el contenido del español previamente. Durante dos sesiones de dos horas cada una en junio de 2023, surgieron cinco textos de este tipo, sin que el investigador los haya solicitado. Es notable que uno de estos textos, ilustrado en (2), está en forma de una conversación, en la que el hablante imagina qué le podría responder un interlocutor imaginario.

(2)

Investigador: *este niño llora mucho* [la frase del estímulo]

a) umiy war jikwan

u-miy	war	jikwan
3-llorar	mucho	niño
'el niño llora mucho'		

b) muyuyi

muyu-yi
por.qué.es-MD
'¿por qué será?'

c) muyum umiy war

muyu-m	u-miy	war
por.qué.es-DS	3-llorar	mucho
'¿por qué llora mucho?'		

d) umiy war

u-miy war
3-llorar mucho
'llora mucho'

e) ñay 'uyaw l'a

ñay '-uyaw l'a
yo 1-saber NEG
'yo no lo sé'

f) muyum

muyu-m
por.qué.es-DS
'¿por qué?'

g) 'uyaw l'a

'-uyaw l'a
1-saber NEG
'no lo sé'

h) urap chwayi tayim

u-rap ch-wa-yi tayim
3-doler 3-estar.sentado-MD cualquiera
'algo le duele'

i) miy urap chwa 'uyaw l'a

miy u-rap ch-wa '-uyaw l'a
pierna 3-doler 3-estar.sentado 1-saber NEG
'[tal vez] le duele una pierna, no lo sé'

j) 'uyaw l'aj

'-uyaw l'aj
1-saber NEG
'no lo sé'

El texto más largo presenta una descripción, de igual manera absolutamente voluntaria, de lo que el colaborador estaba observando mientras estaba trabajando en la sesión de elicitación. Uno de los participantes en la entrevista se levantó de su asiento y fue a caminar un rato en los alrededores de la casa del colabora-

dor, quien se dio cuenta de esto y empezó a reflexionar en ku'alh sobre las posibles causas y consecuencias (3).

(3)

a) ñam yey ulich a

ñam	yey	u-lich	a
ya	corazón	3-estar.mal	MD

'ya está triste'

b) ñam tuwar

ñam	tuwar
ya	cansarse

'ya se cansó'

c) wamp chwa tuwar

w-amp	ch-wa	tuwar
3-caminar	3-estar.sentado	cansarse

'está caminando, se cansó'

d) wamp chwa

w-amp	ch-wa
3-caminar	3-estar.sentado

'está caminando'

e) matx kutxapam wa

matx_kutxap-am	w-a
arroyo-DAT	3-ir

'se fue para el arroyo'

f) matx kutxap sam wa

matx_kutxap	sa-m	w-a
arroyo	DET-DAT	3-ir

'se fue allá al arroyo'

g) ñam kway chyew

ñam	kway	chyew
ya	irse	para.que

'ya se quiere ir'

h) ñam tuwar chayaw

ñam	tuwar	ch-ayaw
ya	cansarse	3-andar

'ya está cansado'

i) ñam tuwar ñam kway chyaw

ñam	tuwar	ñam	kway	chyaw
ya	cansarse	ya	irse	para.que

'ya se cansó, ya se quiere ir'

[una intervención por parte del investigador]

j) ipa waña uyaw chwa

ipa	waña	u-yaw	ch-wa
hombre	camino	3-agarrar	3-estar.sentado

'el hombre agarró el camino'

k) sam wa

sa-m	w-a
DET-DAT	3-ir

'se fue para allá'

l) perder uyaw l'aj

perder	uyaw	l'aj
perderse	saber	NEG

'se va a perder, no conoce [por acá]'

m) chumay 'aniw 'uu l'a

chumay	'-aniw	'-uu	l'a
buscar.PL	1-andar.PL	1-ver	NEG

'lo vamos a buscar y no lo veremos'

n) waña uyaw l'a

waña	uyaw	l'a
camino	saber	NEG

'no conoce el camino'

o) perder wañam chirwiy 'uu l'aj

perder	waña-m	chirwiy	'-uu	l'aj
perderse	camino-DAT	medio	1-ver	NEG

'si se pierde en medio del camino, no lo veremos'

p) matxam perderkis 'uu l'aj

matx-am	perder-kis	'-uu	l'aj
tierra-DAT	perderse-si	1-ver	NEG

'si se pierde en el monte, no lo veremos'

q) matx uyaw l'aj ya we

matx	uyaw	l'aj	ya	we
tierra	saber	NEG	MD	MD

'no conoce [esta] tierra pues'

r) matx uyaw l'aj

matx	uyaw	l'aj
tierra	saber	NEG

'no conoce [esta] tierra'

[una intervención por parte del investigador]

s) kur sam uyiw chayaw

kur	sa-m	u-yiw	ch-ayaw
ya	DET-DAT	3-venir	3-andar

'ya está por allá viniendo'

t) pa'aw chyaw

pa'aw	ch-yaw
pararse	3-estar.parado

'está parado'

u) yaak wuuk chyaw – *está mirando para acá*

yaak	w-uu-k	ch-yaw
este.lado	3-ver-DIR	3-estar.parado

'está mirando para acá'

v) yaak pa'aw wuuk chyaw

yaak	pa'aw	w-uu-k	ch-yaw
este.lado	pararse	3-ver-DIR	3-estar.parado

'se paró y está mirando para acá'

Como se puede ver en (3), el colaborador retoma la narración dos veces, a pesar de que el investigador quería seguir preguntando el material que había preparado para la entrevista.

Como el investigador al principio no estaba preparado para este tipo de contribución textual por parte del colaborador nativohablante, se apegaba más a la dinámica de una sesión de elicitación que a los procedimientos de recolección de textos naturales. No obstante, las interrupciones para pedir repetir la palabra o frase anterior no daban el resultado esperado. Cada vez que el colaborador intentaba volver a lo que ya había dicho anteriormente, las estructuras sintácticas se modificaban y las frases se complementaban con material léxico nuevo, como se puede observar en (4).

(4)

Investigador: *otra vez quemaste el frijol* [la frase del estímulo]

a) kwan mlhimari frijol

kwan m-lhimar-i frijol

ITER 2-quemar-MD frijol

'otra vez quemaste frijol'

b) frijol lhulh mwa mlhimar ñuwaitx a

frijol lhulh m-wa m-lhimar ñuwaitx a

frijol cocer 2-estar.sentado 2-quemar terminarse MD

'estabas cociendo frijol y lo quemaste por completo'

c) lhulh mwa cham mlhimar ñuwaitx a

lhulh m-wa cham m-lhimar ñuwaitx a

cocer 2-estar.sentado todo 2-quemar terminarse MD

'estabas cociendo, y lo quemaste todo por completo'

d) frijol kwan lhulh mwa cham lhap_lhwaitx a

frijol kwan lhulh m-wa cham lhap_lhwaitx a

frijol ITER cocer 2-estar.sentado todo quemarse MD

'otra vez estabas cociendo frijol y todo se quemó'

e) frijol mlhulh mwa

frijol m-lhulh m-wa

frijol 2-cocer 2-estar.sentado

'estabas cociendo frijol'

f) lhap lhwatx

lhap_lhwatx

quemarse

'se quemó'

g) ja pa'am – *falta el agua*

ja pa'am

agua faltar

'faltó el agua'

h) ja pa'am lhap lhwatx sikiyew

ja pa'am lhap_lhwatx sikiyew

agua faltar quemarse por.eso

'faltó el agua, por eso se quemó'

i) ja pa'am sikiyew lhap lhwatx a

ja pa'am sikiyew lhap_lhwatx a

agua faltar por.eso quemarse MD

'faltó el agua, por eso se quemó'

j) ja wuch l'a chwa

ja w-uch l'a ch-wa

agua 3-tener NEG 3-estar.sentado

'no tenía agua'

k) sikiyew lhap lhwatx a

sikiyew lhap_lhwatx a

por.eso quemarse MD

'por eso se quemó'

En el transcurso del fragmento (4), el investigador varias veces interrumpe el discurso pidiendo al colaborador repetir lo que acaba de decir. Sin embargo, en vez de repeticiones exactas el colaborador produce construcciones sinonímicas, como en (4b) y (4d), donde la predicación transitiva "(tú) lo quemaste" se sustituye por una intransitiva: "todo se quemó". En (4e) en la construcción del aspecto progresivo con el verbo auxiliar *wa* 'estar sentado' aparece el prefijo de persona *m-* en el verbo léxico *lhulh* 'cocer', a pesar de que en el mismo contexto en (4b), (4c) y (4d) este prefijo estaba ausente. En (4g) el colaborador agrega la razón por la cual

se podían quemar los frijoles: "les faltaba el agua". Luego, la frase (4g) otra vez se reformula en (4j) mediante el uso de otro verbo y la negación: "no tenían agua".

La naturalidad de estas aportaciones textuales se manifiesta, por ejemplo, en el uso de los marcadores discursivos que normalmente no aparecen en las frases elicítadas que carecen del contexto, tanto lingüístico como metalingüístico. En cambio, en estos pequeños discursos producidos por el colaborador directamente en ku'ah, los marcadores discursivos se emplean de manera sistemática, cuando el contexto imaginario los requiere, como en (5).

(5)

Investigador: *estás contento* [la frase del estímulo]

a) jlikuy mwa

jlikuy	m-wa
estar.contento	2-estar.sentado
'estás contento'	

b) mnmiy mwa – *estás enojado*

m-nmiy	m-wa
2-estar.enojado	2-estar.sentado
'estás enojado'	

c) muyum kunap l'a lhli

muyu-m	kunap	l'a	lhli
por.qué.es-DS	decir	NEG	INT
'¿por qué si no te dije nada?'			

d) muyum mnmiy mwayi me

muyu-m	m-nmiy	m-wa-yi	me
por.qué.es-DS	2-estar.enojado	2-estar.sentado-MD	MD
'¿por qué estás enojado, eh?'			

e) muyuyi

muyu-yi
por.qué.es-MD
'¿por qué será?'

f) muyum mnmiy mwa

muyu-m m-nmiy m-wa
 por.qué.es-DS 2-estar.enojado 2-estar.sentado
 '¿por qué estás enojado?'

g) ñay ch'uch lhli chi l'a – *yo no dije nada*

ñay ch'uch lhli ch=i l'a
 yo nada INT qué=decir NEG
 'yo no dije nada'

h) ch'uch lhli chi l'a ya

ch'uch lhli ch=i l'a ya
 nada INT qué=decir NEG MD
 'no dije nada'

i) ch'uch lhli chi l'a

ch'uch lhli ch=i l'a
 nada INT qué=decir NEG
 'no dije nada'

j) muyum mnmiy mwayi

muyu-m m-nmiy m-wa-yi
 por.qué.es-DS 2-estar.enojado 2-estar.sentado-MD
 '¿por qué estás enojado?'

Los marcadores discursivos tienden a aparecer en abundancia en conversaciones. Se utilizan, sobre todo, para transmitir las expectativas del hablante acerca de la posible reacción del oyente a lo que se le comunica. Hace falta un estudio de índole discursiva y pragmática que pudiera describir las funciones de estos marcadores con mayor precisión.

En (5d) aparecen dos marcadores discursivos de manera simultánea: *-yí* y *me*. Llama la atención que el colaborador repitió la misma frase en (5f) para que el investigador pudiera transcribirla. Sin embargo, como la repetición ya está fuera del contexto de la conversación imaginaria, ambos marcadores discursivos se omitieron.

4. Dificultades y limitaciones

En la sección 3 se presentaron los fragmentos textuales que fueron recopilados durante las sesiones de elicitación en el marco del proyecto de documentación lingüística de la lengua ku'alh. No son resultados de ningún procedimiento específico, ni tampoco hubo un objetivo original para obtener este tipo de datos. Más bien, son productos que surgieron de manera inesperada cuando el colaborador nativohablante tomó la iniciativa de dar seguimiento a la entrevista. Cabe destacar que, para los propósitos de la documentación lingüística, estos fragmentos textuales difícilmente pueden sustituir por completo los textos coherentes en forma de grabaciones "tradicionales" de narraciones y conversaciones. Son demasiado cortos, por lo que no pueden constituir un corpus suficientemente representativo para un estudio de estructuras lingüísticas. Además, a pesar de que representan discurso natural, se generaron en un contexto artificial de elicitación frente a un investigador que no refleja las condiciones reales para el uso de la lengua en la comunidad. Considerando estas limitantes, la utilidad principal de estos textos reside en su potencial para complementar los datos elicitados.

Existen también limitantes metodológicas para la obtención de datos de este tipo. Los fragmentos presentados en la sección anterior surgieron durante el trabajo con un hablante particular, mientras que con los demás este efecto no se ha observado. Otros colaboradores en el proyecto de documentación no han realizado este tipo de intervenciones verbales por su propia iniciativa. Es evidente que la personalidad del colaborador juega un papel fundamental en la producción de textos espontáneos. Como se había mencionado repetidamente en la literatura teórica sobre la documentación lingüística, los hablantes son personas diferentes y tienen personalidades diferentes, y los investigadores deben tomarlo en cuenta al momento de planificar las actividades. Un hablante puede ser bueno para una actividad y el otro, para otra (Chelliah y de Reuse, 2011). Bower (2008: 115-118) menciona que no todos los hablantes se sienten cómodos frente al micrófono, igual que no todos tienen la misma habilidad para imaginar situaciones y describirlas en la lengua con mayor entusiasmo que seguir traduciendo frases aisladas del español. La creatividad verbal del colaborador parece ser el factor esencial para el éxito.

Otro factor importante es el grado en que el hablante está acostumbrado al trabajo colaborativo con lingüistas. Todos los fragmentos que se presentaron en la sección anterior fueron producidos por el hablante que tiene menos experiencia de asesoría lingüística, y es probable que esto no sea una coincidencia. Los demás participantes en el proyecto que ya habían trabajado anteriormente con otros investigadores y que ya están más familiarizados con las dinámicas de elicitación, no han mostrado interés en tomar el protagonismo durante las sesiones de trabajo lingüístico. Posiblemente, esto no se debe a que no tengan la creatividad necesaria, sino a que demuestran respeto a los investigadores y confían en que ellos conocen mejor las estrategias y necesidades para la documentación de su lengua.

El otro problema metodológico tiene que ver con la naturalidad del habla. Los textos que se producen de forma espontánea no se volverán a repetir; por lo menos, en su forma original. En este respecto, los fragmentos presentados en la sección 3 se asemejan al discurso natural. En un inicio esto causaba problemas, ya que el investigador, siguiendo todavía la metodología de elicitación, intentaba apuntar frase tras frase todo el texto que el colaborador estaba produciendo. No obstante, esto era imposible. Como se trata de un discurso natural, la mejor estrategia hubiera sido no interrumpir al hablante y no transcribir el texto de inmediato, sino dejar la transcripción y su verificación para una sesión posterior. Considerando que el colaborador puede empezar a generar estos pequeños fragmentos discursivos de manera inesperada, es específicamente importante grabar toda la sesión de elicitación en audio, desde el principio al final.

5. Conclusiones

Las pautas descritas en este artículo no implican el uso de herramientas específicas y solamente requieren que el investigador sea flexible durante la sesión de elicitación y no se apegue a fuerza al cuestionario preparado con anterioridad, sino que esté atento a las iniciativas del colaborador nativohablante. De esta manera se puede aprovechar de los momentos en los que el colaborador se sienta a gusto para interactuar en la lengua consigo mismo, creando un contexto imaginario a partir de alguna frase elicitada o de la situación metalingüística de la entrevista. Como resultado, se obtiene un fragmento de discurso natural en la lengua que

tiene todos los valores de una narración o una conversación grabada de manera tradicional.

Una de las ventajas de estos fragmentos textuales consiste en la posibilidad de controlar, por lo menos de manera parcial, el contenido léxico. Igual que al recopilar cualquier otro discurso natural, el resultado depende de las cualidades del hablante particular; en este caso, de su imaginación y creatividad. Parece que los hablantes menos apropiados para largas y aburridas sesiones de elicitación son los que más pueden aportar en este tipo de actividades. De esta forma, con una buena planificación del trabajo de campo, el investigador puede obtener los mejores resultados con todos tipos de colaboradores, encontrando actividades más idóneas para cada uno.

La mayor desventaja es que los textos que se producen durante las sesiones de elicitación, por lo general, son demasiado cortos y no pueden reemplazar por completo la recolección de discursos naturales. Sin embargo, los datos textuales obtenidos de esta manera pueden ser específicamente valiosos cuando se trata de la documentación y descripción de lenguas severamente amenazadas que casi no se utilizan en la comunicación diaria. En el contexto sociolingüístico actual, cuando la vitalidad de muchas lenguas originarias en México se está disminuyendo rápidamente y de manera preocupante (García Landa y Terborg, 2011; Velázquez, 2011; Trujillo, 2012; Pfeiler *et al.*, 2014; Sánchez-Fernández y Rojas-Berscia, 2016), estas recomendaciones metodológicas pueden proporcionar una buena estrategia para futuros estudios de lenguas minoritarias y minorizadas.

Abreviaciones

En las glosas se utilizan las siguientes abreviaciones: 1 – primera persona (sujeto), 2 – segunda persona (sujeto), 3 – tercera persona (sujeto), 3/1 – tercera persona (sujeto) y primera persona (objeto), DAT – dativo, DET – determinante, DIR – direccional, DS – distinto sujeto, INT – intensificador, ITER – iterativo, MD – marcador discursivo, NEG – negación, PL – plural.

Referencias bibliográficas

- Austin, Peter K. (2016). Language documentation 20 years on. En Luna Filipović y Martin Pütz (Eds.), *Endangered languages and languages in danger: Issues of documentation, policy, and language rights* (pp. 147-170). John Benjamins.
- Berman, Ruth A. y Dan Isaac Slobin (Eds.) (1994). *Relating events in narrative: A cross-linguistic developmental study*. Lawrence Erlbaum Associates.
- Bochnak, M. Ryan y Lisa Matthewson (2020). Techniques in complex semantic fieldwork. *Annual Review of Linguistics* 6; 261-283.
- Bowern, Claire (2008). *Linguistic fieldwork: A practical guide*. Palgrave Macmillan.
- Burton, Strang y Lisa Matthewson (2015). Targeted construction storyboards in semantic fieldwork. En M. Ryan Bochnak y Lisa Matthewson (Eds.), *Methodologies in semantic fieldwork* (pp. 135-156). Oxford University Press.
- Caldecott, Marion y Karsten Koch (2014). Using mixed media tools for eliciting discourse in indigenous languages. *Language Documentation & Conservation* 8; 209-240.
- Chelliah, Shobhana L. (2001). The role of text collection and elicitation in linguistic fieldwork. En Paul Newman y Martha Ratliff (Eds.), *Linguistic fieldwork* (pp. 152-165). Cambridge University Press.
- Chelliah, Shobhana L. y Willem de Reuse (2011). *Handbook of descriptive linguistic fieldwork*. Springer.
- Clark, Thomas L. (1971). Elicitation in linguistic fieldwork. *American Speech* 46(1/2); 52-57.
- Crane, Thera Marie y Axel Fleisch (2019). Towards a fieldwork methodology for eliciting distinctions in lexical aspect in Bantu. En Lotta Aunio y Axel Fleisch (Eds.), *Linguistic diversity research among speakers of isiNdebele and Sindebele in South Africa* (pp. 129-179). Finnish Oriental Society.
- Dahl, Östen (1985). *Tense and aspect systems*. Blackwell.
- Dale, Ian R. H. (1978). Beyond intuition: The use of questionnaires in linguistic investigation. *Anthropological Linguistics* 20(4); 158-166.
- Evans, Nicholas (2001). The last speaker is dead – long live the last speaker! En Paul Newman y Martha Ratliff (Eds.), *Linguistic fieldwork* (pp. 250-281). Cambridge University Press.

- Everett, Daniel L. (2001). Monolingual field research. En Paul Newman y Martha Ratliff (Eds.), *Linguistic fieldwork* (pp. 166-188). Cambridge University Press.
- Field, Margaret y Amy Miller (2018). *Documentation of the Baja California Yuman languages Kumeyaay and Ko'alh*. Endangered Languages Archive. Recuperado el 18 de febrero de 2025 de <https://www.elararchive.org/dk0357/>.
- García Landa, Laura y Roland Terborg (2011). La vitalidad de las lenguas indígenas de México: un estudio en tres contextos. En Roland Terborg y Laura García Landa (Coords.), *Muerte y vitalidad de lenguas indígenas y las presiones sobre sus hablantes* (pp. 11-28). Universidad Nacional Autónoma de México.
- Hayes, Alfred S. (1954). Field procedures while working with Diegueño. *International Journal of American Linguistics* 20(3); 185-194.
- Himmelman, Nikolaus P. (1998). Documentary and descriptive linguistics. *Linguistics* 36(1); 161-195.
- Himmelman, Nikolaus P. (2006). Language documentation: What is it and what is it good for? En Jost Gippert, Nikolaus P. Himmelman y Ulrike Mosel (Eds.), *Essentials of language documentation* (pp. 1-30). De Gruyter.
- INALI (2010). *Proyecto de documentación inicial de la lengua ku'áhl*. Informe ejecutivo. Instituto Nacional de Lenguas Indígenas. Recuperado el 19 de febrero de 2025 de https://site.inali.gob.mx/pdf/estudios_opiniones/Documentacion_kuahl.pdf.
- Jakobson, Roman (1944). Franz Boas' approach to language. *International Journal of American Linguistics* 10(4); 188-195.
- Joel, Judith (1964). Classification of the Yuman languages. En William Bright (Ed.), *Studies in Californian linguistics* (pp. 99-105). University of California Press.
- Kassian, Alexei; George Starostin; Anna Dybo y Vasiliy Chernov (2010). The Swadesh wordlist: An attempt at semantic specification. *Journal of Language Relationship* 4; 46-89.
- Kinloch, A. M. (1971). The use of pictures in elicitation. *American Speech* 46(1/2); 38-46.
- Krauss, Michael (2007). Classification and terminology for degrees of language endangerment. En Matthias Brenzinger (Ed.), *Language diversity endangered* (pp. 1-8). Mouton de Gruyter.
- Kroeber, Alfred (1943). Classification of the Yuman languages. *University of Califor-*

- nia *Publications in Linguistics* 1; 21-40.
- Lee, Nala Huiying y John Van Way (2016). Assessing levels of endangerment in the Catalogue of Endangered Languages (ELCat) using the Language Endangerment Index (LEI). *Language in Society* 45(2); 271-292.
- Lehmann, Christian (2001). Language documentation: a program. En Walter Bisang (Ed.), *Aspects of typology and universals* (pp. 83-97). Akademie.
- Lewis, M. Paul y Gary F. Simons (2010). Assessing endangerment: Expanding Fishman's GLDS. *Revue Roumaine de Linguistique* 55(2); 103-120.
- Matthewson, Lisa (2004). On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4); 369-415.
- Miller, Amy (2018). Phonological developments in Delta-California Yuman. *International Journal of American Linguistics* 84(3); 383-433.
- Mithun, Marianne (2001). Who shapes the record: the speaker and the linguist. En Paul Newman y Martha Ratliff (Eds.), *Linguistic fieldwork* (pp. 34-54). Cambridge University Press.
- Owen, Roger C. (1960). Baja California: Paipai Indians. *Katunob* 1(3); 19.
- Pfeiler, Barbara, Juana I. Sánchez Hernández y Josué E. Villegas Chin (2014). La vitalidad lingüística del maya en cuatro localidades de la zona metropolitana de Mérida, Yucatán: Cholul, Kanasín, Umán y San José Tzal. En Ricardo López Santillán y Luis Alfonso Ramírez Carrillo (Eds.), *Crecimiento urbano y cambio social: escenarios de transformación de la zona metropolitana de Mérida* (pp. 445-475). Universidad Nacional Autónoma de México.
- Robles Uribe, Carlos (1964). Investigación lingüística sobre los grupos indígenas del Estado de Baja California. *Anales del Museo Nacional de México* 17; 275-301.
- Sánchez-Fernández, Manuel Alejandro (2022). La investigación lingüística de las lenguas yumanas en México (LYUM). *Expedicionario. Revista de Estudios en Antropología* 2(4); 31-43.
- Sánchez-Fernández, Manuel Alejandro y Luis Miguel Rojas-Berscia (2016). Vitalidad lingüística de la lengua paipai de Santa Catarina, Baja California. *LIA-MES: Línguas Indígenas Americanas* 16(1); 157-183.
- Sánchez-Fernández, Manuel Alejandro e Ivette Selene González Castillo (2023). Narraciones y recetas tradicionales pa ipai de Santa Catarina. *Tlalocan* 28(1);

73-113.

TFS Working Group (2011). *The Woodchopper*. Totem Field Storyboards. Recuperado el 18 de febrero de 2025 de <https://www.totemfieldstoryboards.org/stories/>.

Trujillo Tamez, Alma Isela (2012). *La vitalidad lingüística de la lengua ayuk o mixe en tres comunidades: Tamazulapam del Espíritu Santo, San Lucas Camotlán y San Juan Guichicovi*. [Tesis de doctorado, Universidad Nacional Autónoma de México]. Repositorio Institucional de la UNAM.

Velázquez, Virna (2011). El desplazamiento del matlazinca en el Estado de México. En Roland Terborg y Laura García Landa (Coords.), *Muerte y vitalidad de lenguas indígenas y las presiones sobre sus hablantes* (pp. 241-258). Universidad Nacional Autónoma de México.

Vinogradov, Igor y Manuel Alejandro Sánchez-Fernández (2025). El relato de El riaino: una narración ku'alh. *Tlalocan* 30(1); 253-281.

Woodbury, Anthony (2011). Language documentation. En Peter Austin y Julia Sallabank (Eds.), *The Cambridge handbook of endangered languages* (pp. 159-186). Cambridge University Press.

Yegerlehner, John (1955). A note on eliciting techniques. *International Journal of American Linguistics* 21(3); 286-288.