

HACIA LA CONSTRUCCIÓN DE UN DICCIONARIO ELECTRÓNICO DE VERBOS QUECHUA-ESPAÑOL Y ESPAÑOL-QUECHUA

Duran Maximiliano

Université de Franche-Comté Besançon, CRIT, France
Université de Grenoble Alpes, LIG, Francia

duran_maximiliano@yahoo.fr

Recibido el 21 de julio de 2025

Aceptado el 12 de agosto de 2025

Resumen

Este artículo presenta algunos logros en la construcción de un Diccionario Electrónico de Verbos Quechua-Español y Español-Quechua. El objetivo es presentar el vocabulario verbal bilingüe y bidireccional quechua-español de forma exhaustiva y formalizada a fin de desarrollar aplicaciones dentro del Procesamiento Automático de Lenguas (PAL).

Palabras clave: diccionario electrónico, diccionario quechua-español, español-quechua, sufijos verbales, formalización del quechua.

TOWARDS THE CONSTRUCTION OF AN ELECTRONIC DICTIONARY OF VERBS QUECHUA-SPANISH AND SPANISH-QUECHUA

Abstract

This article presents achievements in constructing of an electronic Quechua-Spanish and Spanish-Quechua verb dictionary. The goal is to comprehensively and formally describe the bilingual, bidirectional Quechua-Spanish verb vocabulary, which will enable the development of applications within Automatic Language Processing (ALP)

Keywords: electronic dictionary, Quechua-Spanish dictionary, Spanish-Quechua, verb suffixes, formalization of Quechua

1. Introducción

Dado que no existía un diccionario electrónico de verbos QU-SP, ni SP-QU, nos propusimos construirlo. Para ello nos hemos servido del diccionario digital de verbos quechua-español y su recíproco, el Diccionario de Verbos Español-Quechua construido por M. Duran (2009) y la plataforma lingüística NooJ¹. Entre otras tareas, tuvimos que asociar un modelo flexivo a cada

¹ NooJ es una plataforma de desarrollo lingüístico, que se utiliza tanto para describir lenguajes como para crear aplicaciones de Procesamiento Automático de Lenguas (PAL). NooJ es un software gratuito y de

una de las entradas verbales. Utilizamos para ello los modelos flexivos construidos en gran parte por Duran (2013), (2017). Los diccionarios como el nuestro son herramientas necesarias y útiles para el tratamiento automatizado del quechua, en general y muy particularmente para la traducción automática, que es el objetivo principal de nuestras investigaciones.

En este trabajo, se muestra cómo emplear los diversos recursos que posee el quechua (QU) para generar nuevos verbos por derivación, utilizando un conjunto de sufijos específicos. Como primer paso, hemos inventariado todos los sufijos quechuas que permiten la derivación verbo-verbo y hemos detallado sus correspondientes valores semánticos en español. Este conjunto de sufijos, al que llamo IPS_DRV (Derivadores por Inter-posición); contiene veintisiete elementos. Cada verbo quechua da lugar a por lo menos veintisiete verbos derivados. Así, aplicados a los aproximadamente mil seiscientos verbos simples de la lengua, nos permite obtener cuarenta y tres mil doscientas nuevas formas verbales. Esto nos permitirá obtener la mayoría de las transferencias quechuas de los verbos en español. El diccionario electrónico resultante incluye etiquetas semánticas relevantes para cada verbo quechua y su correspondiente gramática flexiva.

Para complementar este trabajo, presento, finalmente, el diccionario electrónico español-quechua que hemos construido. Este, contiene más de diez mil verbos simples y verbos frasales en español con su respectiva traducción y su respectiva gramática flexiva.

1.1. Un diccionario electrónico de verbos puede ayudar al PAL del quechua

Hay que tener en cuenta que el QU sigue siendo principalmente una lengua oral. Los escasos diccionarios bilingües escritos (QU-SP) presentan varios obstáculos para su uso en PAL: falta de etiquetado de las categorías gramaticales en las entradas (muchos sustantivos, verbos y adjetivos no se presentan en su forma lematizada, sino parcialmente flexionados; bastantes lemas se presentan con ortografía incoherente entre los diferentes dialectos quechua, etc.)

Para una gran parte de nuestro vocabulario verbal, hemos tomado como referencia uno de los primeros y más completos diccionarios bilingües del quechua, el de Diego Gonzales Holguín (16.008), el que es más rico en léxico que el de Fray Tomás de Aquino (1560) recopilado también a mediados del siglo XVI. Se publicó en 1608 en Lima, Perú, bajo el título: “Vocabulario de la Lengua General de todo el Perú llamada lengua Qquichua o del Inca”. Contiene trece mil ciento cincuenta y ocho entradas en quechua. Primero, hemos tenido que transliterar la edición de 1962 de la Universidad de San Marcos (Holguin (1608)) a la ortografía oficial Ayacucho-Chanka para luego ponerlo en formato digitalizado como aparece en la Figura. 1.

Début du premier livre du vocabulaire quechua de la langue générale du Pérou	
a,INT+SP="a"	proveerse"+FLX=V_TR_H
aà,INT+SP=" Dice O"	akakuchkani,V+SP="Estarse
àà Dios ninj,INT+SP="Invocar o exclamar a Dios"	proveyendo"+FLX=V_TR_H
aà dios nichini,INT+SP="Hacer o dejar invocar o exclamar a dios"	akanchani,V+SP="Ensuciar con estiercol"+FLX=V_TR_H
aà dios niykukuni,INT+SP="Invocar interiormente dentro de si"	akarquuni,V+SP="Acavar de proveerse"+FLX=V_TR_H
aà dios niykachay,INT+SP="Invocacion de dios"	aka,N+SP="El orin y escoria del metal"+FLX=N_V_H
aà dios niykachaq sonqo,INT+SP="El exclamador invocador de dios"	akayuq qellay o akaykusqa,V+SP="El hierro o metal con orin"+FLX=V_TR_H
aà,INT+SP="O si deseando algo "	akasapa,A+SP="El metal muy oriento"+FLX=A_V_H
aà,INT+SP="O si o jala "	akannaq,A+SP="El que no lo tiene o Sin orin"+FLX=A_C_H
àà,INT+SP="quespikuiman"	mana akayoq,A+SP="El que no lo tiene o Sin orin"+FLX=A_C_H
aà ninj,INT+SP="Desear ansiosamente"	
aà nipayani,INT+SP="Desear ansiosamente"	
aà nispa munapayani,INT+SP="Desear ansiosamente"	

Fig. 1 Diccionario quechua-español digitalizado de Holguín

Allí podemos observar que los verbos, aparecen conjugados, en la primera persona del tiempo presente, aunque su traducción al español se presenta en infinitivo, por ejemplo:

akanchani, ensuciar con estiércol
waylluni, amar con muestras de afecto amoroso y tierno
rimani, hablar
purini, andar, caminar
mikuni, comer, alimentarse
takini, cantar

Para extraer de este corpus, todas las raíces verbales, hemos aplicado a nuestro diccionario digitalizado el siguiente filtro, usando el formalismo NooJ:

```
NI_q_extr == Buscar/Reemplazar (patrón PERL, ni$|q$, extraer
líneas)
Buscar/Reemplazar (patrón PERL, ni$, filtrar líneas).
```

De esta manera, se han extraído tres mil novecientos noventa y nueve formas verbales escritas en su forma flexiva. Esto no significa que tengamos esta cantidad de verbos simples en el diccionario de Holguín. Esta lista contiene en realidad, muchas redundancias. Por ejemplo, para el verbo *kay* (ser o estar), el morfema *kani* aparece sesenta y ocho veces. Por lo tanto, tuvimos que construir filtros adicionales como el siguiente: VOC-ni_q, NI_q_extr (VOC-H_brut) para obtener un conjunto de formas verbales distintas (636) que potencialmente contienen raíces verbales. Sin embargo, aún contienen diferentes sufijos de derivación para la misma raíz verbal. Tras aplicar nuevos filtros, pudimos aislar todas las raíces verbales en ese diccionario. Sorprendentemente, encontramos apenas noventa y cinco verbos simples diferentes. Por otro lado, el diccionario contiene muchos otros lexemas que muestran otros tipos de formas flexivas, como *chakirqun* [se acaba de secar], que tiene la raíz verbal *chaki-*, flexionada mediante una combinación del sufijo *-rqu* y la terminación *-n* de la tercera persona en tiempo presente. Cabe destacar que el verbo simple correspondiente, *chakiy* «secar», no pudo aparecer en la lista de los noventa y cinco verbos porque no lleva el sufijo *-ni*.

Además, para revelar los radicales verbales que aparecen como sustantivos, resultantes de la derivación de verbos por los sufijos nominalizadores *-q*, *-na*, *-sqa* o la combinación *-rqu-n*, que

Holguin utiliza como marca aspectual para la tercera persona, programamos filtros como los siguientes:

```
Buscar/Reemplazar (patrón PERL, na$, extraer líneas)
Buscar/Reemplazar (patrón PERL, q$, extraer líneas)
Buscar/Reemplazar (patrón PERL, sqa$, extraer líneas)
Buscar/Reemplazar (patrón PERL, rqun$, extraer líneas)
```

Esto nos permitió obtener quinientos treinta y un radicales verbales nominalizados, trescientos veintisiete participios en *sqa*, treinta y tres formas en agentivo con *q* y ciento ochenta y un formas del pasado para la tercera persona 3+s.

Para identificar los verbos que aparecen como verbalizaciones de sustantivos, utilizamos los sufijos de verbalización *-chay*, *-yay*, *-y* como variables y aplicamos búsquedas en NooJ en cascada al diccionario Holguin, como:

```
Buscar/Reemplazar (patrón PERL, cha$, extraer líneas)
Buscar/Reemplazar (patrón PERL, ya$, extraer líneas)
Buscar/Reemplazar (patrón PERL, y$, extraer líneas)
```

Al final de estas operaciones, obtuvimos un poco menos de quinientos verbos simples. Para continuar la búsqueda, estudiamos otras fuentes adicionales, entre ellas algunas contemporáneas como las de Guardia Mayorga (1973), Peroud (1970) y Middendorf (1970).

(1) Guardia Mayorga presenta sus verbos sin ninguna etiqueta que precise la categoría gramatical:

wankuy: vendar, envolver, amancornar.
waqaychay: guardar.
wachay: parir.
waqay: llorar. Fig. dar frutos.
wanquy: palanquear.
waykay: pegar a una persona entre varios.

(2) Los verbos de Perroud se ven así:

chutay. 1. *Estirar, jalar, extender, tirar de, arrastrar, apretar*
2. *Desencoger, estirarse, desperesarse* 3. *Alargarse*

Tampoco contiene etiquetas y traduce el infinitivo *chutay* para las formas pronominalizadas, que debería aparecer como la derivación: *chutakuy*. Esta presentación sugiere la necesidad de añadir etiquetas sintácticas, semánticas como el dominio de uso, el contexto del verbo, el sentido, etc., para las operaciones de desambiguación.

(3) El de Middendorf se ve así:

yuyay==pensar
churay==poner, colocar
iñiy== consentir, creer

Como se puede observar, se muestran varios significados sin etiquetas gramaticales.

Al reunir todos los verbos de estas fuentes y los procedentes de nuestra propia introspección obtuvimos mil seiscientos treinta y un verbos simples con un solo significado, una muestra de los cuales se aprecia en la Fig. 2.

Dictionary contains 1631 entries

```

maywiy, V+SP=columpiar
maywiy, V+SP=oscilar
maywiy, V+SP=vibrar
michay, V+SP=atajar
michay, V+SP=impedir
michay, V+SP=prohibir
michkay, V+SP=sembrar la michka
michuy, V+SP=mezclar una cosa con otra del mismo genero
mikiy, V+SP=ablandar
mikiy, V+SP=humedecer
mikuy, V+SP=alimentarse
mikuy, V+SP=comer
milchiy, V+SP=buscar al tacto los piojos
milchiy, V+SP=espulgar
millmaray, V+SP=cubrirse de hongos un alimento
millqay, V+SP=coger en los brazos

```

Fig. 2 Diccionario de verbos simples QU_SP

En la Fig. 3, mostramos el mismo diccionario enriquecido con la inclusión de etiquetas semánticas y morfosintácticas, y, sobre todo, atribuyendo a cada entrada su correspondiente gramática de flexión y derivación para mil quinientos cuarenta y cuatro verbos.

C:\Users\ASUS\Documents\Noo\qu\Lexical Analysis\SIMPLE VERBS QU_SP.dic
Dictionary contains 1544 entries

```

maywiy, V+itr+SP="oscilar"+FLX=V_ITR_CONJ0
maywiy, V+itr+SP="vibrar"+FLX=V_ITR_CONJ0
michay, V+SP="atajar a uno por que no haya una cosa, prohibir "+FLX=V_TR+DRV=I
michkay, V+SP="sembrar la michka"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NA
michuy, V+SP="mezclar una cosa con otra del mismo genero"+FLX=V_TR+DRV=IS:NCO
mikiy, V+SP="ablandar"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NAS:NVOCAL+DR
mikiy, V+SP="humedecer"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NAS:NVOCAL+D
mikuy, V+PRO+SP="alimentarse"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NAS:NV
mikuy, V+tr+SP="comer"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NAS:NVOCAL+DR
milchiy, V+SP="buscar al tacto los piojos"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCON
milchiy, V+SP="espulgar"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=NAS:NVOCAL+
millmaray, V+itr+SP="cubrirse de hongos un alimento"+FLX=V_ITR_CONJ0
millqay, V+SP="coger en los brazos"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV=
millqay, V+tr+SP="llevar algo en la falda"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCON
miniy, V+SP="estar en ayunas de comidas y bebidas"+FLX=V_TR+DRV=IS:NCONSO+DRV
mintuy, V+itr+SP="acrecerntar"+FLX=V_ITR_CONJ0
miqlay, V+tr+CZ+SP="llevar algo en la falda"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:N
mirararay, V+SP="castigar al juez aporrear"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCO
mirararay, V+SP="voltear al vencido"+FLX=V_TR+DRV=IS:NCONSO+DRV=QS:NCONSO+DRV
miray, V+itr+SP="acrecerntar"+FLX=V_ITR_CONJ0

```

Fig. 3 Versión electrónica del diccionario QU_SP de verbos simples

Sin embargo, estos mil quinientos cuarenta y cuatro verbos traducidos resultan todavía insuficientes con relación a los más de once mil verbos del español que figuran en los diccionarios impresos. Lo esperanzador, frente a esta disparidad numérica, es que este vocabulario pueda ser agrandado gracias a la morfología de la lengua, tal como pasamos a ver enseguida.

2. La morfología quechua ayuda a generar nuevos verbos

Ya hemos encontrado algunos verbos obtenidos por derivación de un verbo simple, por ejemplo, *akakuchkani* [estar proveyendo] o *chakirqun* [acaba de secarse]. Proceden del diccionario de Holguín: el verbo *akay* [defecar], está derivado por el sufijo *-chka*, y el verbo *chakiy* [secar], con el sufijo *-rqu*. Como hemos señalado, ese diccionario contiene tres mil novecientos noventa y nueve verbos derivados, de los cuales solo noventa y un son simples y la mayoría son derivados. Los siguientes ejemplos muestran otros sufijos verbales que producen nuevos verbos a partir del verbo *maskay* [buscar]:

maska**kuy** buscarse a uno mismo [rebuscar para sí]
 maska**ykuy** buscar con determinación [buscar cuidadosamente]
 maska**pay** espigar [rebuscar]
 maska**riy** buscar superficialmente [buscar superficialmente]
 maska**chiy** hacer buscar a alguien [ordenar la búsqueda]
 maska**payay** hacer investigación [investigar]

Podemos obtener hasta veintisiete verbos nuevos, lo que sugiere la existencia de un conjunto de sufijos especializados que nos permitirán obtener esta derivación productiva V-V, lo que puede ayudarnos a mejorar considerablemente el léxico verbal inicial.

Hemos inventariado todos los sufijos verbales que intervienen en la derivación y la flexión de un verbo. Durán (2017), (2023) ha propuesto cuatro clases²: las desinencias de la conjugación (DC), los sufijos de interposición (SIP), los sufijos pos posicionales (SPP) y el conjunto de sufijos nominalizadores verbales (VN_S).

En efecto, en el proceso de la conjugación de un verbo, existe una clase específica de sufijos que se interponen entre la raíz verbal y la desinencia de la conjugación (en el tiempo presente o futuro), estos sufijos interpuestos (SIP) son:

$SIP = \{-chaku, -chi, -chka, -kacha, -kamu, -kapu, -ku, -lla, -mpu, -mu, -naya, -pa, -paya, -pu, -rqa, -raya, -ri, -rpari, -rqu, -ru, spa, -sqa, -tamu, -wa, -ykacha, -ykachi, -ykamu, ykapu, -ykari, -yku, -ysi\}$ (27)

Podemos expresar esta derivación mediante la fórmula:

$Dsip(V1root) = V1root + sip + y == V2$ (donde D: derivación, sip: sufijo inter posicional, root: raíz verbal)

Esto significa que la función $Dsip$ aplicada al verbo V1 generará el verbo derivado V2, como aparece en los siguientes ejemplos:

```

    asiy [reirse]
    Dri(asi) = asiy+riy = asiriy, aquí el sufijo es-ri
    asiriy =[sonreír]
    samay [descansar]
    Drqu(sama) = sama+rqu+y = samarquy, aquí el sufijo es-rqu
    samarquy [reposar]
  
```

O para el verbo *maskay*/buscar, tendremos:

maska**kuy** [buscarse a sí mismo]
 maska**ykuy** [buscar con determinación]
 maska**pay** [buscar con celo]
 maska**riy** [buscar superficialmente]
 maska**chiy** [hacer que alguien busque]

² Los sufijos interpuestos (SIP): aparecen entre la raíz verbal y la desinencia de la conjugación. En contraposición a los sufijos que aparecen después de la desinencia de la conjugación, a los que llamamos Sufijos Pospuestos (SPP): $SPP = \{-ch, -chá, -chik, -chiki, -chu, -chu(\?), -chusina, -m, -mi, -má, -man, -ña, -pas, -puni, -qa, -raq, -s, -si, -taq, -yá\}$ (20)

El conjunto de desinencias de conjugación (DC) utilizadas para el Tiempo Presente (PR) y para las utilizadas para la conjugación en Tiempo Futuro (F):

$S_DC = \{i, nki, n, nchik, niku, nkichik, nku, saq, nki, nqa, sunchik, saqku, nkichik, nqaku\}$ (14)

Desde el punto de vista formal, estos SIP se aglutan a una raíz verbal y generan un verbo derivado, según el siguiente paradigma desarrollado por Duran (2013):

```
SIP1_CONJ = <B>(:CHAKU | :CHI | :CHKA | :YKACHA | :YKACHI |  
:YMANA | :YKAMU | :YKAPU | :YKARI | :YKU | :YSI | :KACHA |  
:KAMU | :KAPU | :KU | :LLAV | :MU | :NAYA | :APV | :PAYA |  
:PU | :RAYAV | :RIV | :RPARI | :RQU | :RU | :TAMU) (y/INF)
```

Esta derivación generará veintisiete nuevos verbos que contienen un sufijo SIP, como se muestra con los siguientes ejemplos en los cuales podemos observar que se han interpuesto, respectivamente, los sufijos *ri*, *ku*, *nay* y *rqu*:

<i>asiy</i>	[reir]	>	<i>asiriy</i>	[sonreir]
<i>ripuy</i>	[viajar]	>	<i>ripukuy</i>	[mudarse]
<i>rakiy</i>	[separar]	>	<i>rakinayay</i>	[deseo de divorciar]
<i>samay</i>	[descansar]	>	<i>samarquy</i>	[reposar]

De manera más general, cuando aplicamos la gramática SIP1_CONJ al conjunto de mil trescientos cincuenta y seis raíces verbales simples, obtenemos treinta y seis mil seiscientos veintidós nuevos verbos derivados, como se muestra en la Fig. 4, entre los que encontraremos, muchas de las traducciones de los verbos españoles que buscábamos.



Fig. 4 Diccionario obtenido por derivación mono-sufijal del conjunto de verbos simples (36.622)

Otra fuente numerosa para la obtención de nuevos verbos se halla en el conjunto de verbos derivados obtenidos mediante la aglutinación de dos SIP. Los SIP pueden combinarse entre sí, siguiendo la matriz booleana propuesta por Durán (2023), que marca con “1” cuando la combinación es gramaticalmente válida y con “0” si no lo es. Las mil seiscientas raíces verbales simples originales, al aglutinarse con esas combinaciones bi-sufijales, van a generar 268.827 nuevos verbos, entre los cuales aún se pueden encontrar numerosas transferencias para los verbos compuestos del español. Véase un ejemplo en la figura 5.

Dictionary contains 268827 entries
mikutamuykapuy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+AEP+SOIN3+INF mikutamuykamuy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+AEP+PREAT+INF mikutamuykachiy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+AEP+POLI+INF mikutamuchkay, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+AEP+PROG+INF mikutamuchiy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+AEP+FACT+INF mikurullay, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+PRES+POL1+INF mikuruchkay, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+PRES+PROG+INF mikurgullay, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+PAPT+POL1+INF mikurquysi, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+PAPT+COLL+INF mikurquchkay, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+PAPT+PROG+INF mikurparirquy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+ASUR+PAPT+INF mikurparipuy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+ASUR+APT+INF mikurparimuy, mikuy, V+PRO+SP="alimentarse"+FLX=V_TR_INF+ASUR+ACENT+INF

Fig. 5 Verbos derivados de QU-SP que llevan dos sufijos IPS (268,827)

3. Del diccionario electrónico de verbos francés-quechua al diccionario español-quechua

El procesamiento automático del par de lenguas quechua-español requiere la construcción de un diccionario electrónico de verbos. El diccionario QU_SP que se muestra en la Fig. 2 nos ha servido de base para construir nuestro primer diccionario quechua-español que contenía cinco mil seiscientos verbos simples y frasales.

Para mejorarlo, he aprovechado el trabajo previo de Duran (2017), en el cual obtuvo manualmente la traducción de ocho mil seiscientos setenta y siete verbos franceses al quechua, procedentes del diccionario francés de Dubois y Dubois-Charlier (1994). En primer lugar, se obtuvo la traducción de los verbos franceses al español y luego se pusieron en línea los verbos quechua con las correspondientes traducciones castellanas. Tras esta etapa, para obtener la traducción de más verbos españoles entre los treinta y seis mil seiscientos veintidós nuevos verbos derivados, seguimos de cerca el siguiente esquema:

Si tomamos un verbo simple como *asiy* [reír], y aplicamos el sufijo SIP «ri», obtenemos el nuevo verbo *asiriy* [sonreír], el que se puede encontrar en algunos diccionarios impresos. Por el contrario, el verbo derivado *asichakuy*, derivado por el sufijo SIP «chaku», no está presente en ninguno de esos diccionarios, y nos daría el verbo frasal «reír ridículamente». Luego tomamos, uno a uno los otros sufijos SIP y tratamos de asignarle su traducción castellana.

Tomemos otros ejemplos de verbos derivados que no aparecen en los diccionarios impresos; para obtener su traducción, tuvimos que deducirlos de la semántica del sufijo “*kui*” empleado:

ripuy [irse], que deriva en *ripukuy* [mudarse];
maskay [buscar] > *maskakuy* [buscar para sí];

o con otros sufijos como: *naya*, *yku*, *pa*, *rqu*, obtuvimos:

rakiy [separar] > *rakinayay* [deseo de separar];
maskay [buscar] > *maskaykuy* [estudiar minuciosamente];
maskay [buscar] > *maskapay* [investigar]
samay [descansar] > *samarquy* [reposar];

Este trabajo manual sostenido durante más de una década (y no está completamente terminado) nos ha permitido obtener finalmente el diccionario electrónico de verbos quechua-español el que contiene los dieciséis mil doscientos ochenta y siete verbos que se presentan en la Fig. 5.

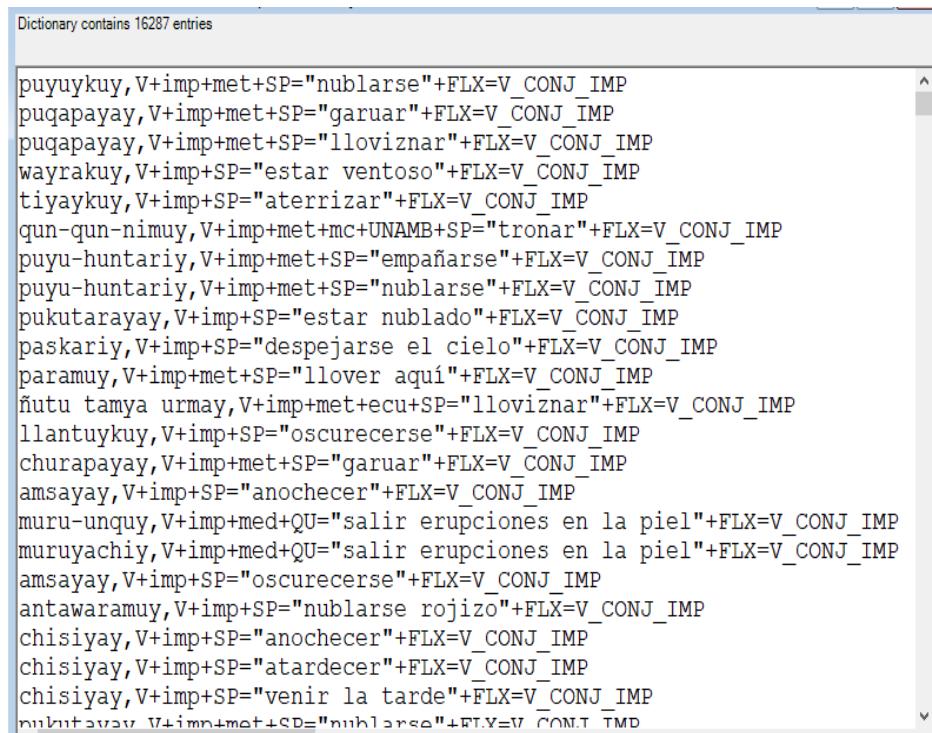


Fig. 5 Diccionario electrónico de verbos quechua-español (contiene 16.287 verbos)

4. El diccionario electrónico de verbos en español-quechua

Lograda una primera parte del proyecto de traducción automática (TA) en lo concerniente al léxico, necesitábamos complementarla con un diccionario electrónico de verbos español-quechua. Después de indagar sobre la disponibilidad de un diccionario electrónico español, notamos que pocos diccionarios monolingües de español de código abierto posean una configuración digital. Entre esos pocos, encontramos en el trabajo de A. Ríos (2008). Allí, ella utiliza un diccionario electrónico español-quechua considerando, para el español, un conjunto de etiquetas que simbolizan las categorías gramaticales dadas por la Real Academia Española (RAE) y, para el quechua, otro conjunto de etiquetas correspondientes a la variante escrita del quechua del Cuzco. De nuestra parte, en el presente trabajo, abordamos la variante quechua de Ayacucho.

Es probable que aplicaciones en línea, como Google u otros programas de traducción recientes, hayan creado sus propios diccionarios electrónicos, pero no podemos utilizarlos porque no son de código abierto. Entonces nos propusimos construir nuestro propio diccionario electrónico español-quechua, variante quechua de Ayacucho.

Dentro de la biblioteca de la plataforma NooJ, encontramos, integrado en el módulo de español, el diccionario electrónico de la RAE (2009). Fue desarrollado inicialmente en la Universidad de Barcelona y luego se implementó con lexemas específicos de las variantes argentino-rioplatense y chilena del español. Contiene 68.783 entradas y es totalmente compatible con el formalismo de NooJ. Por lo tanto, decidimos tomarlo como recurso inicial.

Tabla 1. Etiquetas equivalentes correspondientes entre SP y QU

Abreviaciones seguido de su descripción en SP y QU			
Equivalencias simbólicas esenciales entre el QU y el SP			
SP	QU	FLEXIÓN en Castellano	FLEXIÓN en Quechua
inf	INF	inf infinitivo	INF Infinitivo
ppio	GER1	ppio participio indefinido	GER1 gerundio flexionado
ger	GER	ger gerundio	GER gérondíf -spa simultanée
pres	PR	pres presente	PR presente
ppp	PPA	ppp había comido	mikusjani SQA_PR_PLU
pps	PRES	pps pasado simple cantaste	PRES pasado simple RU takirunki
fut	F	fut futuro	F futuro
cond	C	cond condicional	C conditional
pi	PASS	pi pretérito perfecto indefinido cantabas	PASS RA passé simple accompli cantabas
	PASSA	había rápidamente	PASSA RQUSQA_PRM2
subj	SUBI	subj subjuntivo fut+subj	SUBI subjuntivo imperfecto SPAJA
subj	SUBIPAS	subj subjuntivo fut+subj	SUBI subjuntivo imperfecto SQA_PR_SS subj
subj	SUBI	subj subjuntivo fut+subj	SUBI subjuntivo imperfecto SQA_PR_SS subj
imp	IP	imp imperativo	IP imperativo directo
ppcomp	ppi	pretérito perfecto rja	PPI pretérito perfecto indefinido rja
la	1+s	la primera persona singular	1+s primera persona singular

Después de extraer los verbos contenidos en este diccionario, comenzamos el largo proceso de traducirlos manualmente. Hasta la fecha, hemos obtenido la traducción al quechua de diez mil novecientos doce verbos españoles, como se muestra en la Fig. 6.

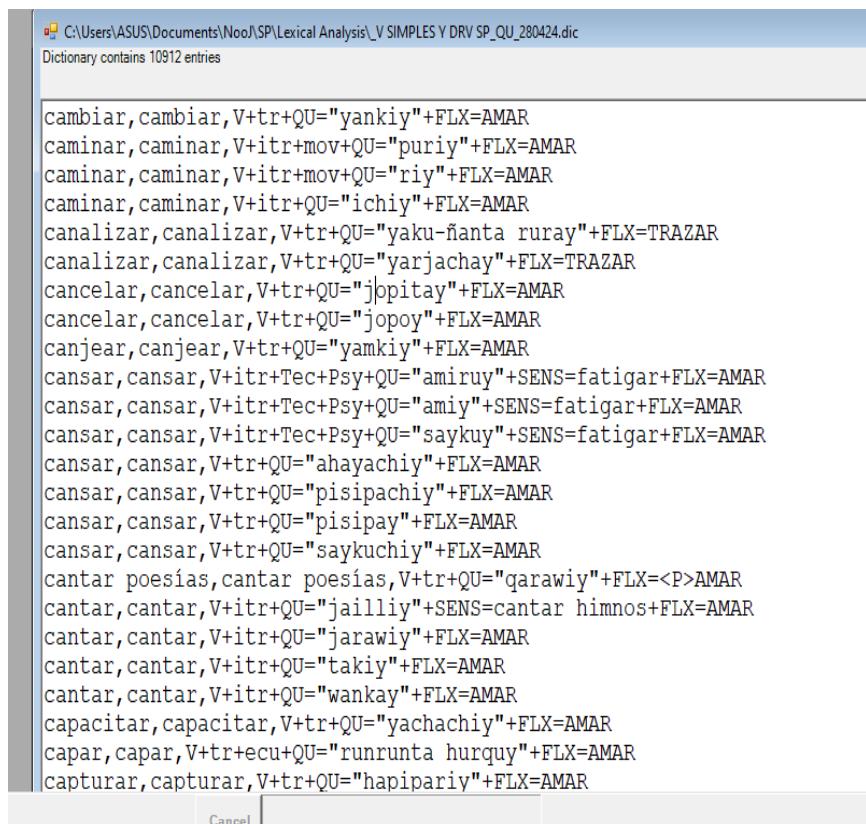


Fig. 6 Diccionario electrónico de verbos español-quechua (contiene 10912 verbos)

Este diccionario incluye verbos simples, así como verbos fraséales en español, tales como:

andar a ciegas [taplaykachay], derivado del verbo taplay [divagar]

hacer algo con alegría [hapachay], derivado del verbo hapay [alegrarse]
poner algo en pie [tiqnuchiy], derivado del verbo tiqnuuy [estar de pie]

En él, podemos observar que:

1. Cada entrada incluye varios morfemas que representan etiquetas morfosintácticas y semánticas como V = verbo, tr = transitivo, itr = intransitivo, mov = movimiento, Psy= psicológico, etc.
2. La traducción al quechua correspondiente. En caso de polisemia, se incluye una línea para cada significado; por ejemplo, caminar tiene tres traducciones al quechua.

caminar,caminar,V+itr+mov+QU="puriy"+FLX=AMAR
caminar,caminar,V+itr+mov+QU="riy"+FLX=AMAR
caminar,caminar,V+itr+QU="ichiy"+FLX=AMAR

3. Cada entrada contiene una gramática flexiva FLX, lo que permite generar las flexiones. La gramática FLX generará todas las formas conjugadas y derivadas del verbo. Por ejemplo, la fórmula FLX=AMAR que aparece con el verbo **caminar** tiene la fórmula siguiente:

AMAR = (:TER | :TER1) | (<B2> (:TER2 | :TER3 | :TER4 | :TER5));

y cada paradigma, componente de la sub-fórmula TER se presenta como sigue:

TER1 = (é/fut+ind+1a+sg | ás/fut+ind+2a+sg | á/fut+ind+3a+sg |
emos/fut+ind+1a+pl | éis/fut+ind+2a+pl | án/fut+ind+3a+pl |
ía/cond+ind+1a+sg | ías/cond+ind+2a+sg; etc.

5. Conclusión

Hemos logrado construir un diccionario electrónico quechua-español. Cada entrada tiene su gramática flexiva específica y diferentes etiquetas semánticas y morfosintácticas. Contiene actualmente sesenta y tres mil veinte entradas con sus respectivos parámetros lingüísticos.

Paralelamente, partiendo del diccionario monolingüe de verbos en español del módulo SP NooJ, lo hemos completado para convertirlo en un diccionario electrónico bilingüe que reúne en la actualidad diez mil novecientos doce verbos españoles con su traducción al quechua.

Nuestro trabajo futuro se enfoca en un sistema que nos ayude a poner en línea un millón cuatrocientas sesenta y siete mil formas verbales conjugadas del español con las correspondientes traducciones al quechua, seleccionadas de las tres millones doscientos treinta y seis mil formas conjugadas del QU.

6. Referencias

Dubois, J. & F. Dubois-Charlier. (a). Dictionnaire électronique des verbes français. Table Excel, 26.610 entrées. Téléchargeable sur <http://www.modyco.fr/fr/Ressources/ldlvf.html>. Version XML sur <http://rali.iro.umontreal.ca/rali/?q=fr/lvf> (1994).

Duran, Maximiliano, Diccionario Quechua-Castellano-Quechua, Editions HC. Paris, (2009).

Duran, Maximiliano, Formalizing Quechua verbs Inflexion. Proceedings of the NooJ 2013 International Conference, Saarbrücken, (2013).

Duran, Maximiliano, Dictionnaire électronique français-quechua des verbes pour le TAL. Thèse Doctorale. Université de Franche-Comté. Mars 2017 (2017).

Duran, Maximiliano, Gramáticas morfo-sintácticas para el reconocimiento automático de raíces verbales en quechua. Aprendo con NooJ. Universidad de Rosario. Argentina. (2023)

Fray Domingo de Santo Tomas, Lexicom, o Vocabvlario de la Lengua General del Perv. Impresso en Valladolid par Francisco Fernandez de Cordova, (1560).

Gonçalez Holguin, Diego, Vocabulario de la Lengva General de todo el Perv llamada Len-gua Qquichua o Lengua del Inca. Impresa en la Ciudad de los Reyes. Por Francisco del Can-to, (1608).

Guardia Mayorga, César, Gramatica Kechwa, Ediciones Los Andes. Lima Peru, (1973).

Middendorf, Ernst W., Gramatica Keshua. Trad. d'allemand [Das Runa Simi oder die Keshua Sprache, wie sie gegenwärtig in der Provinz von Cusco gesprochen wird] par Ernes-to More. Aguilar, Madrid, (1970).

Perroud, Pedro Clemente, Diccionario castellano kechwa, kechwa castellano. Dialecto de Ayacucho. Santa Clara, Peru. Seminario San Alfonso, (1970).

Real Academia Española: Asociación de Academias de la Lengua Española: Nueva Gtamáti-ca de la Lengua Española. Espasa-Calpe, Madrid, España (2009)

Rios (Annette Rios, Anne Gohring, and Martin Volk. A quechua-spanish parallel treebank. LOT Occasional Series, 12:53–64. (2008)

Silberstein Max, La formalisation du dictionnaire LVF avec NooJ et ses applications pour l'analyse automatique de corpus. Langages 3/2010 (n° 179-180), pp. 221-241. (2010).