
Entretextos - Artículos/Articles

Revista de Estudios Interculturales desde Latinoamérica y el Caribe
Facultad Ciencias de la Educación. Universidad de La Guajira. Colombia
ISSN: 0123-9333. Año: 14 No. 26 (ENERO-JUNIO), 2020, pp. 141-163

Un proyecto pionero inconcluso: La base de datos lingüísticos y cartografiables del Centro Colombiano de Estudios de Lenguas Aborígenes -CCELA-, década de los 90⁶⁴

An unfinished pioneering project: The Linguistic and Cartographic Database of the Colombian Center for Aboriginal Language Studies, CCELA, 1990s

*Wanee a'yatawaa maakajaainyülia: Jiichajia anüikii akaraloktajuushi
jajapulujutu Laülaapia*

Kasanaaratta Napüütchin Kusina julu'u Kolompia CCELA 90 tou'upüna.

María Emilia Montes Rodríguez⁶⁵
Docente Universidad Nacional de Colombia

Resumen

Se describe lo acontecido entre 1991 y 1998 con el proyecto de la base de datos lingüísticos y cartografiables de 32 lenguas indígenas y criollas, tarea colectiva de largo alcance que se propuso en un centro de investigación colombiano. Se presentan los objetivos de la tarea, los aspectos metodológicos y financieros, las etapas, aspectos generales del diseño de los cuestionarios (fonológico, léxico y gramatical). Ilustramos la exposición con elementos y datos de la lengua amazónica *tikuna*. Hacemos una reflexión más detallada sobre el aporte del componente léxico de la tarea. Resaltamos los hallazgos eminentemente cualitativos sobre la conceptualización del cuerpo humano en las lenguas consideradas. La discontinuidad en las políticas institucionales y de apoyo fue un factor problemático en este tipo de tareas. La rápida obsolescencia de equipos y programas informáticos fue otro factor que incidió en estos proyectos de mediano y largo plazo. Proponemos algunos replanteamientos metodológicos y algunas ideas para

64 Artículo derivado de la participación como miembro del equipo de investigadores de este gran proyecto colectivo (1995 y 1997). Homenaje al profesor Francisco Pérez, egresado de la primera promoción de la Maestría en Etnolingüística (1984-1987), miembro del C.C.E.L.A. y participe del proyecto Base de Datos.

65 Comunicadora social, Universidad Javeriana, Magister en Etnolingüística, Universidad de los Andes. Doctora en Lingüística, Universidad de la Sorbona, Francia. E-mail: memontesr@unal.edu.co

retomar -desde las condiciones actuales- algo de esa valiosa idea pionera que sigue siendo pertinente y necesaria. La difusión digital y en red de datos en bruto no fue posible en su momento; algunos elementos en formatos básicos son recuperables. La historia de la lingüística de lenguas amerindias y criollas en Colombia amerita un espacio que considere hitos como este ambicioso proyecto.

Palabras claves: lexicografía (lenguas indígenas), lenguas nativas (Colombia), informática y lingüística, bases de datos lingüísticas, historia de la lingüística (Colombia), cuestionarios lingüísticos, C.C.E.L.A.

Abstract

We describe here what happened between 1991 and 1998 with the project of the linguistic and cartographic database of 32 indigenous and Creole languages, a long-range collective task that was proposed in a Colombian research center. The objectives of the task, the methodological and financial aspects, the stages, and general aspects of the design of the questionnaires (phonological, lexicon and grammar) are presented. We illustrate the exposition of this database with elements and data from the Amazonian Tikuna language. We make a more detailed reflection on the contribution of the lexical component of the task. We highlight the eminently qualitative findings on the conceptualization of the human body in the languages considered. The discontinuity in institutional and support policies is another problematic factor in this type of task. The rapid obsolescence of computer hardware and software is another factor to consider in medium and long-term projects like this one. We propose some methodological rethinking and some ideas to take up -from the current conditions- something of that valuable pioneering idea that is still pertinent and necessary. Digital and webcasting of raw data was not possible at the time, but some items in basic formats are recoverable. The history of the linguistics of Amerindian and Creole languages in Colombia deserves a space that considers milestones such as this ambitious project that is still relevant.

Key words: lexicography (indigenous languages), native languages (Colombia), computing and linguistics, linguistic databases, history of linguistics (Colombia), linguistic questionnaires.

Aküjuushi palitpüchiru'u

AKüjüna achikirü tü alatakainka 1992 1998 jünainjee karaokta pütchimaajatü jee kejeetsereein apünüin jikii piamaüin nanüiki kusina jee na kepiajiraaka namaa, a'yatawaa müle'u jüikale'erala laülaapia kolonpiaana. Eiyatünüsü tü ju'unajirüinjatüka, jamüinjatüi jukua'ipa jü'yataaya, jütsüin yanamaaka, jerainjatüin ji'ipapa'a, jünüikimaainjatüka tü asakiraakaa (jümüralu'u pütchikaa, jalajeewoliün pütchikaa jee jüshajia) weiyatüin aküjalaakaa jünainjee jukua'ipa nanüikipala kusina amasoonaje'ewolii Tikuuna münüshii. Waikale'erüinjaa wekiii jüpúla anashaatainjatüin jüküjipala pütchi jüchejaaka a'yatawaka. Weiyatashaatain tü oustalaaka julu'ujee jüsanaayashaatain jukua'ipapala tü aküjalaakaa eiyalaaka nanainjee na wayuu kanüikikaa. Moloiriraain na'yataayapala jee nakaalijiapala na alijunakana, apounajaasü a'yatawaka. Jülatamaatüin jukua'ipa wanee kachuweera eeka jüntüin, laüla'alaajaasü nnojülüwa'aya

jajattayaain wanee a'yatawaa eekai múleuin. Wawanajainja'a jukua'ipa a'yatawaaaka jumaayale juu'utku ekiiru'ujeejatka paala jüpüla aikale'era yaajeejachiki juchukuwa'aya tü anasü ekiiru'uyutkaa. Juwalakajia kachuweeralu'ujee jee jülatiruushin makaaya nnojotsü asawatünün paalainka, otta juu'utku a'yatawaaaka eekai kate'ejanain atumawaa paala ouyante'ennüsü juchukua'aya. Jukua'ipa jü'yataayapala nanüiki kusina naa na kepiajiraakana namaa julu'u Kolonpia, kapülainjeesü jümüiwa'a wanee múle'u kasa maaka jaa'in a'yatawaaaka tüü.

Pütchi katsüisükat: Nanüikima'a kusina, anüikii jipialujutu (Kolonpia), kachuweera akanüijuushi, jikii jukua'ipa anüikii, júchikua pütchimaajatü (Kolonpia), jüsakiria juchiirua pütchi C.C.E.L.A.

Hoy en día es una exigencia de la investigación lingüística tener datos procesados con herramientas informáticas y disponibles digitalmente. La lingüística del corpus que se hace sobre lenguas mayoritarias (español, inglés, portugués) trabaja con enormes volúmenes de información procesados a través de la computación. La documentación lingüística recoge –con formatos o métodos prediseñados– grandes cantidades de material sobre lenguas minoritarias y los publica digitalmente para que se acceda a ellos de manera ágil. La sociolingüística procesa también registros voluminosos de la mayor cantidad posible de hablantes y hace tratamiento estadístico de los mismos.

En los 90 del siglo XX, la lingüística de lenguas aborígenes (indígenas y criollas) despuntaba en Colombia con un enfoque eminentemente descriptivo y etnográfico. El antecedente más cercano en este tipo de tareas era la mezcla de lingüística y evangelización propiciada por varias décadas de trabajo del Instituto Lingüístico de Verano –I.L.V.– El contexto de esa década imponía una mirada diferente, impulsada desde sectores de la academia nacional y acorde además con las reivindicaciones de respeto y reconocimiento de la diversidad territorial, cultural y lingüística. En ese ambiente de trabajo científico, la informática aplicada a la lingüística era aún algo novedoso y poco desarrollado en nuestro país.

1. El nacimiento de una idea novedosa

El equipo de trabajo que se formó en la Universidad de los Andes desde 1984 en torno al posgrado de Etnolingüística (iniciativa financiada por la cooperación francesa y el gobierno colombiano⁶⁶) llevó a cabo durante varios años un proyecto que apuntó a sistematizar y procesar con herramientas informáticas los datos de las lenguas que se habían investigado a lo largo del proceso de formación e investigación.

66 Desde 1983 se gestó en un laboratorio de investigación del CNRS la idea del montaje del programa de Maestría en Etnolingüística en Colombia y luego se constituyó el C.C.E.L.A. agrupando profesores franceses y egresados en 1989, después de haber concluido la segunda promoción del posgrado (Landaburu, 1990).

La idea inicial proviene de una conversación entre Jon Landaburu y Christian de Muizon, director del IFEA y profesor del *Museum National d'Histoire Naturelle* en 1990 en la que de Muizon sugiere una base de datos cartografiada. Landaburu confronta la obra de los “*Atlas linguistiques des régions de France*” (para los cuales fue miembro del comité de evaluación ante dirección del *Centre National de la Recherche Scientifique*, C.N.R.S.) y conoció el trabajo del “Atlas Lingüístico Etnográfico”, A.L.E.C. del Instituto Caro y Cuervo (1950-1982), ante lo cual consideró que era una idea con múltiples aplicaciones no sólo lexicográficas⁶⁷. El apoyo del I.F.E.A. materializó la idea que se discutió con Tulio Rojas en París y luego, en septiembre de 1991, con Miguel Ángel Meléndez y todo el equipo del Centro Colombiano de Estudios de Lenguas Aborígenes, C.C.E.L.A. (Jon Landaburu, comunicación personal, correo electrónico de marzo de 2020).

De acuerdo con Rojas, 2001, la conceptualización del proyecto inicia en 1991. Es decir, que su origen es paralelo a la formulación de la nueva Constitución Política colombiana que reconoce explícitamente la diversidad lingüística del territorio. Esa diversidad era uno de los focos de interés de la propuesta comparativa, tipológica que estuvo tras este proyecto.

En 1996 se expresaban así los objetivos del proyecto de Base de Datos:

- Acceso rápido a datos lingüísticos estandarizados y numéricamente importantes relativos a la diversidad lingüística considerada.
- Representación en mapas de la distribución geográfica de estos datos para facilitar la investigación tipológica, comparativa e histórica.
- Publicación o difusión de parte de estos datos, cartografiados o no, en función de las necesidades de las investigaciones (lingüística, arqueológica, histórica, antropológica, etc.) de las organizaciones indígenas, de las agencias del Estado, de los programas educativos, etc. (CCELA, 1996: 7, retomado en Rojas, 2001: 175-176).

La financiación de las fases iniciales de este proyecto de largo aliento tuvo la participación de entidades francesas: el C.N.R.S., el I.F.E.A., el *Institut de Recherche pour le Développement*, I.R.D. (O.R.S.T.O.M.). a través de las acciones de cooperación que estas entidades establecieron con el C.C.E.L.A. de la Universidad de los Andes. A partir de 1996 la fase de trabajos de campo para el cuestionario léxico tuvo el apoyo financiero de Colciencias, el Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología Francisco José de Caldas (Rojas, 2001: 176).

67 La tradición dialectológica de los atlas lingüísticos que se hicieron diversos puntos del mundo es pues un antecedente de esta idea. De manera indirecta también lo es la compilación sobre lenguas indígenas del Instituto Caro y Cuervo (González y Rodríguez de Montes, M.L. 2000). Dicha obra se concibió inicialmente hacia 1985 como un ‘estudio preliminar’ para un atlas, que sería además una tarea común a varios países del Convenio Andrés Bello (según el Prólogo de M. L. Rodríguez en dicha obra). Dado que esto no fue posible, la compilación del 2000 quedó como un testimonio del estado de la investigación en el momento.

Otro antecedente de la Base de Datos fue el proyecto exploratorio de inteligencia artificial que se desarrolló en el C.C.E.L.A. hacia 1989, sobre la familia Guahibo (Landaburu, 1990), bajo la dirección de Francisco Queixalós. Este proyecto tenía interés en hacer comparatismo lingüístico con apoyo informático.

2. Metodología de trabajo

El equipo permanente de este proyecto estuvo integrado por el director científico del mismo (Jon Landaburu del C.N.R.S. de Francia) quien junto con otros dos investigadores del mismo C.N.R.S. (Francisco Queixalós y Elsa Gómez, también miembros del equipo de la Base de Datos) fue el impulsor y fundador del posgrado de Etnolingüística que formó los investigadores del C.C.E.L.A. El coordinador académico de este proyecto (Tulio Rojas) y un investigador (Miguel Meléndez, sub-director del C.C.E.L.A.) estuvieron en las diversas etapas del proyecto. Un total de 32 investigadores –incluidos los ya nombrados– participaron en diversos momentos, apoyaron el diseño de los cuestionarios, aplicaron encuestas, asistieron a seminarios de trabajo y escribieron ponencias y artículos sobre el tema lexicográfico. Hubo además un comité de seguimiento y elaboración de los cuestionarios integrado por un número variable de ocho a diez personas en las diferentes etapas y se contó con asesoría de investigadores expertos invitados para temas específicos (Rojas, 2001: 176).

Las 32 lenguas abordadas no son la totalidad de las lenguas y variantes dialectales presentes en el territorio colombiano (más de 60), pero sí hubo una importante cobertura de regiones y familias lingüísticas. Se destaca la presencia de diversos puntos de encuesta para las lenguas *embera* de la familia Chocó, la presencia de datos de lenguas extintas (o ‘dormidas’) como el *tinigua* y el *carijona*, la presencia de lenguas poco documentadas como el *yuko* y el *chimila*.

Tabla 1. Lenguas, familias, puntos e investigadores participantes en el cuestionario fonético-fonológico y/o en los cuestionarios léxico y gramatical. Tomado y adaptado de CCELA (1996: 8) y de Rojas (2001: 177).

Región	Lenguas	Familia lingüística	Puntos de encuesta	Investigador
Norte	wayuu	Arawak	Santa Clara, La Guajira	Francisco Justo Pérez
	kogui	Chibcha	Santa Rosa, Yinkuamero, Sierra Nevada de Santa Marta	Carolina Ortiz
	damana	Chibcha	Abingüe, Sierra Nevada de Santa Marta	María Trillos Amaya
	chimila	Chibcha	Issaoristunna, Magdalena	María Trillos Amaya
Norte	ika	Chibcha	Séynimin, Sierra Nevada de Santa Marta	Rubiel Zalabata

Región	Lenguas	Familia lingüística	Puntos de encuesta	Investigador
Nororiente	barí	Chibcha	Birikaira, Norte de Santander	María Cristina Mogollón
	uwa (tunebo)	Chibcha	Angosturas, Arauca	Clara Casilimas
	yuko	Caribe	Iroka, Cesar	Camilo Robayo
Occidente	embera y embera-chamí	Chocó	Cristianía, Garrapatas, Sambú. Antioquia, Valle, Chocó	Daniel Aguirre
	embera sur	Chocó	Esmeraldas Ecuador	Mario Hoyos Benítez
	embera	Chocó	Jaidukama y Andágueda, Antioquia y Chocó	Rito Llerena
	embera-saija	Chocó	Guangüí Cauca	Mercedes Prado
	kuna	Chibcha	Caimán Nuevo, Antioquia	Rito Llerena
Suroccidente	páez (nasa yuwe)	Páez	Caldono, Cauca	Rocío Nieves
	páez (nasa yuwe)	Páez	Muchique-Tigres, Cauca	Tulio Rojas
	guambiano (namtrik)	Barbacoa	Guambía, Cauca	Beatriz Vásquez
	guambiano (namtrik)	Barbacoa	Quizgó, Cauca	Lilia Triviño
	totoró	Barbacoa	Pedregal, Cauca	Martha Pabón
	kamëntsá	Kamëntsá	Sibundoy, Putumayo	Narciso Jamioy
	kofán	Kofán	Santa Rosa de Sucumbíos Putumayo	María Elena Tobar
Orinoquia	achagua	Arawak	Umapo, Meta	Miguel Meléndez
	sáliba	Sáliba-Piaroa	El Duya Casanare	Hortensia Estrada
	piapoco	Arawak	Laguna colorada, Guainía	Andrés Reinoso
	cuiba-maibén	Guahibo	S.J. de Ariporo, Casanare	Ana J. Merchán
	guayabero	Guahibo	Barrancocolorado, Guaviare	Nubia Tobar Ortíz
	tinigua	Tinigua Pamigua	Macarena, Meta	Nubia Tobar Ortíz
	sikuani	Guahibo	Santa Rita, Vichada	Francisco Queixalós
Amazonia	puinave	Puinave	Puerto Valencia, Guainía	Jesús Mario Girón
	tatuyo y barasana	Tukano Oriental	Pirá-Paraná, Vaupés	Elsa Gómez-Imbert
	kubeo	Tukano Oriental	Kuduyarí, Vaupés	Simón Valencia
	tanimuka	Tukano Oriental	Guacayá, Amazonas	Natalia Eraso

Región	Lenguas	Familia lingüística	Puntos de encuesta	Investigador
Amazonia	tanimuka	Tukano Oriental	Guacayá, Amazonas	Natalia Eraso
	andoke	Andoke	Quebrada Aduche, Amazonas	Jon Landaburu
	carijona	Caribe	Miraflores, Guaviare	Camilo Robayo
	bora	Muinane-Bora	Las Palmas, Amazonas	Rosa A. Escobar
	muinane	Muinane-Bora	Monochoa, Amazonas	Consuelo Vengoechea
	ticuna	Yurí-Tikuna	Amacayacu, Amazonas	María Emilia Montes
San Andrés Islas	criollo	Criollos	San Andrés	Carol O'Flynn
Orinoquia	achagua	Arawak	Umapo, Meta	Miguel Meléndez
	sáliba	Sáliba-Piaroa	El Duya Casanare	Hortensia Estrada
	piapoco	Arawak	Laguna colorada, Guainía	Andrés Reinoso
	cuiba-maibén	Guahibo	S.J. de Ariporo, Casanare	Ana J. Merchán
	guayabero	Guahibo	Barrancocolorado, Guaviare	Nubia Tobar Ortíz
	tinigua	Tinigua Pamigua	Macarena, Meta	Nubia Tobar Ortíz
	sikuani	Guahibo	Santa Rita, Vichada	Francisco Queixalós
Amazonia	puinave	Puinave	Puerto Valencia, Guainía	Jesús Mario Girón
	tatuyo y barasana	Tukano Oriental	Pirá-Paraná, Vaupés	Elsa Gómez-Imbert
	kubeo	Tukano Oriental	Kuduyarí, Vaupés	Simón Valencia
	tanimuka	Tukano Oriental	Guacayá, Amazonas	Natalia Eraso
Amazonia	andoke	Andoke	Quebrada Aduche, Amazonas	Jon Landaburu
	carijona	Caribe	Miraflores, Guaviare	Camilo Robayo
	bora	Muinane-Bora	Las Palmas, Amazonas	Rosa A. Escobar
	muinane	Muinane-Bora	Monochoa, Amazonas	Consuelo Vengoechea
	ticuna	Yurí-Tikuna	Amacayacu, Amazonas	María Emilia Montes
San Andrés Islas	criollo	Criollos	San Andrés	Carol O'Flynn

3. Fase del cuestionario fonético-fonológico

La primera parte del proyecto entre 1993 y 1995, tuvo un componente fonético-fonológico. Para éste se diseñó un cuestionario estandarizado que en 1995 ya se había aplicado en 23 de los 32 puntos de encuesta y cuyos datos se habían ingresado en un sistema de gestión. Posteriormente, en 1997, se involucraron más lenguas e investigadores para llegar a 32 lenguas e investigadores y 38 puntos de encuesta como lo describe la Tabla 1 (Rojas, 2001: 177).

No hubo mayores dificultades en el diseño de los cuestionarios ni en la información que dieron los investigadores, pues todos habían tenido que hacer una propuesta de fonología de las lenguas estudiadas⁶⁸. Se solicitaba dar un inventario de fonemas, de alófonos, de procesos fonológicos y de estructuras silábicas, así como datos sobre prosodias.

4. Cuestionario gramatical y guía para la ficha de gramemas

Con un criterio similar al del cuestionario fonético-fonológico, se diseñó desde 1993 un cuestionario gramatical que incluía un listado exhaustivo de “gramemas” (prefijos, sufijos, clíticos) con su correspondiente identificación. Avanzar hacia los temas gramaticales generales, fue una tarea de mayor complejidad debido a los supuestos teóricos a considerar. La aplicación de este cuestionario fue una tarea posterior a la del cuestionario léxico. No hubo evaluación de lo ocurrido con la aplicación de este cuestionario. Elementos claves que sustentaron el diseño de este cuestionario, así como interrogantes iniciales del proceso de diseño, estaban orientados por las opciones descriptivas de tipo estructuralista, con énfasis en la semántica gramatical, una línea que caracterizó las tesis del posgrado de etnolingüística y las subsiguientes publicaciones derivadas de la actividad del centro de investigación⁶⁹:

De esta manera se perfila un engranaje teórico-descriptivo que recoge, por un lado, la “lista de gramemas” (integraría aspectos como: forma básica del gramema, alomorfos, caracterización morfológica, interpretación semántico-funcional, etc.), y de otro lado, “temáticas gramaticales” específicas como (...) sintaxis de la predicación, estructuras actanciales, tipos de oraciones), las cuales podrían desglosarse en apartados como, por ejemplo, esquemas de oraciones mínimas, diferentes tipos [de oración], orden en una oración simple, expresión de la predicación de ubicación, de la predicación de propiedad, expresión del agente, expresión del paciente, manifestaciones por índices intraverbales, por nominales, etc. (Rojas, 2001: 199).

68 Aunque ésta fue la primera tarea y la que se cumplió de manera más completa no disponemos de archivos digitales ni en papel para la lengua *tikuna*.

69 El Centro Colombiano de Estudios de Lenguas Aborígenes C.C.E.L.A. publicó nutridamente desde 1988 una serie de Descripciones, otra de Memorias, otra de Diccionarios y una más tardía de Tradiciones orales. Estas publicaciones están en su mayoría agotadas y aún no tienen edición digital, pero han sido la fuente citada de estudios posteriores.

En el caso del *tikuna* se cumplió la tarea del cuestionario gramatical, que consistía básicamente en una ficha de gramemas con una identificación muy detallada de cada ítem. En el momento, hacia 1997, se identificaron un poco menos de 100 gramemas para esta lengua. El inventario de formas tiene de por sí un valor descriptivo y un potencial comparativo. Los criterios usados en la ficha de gramemas son un valioso aporte metodológico para la construcción de descripciones morfosintácticas de las lenguas de tradición oral. Atestiguan un conocimiento y manejo empírico de una gran diversidad de fenómenos presentes en las lenguas estudiadas desde el C.C.E.L.A. y a lo largo del posgrado de Etnolingüística.

Para llenar esta extensa ficha se dieron una serie de indicaciones que están contenidas en la Guía (Landaburu y C.C.E.L.A., s.f.), un documento cuyas principales características resumimos a continuación. En la segunda parte de esta ficha, la clasificación formal del gramema, hay asuntos relativos a la morfología, identificación del carácter libre, ligado o clítico; de la manifestación del morfema como forma aditiva, forma supletiva, forma cero; de sus características como morfema radical o vinculado a una determinada clase (nombre, verbo, adjetivo, otras). Se anotan particularidades a tener en cuenta:

Apunte cuidadosamente las dificultades de adscripción de gramemas que haya podido tener y los inconvenientes derivados de la adaptación a una taxonomía de clases de palabras distinta a la que ha podido establecer para su lengua. En particular, si tiene clases de palabras que no están en la lista, indíquela en las casillas vacías (por ejemplo “locativos”). Antes de hacerlo fíjese bien en la definición que se da aquí de cada clase de palabras para que la diferencia no sea un simple problema de designación (“partícula” en vez de “adverbio”, ‘cualificativo’ en vez de “adjetivo”, etc.). Si una de sus clases no está en la lista pero que equivale a la suma de varias de las presentadas tampoco vale la pena reseñarla (por ejemplo, si usted tiene una clase “palabra predicativa” que recoge “nombre”, “verbo” y “adjetivo”, o una palabra ‘cualificativo’ que recoge “adjetivo” y “adverbio”). Pensamos en este caso que la pertinencia de su clasificación, por ser sintáctica, no afecta la validez morfológica de los datos presentados en un marco más diferenciado. De todas maneras, queda abierta la clasificación para una ulterior confrontación de datos (Landaburu y C.C.E.L.A., s.f.)⁷⁰.

Se definen criterios para establecer la pertenencia de una forma radical a alguna de las clases más usuales, los verbos, los nombres, los adjetivos, los adverbios. Se establecen definiciones para los elementos gramaticales como pronombres, relatores espaciales, cuantificadores, verbos auxiliares, identificadores, conectores.

En la cuarta parte sobre la función sintáctica del gramema se esperaba establecer hechos relativos a las relaciones de dependencia dentro de la palabra, concernía a los morfemas gramaticales ligados. Debía establecerse si el morfema en cuestión era ‘derivativo’ o ‘categorial’, advirtiendo dificultades metodológicas en la clara diferenciación de ambos tipos:

70 Agradecemos al doctor Jon Landaburu, director científico del proyecto Base de Datos del C.C.E.L.A. haber compartido este documento.

No es porque, por ejemplo, el número o el sexo son (o pueden ser) categoriales en castellano que tienen que serlo en su lengua. Al revés, un clasificador puede ser categorial en su lengua cuando su expresión sería derivativa o composicional en castellano (Landaburu y C.C.E.L.A., s.f.).

En la quinta parte se trataba de establecer la función semántico-gramatical del gramema (propuesta provisional). Se anotaba el enfoque eminentemente funcionalista de autores como Dik, Foley, Van Valin, Givon, Searle y Halliday. Para las relaciones semánticas se seguían criterios de Pottier, Lyons, Greenberb, Hagège, Shopen, Comrie, entre otros.

En la identificación de esos asuntos subyace una concepción de la construcción por niveles de la oración y de lo que el locutor hace en el acto de habla que instaura:

a) [el locutor] expresa en su lengua las entidades y lugares sobre los cuales quiere decir algo (paso A: construcción del término), b) expresa en su lengua lo que les pasa, la relación que mantienen o la característica que tienen (paso B: construcción de la función predicativa), c) organiza en su lengua según los patrones de ésta la adecuación de 1 con 2. El resultado expresa un estado de cosas en un mundo posible (paso C: construcción de la predicación), c) expresa en su lengua dónde, cuándo y cómo tiene lugar el estado de cosas constituido por esas entidades y esa relación que los une (paso D : construcción de la proposición), d) expresa en su lengua lo que hace con este estado de cosas en relación a su interlocutor (se lo cuenta, se lo pregunta, se lo ordena, lo solicita, etc. (paso E : construcción de la oración) (Landaburu y C.C.E.L.A., s.f.).

Sobre los ‘nombres’ por ejemplo, se establecían las determinaciones cualitativas y cuantitativas más usuales y los rasgos allí presentes. Esto se repetía para cada uno de los tipos de expresión de manera exhaustiva y detallada, de modo que la ‘Guía’ constituye una completa síntesis de los temas por los que debe transitar toda descripción gramatical completa: tipos de argumentos, relaciones gramaticales, sistemas y marcas de caso, cambios en las relaciones gramaticales, operaciones relativas a la jerarquización de la información (tópico, foco); también se daban orientaciones sobre la identificación de diversos tipos de nociones aspectuales, espaciales y temporales en los predicados. Para la expresión de la ‘proposición’ se consideraron diversas posibilidades de categorías y operaciones modales (lo real, lo virtual, lo óptico, lo epistémico, entre otras). La indicación de fuente de conocimiento es un rasgo característico de algunas lenguas amerindias y así lo precisaba la guía: “No se caracteriza el valor de verdad de mi conocimiento ni mi compromiso con él sino cómo conozco el estado de cosas de referencia. Las posiciones más frecuentes son: reportado (o citativo), inferido, conocido de vista, conocido por otros sentidos, etc.” (Landaburu y C.C.E.L.A., s.f.).

En cuanto a los modos oracionales se citaban -en el mismo documento- como frecuentes los siguientes: “declarativo, interrogativo, dubitativo (puede ser objetivo o subjetivo), imperativo, iusivo, exhortativo, optativo, permisivo, preventivo, exclamativo, mirativo (sorpresa), etc.”.

**Esquema de los campos contemplados en el formato de ficha de gramemas
Base de datos del Centro colombiano de estudios de lenguas aborígenes -C.C.E.L.A.-**

Primera parte: Identificaciones
Nombre de la lengua
Lugar de recolección
Nombre del investigador
Número asignado al morfema
Forma básica en escritura AFI
Alomorfos listados alfabéticamente

2º parte: Clasificación formal del gramema
Autonomía y forma
Tipo de manifestación
Lugar de manifestación

3º parte: Adscripción del gramema
Palabras léxicas
Palabras gramaticales

4º parte: Función sintáctica del gramema en la palabra
Radical
Derivativo (homogéneos, semantismo)
Derivativo heterogéneo (clase de palabras a la que se aplica, clase que se obtiene; semantismo)
Categorial

5º parte: Función semántico-gramatical del gramema Categorías presentes en la expresión de la entidad
Determinaciones cualitativas
Determinación cuantitativa
Referenciación
Determinación de posesión o dependencia
Identificación/comparación
Identificadores o Especificadores
Función casual
Jerarquización a nivel de la predicación
Jerarquización a nivel de la oración

Categorías presentes en la expresión de la función predicativa
Determinaciones ligadas a la valencia
Determinaciones ligadas al cambio interno (aspecto)
Determinaciones ligadas a la cantidad
Determinaciones ligadas al espacio
Categorías u operaciones presentes en la expresión de la predicación
Sujeto
Objeto
Objeto indirecto, oblicuo
Categorías u operaciones presentes en la expresión de la proposición
Determinación óptica de la predicación
Localización temporal de la predicación
Localización espacial de la predicación
Indicación de la fuente de conocimiento
Categorías presentes en la expresión de la oración
Declarativo
Interrogativo
Dubitativo (puede ser objetivo o subjetivo)
Imperativo
Iusivo
Exhortativo
Optativo
Permisivo
Preventivo
Exclamativo
Mirativo (sorpresa), etc..

Figura 1. Esquema de los campos contemplados en la “Ficha de gramemas del Centro Colombiano de Estudios de Lenguas Aborígenes, C.C.E.L.A. 1991-1998.

Fuente: Tomado del original exportado de *File Maker 6 2002* para *Mac OS X 10.4.11* a PDF⁷¹ (5 páginas) y de la guía de Landaburu y C.C.E.L.A. s.f.)

71 Selección y síntesis de M.E. Montes, 2020.

5. El cuestionario léxico y la perplejidad en el análisis de fenómenos del léxico

Se diseñó un cuestionario que tenía una “visión enciclopédica” con apartados definidos con criterios semánticos. El total de ítems de este cuestionario fue de 3.360 palabras agrupadas en un capítulo sobre el conocimiento de la naturaleza, otro sobre el mundo humano y un tercero sobre la acción del hombre sobre la naturaleza (Rojas, 2001: 182). El diseño estuvo precedido de una importante tarea de discusión colectiva y seminarios especiales. Se adelantaron encuestas de terreno específicas para completar los datos faltantes y para adelantar segmentación. La tarea de compilación de léxico generó publicaciones y eventos en los que se analizaron los resultados obtenidos en esa indagación y se discutió la metodología utilizada.

Desde las primeras reuniones de investigadores se detectaron asuntos teóricos y metodológicos problemáticos:

Lengua		Ticuna	Familia	Ticuna	Investigador	M. MONTES
<p>Capítulo 2. EL MUNDO HUMANO</p> <p>Sub-capítulo 2.1 Cuerpo</p> <p>X.x.x 2.1.1 Partes y órganos</p> <p>X.x.x.x</p> <p>X.x.x.x.x</p>						
Nº del ítem	51	Castellano	Hueso			
Ítem leng. ind.	/da\$ cida\$a\$! /				Fiabilidad del dato	Certero
Tipo de unidad	Palabra					
Clase sintáctica	Nombre					
División silábica	/da\$ -ci-da\$a\$! /					
Estructura silábica	CV-CV-CV-V					
División morfológica	//da\$ - cida\$a\$! //					
Estructura morfológica	//CV-CVCVV //					
Glosa morfológica	//su de él-hueso //					
Lexema (1)	- cida\$a\$!	Lexema (2)				
Lexema (3)		Lexema (4)				
Véase también	Esqueleto 211, 44					

Figura 2. Ejemplo de fragmento de ficha en *File Maker 6 2002* para *Mac OS X 10.4.11* (convertida en 2014 a PDF de archivo original del C.C.E.L.A.), cuestionario léxico *tikuna* de M.E. Montes. Ítem: /da-cíd 'a / “esqueleto, hueso(s)” un nombre inalienable que exige prefijo posesivo neutro de tercera persona.

Un primer balance del trabajo con este cuestionario nos ha revelado que una buena parte de las preguntas ha quedado sin respuestas y que algunos campos semánticos tienen una gran complejidad. Es evidente que tenemos un choque entre nuestra concepción y la concepción de las comunidades. Partimos de una organización occidental (enciclopédica) de la experiencia y procuramos acercarnos a la organización de las comunidades aborígenes. (Rojas, 2001: 186).

5.1 Conocimiento léxico y conocimiento cultural y ambiental

A pesar del obstáculo estructural percibido por los investigadores (que implica un cuestionamiento metodológico de fondo al punto de partida que fue el cuestionario) se logró graficar en mapas algunos ítems que podían ser agrupados por regiones, familias o grupos de lenguas. En Rojas, 2001: 196, se presenta como ejemplo el mapa de la palabra ‘sal’ (sin diferenciar aún entre términos para ‘sal marina’, ‘sal gema’, ‘sal vegetal’ ‘fuente de sal’ como lo preveía el cuestionario) en la que es fácil detectar por ejemplo un área de préstamos del Tupí en la región amazónica con diferentes formas cognadas como *júkiúra* del tikuna, *iukira* de la ‘*língua geral*’, *yukira* del tanimuka, *jukira* del cubeo. También se detectan posibles préstamos del castellano (?) en el chimila *salí* y quizá en las formas encontradas en embera *thã* y *tã* (?). Subsiste entre los Uitoto la tradición de elaborar sales vegetales a partir de cortezas de ciertos árboles y palmas, sales que se mezclan con la pasta de tabaco para hacer el *ambil*. Entre los Tikuna -como entre otros pueblos amazónicos- se reconocen los ‘salados’ *na-tüü*⁷² o afloramientos naturales de sal donde van a beber los animales de la selva, así como los ‘salados’ de árbol donde los loros y otras aves buscan esta sustancia. La palabra *tikuna* tomada del Tupí sirve para designar la sal mercancía que se compra hoy, la que pudo llegar con los conquistadores. La otra palabra que designa los lugares de la tierra o de la vegetación donde los animales buscan sal es un radical nominal, un nombre ligado o inalienable que tiene un prefijo genérico de tercera persona, *na-tüü*. La dificultad que señalaba Rojas se deriva así de la complejidad del conocimiento cultural, ambiental, social y antropológico que supone la indagación de este cuestionario léxico. Este único mapa disponible nos plantea tanto problemas metodológicos como potenciales hipótesis sobre difusión de términos en áreas geográficas, como lo preveía el objetivo de tener una cartografía.

Incluimos a título de ejemplo algunos ítems de la primera parte del extenso cuestionario léxico sobre el mundo natural, y en particular sobre el mundo físico, apartado 1.1.

72 Escrito en ortografía. La vocal /uu/ se representa con ü

Tabla 2. Extracto de ítems del cuestionario léxico: Mundo Natural, Mundo Físico. Elaborado por el Centro Colombiano de Estudios de Lenguas Aborígenes –C.C.E.L.A.– 1991-1998. Dirección científica: Jon Landaburu, CNRS. Coordinación general: Tulio Rojas Curieux. Tomado de archivo digital exportado de *File Maker 6* 2002 para *Mac OS X 10.4.11* a *Excel[®]*³

1.1 Mundo Físico	
1.1.1 Estados de la materia	
1.0	Aceite
2.0	Cristal
3.0	Engrudo
4.0	Fluido
5.0	Fuego
6.0	Gaseoso
7.0	Gastado
8.0	Goma
9.0	Grasa
10.0	Humo
11.0	Impuro
12.0	Inerte
13.0	Líquido
14.0	Masa
15.0	Materia
16.0	Metal
17.0	Pasta
18.0	Polvo
19.0	Protuberancia
20.0	Puro (no mezclado)
21.0	Reverso
22.0	Ruina
23.0	Sólido
24.0	Substancia
25.0	Suciedad
26.0	Superficie
27.0	Tenue
28.0	Tornarse (volverse)
29.0	Universo / Mundo
30.0	Vapor
31.0	Volátil
Otros	

1.1.2 Tierra	
1.1.2.1 Substancias y minerales	
1.0	Tierra
2.0	Tierras (tipos: negra, roja, amarilla, etc)
3.0	arcilla
4.0	arena
5.0	azufre
6.0	barro
7.0	caliza
8.0	carbón: tipos
9.0	cobre
10.0	cuarzo
11.0	esmeralda
12.0	estaño
13.0	fuelle de agua salada
14.0	gravilla - cascajo
15.0	hierro
16.0	Mármol
17.0	Mercurio
18.0	Mina
19.0	onix
20.0	oro
21.0	petróleo
22.0	plata (mineral)
23.0	platino
24.0	plomo
25.0	roca
26.0	rubí
27.0	sal gema
28.0	sal marina
29.0	Sal vegetal
30.0	zinc
31.0	fundir
32.0	Lavar oro
33.0	incandescente
Otros	

73 Selección y conversión de M.E. Montes, 2020.

5.2 El léxico del cuerpo humano

Ante la dificultad metodológica que parece haber suscitado la aplicación de diversos apartados del cuestionario en los apartados de ‘Mundo natural’, se optó por reflexionar conjuntamente sobre el léxico del cuerpo humano, dentro del apartado ‘Mundo humano’ quizá porque permitía evadir los problemas de no comparabilidad evidentes en los campos de medio ambiente y naturaleza, donde cada ecosistema imponía un inventario propio atravesado por los conceptos culturales.

De todas maneras, el tema aparentemente simple y concreto del léxico del cuerpo humano derivó hacia la cultura, la naturaleza y lo metafórico, asuntos no fácilmente capturables en este tipo de herramientas como el cuestionario. Así lo muestran dos compilaciones que resumieron jornadas conjuntas de trabajo y discusión. Una de ellas hecha en 1998 se orientó específicamente a la tarea de analizar el léxico del cuerpo humano a través de la gramática y de la semántica y se derivó directamente de la base de datos; otra posterior, en el 2001, es más abierta en temáticas y enfoques, pero continúa la línea de preguntas suscitadas por la tarea inicial del cuestionario léxico.

En los textos de Aguirre 1998a y b, Llerena 1998, Meléndez 1998 y 2001, Montes, 1998 y 2001, Ortiz 1998 y 2001, Pabón, 1998 y 2001, Reinoso, 1998, Rojas, 1998a y b, Trillos, 1998, Vásquez, 1998 y Yule, 2001 se abordan los temas clásicos de la lexicografía, como la descripción de procesos de derivación, composición y morfosintaxis de las unidades léxicas, la partonomía; pero también temas definitivamente etnolingüísticos como cuerpo y espacialidad, cuerpo y territorio, cuerpo y conceptualización cultural, cuerpo como fuente de metáfora, terminología de los olores, espacialidad en el cuerpo humano, entre otras. En la compilación de 1998 aparecen además otros temas lexicográficos de gran valor como los nombres comunes, los nombres de lugares, los nombres propios, los procesos de diccionarización y el estudio histórico de los vocabularios. Podemos afirmar que gran parte de la tarea de análisis cualitativo e interpretación de los universos semánticos que sugieren estos vocabularios está aún por hacerse. La elaboración de diccionarios o glosarios por campos semánticos en lenguas nativas podría iniciar con la información compilada para ser complementada con equipos amplios de hablantes nativos.

6. El sistema de gestión de datos: La rápida obsolescencia de las herramientas informáticas

En la época se trabajó con equipos y plataforma *Mac Intosh* y con aplicaciones compatibles con estos equipos. En particular, se acudió a los programas *Hyper Card* y *File Maker*, dos *softwares* que sirvieron para la entrada y sistematización preliminar de datos. Se elaboraron varios mapas a manera de piloto, incluidos mapas de familias lingüísticas y de zonas geográficas (para más detalle ver Rojas, 2001: 196-198).

Si bien hubo un avance notable en la sistematización de información y en la compilación y comparación de datos con diversos criterios, no se publicaron los datos brutos ni en red ni en físico. No era fácil tener acceso rápido a sistemas y recursos como los que hoy posibilitan compartir y publicar ágilmente resultados parciales del procesamiento que se llevaba a cabo.

La detención temporal del proyecto por razones logísticas y financieras⁷⁴ y la rápida obsolescencia de los equipos y *softwares* utilizados condujo a que fuese inviable darle continuidad. Las fuentes especiales que se usaron para digitar los símbolos usados en la escritura fonética y fonológica de las diversas lenguas funcionaban en equipos *Mac* pero no eran fácilmente recuperables y la conversión de esos resultados a otras plataformas no se logró. Esto que parece un impase menor es uno de los cuellos de botella de la información compilada⁷⁵. Si fuera necesario digitar o convertir todos los datos de la Base la recuperación de la información sería excesivamente dispendiosa. En el caso del tikuna se perdieron los caracteres especiales y diacríticos que se usaron para representar la nasalidad y el tono.

7. Balance y replanteamientos necesarios

Las enormes dificultades que implicaba mantener actualizado el sistema informático de gestión de datos, la rápida obsolescencia de equipos y *software* y algunos problemas logísticos, condujeron a que no fuera posible lograr los resultados previstos en términos de la publicación y comparación masiva de datos originales.

En la época, los datos fueron compilados en papel inicialmente a través de formatos en los cuales cada investigador escribía a mano para ser luego digitados con las fuentes especiales en el sistema de gestión de datos. En algunos casos, los investigadores tuvieron luego copia digital en diskette de esos datos.

Finalmente, la parte informática del proyecto fue abandonada y una parte de los archivos digitales disponibles en medio magnético se dispersó o se deterioró irremediamente. Los resultados hoy en día son las publicaciones y análisis conjuntos e individuales que hicieron los investigadores participantes en diversas etapas de esta tarea científica colectiva y el diseño metodológico comprehensivo para captar diversos aspectos de la estructura de las lenguas, desde los sistemas sonoros hasta la estructuración del léxico.

74 La cooperación francesa era entendida como el fomento de una tarea que debía continuarse en el país. Una vez finalizada esta fase, los recursos aportados por la universidad privada que alojó el proyecto no fueron suficientes para mantener el ritmo de trabajo exigido; tampoco hubo continuidad en la financiación estatal a las tareas de investigación básica sobre lenguas indígenas. El panorama político cambió y hubo otras prioridades. El equipo de trabajo se disolvió.

75 En las fases iniciales de la investigación sobre lenguas amerindias, la escritura de símbolos especiales (para la notación fonética, fonológica e incluso ortográfica) representaba enormes dificultades. Los tipos disponibles en imprentas eran los de los hispanistas. La obra del Instituto Caro y Cuervo (González y Rodríguez, 2000) por ejemplo, se trabajó en *Word Perfect* con macros que permitían crear caracteres (comunicación personal de M.L. Rodríguez). Las publicaciones iniciales del C.C.E.L.A. en Word de Mac Intosh también acudían a la creación de caracteres y al uso de fuentes especiales diseñadas por laboratorios de lingüística africanista en Francia.

7.1 Otra metodología para tareas futuras

En cuanto al cuestionario léxico, el más productivo y al mismo tiempo el que implicaría mayores retos, es claro que se requieren grandes replanteamientos:

En primer lugar, la interdisciplinariedad. Es imposible para un lingüista preguntar, por ejemplo, por el ‘mundo natural’ sin apoyarse en expertos y en los estudios previos sobre medio ambiente de las regiones y ecosistemas: botánica, zoología, ecología, limnología, geología, por citar algunas de las disciplinas que permitirían un mejor acercamiento al complejo ‘mundo natural’ de las lenguas nativas.

En segundo lugar, las formas de indagación. En un asunto como el conocimiento cultural sobre el medio natural, la indagación debería hacerse en etapas sucesivas e intensivas con equipos más amplios de hablantes nativos a través de talleres. Igual ocurre con otras áreas de indagación sobre parentesco, sentimientos, sensaciones, comportamientos⁷⁶. Es muy difícil para quien no tenga una mirada antropológica y/o endógena, entender el porqué de los comportamientos del léxico. En nuestro caso, trabajamos individualmente con varias fuentes orales, pero ello no garantiza en ciertos temas la certeza del dato.

En tercer lugar, la agilidad en la gestión, sistematización y publicación de los resultados que deben ser compartidos y editados simultáneamente por los equipos de trabajo. Esto es posible hoy a través de programas informáticos y aplicaciones que facilitan la tarea que en la década de los 90 tuvo que hacerse separadamente y secuencialmente.

7.2 Perspectivas: Continuidad de los temas léxicos y etnolingüísticos

Las preguntas sobre las particularidades del abordaje del léxico de una lengua que expresa una cosmovisión particular y una forma específica de organizar la experiencia, se detectaban ya en la fase de discusión de resultados:

Con lenguas que tienen términos comunes para ‘piel, pellejo, cáscara, corteza, cuero’, ¿estamos ante un semantismo polifuncional (...) o ante la presencia de verdaderos hiperónimos que pueden ser comprendidos, en este caso, por ejemplo, bajo el concepto general de ‘cobertura’? (Aguirre, 200 1a: 7).

El avance de los investigadores en el conocimiento de las lenguas y las culturas que estudian desde hace décadas ha permitido también un avance en el mejor entendimiento de conceptualizaciones reflejadas parcialmente en el léxico. En nuestro caso, el trabajo sobre el ‘mundo vegetal’ y en particular sobre los nombres de las plantas en la lengua tikuna nos condujo a identificar elementos de la clasificación nominal, de la conceptualización cultural sobre los ecosistemas y de las relaciones entrettejidas de la gente, los animales, los árboles,

⁷⁶ Se requiere un diseño metodológico similar al que sirvió para tareas de traducción colectiva de apartes de la Constitución de 1991 a algunas lenguas, metodología que se replicó luego en diversas experiencias adelantadas por investigadores en la producción de otros materiales.

los insectos. La búsqueda de taxonomías y campos semánticos en el ‘mundo vegetal’ nos llevó a re-pensar el rol del pensamiento metafórico expresado a través del léxico (Montes, 2001 y 2018).

Otro tema que requirió una investigación en profundidad y un replanteamiento metodológico radical fue la indagación sobre términos relativos al espacio y a la orientación espacial. La complejidad de esta indagación es contraria a cualquier acercamiento con encuesta simple, se requiere de etnografía y de experimentación, así como de la mirada endógena del lingüista indígena. Diversos trabajos posteriores han continuado esa línea.

En el momento actual es posible el diseño de diccionarios en línea⁷⁷, colaborativos y con la participación simultánea de varios aportantes. Un ejemplo de este tipo de tareas en Colombia es el diccionario en línea de la lengua sáliba⁷⁸, del cual se anota que es “un diccionario en construcción permanente”. Un antecedente de la metodología usada en el diccionario electrónico del sáliba fue el diccionario en línea de la lengua muysca (Gómez Aldana) y la página de documentación del grupo interdisciplinario *Muyscubun*, grupo interdisciplinario que ha apoyado seriamente la recuperación de la lengua muisca en zona cundiboyacense. Implican un enorme trabajo de discusión y argumentación en la toma de decisiones, como se hizo en su momento en el C.E.L.A. La tarea de revitalización, fortalecimiento, mantenimiento o rescate de las lenguas nativas pasa siempre por la búsqueda de las palabras olvidadas y de los conocimientos que se difuminan con esas palabras.

Diversos temas lexicográficos y etnolingüísticos serían posibles a partir de los datos aún disponibles de esa indagación colectiva de los 90: expresión de sentimientos, emociones y sensaciones; léxico de elementos de cultura material; terminología de parentesco y formas de tratamiento; léxico de tipos de acciones; nombres de las estrellas y los astros, entre muchos otros.

En la etapa de actividad del proyecto de la Base de Datos, se enunció la necesidad de elaborar un cuestionario sociolingüístico pero la tarea no se inició. De cierta forma, esta idea sí tuvo espacio en otras iniciativas posteriores, adelantadas en muy diversos ámbitos: El muy completo Atlas Sociolingüístico de América Latina y el Caribe (Sichra, ed. 2009), es una buena muestra de esa tarea en la macro región. Así mismo, las tareas y documentos del Plan de Protección a la Diversidad Etnolingüística que se adelantaron con el Ministerio de Cultura (2009), cumplieron con el objetivo de establecer un panorama de contextos, uso y vitalidad de las lenguas nativas a través de los “autodiagnósticos”⁷⁹. Ha sido una constante en las universidades y centros de investigación hacer diagnósticos puntuales y descriptivos sobre la situación de las lenguas nativas, sobre los fenómenos de contacto, así como sobre

77 En la serie de Diccionarios el C.C.E.L.A. publicó 2 volúmenes: el de la lengua *sikuani*, de autoría de Queixalós y el de la lengua achagua de Meléndez, 2011.

78 Publicado por el Instituto Caro y Cuervo a partir de la investigación y compilación de la lingüista Hortensia Estrada y con la participación de hablantes nativos y un equipo asesor en temas de informática y lexicografía.

79 Estos autodiagnósticos elaborados participativamente con varias comunidades indígenas están inéditos. Han servido como guía para los pueblos nativos en la formulación de sus planes de trabajo en temas de educación.

las iniciativas de revitalización. Esta información –dispersa aún– proporciona una serie de miradas complementarias sobre el tema sociolingüístico que el C.C.E.L.A. no alcanzó a concretar en su momento.

7.3 Publicación y difusión de los datos

La información compilada en este proyecto colectivo está dispersa, en los archivos personales de los investigadores participantes y parcialmente en la Universidad de Los Andes⁸⁰. Ahora es mucho más fácil pensar en poner los materiales en línea para consulta a través de la red. Es importante visibilizar la tarea del C.C.E.L.A. y en particular el aporte metodológico que está tras el diseño de los tres cuestionarios.

Necesidades actuales:

- Publicar los formatos y campos considerados para el diseño de los cuestionarios fonológico, léxico y gramatical como aporte a la metodología del trabajo de campo lingüístico.
- Publicar los resultados disponibles en papel⁸¹ de los cuestionarios fonológicos, gramaticales y léxicos, escaneados en su estado actual, pero con metadatos que permitan su clasificación y con la necesaria advertencia sobre la fecha de recolección.
- Recuperar algunos de los archivos digitales, exportar los datos y convertirlos a formatos actualizados, publicarlos con los requerimientos de la documentación lingüística. Eventualmente, publicarlos junto con los materiales sonoros en formato análogo que acompañaban los cuestionarios⁸².

Estas tareas implican un trabajo de ‘documentación retrospectiva’, o sea, de recuperación de datos en formatos obsoletos. Requieren de un apoyo y financiación institucional estable de mediano plazo. Requieren así mismo la constitución de un equipo interdisciplinario que involucre expertos en tareas de informática aplicada a la lingüística. Todo lo anterior debería hacerse a la menor brevedad pues no podemos esperar el deterioro que ocasiona el tiempo en la información aún disponible.

La tarea realizada desde el C.C.E.L.A. durante varias décadas fue un aporte vital para el estudio de las lenguas indígenas y criollas en Colombia, para visibilizar la importancia científica y patrimonial de las mismas y para apoyar una consideración respetuosa para estas lenguas y sus hablantes. El entrenamiento de equipos de lingüistas que a su vez continuaron la tarea de formación ha permitido ampliar el horizonte a otras regiones y a otras instituciones. La nueva generación de lingüistas formados en distintas universidades ha aportado otra manera

80 Algunas de las tareas aquí enunciadas y otras relativas a la publicación digital de las series Descripciones y Memorias están en curso en la Universidad de los Andes (comunicación personal de D. Aguirre-L.)

81 Es incierta aún la cantidad y calidad de los cuestionarios en papel que habrían sobrevivido a sucesivos cambios de oficinas y al deterioro del tiempo, pero de todas formas parece ser que es mejor la situación de los archivos en papel que la de los archivos digitales, algunos de los cuales sufrieron daños irreparables.

82 Las cintas de cassette en las que recogimos datos de audio que ilustraban los cuestionarios, son especialmente frágiles y muchas se deterioraron.

de trabajar y otras metodologías y subdisciplinas como la documentación y la informática. En este momento es pertinente hacer historia de nuestra disciplina y recuperar algunos de los principales aportes metodológicos que se construyeron colectivamente durante el tiempo de actividad del C.C.E.L.A. El diseño y aplicación de los cuestionarios implicaron una valiosa reflexión conjunta que vale la pena visibilizar.

Referencias bibliográficas

- Aguirre, Daniel (1998a). “Cosmos, naturaleza y cuerpo humano entre los embera” (pgs. 113-120). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- _____ (1998b). “Seminario sobre el léxico del cuerpo humano” (pgs. 7-14). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Centro Colombiano de Estudios de Lenguas Aborígenes -C.C.E.L.A.- (1996). *Informe de actividades 1995 presentado ante las autoridades del convenio Universidad de Los Andes-Instituto Francés de Estudios Andinos*. Universidad de Los Andes. Impreso.
- _____ (1997). *Informe de actividades 1996. Estudios e investigaciones de lenguas amerindias y criollas en las áreas científica, cultural y educativa*. Universidad de los Andes. Impreso.
- Gómez Aldana, Diego F. *Diccionario muisca - español*. Fecha de consulta: febrero de 2020. Publicación digital en muisca.cubun.org/Categoría: Diccionario.
- González, M.S. y Rodríguez, M.L. (2000). *Lenguas indígenas de Colombia. Una visión descriptiva*. Bogotá: Instituto Caro y Cuervo.
- Hortensia Estrada Ramírez. Indígenas sálibas (compilación e investigación lingüística). Instituto Caro y Cuervo. *Diccionario electrónico sáliba-español: una herramienta interactiva para la documentación de la lengua y la cultura sáliba*. Disponible en: http://saliba.caroycuervo.gov.co/index.php/P%C3%A1gina_principal. Acceso en febrero de 2020.
- Landaburu, Jon (1990) «Un programme de développement de la linguistique amérindienne en Amérique Latine : le Centre Colombien d'Études en Langues Aborigènes (C.C.E.L.A.)» En: *Amérindia, revue d'Ethnolinguistique Amérindienne*, N° 15.
- _____ y Centro Colombiano de Estudios de Lenguas Aborígenes, C.C.E.L.A. (S.F) *Guía para la ficha de gramemas de la Base de Datos del C.C.E.L.A.* impreso (archivo de J. Landaburu, director científico del proyecto Base de Datos).
- Llerena, Rito (1998). “Léxico del cuerpo humano en la lengua ñp̄ra del Alto Sinú y el río San Jorge” (pgs. 121-128) En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Meléndez, Miguel (1998). “Productividad y extensión de las denominaciones del cuerpo humano en achagua” (pgs. 145-154). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.

- _____ (2001). “Áreas semánticas y clases de palabras cubiertas por los prefijos ka- ‘posesivo’ y ma-‘privativo’ en achagua” (pgs. 171-182). En: Meléndez, Miguel y Montes, María E. (1997) *Diferencias y similitudes en las estructuras del léxico de lenguas aborígenes*. Memorias 7. Colección Lenguas Aborígenes de Colombia. Bogotá: Universidad de Los Andes-C.C.E.L.A.-Colciencias.
- _____ (2011). *Diccionario achagua-español, español-achagua*. Bogotá: C.C.E.L.A. Universidad de Los Andes.
- Ministerio de Cultura. *Plan de Protección a la Diversidad Lingüística, PPD.E.* (2009). Disponible en https://www.mincultura.gov.co/prensa/noticias/Paginas/2008-01-23_6081.aspx. Acceso en febrero de 2020.
- Montes, María E. (1998). “Posibilidades de análisis en el léxico del cuerpo humano de la lengua ticuna” (pgs. 155-172). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- _____ (2001). “La terminología de los olores y de los sabores en la lengua ticuna” (pgs. 113-138). En: Meléndez, Miguel y Montes, María E. (1997) *Diferencias y similitudes en las estructuras del léxico de lenguas aborígenes*. Memorias 7. Colección Lenguas Aborígenes de Colombia. Bogotá: Universidad de Los Andes-C.C.E.L.A.-Colciencias.
- _____ (2001). “Los nombres de las plantas, sus partes y sus espacios de crecimiento. Aproximación etnolingüística a partir de datos de la lengua tikuna. Amazonia colombiana”: 523-555. En: Zárate y Franky (eds.) *Imanimundo. Estudios en la Amazonia colombiana*. Bogotá: Universidad Nacional de Colombia. Disponible en línea en <http://bdigital.unal.edu.co/3740/2900/nombresdelasplantassuspartesyusespaciosdecrecimiento.pdf>. Acceso en febrero de 2020.
- _____ (2018). “Taxonomías y cadenas de asociaciones: recorrido por temas etnolingüísticos en tikuna (yurí-tikuna) de la Amazonía colombiana”: 191-216. *Boletín de Antropología*. Universidad de Antioquia, Medellín, vol. 33, N.º 55.
- Ortiz, Carolina (1998) “El léxico del cuerpo humano en la lengua kogui: hacia una clasificación de los formantes nominales” (pgs. 15-34). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- _____ (2001). “Lexicalización o gramaticalización: ¿Se incluyen o no los derivativos en un léxico de la lengua kogui?” (pgs. 139-154). En: Meléndez, Miguel y Montes, María E. (1997) *Diferencias y similitudes en las estructuras del léxico de lenguas aborígenes*. Memorias 7. Colección Lenguas Aborígenes de Colombia. Bogotá: Universidad de Los Andes-C.C.E.L.A.-Colciencias.
- Ospina, Ana M. (editora) (2013). *Expresión de nociones espaciales en lenguas amazónicas*. Bogotá: Universidad Nacional de Colombia - Instituto Caro y Cuervo
- Pabón, Marta (2001). “Atrás y adelante: ¿Otra dimensión del cuerpo humano? El caso del idioma de Totoró, Cauca, Colombia” (pgs. 197-214). En: Meléndez, Miguel y Montes, María E. (1997) *Diferencias y similitudes en las estructuras del léxico de lenguas aborígenes*. Memorias 7. Colección Lenguas Aborígenes de Colombia. Bogotá: Universidad de Los Andes-C.C.E.L.A.-Colciencias.
- _____ (1998). “Características sintácticas del léxico del cuerpo humano de la lengua totoró (Cauca, Colombia)” (pgs. 99-112). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.

- Queixalós, Francisco (s.f.) *Diccionario sikuani-español*. Bogotá: C.C.E.L.A. Universidad de Los Andes- Programa de Etnoeducación del Ministerio de Educación.
- Reinoso, Andrés (1998). “El cuerpo humano en la cultura piapoco” (pgs. 129-144) En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Rojas, Tulio (1998a). “El proceso de construcción de una base de datos lingüísticos y cartografiados del C.C.E.L.A.” (pgs. 173-204). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- _____ (1998b). “Kl^yuno o el polo conceptual que evade todo intento de aprehensión” (pgs. 59-80). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Sichra, Inge (Ed.) (2009). *Atlas sociolingüístico de los pueblos indígenas en América Latina*. Bolivia: UNICEF-PROEIB-Andes.
- Trillos, María (1998). “Léxico del cuerpo orgánico en damana” (pgs. 35-58). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Vásquez, Beatriz (1998). “El cuerpo humano en guambiano” (pgs. 81-98). En: Eraso, Natalia (comp.) *El léxico del cuerpo a través de la gramática y la semántica*. Memorias 5. Colección Lenguas Aborígenes de Colombia. Bogotá: C.C.E.L.A. Universidad de Los Andes-Colciencias.
- Yule, Marcos (2001). “Las expresiones léxicas en la organización del universo socio-cultural en los nasa (páez)” (pgs. 85-92). En: Meléndez, Miguel y Montes, María E. (1997) *Diferencias y similitudes en las estructuras del léxico de lenguas aborígenes*. Memorias 7. Colección Lenguas Aborígenes de Colombia. Bogotá: Universidad de Los Andes-C.C.E.L.A.-Colciencias.