



Transformations and Paraphrases for Quechua Sentiment Predicates

Maximiliano Duran^{1,2} (✉)

¹ Université de Franche-Comté, Besançon, France
duran_maximiliano@yahoo.fr

² LIG, UGA, Grenoble, France

Abstract. In this paper, I present a study on the automatic generation of paraphrases of Quechua sentiment predicates. Using the transformational NooJ engine, I first build the rules corresponding to the elementary transformations for the Quechua language. I then describe in detail grammars performing pronominalization, reduction, and permutation of the arguments, passivation, and some others. Then, I show how they can combine with one another, respecting certain syntactic constraints, in order to obtain complex transformations. I present the electronic dictionary of Quechua sentiment verbs that I have built. Finally, I construct a particular subclass of transformations that will automatically generate paraphrases of a Quechua sentiment predicate.

Keywords: NooJ · Quechua · Syntactic analysis · Transformational analysis · Transformational grammar · Sentiment verbs · Sentiment predicates in quechua · Paraphrase · Machine translation

1 Introduction

Our goal is to construct a syntactic grammar that recognizes all the grammatically correct paraphrases of a Quechua sentiment sentence like {*Gervasio Romildata kuyan*/Gervasio loves Romilda}. As Languella [10], I adopt Harris's concept of paraphrase, which Harris [9] borrowed from the concept of morphism in mathematics. A morphism is a structure-preserving map from one mathematical set of objects to another. According to Harris, a sentence is a paraphrase of another one if a change occurs in the morphophonemic shape of the transformed sentence while preserving the original lexical morphemes and meaning. Paraphrasing has been used by many authors, such as Barreiro [1, 2], who used it to prepare texts for machine translation; Ben et al. [3], who used it for bilingual MT for Arabic-English translation of relative clauses; or Fehri et al. [8], who used it for Arabic-French translation of named entities.

Silberztein [11, 12] shows how, by combining a parser and a generator, and applying them to a syntactic grammar, one can build a system that takes one sentence as its input, and produces all the sentences that are morpho-syntactically or semantically related to the original sentence, or share the same lexical material with it. Figure 1 shows our NooJ grammar for the passive transformation:

- [Passive]{*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Romildam Gervasiopa kuyasqan*/Romilda is loved by Gervasio}.

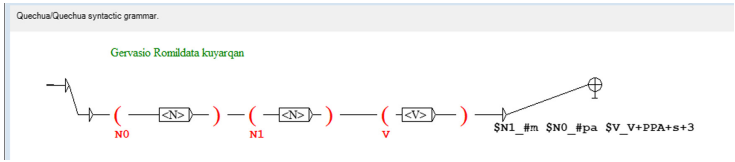


Fig. 1. [Passive]: {*Romildam Gervasiopa kuyasqan*/Romilda is loved by Gervasio}

The graph of Fig. 1 uses three variables, \$N0 (the subject S), \$N1 (the object O) and \$V (the verb V). When parsing the sentence {*Gervasio Romildata kuyan* (N0 N1 V)/Gervasio loves Romilda}, the variable \$N0 stores the word {Gervasio}, \$N1 stores the word {Romilda} and \$V stores the word {*kuyan*/loves}.

The grammar output “\$N1_m \$N0_#pa \$V_V+PPA+s+3” produces the string {*Romildam Gervasiopa kuyasqan*/Romilda is loved by Gervasio}, where V+PPA+s+3 symbolizes the conjugated form past participle of the verb {*kuyay*/to love}, third person singular.

2 Syntactic PoS Transformations

2.1 Noun Transformations

In order to write a grammatical transformation of a sentence in Quechua, an agglutinative language, one needs to take into account the morpho-syntactic transformations of each PoS included in the sentence. This is because each PoS may be inflected or transformed by agglutination of one or more suffixes, acting as operators, coming from their corresponding class of suffixes.

For instance, *Calisto*, a proper name, may be inflected, as follows: {*Calistom*/it is Calisto} (-*m* is the suffix of assertion), {*Calistos*/people say that it is Calisto} (-*s* is the suffix of hearsay), {*Calistoqa*/It’s Calisto} (-*qa* is the suffix of topic), {*Calistochá*/It is probably Calisto} (-*chá* is the suffix of uncertainty), etc.

All these inflected forms are obtained applying N-suff¹ operators using the NooJ grammar of Duran [7]. Many binary or ternary combinations of these suffixes may also generate grammatical transformations of (N). Figure 2 shows part of the grammar that generates the mono-suffix transformations of the proper name Calisto.

Let us consider the following direct sentence: {*Calisto wasinta llimpin*/Calisto paints his house}. Replacing the proper noun Calisto by some of its inflections, we obtain transformed sentences like:

¹ Suffixes for the inflection of proper nouns ending in a vowel: N-suff = {-*ch*, *chá*, -*cha*, -*chik*, -*chiki*, -*chu*, -*chu?* -*hina*, -*kuna*, -*lla*, -*má*, -*man*, -*manta*, -*m*, -*ntin*, -*niraq*, *niray*, -*ña*, -*p*, -*pa*, -*paq*, -*pas*, -*poss*(7v), -*puni*, -*qa*, -*rayku*, -*raq*, -*ri*, -*s*, -*ta*, -*taq*, -*wan*, -*ya!*, -*yá*, -*yupa*}.

- {*Calistos wasinta llimpin*/they say that Calisto paints his house}
- {*Calistoqa wasinta llimpin*/It's Calisto that paints his house}
- {*Calistochá wasinta llimpin*/it is probable that Calisto paints his house}

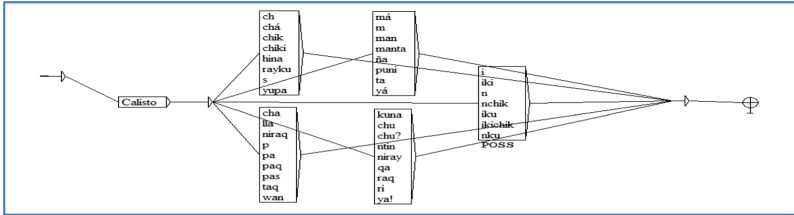


Fig. 2. Mono-suffix transformation of a subject

But this grammar may also generate some ungrammatical forms, such as {**Calistop wasinta llimpin*/*Calisto's paints his house}, erroneous because *Calistop* is in the genitive case. That's why we need to add syntax constraints to this grammar in order to obtain more accurate outputs.

Furthermore, in Quechua, we may also get transformations by combining two or more N-suff suffixes. For instance, according to Duran [5 p. 73], using two agglutinated nominal suffixes, we obtain more than 320 grammatical transformations of a proper noun, like: {*Gervasio-cha-m llimpin*/little Gervasio does paint}; {*Gervasio-raq-mi llimpin*/Gervasio paints first}; {*Gervasio-lla-m llimpin*/only Gervasio paints}; etc.

For three agglutinated nominal suffixes, we obtain more than 720 grammatical transformations, such as {*Gervasio-cha-lla-m llimpin*/it's only little Gervasio who paints}; {*Gervasio-cha-lla-s llimpin*/they say that it's only little Gervasio who paints}; {*Gervasio-nchik-lla-s llimpin*/they say that it's only our little Gervasio who paints}; etc.

2.2 Pronoun Inflections

For a pronoun like {*pay* (third person singular)} we will have the following transformations:

- {*paymi*/it is him (assertion)}, {*paysi*/people say that it is him (hearsay)},
- {*payqa*/concerning him (topic)}, {*paychá*/It is probably him (uncertainty)}.

The class of pronominal suffixes contains 39 elements, which implies that we may obtain 39 mono-suffix inflections. However, only 28 of these suffixes generate transformed nouns that can be used as the subject of the transformed sentence.

If we take again the sentence: {*Calisto wasinta llimpi n*/Calisto paints his house} and replace the proper noun Calisto by some of these transformed pronouns, we obtain sentences like these:

- {*Paysi wasinta llimpin*/they say that he paints his house}
- {*Paychá wasinta llimpin*/it is probable that he paints his house}.

2.3 Verb Derivation

For any Quechua verb, we have two classes of suffixes which derivate verbs: interposition suffixes (IPS)² and postposition suffixes (PPS)³.

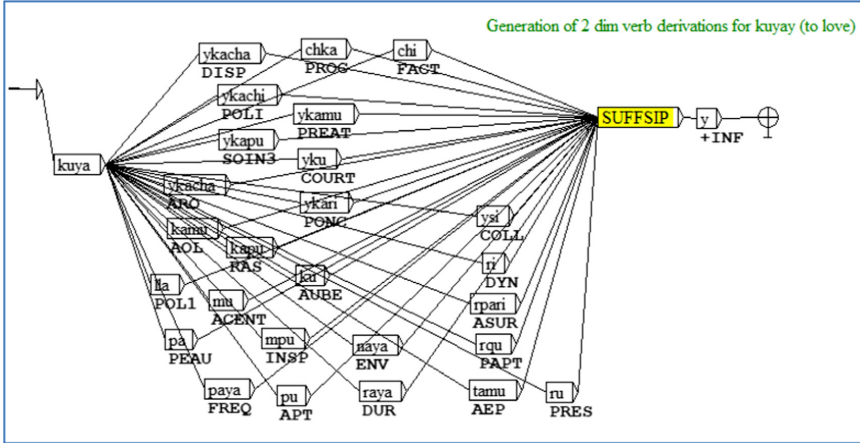


Fig. 3. *Kuyay* (to love) derived by two-dimensional combinations of IPS suffixes

Taking the sentiment verb {*kuyay*/to love}, and parsing it with the grammar of Fig. 3, we obtain 624 new derived verbs with {*kuya*} as the lemma and a combination of two SIP suffixes. They look like in the following sample:

```
# Dictionary generated automatically
kuyarrichkay, V+FR="to start loving"+FLX=V_SIP_INF+DYN+CHKA
+INF;
kuyappayarquy, V+FR="to love repeatedly in a short
time"+FLX=V_SIP_INF+FREQ+RQU+INF;
kuyappayariy, V+FR="to love repeatedly in a delicate man-
ner"+FLX=V_SIP_INF+FREQ+RI+INF;
kuyayykachamuy, V+FR="to love in a dispersed manner"+FLX
=V_SIP_INF+ARO+MU +INF; ...
```

² IPSsuff = {*chaku, chi, chka, ykacha, ykachi, ykamu, ykapu, ykari, yku, ysi, kacha, kamu, kapu, ku, lla, mpu, mu, naya, pa, paya, pu, ra, raya, ri, rpari, ru, tamu, rqa, rqu, spa, sqa, na, pti,stin, wa*}. Only the first 27 of them serve to derivate into a verb.

³ PPSsuff = {*ch, chá, chik, chiki, chu(?)*, *chu, chusina, má, man, m, mi, ña, pas, puni, qa, raq, s, si, taq, yá(!)*}. A total of 28 suffixes.

For three-layer combinations of IPS suffixes, we apply the algebraic grammar $V_SIP_INF = :V_SIP1_INF:V_SIP2_INF:V_SIP3_INF$; for details, see Duran [5], which generates 3,006 new derived verbs, in the infinitive form, ready to be derived or conjugated, as shown in the sample below:

kuyap**paruchkay**, V+EN="to keep loving in an intensive manner"+FLX=V_SIP_INF+PEAU+PRES+PROG+INF;
 kuyay**kachapayamuy**, V+EN="to love repeatedly in a dispersed manner"+FLX=V_SIP_INF+DISP+FREQ+ACENT+INF;
 kuyap**payaykarikuy**, V+EN="to love mutually repeatedly and with care"+FLX=V_SIP_INF+FREQ+PONC+AUBE+INF;

3 The Verbal PPS Transformations

For obtaining one layer of PPS inflections of a verb, we apply the grammar of Fig. 4.

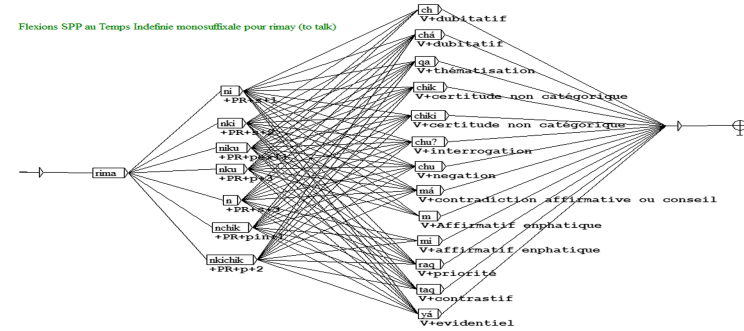


Fig. 4. rimay (to talk) inflected by one-layer postposition suffixation (PPS)

To generate all the binary PPS transformations of a verb, following the guidelines given in the work of Duran [5, pp. 255–257], I have constructed the NooJ grammar

- $V_PPS2 = : PPS2 | : PPS2_F$;

where **PPS2** contains 72 agglutinations for the present tense, such as the following: {chusinam | manchá | manchik ... | punis | punitaq | ... raqchik | raqchu? | taqchu | taqmá | taqmi | taqsi | taqyá}. **PPS2_F** corresponds to the future paradigms. This grammar produces 1,008 binary PPS transformations of the verb *kuyay*, such as those appearing in the following sample:

kuyankuta**taqchá**, kuyanik**kuchusinam**, kuyanik**kuraqchá**, kuyanki **puni chu**, kuyanki**punichu?**, kuyanki**rajsi**, kuyani**ñachik**, kuyani**punimá**, kuyai**raqyá**, kuyaik**manraq**,...

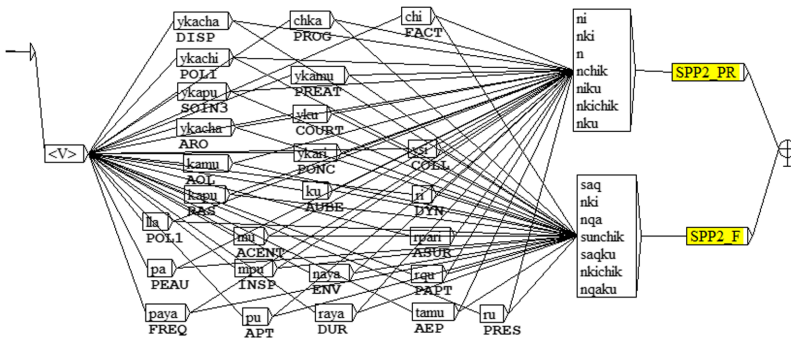
One of the important properties of verb transformation in Quechua is that one can compose IPS and PPS transformations in order to obtain mixed transformations like this one:

- {*rima-yku-ni-raq-mi* (*rimaykuniraqmi*/First of all, I presented my greetings)}

where we have one IPS (-*yku-*) between the lemma and the ending *ni*, the first person singular and two PPS (-*raq, -mi*) after the ending.

The grammar of Fig. 5, constructed following Duran [5, pp. 257–258] and Duran [4], generates mixed transformations. In the example below, I apply a mixed-grammar composed of two-dimensional PPS and one-dimensional IPS transformations, to the verb {*rimay*/to talk}. We use the grammar V_MIX12, which generates all the transformations containing 1 IPS transformation and two PPS transformations. After parsing, we obtain 17,280 mixed transformed verbal forms like in the sample below:

- *rimakamunraqmi*, V+EN="to talk"+FLX=V_MIX12+AOL+s+3
- *rimakachankichikmanpas*, V+EN="to talk "+FLX=V_MIX12+ARO+p+2
- *rimakachankichikpaschá*, V+EN="to talk "+FLX=V_MIX12+ARO+p+2
- *rimakachankichikpaschik*, V+EN="to talk "+FLX=V_MIX12+ARO+p+2



Generation of mixed transformations SIP1_SPP2 of a verb

Fig. 5. A mixed IPS-SPP transformation grammar for a direct transitive sentence

Let's consider some examples of transformations including these mixed inflections of the verb to talk {*Calisto mamanta rimaykunraq*/Calisto talks to his mother}:

- {*Calistos mamanta rimaykunraq*/they say that, first of all Calisto talked to his mother}
- {*Calistoqa mamanta rimaykullanraq*/concerning Calisto, he first of all talked to his mother with much respect}
- {*Calisto mamanta rimariykunqaraq*/Calisto will talk whispering first to his mother}.

In Fig. 6, we show the syntactic grammar, which gathers these combined transformations. It is capable of recognizing more than one million transformed sentences of a direct transitive sentence such as { *Gervasio Romildata kuyan*/Gervasio loves Romilda }.

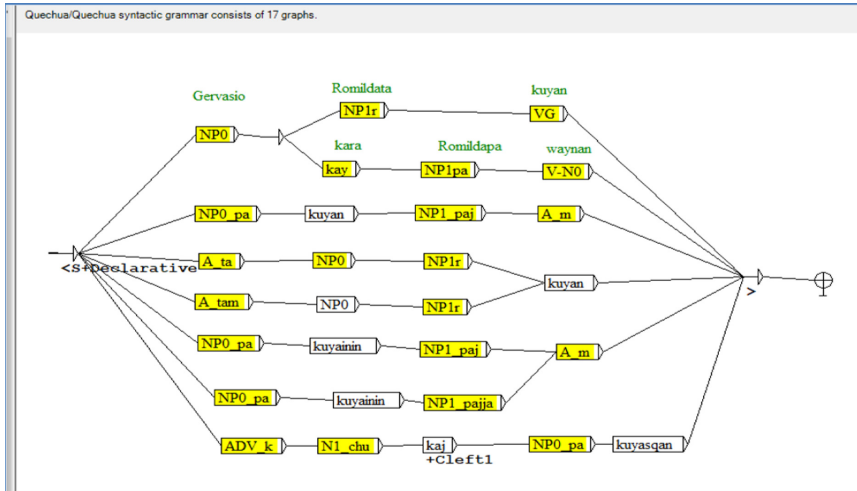


Fig. 6. A grammar which recognizes the transformations of “*Gervasio Romildata kuyan*”

3.1 The Class of Sentiment Verbs

Before introducing our study on paraphrases of sentiment predicates, we will present our dictionary of sentiment verbs. The multilingual electronic dictionary (QU-FR-QU) that we have been building during the last eight years has been enhanced with new linguistic information (namely, inflectional, derivational and morpho-syntactic properties) as well as with some semantic relations. This permits, among other linguistic phenomena, the generation of paraphrases. This dictionary includes the class of sentiment verbs, marked by “+sent”. Figure 7 shows an excerpt of this class of verbs (QU-FR).

3.2 Elementary S-Transformation⁴ of a Sentiment Sentence

Let us consider the declarative initial sentence of the type N0N1V:

$$\{Gervasio Romildata kuyan/Gervasio loves Romilda\} \tag{1}$$

which we want to paraphrase. We present some elementary transformations of a SOV sentence.

- (1) The permutation N1_V: [PermN1_V]

⁴ S-Transformation or Sentence-Transformation: “is an operator that links sentences that share common semantic material” – Harris [9].

```

# Language is: qu
# extract "+sent"
amirparisja-kay,V+Vq+itr+sent+DOM=PSY+FR="être dégoûté"+FLX=V_I TR_CONJ0
amisja-kay,V+sent+DOM=PSY+DOM=PSY+FR="en avoir marre"+FLX=V_I TR_CONJ0
anayniy,V+Vq+tr+sent+FR="admirer"+SENS=encenser_admirer+DOM=AAA+FLX=V_TR_CONJ0
afayniy,V+Vq+tr+sent+FR="louer"+SENS=BBB+DOM=AAA+FLX=V_TR_CONJ0
anchay-anchay_kay,V+Vq+itr+sent+DOM=PSY+FR="être arrogant"+FLX=V_I TR_CONJ0
anchay-anchay_kay,V+Vq+itr+sent+DOM=PSY+FR="être hautain"+FLX=V_I TR_CONJ0
anchayupay,V+Vq+tr+sent+sent+FR="apprécier"+SENS=BBB+DOM=AAA+FLX=V_TR_CONJ0
asiriy,V+Vq+tr+sent+FR="rire doucement"+SENS=avoir expression_gaie+DOM=SOM+FLX=V_TR_CONJ0
atipasja-kay,V+Vq+itr+sent+DOM=PSY+FR="être abbatu"+FLX=V_I TR_CONJ0
atipasja-kay,V+Vq+itr+sent+DOM=PSY+FR="être impuissant"+FLX=V_I TR_CONJ0
atoj-kay,V+Vq+itr+sent+DOM=PSY+FR="être spiègle"+FLX=V_I TR_CONJ0
auki_kay,V+Vq+itr+sent+DOM=PSY+mc+UNAMB+FR="être fougueux"+FLX=V_I TR_CONJ0
aynikuy,V+Vq+tr+sent+DOM=PSY+DOM=PSY+FR="être solidaire"+FLX=V_I TR_CONJ0
ayniy,V+sent+DOM=PSY+DOM=PSY+FR="coopérer"+FLX=V_I TR_CONJ0
chikij-kay,V+Vq+itr+sent+DOM=PSY+FR="être envieux"+FLX=V_I TR_CONJ0
chignichiy,V+Vq+tr+sent+DOM=PSY+FR="semer la zizanie"+SENS=semer la zizanie+DOM=AAA+FLX=V_TR_CONJ0
chuya-kay,V+Vq+itr+sent+DOM=PSY+FR="être clair"+FLX=V_I TR_CONJ0
ima-jukuj-kay,V+Vq+itr+sent+DOM=PSY+mc+UNAMB+FR="être indifférent"+FLX=V_I TR_CONJ0
ima-jukuyakuy,V+Vq+itr+sent+DOM=PSY+mc+UNAMB+FR="devenir indifférent"+FLX=V_I TR_CONJ0
quchukuy,V+sent+DOM=PSY+DOM=PSY+FR="devenir joyeux"+FLX=V_I TR_CONJ0
jochukuy,V+Vq+itr+sent+DOM=PSY+FR="se rejouir"+FLX=V_I TR_CONJ0
kaminakuy,V+Vq+tr+sent+FR="se insulter"+SENS=s'insulter_réciproquement+DOM=AAA+FLX=V_TR_CONJ0
kamiy,V+Vq+tr+sent+FR="gronder"+SENS=menacer+DOM=SOC+FLX=V_I TR_CONJ0
kusikuy,V+sent+DOM=PSY+DOM=PSY+FR="devenir heureux"+FLX=V_I TR_CONJ0
kusinaj-kay,V+Vq+itr+sent+DOM=PSY+FR="être mécontent"+FLX=V_I TR_CONJ0
kuyay,V+Vq+tr+sent+FR="aimer"+SENS=aimer_chérir+DOM=PSY+FLX=V_I TR_CONJ0
llakikuy,V+Vq+itr+sent+FR="s'attrister"+SENS=se chagriner_désoler+DOM=PSY+FLX=V_I TR_CONJ0

```

Fig. 7. Excerpt of the class of QU sentiment verbs.

In QU it is possible to permute the verb and the object without modifying the semantics of the sentence:

- [PermN1_V] {*Gervasio Romildata kuyan/Gervasio loves Romilda*} = {*Gervasio kuyan Romildata/Gervasio loves Romilda*},
which can be symbolized as: [PermN1_V] (N0N1V) = N0 V N1

(2) The permutation N0_V: [PermN0_V]

It is also possible to permute the verb and the subject without modifying the semantics of the sentence:

- [PermN0_V] {*Gervasio Romildata kuyan/Gervasio loves Romilda*} = {*Kuyan Gervasio Romildata/Gervasio loves Romilda*},
which can be symbolized as: [PermN0_V] (N0N1V) = V N0 N1

(3) Pronominalize the subject N0: [ProN0]. We will have:

- [ProN0] {*Gervasio Romildata kuyan/Gervasio loves Romilda*} = {*pay Romildata kuyan/He loves Romilda*},
which can be symbolized as: [ProN0] (N0NV) = pay N1 V

(4) Pronominalize the object N1: [ProN1]. We will have:

- [ProN1] {*Gervasio Romildata kuyan/Gervasio loves Romilda*} = {*Gervasio payta kuyan/Gervasio loves her*},
which can be symbolized as: [ProN1] (N0N1V) = N0 payta V

(5) Pronominalize both the subject and the object N0, N1: [ProN0N1]

We will have:

- [ProN0N1] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Pay payta kuyan*/He loves her},
which can be symbolized as: [ProN0N1](N0N1V) = pay payta V

(6) Nominalize the verb V: [Vnom_i]. We will have:

- [Vnom_i] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Romildam Gervasiopa kuyainin*/Romilda is Gervasio's love},
which can be symbolized as: [Vnom_i](N0N1V) = N1m payta V

(7) Nominalize the verb V: [Vnom_j]. We will have:

- [Vnom_j] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Romildam Gervasiopa kuyaqnin*/Romilda is Gervasio's lover},
which can be symbolized as: [Vnom_j](N0N1V) = N1m N0pa V+NV+POSC_c+3+s,
where *NV+POSC_c+3+s* stands for: nominalized verb as agentive and possessive, in the third singular person.

(8) Passive of verb V: [Passive] or The Passive transformation. We will have:

- [Passive] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Romildam Gervasiopa kuyasqan*/Romilda is loved by Gervasio},
which can be symbolized as: [Passive](N0N1V) = N1m \$N0pa \$V_V+PPA+s+3

(9) Cleft operator: [Cleft_0]

Here, a single message is divided into two classes. We will have:

- [Cleft_0] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Gervasiom kachkan Romilda kuyaq*/It is Gervasio who loves Romilda},
which can be symbolized as: [Cleft_0](N0N1V) = N0m kachkan N1 V+NOM_V+QS

(10) Cleft: [Cleft_1].

- [Cleft_1] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Romildatam Gervasioga kuyan*/It is Romilda that Gervasio loves},
which can be symbolized as: [Cleft_1] (N0N1V) = N1tam N1qa V+PR+3+s

(11) Cleft: [Cleft_2]

- [Cleft_2] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Gevasiom Romilda kuyaqqq*/Gervasio is the one who loves Romilda},
which can be symbolized as: [Cleft_2] (N0N1V) = N0m N1
V+V+NOM_V+QS+THE

(12) Adverb: [ADV_V]

We will have:

- [ADV_V] {*Gervasio Romildata kuyan*/Gervasio loves Romilda} = {*Gevasio Romildata achkallataña kuyan*/Gervasio loves Romilda really a lot},
which can be symbolized as: [[ADV_V]] (N0N1V) = N0 N1ta achkata V+PR+3+s.

Table 1 shows a partial list of these transformations.

Table 1. Elementary transformations of {*Gervasio Romildata kuyan*}

[PermN1_V]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasio kuyan Romildata</i> /Gervasio loves Romilda}
[PermN0_V]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Kuyan Gervasio Romildata</i> /Gervasio loves Romilda}
[Pron-O]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Pay Romildata kuyan</i> /He loves Romilda}
[Pron-1]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasio payta kuyan</i> /Gervasio loves her}
[Pron-2]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasio Romildata kuyaita qun</i> /Gervasio gives Romilda love}
[Cleft-0]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasiom kachkan Romilda kuyaqqin</i> /It is Gervasio who loves Romilda}
[Cleft- 1]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Romildatam Gevasioqa kuyan</i> /It is Romilda that Gervasio loves}
[Nom-O]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasio Romildawan kuyanakun</i> /Gervasio is in love with Romilda}
[Nom-V]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Gervasio kuyaita qatallin Romildapaq</i> /Gervasio feels love for Romilda}
[Nom-1]	{ <i>Gervasio Romildata kuyan</i> } = { <i>Romildam Gervasiopa kuyainin</i> /Romilda is Gervasio's love}
[Passive]	{ <i>Gervasio Romildata kuyan</i> /Gervasio loves Romilda} = { <i>Romildam Gervasiopa kuyasqan</i> /Romilda is loved by Gervasio}

For many of these elementary transformations, it is possible to construct grammars that perform the reverse operation, as we can see in Fig. 8.

Quechua/Quechua syntactic grammar consists of 2 graphs.

\rightarrow (GN) (GN) (<V+PR+3+s>) [Passive_inv] \$0_#m \$\$_#pa \$V_V+PPA+s+3

Gervasio Romildata kuyan >
 (passive_inv) Romildam Gervasiopa kuyasqan

Enter expression: Gervasio Romildata kuyan

54 analyses. Click a solution below to display the corresponding path: Perfect Match Pa

Paths	Outputs
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyaysqan
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyachakusqan
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyachisqan
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyachkasqan
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyaykachusqan
("Main" ("GN" Gervasio) ("GN" Romildata) kuyan)	[Passive_inv]Romildam Gervasiopa kuyaykamusqan

Fig. 8. Grammar for the passive-inv operation of {*Gervasio Romildata kuyan*} → *Romildam Gervasiopa kuyasqan*

4 Composition of Elementary Paraphrasers

In order to obtain composed paraphrasers C_PARPH, it is possible to construct a sequential composition of two or more elementary paraphrasers. They should, of course, respect certain syntactic constraints. Below are some examples.

- (1) As said before, one or more of the PoS contained in each of these elementary transformations may be derived or inflected separately. For example, for (12) we may get {*Gevasio Romildata achkallataña kuyan* / *Gervasio loves Romilda really a lot*}, where the noun and the verb have been derived.
- (2) The 12 operators presented earlier can be sequentially applied to a sentence, respecting the syntactic rules, in order to obtain composed paraphrasers, such as the following:

PermN1_V+ProN0
PermN1_V+ProN1
PermN1_V+ProNON1
PermN1_V+ADV_achka
PermN0_V+PermN1_V
ProN0+PermN1_V
ProN1+PermN1_V
ProNON1+PermN1_V...

To apply the first composed paraphraser of the list, [PermN1+ProN0], we apply first [ProN0]:

$$[ProN0] \{ Gervasio Romildata kuyan \} = \{ pay Romildata kuyan \}$$

Then we apply PermN1_V:

$$[\text{PermN1_V}] \{ \text{pay Romildata kuyan} \} = \{ \text{pay kuyan Romildata} \}$$

We may now compose three elementary transformations, such as below:

$[\text{PermN1_V} + \text{ADV_achka} + \text{ProN0}] \{ \text{Gervasio Romildata kuyan} \}$. We will have the following sequence of results:

- ProN0: *pay Romildata kuyan*
- ADV_achka: *pay Romildata achkata kuyan*
- PermN1: *pay Romildata kuyan achkata*

$$[\text{PermN1_V} + \text{ADV_achka} + \text{ProN0}] \{ \text{Gervasio Romildata kuyan} \} = \{ \text{pay Romildata kuyan achkata} \}.$$

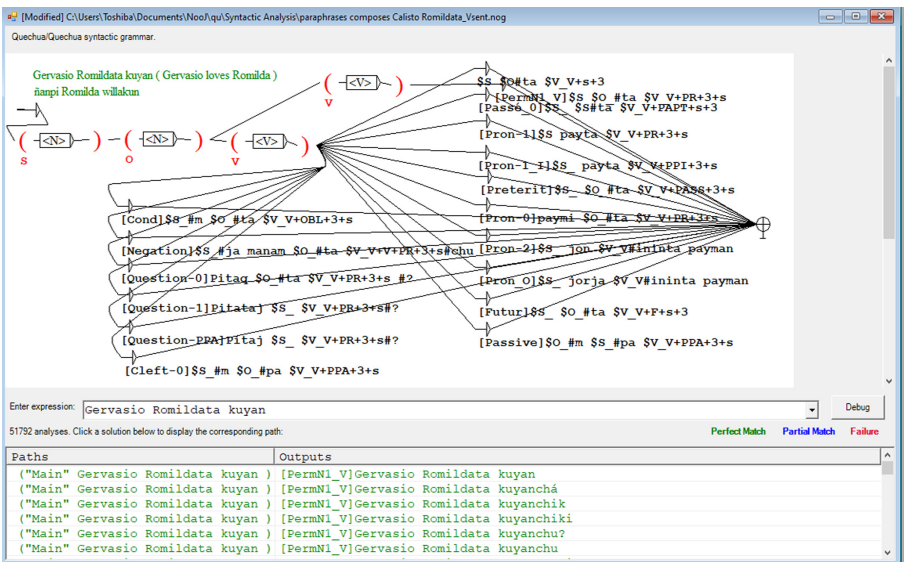


Fig. 9. Elementary paraphraser for an SOV sentence

The grammar of Fig. 9 concatenates these compositions. When applied, it generates 2,940 paraphrases like those appearing under the graph.

5 Conclusion

In this paper, I have shown several NooJ grammars capable of recognizing and producing a large number of sentences that are generated by transformations from an initial direct

sentence. In particular, I presented grammars corresponding to elementary transformations of any direct sentiment predicate like {*Gervasio Romildata kuyan*/Gervasio loves Romilda}. By composing elementary paraphrasers, I have built a complex paraphraser which generates a large number of paraphrases.

I plan to construct a more comprehensive set of transformations and paraphrasing grammars, which I hope will help me in the implementation of the resources for our machine translation project for Quechua.

References

1. Barreiro, A.: ParaMT: a paraphraser for machine translation. In: Teixeira, A., de Lima, V.L.S., de Oliveira, L.C., Quaresma, P. (eds.) PROPOR 2008. LNCS (LNAI), vol. 5190, pp. 202–211. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85980-2_21
2. Mota, C., Carvalho, P., Raposo, F., Barreiro, A.: Generating paraphrases of human intransitive adjective constructions with Port4NooJ. In: Okrut, T., Hetsevich, Y., Silberztein, M., Stanislavenka, H. (eds.) NooJ 2015. CCIS, vol. 607, pp. 107–122. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42471-2_10
3. Ben. A., Fehri, H. Ben, H.: Translating arabic relative clauses into english using the NooJ platform. In: Monti, J., Silberztein, M., Monteleone, M., di Buono, M.P. (eds.) Formalizing Natural Languages with NooJ 2014, pp. 166–174. Cambridge Scholars Publishing, Newcastle (2015)
4. Duran, M.: Morphology of MWU in Quechua. In: Proceedings of The 3rd Workshop on Multi-Word Units in Machine Translation and Translation Technology (MUMTTT 2017), pp. 32–42. Editions Tradulex, Geneva (2018)
5. Duran, M.: Dictionnaire électronique français-quechua des verbes pour le TAL. Thèse Doctorale. Université de Franche-Comté. Mars 2017 (2017)
6. Duran, M.: The annotation of compound suffixation structure of quechua verbs. In: Okrut, T., Hetsevich, Y., Silberztein, M., Stanislavenka, H. (eds.) NooJ 2015. CCIS, vol. 607, pp. 29–40. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42471-2_3
7. Duran, M., Morphological and syntactic grammars for recognition of verbal lemmas in Quechua. In: Proceedings of the 2014 NooJ International Conference and Workshop, Sassari (2014). Cambridge Scholars Publishing, Newcastle (2015)
8. Fehri, H., Haddar, K., Ben Hamadou, A.: Integration of a transliteration process into an automatic translation system for named entities from Arabic to French. In: Proceedings of the NooJ 2009 international Conference and Workshop, pp. 285–300. Centre de Publication Universitaire, Sfax (2010)
9. Harris, Z.: Mathematical Structures of Language. Interscience, New York (1968)
10. Langella, A.M.: Paraphrases for the Italian communication predicates. In: Barone, L., Monteleone, M., Silberztein, M. (eds.) NooJ 2016. CCIS, vol. 667, pp. 196–207. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-55002-2_17
11. Silberztein, M.: Automatic transformational analysis and generation. In: Gavriilidou, Z., Chatzipapa, E., Papadopoulou, L., Silberztein, M. (eds.) Proceedings of the NooJ 2010 International Conference and Workshop, pp. 221–231. University of Thrace, Komotini (2011)
12. Silberztein, M.: Language Formalization: The NooJ Approach. Wiley, Hoboken (2016)