

MORFOLOGÍA Y DICCIONARIO ELECTRÓNICO DE NOMBRES EN QUECHUA

Maximiliano Duran
(UFC, Université de Franche-Comté, Besançon, Francia)
duran_maximiliano@yahoo.fr

Resumen

Presentamos aquí dos diccionarios electrónicos de nombres en quechua, uno con traducción al español y otro, al francés. Cada lexema incluye sus propiedades sintáctico-semánticas y su correspondiente paradigma flexional FLX.

Después de hacer un inventario de los sufijos nominales, mostramos la manera cómo hemos construido la matriz fundamental que rige la morfología nominal del quechua, comenzando por aglutinación binaria de estos sufijos. Luego presentamos el método para obtener las gramáticas racionales que nos permitan generar, automáticamente, usando la plataforma lingüística NooJ, todas las formas nominales conteniendo dos o más sufijos.

Luego mostramos cómo, aplicando a un pequeño texto quechua, interrogaciones en formato NooJ se obtienen las correspondencias de las formas nominales reconocidas en ese texto, flexionadas o no, y cómo, enseguida se las puede visualizar gráficamente, en colores, sobre ese texto, obteniéndose un alto índice de reconocimiento (94%), que prueba el grado de precisión de nuestras gramáticas.

Palabras clave

diccionario electrónico, quechua, forma nominal, sufijo nominal, morfología quechua

ELECTRONIC DICTIONARY OF QUECHUA NAMES AND MORPHOLOGY

Abstract

We present here two electronic dictionaries of Quechua names, one with a translation into Spanish and the other into French. Each lexeme includes its syntactic-semantic properties and its corresponding FLX inflectional paradigm.

After making an inventory of the nominal suffixes, we show how we have constructed the fundamental matrix that governs the nominal morphology of Quechua, starting with binary agglutination of these suffixes. Then we present the method to obtain the rational grammars that allow us to generate, automatically, using the NooJ linguistic platform, all the noun forms containing two or more suffixes.

Then we show how, applying to a small Quechua text, interrogations in NooJ format, the correspondences of the nominal forms recognized in that text, bent or not, are obtained, and how, immediately they can be displayed graphically, in colors, on that text, obtaining a high recognition rate (94%), which proves the degree of precision of our grammars.

Keywords

electronic dictionary, quechua, nominal form, nominal suffixes, quechua morphology

1. Introducción

1.1. Generalidades sobre el quechua

El idioma quechua fue el idioma oficial de la civilización Inca. Nació probablemente durante la civilización Huari, cuya capital, del mismo nombre, se encuentra en Ayacucho, a mediados del primer milenio de nuestra era, en la región central de los Andes.

Según los primeros cronistas europeos (s. XVI), se hablaba en todo el territorio inca que comprendía el sur de Colombia (Pasto), todo el territorio de Ecuador y Perú, parte de Bolivia y el norte de Argentina y Chile.

Unas décadas después del establecimiento de la colonia, la población de habla quechua ha seguido disminuyendo respecto a la población total. De casi el 100% de las personas que hablaban y entendían el idioma (estimado alrededor de 12 millones), ahora son menos de 6.280.000 (censo de 2010) es decir, 13% de la población de la misma región. Es se debe, principalmente, a la falta de escolarización en quechua, en esta región.

Con respecto a la talla de la población de quechua-hablantes en las próximas décadas, si las políticas educativas continúan como hasta hoy, no sólo le vemos un futuro incierto, sino sombrío. De ahí la calificación como 'lengua en peligro' de desaparición declarada por la UNESCO, en el Atlas Mundial de las Lenguas publicado en 2009 en París.

1.2. ¿Cómo salvar al quechua?

En este contexto, todos los esfuerzos para salvar esta "especie" lingüística, patrimonio de la humanidad, en peligro de desaparecer, son bienvenidos. La principal es la educación obligatoria en quechua en los territorios de habla quechua.

La publicación de diccionarios en papel o electrónicos (léxico general, científicos y técnicos, diccionarios etimológicos, topónimos, de nombres propios, palabras compuestas, metáforas, refranes, expresiones fijas, sinónimos, etc.).

El desarrollo de programas informáticos para la traducción de textos asistida por ordenador, etc.

2. Los diccionarios electrónicos QU-SP y QU-FR

Teniendo en cuenta esta motivación, uno de nuestros primeros recursos lingüísticos que debíamos desarrollar, fue la construcción de un diccionario electrónico general quechua-francés (DGQ-FR) y también un diccionario electrónico general quechua-español (DGQ-SP).

En la Figura 1, podemos ver un extracto de nuestros dos diccionarios electrónicos DGQ-SP y DGQ-FR, los cuales hemos venido construyendo desde los años 1990 y hoy, hacen parte de los recursos lingüísticos integrados en nuestro robot quechua YACHAJ, destinado a ser un traductor automático de textos.

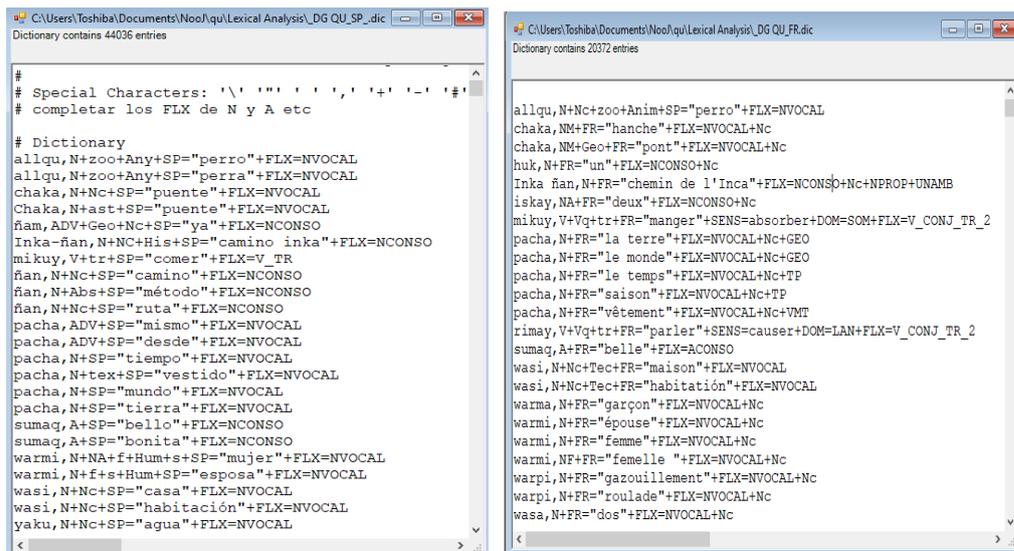


Figura 1 Extractos de los diccionarios generales QU-SP y QU-FR.

Dentro de nuestros diccionarios, en conformidad con sus atributos lingüísticos, las palabras integran diferentes clases y subclases de ontologías jerárquicas. Para muchas entradas, también hemos proporcionado propiedades sintáctico-semánticas y su traducción al francés o al español. Notemos que una entrada QU corresponde a una transferencia FR¹ (o una transferencia SP respectivamente), lo que permite obtener palabras desambiguadas. Por ejemplo, vemos que, para el francés, *pacha* = temps (tiempo) y también *pacha* = vêtement (ropa), el primero corresponde al tiempo (Tp) y el segundo define un objeto concreto, hecho de tela o cuero u otro material y que sirve para vestir. Es por eso que intentamos tener tantas entradas como significados haya para la misma palabra en el idioma de origen.

3. El diccionario electrónico de nombres DGQ-SP

Partiendo de nuestro diccionario bilingüe general quechua-español DGQ-SP, que editamos en versión papel, en 2003, ver Duran (2009), y aplicando en NooJ (Silberztein, M. (2016)), la interrogación (, N +), ver Figura 2, podemos extraer todas las entradas marcadas del nombre de la categoría gramatical (N), más de 16.000 nominales unidades lingüísticas atómicas.

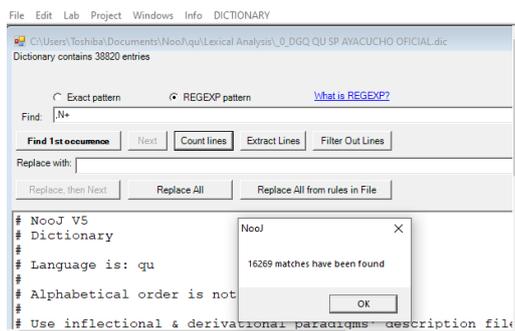


Figura 2 Demanda de extracción de nombres a partir del Diccionario General QU-SP.

3.1. Propiedades del diccionario

Comentemos las propiedades semánticas que acompañan a las entradas de la muestra presentada en esta figura y otras marcas.

¹ Hemos tomado de [Tier \(2011\)](#) un cierto número de vocablos traducidos al francés.

A este diccionario extraído de nombres, que contiene indistintamente nombres simples y declinados, aplicamos el algoritmo, "extraer lemas nominales". Ello nos produce el léxico de nombres simples quechua presentado en la Figura 3. Este diccionario electrónico contiene 1,478 entradas. No incluye nombres propios.

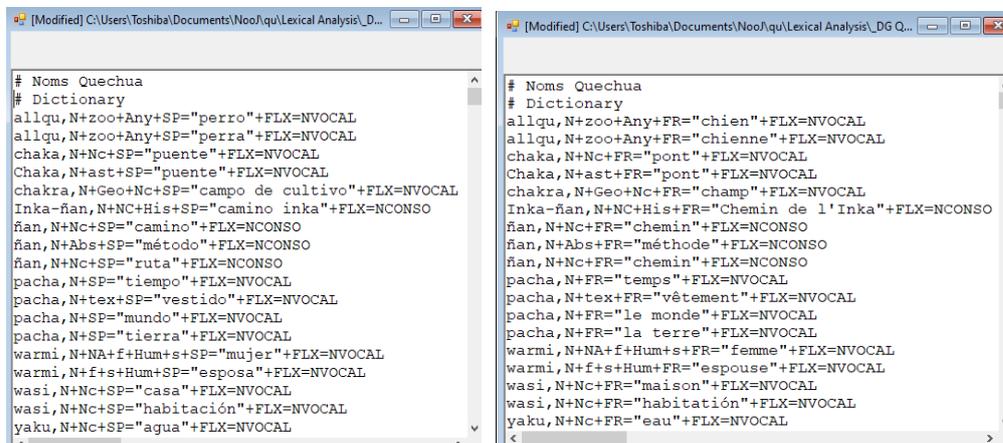


Figura 3 Extractos de diccionarios QU-SP y QU-FR.

Todas las entradas tienen asignado un paradigma flexional, representado por FLX =. Por ejemplo, la unidad de entrada wasi (casa) se flexiona de acuerdo con la clase de paradigma NVOCAL, por lo que aparece la entrada: wasi, N + FR = “maison” + FLX = NVOCAL. Los paradigmas flexionales son modelos de patrones estándar o prototipos basados en reglas de sufijación morfológica. Estas reglas cubren la variación en el número, los diminutivos y superlativos, verbalizaciones, casos.

1. *allqu* (SP= perro FR=chien) (dog) lleva la marca (N) para nombre, como nombre común (Nc), pertenencia a la *zoología* (*zoo*) y que se flexiona siguiendo el paradigma NVOCAL (FLX=NVOCAL);
2. *chakra* un sustantivo (N) que se declina según el paradigma morfológico NVOCAL (FLX = NVOCAL) para sustantivos que terminan en vocal, correspondiente al sustantivo SP “campo cultivado”, con propiedades semánticas como nombre común (Nc en SP), un emplazamiento geográfico (Geo en SP), y puede derivar en verbo por el sufijo -yay;
3. *ñan* (SP= camino) está clasificada como *Nombre común o concreto* (Nc), y la otra entrada *ñan* (SP= método) (*Abs*), que es la marca de abstracto, las dos se declinan según el paradigma morfológico NCONSO (FLX = NCONSO);
4. *Inka ñan* un sustantivo (N) que se declina según el paradigma morfológico NCONSO (FLX = NCONSO) para sustantivos que terminan en consonante, un sustantivo que viene de la Historia (His) correspondiente al sustantivo SP (Camino Inca);
5. *pacha* un sustantivo (N) que se declina según el paradigma morfológico NVOCAL (FLX = NVOCAL) para sustantivos terminados en vocal, correspondiente al sustantivo SP (la tierra), con propiedades semánticas como nombre común (Nc), un emplazamiento geográfico (Geo), que difiere de la otra entrada, su homónimo siguiente,
6. *pacha* un sustantivo (N) que se declina según el paradigma morfológico NVOCAL (FLX = NVOCAL) para sustantivos que terminan en vocal, correspondiente al sustantivo SP “tiempo”, con propiedades semánticas como nombre común (NC), esta forma tiene el siguiente el tercer homónimo,
7. *pacha* un sustantivo (N) que se declina según el paradigma morfológico NVOCAL (FLX = NVOCAL) para sustantivos que terminan en vocal, correspondiente al sustantivo SP “vestido”, con propiedades semánticas como nombre común (Nc),

El género gramatical, f femenino, m masculino, no es propio del quechua, por lo cual no figura sistemáticamente en los nombres. El género natural se diferencia en los mismos lemas, p. Ej. maqta muchacho, pasña muechacha o por medio de frases nominales, p. ej. perro (macho) *urqu allqu*, perra (hembra) *china allqu*. La mayoría de las formas de parentesco en quechua, realizan la distinción del sexo de ambas personas involucradas en la relación en el mismo léxico:

| | |
|---|------------------------------------|
| hermano de un hombre <i>wauji (wawqi)</i> | hermano de una mujer <i>turi</i> . |
| hermana de una mujer <i>pani (pana)</i> | hermana de una mujer <i>ñaña</i> . |

4. Los sufijos nominales quechua

Desde el punto de vista de la gramática generativa, son partículas que representan morfemas gramaticales que poseen la propiedad de combinarse con lexemas. Dan origen, para los sustantivos, a todos los casos gramaticales, la pluralización, el número, derivacionales (que tiene por efecto un cambio de categoría) y para los verbos, los tiempos, el aspecto, participio, infinitivo, nominalizadores.

Existen tres tipos:

- Aquellos que influyen o concretan el significado del tema que les precede sin cambiar de categoría (*cha, kuna, pas, paq, ...*)
- Aquellos que provocan un cambio de categoría gramatical. Son los que adjetivan un sustantivo (*niraq, nirai, niq*), o verbalizadores, que transforman un sustantivo en verbo (*y, chay, yay*).
- Aquellos que impiden toda nueva aglutinación. Para los sustantivos, son los posesivos (*i, iki, n, ...*), los dubitativos (*ch, chá, s, si, ...*), la negación (*chu*), la interrogación (*chu ?*)

Por ejemplo, para formar el plural de un sustantivo en quechua simplemente agregamos el sufijo “*kuna*” al final de la entrada. Para formalizar esta regla, como paradigma en NooJ, la escribimos de la siguiente manera

KUNA = <E>/N+s | kuna/N+p

Aquí <E> es la cadena vacía de la concatenación, y "|" es el operador de la disyunción, la que nos permite seleccionar secuencias alternas.

Este paradigma establece que, si agregamos una cadena vacía a la entrada lexical, por ejemplo, *wasi* (casa), obtenemos el caso nominativo del sustantivo *wasi* en singular y si le agregamos “*kuna*”, obtenemos la forma plural *wasikuna*, las casas.

Para formar el beneficiario de un sustantivo, agregamos el sufijo “*paq/para*” al final de la entrada, teniendo el siguiente paradigma: PAQ = paq/N+BEN

El que genera la inflexión *wasi-paq* «para la casa» (*paq*: indica el benefactor).

En el quechua de Ayacucho-Chanka encontramos el siguiente conjunto de sufijos nominales:

SUF_N = { *-ch*, *-chá*², *-cha*, *-chik*, *-chiki*, *-chu*, *-chu(?)*³, *-hina*, *-kama*, *-kuna*, *-lla*, *-má*, *-man*, *-manta*, *-m*, *-mi*, *-mpa*, *-nimpa*, *naq*, *-nta*, *-ninta*⁴, *-nintin*, *-ntin*, *-niraq*, *-niyuq*, *-niq*, *-ña*, *-p*, *-pa*, -

² Los sufijos *chá, má, ...* son homónimos de los respectivos sufijos verbales pospuestos (SPP), como los describe Duran (2017). Se aplican a sustantivos y verbos. Por eso, algunos autores los llaman “sufijos ambivalentes”, Soto (1976). En realidad, éstos, indican significados específicos en cada categoría. Para que el procesamiento informático de las inflexiones sea más consistente, nosotros mantendremos la separación de estos sufijos como nominales (Suf_N) y verbales (SPP).

³ Los signos de interrogación y de exclamación no forman parte del sufijo, los usamos para transcribir la presencia de entonación ascendente seguida de una pausa para interrogar y descendente seguida de una pausa para el signo de exclamación. Estas entonaciones quizás no existieron antes del contacto de lenguas con el castellano.

⁴ *ninta* es de hecho una aglutinación de la partícula de apoyo “*-ni-*” y el sufijo “*-nta*”, aplicable a las formas que terminan en consonante. Esta misma partícula “*ni*” interviene en las aglutinaciones *ninka, ninta, nintin*, etc.

paq, -pas, -pi, -poss(7v+7c), -puni, -pura, -qa, -rayku, -raq, -ri, -s, -si, -sapa, -su, -ta, -taq, -wan, -y(!), -ya(!), -yá, -yupa, -yuq}⁵ (50+7v+7c)

Donde *poss (7v+7c)* es el conjunto (*-i, -iki, -n, -nchik, -iku, -ikichik, -inku; -nii, -niiki, -nin, -ninchik, -niiku, -niikichik, -ninku*) de sufijos posesivos.

Los sufijos nominalizadores de verbos:

S_VN = {*-y, -na, -q, -sqa*}

Los sufijos verbalizadores de nombres:

S_NAV = {*y, yay, chay*}

He aquí algunos ejemplos de inflexiones del nombre *wasi/ casa*, que usa un sufijo de la lista:

wasi-cha « casa pequeña » (*cha* : indica el diminutivo)⁶

wasi-pas « la casa también » (*pas*: indica el inclusivo)

wasi-kuna « las casas » (*kuna*: es el pluralizador)

wasi-paq « para la casa » (*paq* : indica el benefactor)

wasi-man « hacia la casa » (*pas* : indica el direccional)

Podemos escribir las gramáticas que permiten obtener las formas del posesivo aplicando el paradigma:

POSSV = *i/POS+s+1 | iki/POS+s+2 | n/POS+s+3 | nchik/POS+PIN+1 | iku/POS+PEX+1 | ikichik/POS+p+2 | nku/POS+p+3*;

En el cual, (*-i, -iki, -n, -nchik, -iku, -ikichik, -nku*) es el conjunto de sufijos posesivos, para los nombres que terminan en vocal

POSSC = *nii/POS+s+1 | niiki/POS+s+2 | nin/POS+s+3 | nchik/POS+PIN+1 | niiku/POS+PEX+1 | niikichik/POS+p+2 | ninku/POS+p+3*;

Donde (*-nii, -niiki, -nin, -ninchik, -niiku, -niikichik, -ninku*) es el conjunto de sufijos posesivos, para los nombres que terminan en consonante o en *ai, ui, ii, au*.

4.1. Los sufijos bloqueadores.

Son aquellos que no aceptan ninguna combinación posterior con otro sufijo. Son los siguientes (15) :

-ch, chá, -chiki, -chus, -chusina; expresan la duda, la incertidumbre en diversos grados

-chu ; expresa la negación *wasichu* / no es la casa

-chu ?; interrogación, ¿*taitaikichu ?* / ¿es tu padre?

-m, -mi, asertivo o enfatizador *taytaimi* :sí, es mi padre

-má, -qa ; son contrastivos en diferentes grados

-puni, asertivo absoluto *waujei puni* / mi propio hermano

-qa, contrastivo para los nombres; *rumiqa llasaq/* en cambio, la piedra es pesada

-y !, ya !; el vocativo, invocación o exclamación: *mama-lla-y !* oh mi madre, *mama-lla-ya!* / Ah mi madre

-yá, asertivo evidencial : *atoq-yá* / te digo que sí es el zorro

4.2. Las gramáticas flexionales generativas

La gramática racional que permite generar todas las formas de un sustantivo que terminan en vocal es:

N_V_1 = :CH |:CHA |:CHAA |:CHIKI |:CHU |:CHUS |:CHUI|: GEP |: GEPA |:KAMA |:KAQLLA |:KUNA |:LLA |:MAA |:MAN |:MANTA |:MASI |:MM |:MPA |:NAQ |:NTA |:NTIN |:NIRAQ |:NIYUQ |:ñA |:NIQ |:PAQ |:PAS |:PI |:POSSIF |:PUNI |:PURA |:QA

⁵ Ver Guardia Mayorga (1973), Perroud (1970), Pino (1980).

⁶ Para una descripción detallada de las características semánticas inducidas, al sustantivo, por cada sufijo nominal, véase el capítulo dos del trabajo de Duran (2017).

|:QINA |:RAIKU |:RAQ |:RI|:RII |:SAPA |:SSIV |:SU |:TA |:TAQ |:WAN |:YA |:YY |:YAA |:YYAA |:YUPA |:YUQ⁷;

Si aplicamos esta gramática usando el motor de inflexión de NooJ a un sustantivo, "perro" por ejemplo, genera formas nominales, cuyo tema nominal *allqu* permanece invariable como se ve en el siguiente extracto:

allqu,allqu,N+ani+zoo+SP="perro "+FLX=NVOCAL+Nc
allquch,allqu,N+ani+zoo+SP="pued~~a~~e que sea el perro"+FLX=NVOCAL+DPRO
allquchu,allqu,N+ani+zoo+ SP ="no es el perro"+FLX=NVOCAL+NEG
allqucha,allqu,N+ani+zoo+ SP ="el perro pequeño "+FLX=NVOCAL+DIM
allqukuna,allqu,N+ani+zoo+ SP ="los perros"+FLX=NVOCAL+PLU
allquman,allqu,N+ani+zoo+ SP ="en dirección al perro "+FLX=NVOCAL+DIR
allqumanta,allqu,N+ani+zoo+ SP ="a propósito del perro "+FLX=NVOCAL+ORIG
allquantin,allqu,N+ani+zoo+ SP ="incluyendo al perro "+FLX=NVOCAL+TOT

La representación gráfica de la misma gramática de dimensión 1 (solo interviene un sufijo), aparece en la Figura 4.

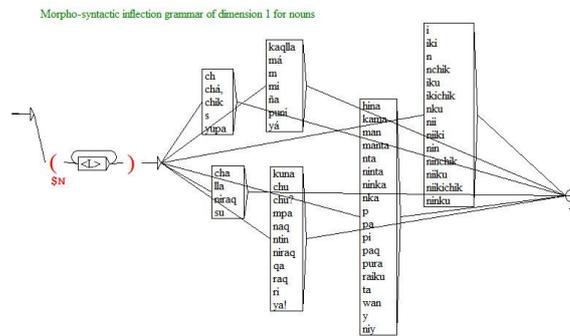


Figura 4 Gramática local de la inflexión nominal de una dimensión.

4.3. La matriz M_Suf_N

Notamos que, en el idioma, es corriente encontrar inflexiones con dos sufijos:

wasi-cha-paq / "para la casita"
wasi-iki-not / "tu casa también"

Para obtener una gramática capaz de generar automáticamente todas las inflexiones con dos sufijos, hemos construido, durante 2011-2012, una matriz de valores booleanos, 0 y 1, a partir de la tabla, cuyo extracto aparece en la Figura 5, a la que denominaremos M_Suf_N. Contiene 4096 casillas.

El valor "1" que encontramos allí, corresponde a aglutinaciones gramaticalmente válidas para dos sufijos. Estos valores se obtuvieron por introspección o en el corpus.

⁷ Donde, CH = ch/DINT; CHA = cha/DIM; CHAA = chá/PRO; CHIK = chik/RESG; CHU = chu/NEG; etc. Para la lista completa ver Duran (op.cit.)

| SUF_N | CH | CHAA | CHIK | CHIKI | CHUI | CHUN | DCHA | GEP | GEPACH | HINA | KAMA | KUNA | LLA | MAA | MAN | MANT | MASI | MM | MPA | NIMP | NAQ | NKA | NINK | NTA | NINTA | NTIN | NINTI | NIQ | NIRA | NIYUC | ÑA | PAJ | PAS | PI | POSV | Pi |
|--------|----|------|------|-------|------|------|------|-----|--------|------|------|------|-----|-----|-----|------|------|----|-----|------|-----|-----|------|-----|-------|------|-------|-----|------|-------|----|-----|-----|----|------|----|
| CH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| CHAA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CHIK | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CHIKI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CHUI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CHUN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| DCHA | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| GEP | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| GEPACH | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HINA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| KAMA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| KUNA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| LLA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | |
| MAA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MAN | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MANTA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MASI | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MPA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NIMP | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NAQ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NKA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NINKA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NTA | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NINTA | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Figura 5 Tabla booleana de valores para combinaciones válidas de dos sufijos nominales.

Hay quinientas setenta y un casillas con el valor 1. Por lo tanto, tendremos tantas inflexiones válidas de dos sufijos de *wasi*. El punto (MAN, CHUN) de esta tabla nos dará la inflexión:

Wasi-man-chu / no es hacia la casa (*man* / direccional y *chu* / negación).

Esta matriz, luego de algunas operaciones matriciales, nos permitió construir la gramática racional, que contiene 571 paradigmas de dos sufijos, lo que permite flexionar un sustantivo que termina en vocal, parte del cual aparece a continuación:

$N_V_2 = :DCHACH \quad |:DCHACHAA \quad |:DCHACHUI \quad |:DCHACHUN \quad |:DCHAGEPA$
 $:|:DCHAKAMA \quad |:DCHAKUNA \quad |:DCHAMASI \quad |:DCHAMM \quad |:DCHAÑA \quad |:DCHANAQ$
 $:|:DCHANIQ \quad |:DCHANIRAQ \quad |:DCHAPAQ \quad |:GEPACH \quad |:GEPACHAA \quad |:GEPACHUN \quad \dots$
 $:|:GEPAKAMA \quad |:|:NAQGEPA \quad |:NIQGEPA \quad |:KUNANTIN \quad |:NTINMI \quad |:NTINCHUI \quad |:NTINCHUN$
 $:|:NTINMAA \quad |:NTINMI \quad |:PURANTIN \quad |:HINANTIN \quad |:NIQMAA \quad |:NIQMAN \quad |:NIQMANTA$
 $:|:NIQMI \quad |:NIQNINTA \quad \dots \quad |:SUCHAA \quad |:SUCHUI \quad |:SUCHUN \quad |:SUGEPA \quad \dots \quad |:SUPAQ \quad |:TACH$
 $:|:TACHAA \quad |:TACHUN \quad |:WANCHUI \quad |:WANCHUN \quad \dots \quad |:RAYKUSISVO \quad |:TASISVO$
 $:|:YUPASISVO;$

Aplicándole el motor de flexión de NooJ, obtenemos quinientas setenta y una formas nominales de *wasi*, como se puede ver en la Figura 6, con dos sufijos aglutinados. A continuación, mostramos un extracto:



Figura 6 Muestra de flexiones sufijales de *wasi*.

wasichaman, wasi, N+FR="maison"+FLX=N_V_2+DIM+DIR
wasichamasi, wasi, N+FR="maison"+FLX=N_V_2+DIM+PAR
wasicham, wasi, N+FR="maison"+FLX=N_V_2+DIM+ASS
wasichaña, wasi, N+FR="maison"+FLX=N_V_2+DIM+TRM
wasichapaj, wasi, N+FR="maison"+FLX=N_V_2+DIM+BEN
wasipach, wasi, N+FR="maison"+FLX=N_V_2+GEN+DPRO
wasipachá, wasi, N+FR="maison"+FLX=N_V_2+GEN+DPRO
wasipachu?, wasi, N+FR="maison"+FLX=N_V_2+GEN+ITG
wasipachu, wasi, N+FR="maison"+FLX=N_V_2+GEN+NEG
wasipakama, wasi, N+FR="maison"+FLX=N_V_2+GEN+LIM
wasipamá, wasi, N+FR="maison"+FLX=N_V_2+GEN+CTR
wasipaniraj, wasi, N+FR="maison"+FLX=N_V_2+GEN+ASI
wasipaniraj, wasi, N+FR="maison"+FLX=N_V_2+GEN+ASI
wasipañiraj, wasi, N+FR="maison"+FLX=N_V_2+GEN+ASI

wasipanirai, wasi, N+FR="maison"+FLX=N_V_2+GEN+ASI

4.4. Flexiones tri sufijales

Estas inflexiones se denominan así porque contienen tres sufijos aglutinados. Para encontrar todas esas combinaciones, primero identificaremos, entre las 571 combinaciones de dos sufijos, aquellas que son bloqueadoras. En efecto, este conjunto incluye los sufijos que llevan como segundo componente, los que inducen valores de duda (-*ch*, -*chá*, -*s*, -*si*), la vacilación (-*chik*, -*chiki*), la negación (*chu*), la interrogación (*chu* ?), el vocativo (*y!*), entre otros. En conformidad con la sintaxis quechua, los podemos dejar de lado, porque ya no participarán en nuevas combinaciones, quedan así 402 hábiles.

La nueva tabla contendrá $402 \times (64-15) = 19\ 698$ casillas para rellenar a mano con 0 o 1. Aplicamos algunas operaciones matriciales (multiplicación escalar, transposiciones, entre otras), para aligerar el trabajo de validación de ceros y unos (los detalles de estas *los* podemos encontrar en Duran (2013, 2014, 2017)), obteniendo así 1784 aglutinaciones capaces de generar flexiones tri-sufijales gramaticalmente válidas. A continuación, podemos ver una muestra de paradigmas de esas combinaciones:

```
CHAGEPQINA = :CHA :GEP :QINA;  
CHAGEPACHUI = :CHA :GEPA :CHUI;  
KUNAYUQPAQ = : KUNA :YUQ :PAQ;  
KUNAYUQPAS = : KUNA :YUQ :PAS;  
KUNAYUQPURA = : KUNA :YUQ :PURA;  
KUNAYUQRAIKU = : KUNA :YUQ :RAIKU;  
NIRAQLLAMA = : NIRAQ :LLA :MAA;  
NIRAQLLAMANTA = : NIRAQ :LLA :MANTA;  
PIÑACHUS = : PI :ÑA :CHUS;  
POSSVMANKAMA = : POSSV :MAN :KAMA;
```

Ahora, ya podemos construir la gramática racional *N_V_3*, bastante compleja, ya que contiene estos mil setecientos ochenta y cuatro paradigmas. La Figura 7 nos muestra un extracto de estas formas flexionadas tri-sufijales, obtenidas para el nombre *wasi*:

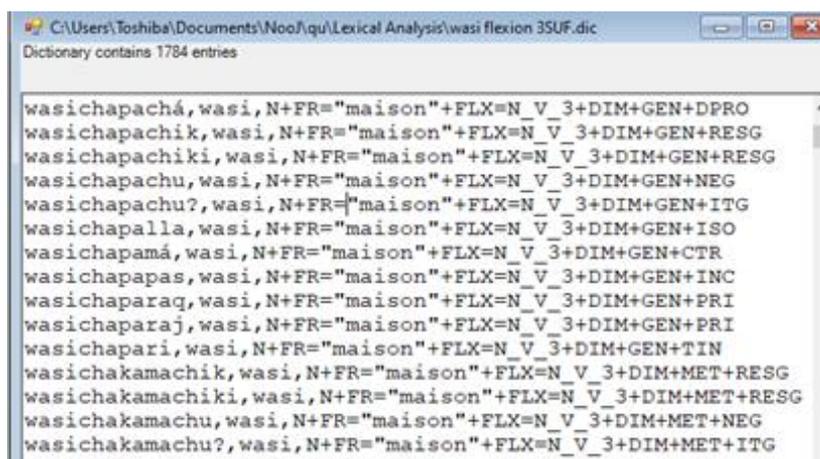


Figura 7 Formes flexionadas de *wasi* conteniendo tres sufijos.

Siguiendo prácticamente el mismo algoritmo, construimos la gramática que permite generar, automáticamente, todas las formas válidas que contienen la combinación de cuatro sufijos (que demanda llenar a mano ochenta y siete mil cuatrocientas dieciséis casillas con ceros o unos (que

será seguido de gramáticas para cinco, seis y siete sufijos), un extracto del cual se muestra en la Figura 8.

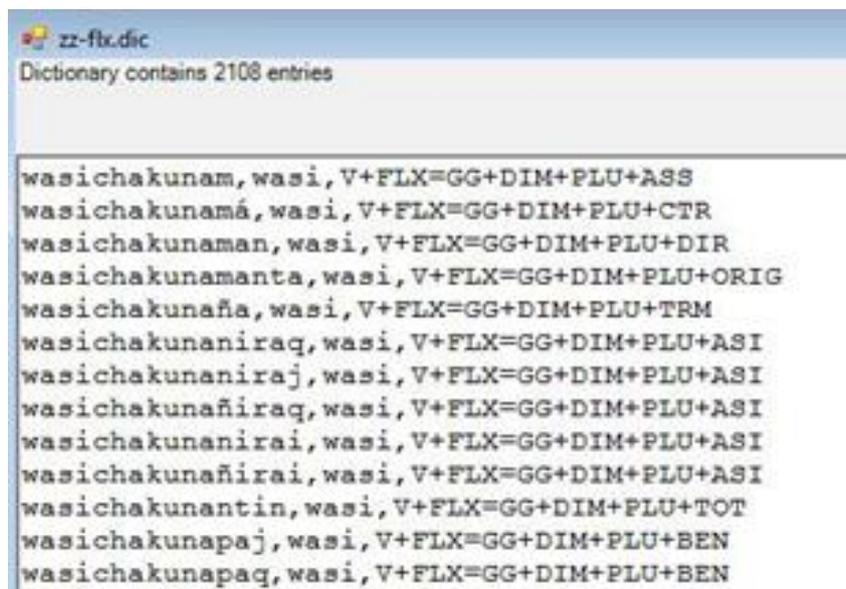


Figura 8 Flexiones quadri sufijales de wasi.

Las combinaciones gramaticalmente válidas de 4, 5, sufijos son muy frecuentes y numerosas en la comunicación cotidiana. Las combinaciones de 6 o 7 sufijos son menos numerosas. ej.

KUNALLAPAQRAQ = :KUNA :LLA :PAQ :RAQ ; (*Wasi-kuna-lla-paq-raq/* en principio, solo para las casas) (4 suff.).

DCHAGEPAHINAKAMACH = : DCHA :GEPA :HINA :KAMA :CH ; (*wasi-cha-pa-hina-kama-ch/* hasta aquellas que se parecen a la casa pequeña) (5 suff.)

CHALLAIKICHIKPAQHINARAQ= :CHA :LLA :IKICHIK :PAQ :HINA :RAQ; *wasi-cha-lla-ikichik-paq-hina-raq/* en principio, como si fuera algo para vuestra casa pequeña (6 suff.)

CHA-KUNA-LLA-PAQ -RAQ-CHU-SISV = : CHA :KUNA :LLA :PAQ :RAQ :CHU :SISV ; *wasi-cha-kuna-lla-paq-raq-chu-s/* me pregunto si, en principio, es para vuestras casas pequeñas.

Para obtener automáticamente, la totalidad de formas flexionadas de nombres quechua, tomamos nuestro diccionario electrónico de mil cuatrocientos setenta y ocho sustantivos simples, luego le aplicamos las siguientes demandas: LAB> Dictionary> set (SINGLE NAMES QU-FR.dic)> inflect. Hemos obtenido más de 1.600.000 formas flexionadas, con 1, 2 o 3 sufijos, como se puede ver en la Figura 9.

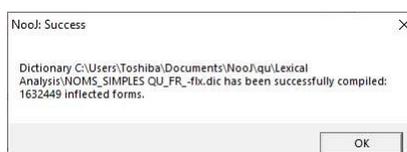


Figura 9 Total de formas nominales generadas a partir de NOMS SIMPLES QU-FR.dic.

5. Análisis Lingüístico

Los paradigmas y gramáticas que hemos presentado pueden ser utilizados para efectuar el análisis lingüístico de textos.

Tomemos para esto, como ilustración, el pequeño cuento andino antiguo: *Pakpaku chayñachawan rimanakun*, recopilado por Guardia Mayorga (1973) que incluye 272 palabras. Hemos aplicado algunas interrogaciones. La interrogación de concordancia <N+DIM>, que va a buscar identificar todas las concordancias que sean sustantivos flexionados al diminutivo, ha reconocido, en el cuento, 8 de esas formas flexionadas, las cuales aparecen en la Figura 10.

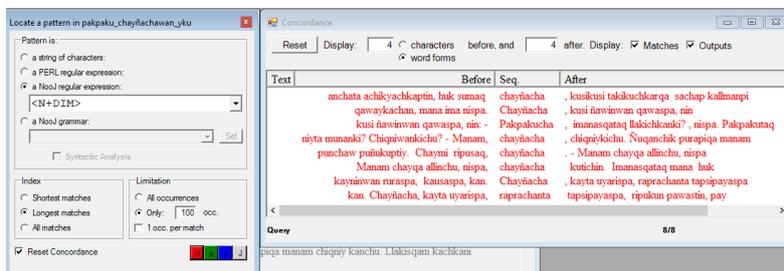


Figura 10 Concordancias correspondientes a la demanda <N+DIM>.

Después, hemos aplicado la operación “Color matching” y “coloración en el texto”, en Análisis Lingüístico, dentro de NOOJ y obtuvimos lo que aparece en la Figura 11, donde podemos observar, coloreados en rojo, que los sustantivos flexionados han sido identificados correctamente (94%).

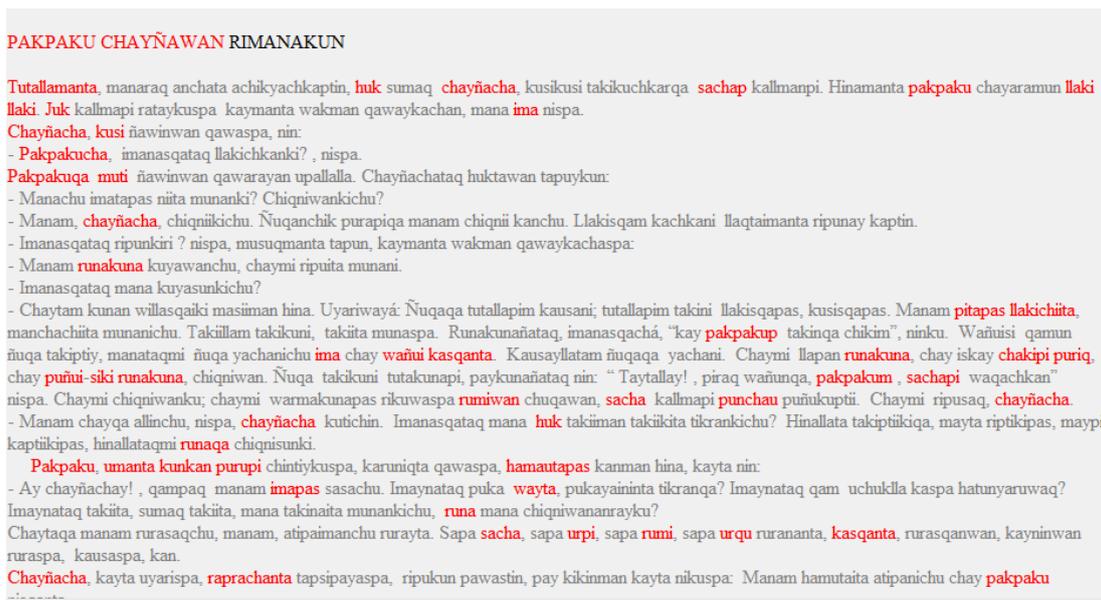


Figura 11 Visualización gráfica de las formas flexionadas identificadas por el operador <N>.

Conclusión

Hemos presentado nuestros dos diccionarios electrónicos de nombres quechua QU-SP y QU-FR, con la traducción respectiva al español y al francés. Cada lexema incluye sus propiedades sintáctico-semánticas y su correspondiente paradigma flexional FLX.

Hemos mostrado, la matriz fundamental de valores de las combinaciones binarias compatibles, de los 64 sufijos nominales (y ulteriormente de más sufijos). Luego hemos presentado el camino seguido para obtener las correspondientes gramáticas racionales que permiten obtener, automáticamente, combinaciones de dos o más sufijos que van a producir formas nominales multi-sufijales.

Luego, aplicando ciertas interrogaciones en formato NooJ, a un pequeño texto quechua, se ha podido identificar en él, las formas nominales correspondientes a los paradigmas, aunque aparezcan en forma flexionada en el texto. Hemos obtenido un alto índice de reconocimiento de formas nominales flexionadas (94%), probando así el grado de precisión de nuestras gramáticas.

Tenemos en perspectiva, la construcción de un léxico detallado conteniendo las traducciones al francés y al español del millón seiscientos mil formas nominales generadas, para introducirlo entre los recursos lingüísticos de nuestro robot YACHAJ. Es una tarea que estamos comenzando.

Referencias

- Duran, M. (2009). Dictionnaire Quechua-Français-Quechua, Editions HC. Paris.
- Duran, M. (2013). Formalizing Quechua Noun Inflexion. Formalizing Natural Languages with NooJ. Edited by A. Donabédian, V. Khurshudian and M. Silberztein. Cambridge scholars. Newcastle upon Tyne.
- Duran, M. (2014). Morphological and Syntactic Grammars for the Recognition of Verbal Lemmas in Quechua. Proceedings of the NooJ 2014 International Conference, Sassari.
- Guardia Mayorga, C. (1973). Gramática Kechwa, Ediciones Los Andes. Lima Peru.
- Itier, C. (2011). Dictionnaire Quechua-Français, Paris. L'Asiathèque. Paris.
- Perroud, P. C. (1970). Diccionario español kechwa, kechwa español. Dialecto de Ayacucho. Santa Clara, Perú. Seminario San Alfonso.
- Quechua Module available at: <http://www.nooj-association.org/resources.html> [last visit 01- 05- 2021].
- Silberztein, M. (2015). La formalisation des langues, l'approche de NooJ. Iste Ediciones, London.
- Silberztein, M. (2016). Formalizing Natural Languages: The NooJ Approach. Iste Ediciones. London.
- Spanish Module available at: <http://www.nooj-association.org/resources.html> [last visit 01- 05- 2021].
- Soto Ruiz, C. (1976). Gramática quechua: Ayacucho-Chanca. Lima: Ministerio de Educación, Instituto de Estudios Peruanos.