Pragmatic Structures in Aymara

Petr Homola
Codesign, s.r.o.
Czech Republic
phomola@codesign.cz

Matt Coler
INCAS[3]
The Netherlands
MattColer@incas3.eu

Abstract

The paper describes overt marking of information structure in the indigenous Andean language Aymara. In this paper we show that although the word order is free, Aymara is not a discourse-configurational language (Kiss, 1994); instead, information structure is expressed only morphologically by pragmatic suffixes. The marking of 'topic' is more flexible that the marking of 'focus' (be it at the clause level or NP-internal). Since overt marking of information structure is partial, this paper also devotes considerable attention to the resolution of underspecification in Aymara.

1 Introduction

Chomsky's original approach to formal syntax assumes that sentences consist of constituents and that the type and order of these constituents define configurations that specify grammatical relations. As has been shown by Hale (1983), there are languages in which word order has no or limited relevance for grammatical relations and the respective constituent trees are flat. In most such languages, word order is said to specify information structure (hence the name discourse-configurational languages coined by Kiss (1994), as opposed to syntax-configurational languages). We show that Aymara is neither syntax-configurational nor discourse-configurational. In this language, information structure is expressed solely morphologically. This is very rare and therefore Aymara is very important for the research of information structure in general. As Bossong (2009) puts it (referring to Aymara and Quechua): "Cross-linguistically, grammemes explicitly expressing the function of theme are not very frequent; grammemes expressing the function of the rheme are a highly marked typological rarity [...]. Still more idiosyncratic is the combination of the thematic marker with a grammeme combining the two functions of *question* and *negation*."

We are not aware of any other language (except for, to some extent, Quechua) in which word order is irrelevant to information structure. Hardman (2000) says about Aymaran languages that "[t]he structural elements of a sentence may occur in any order and are at the disposal of the speaker for stylistic play." Hardman et al. (2001) say that word order in Aymara "affects only style, not the grammar nor basic semantics." Bossong (1989) says about Quechua (which is typologically very similar) that "word order is not only *free* but it is not primarily used as a means for expressing pragmatic functions as such." A detailed analysis of morphological information structure marking in Aymara can thus shed light on this module of grammar which is not expressed morphologically in most languages.

Section 2 presents basic facts about Aymara. Section 3 describes the overtly marked discourse categories in Aymara. Section 4 describes the means of morphosyntactic, semantic and pragmatic referent identification. Section 5 treats pragmatic marking in complex predicates. Finally, we conclude in Section 6.

2 Basic Facts about Aymara

Aymara is spoken by communities in a region encompassing Bolivia, Chile and Peru, extending north of Lake Titicaca to south of Lake Poopó, between the western limit of the Pacific coast valleys and eastward to the Yungas valleys. The language is spoken by roughly two million speakers, over half of whom are Bolivian. The rest reside in Peru and Chile. The Aymaran family (comprised of Aymara, Jaqaru, and Kawki) is a linguistic isolate with no close relative. Aymara is

an affixal polysynthetic[1] language (according to Mattissen's (2006) classification) with a rich morphology. It is SOV with modifier-head word order. Aside from the morphologically unmarked subject, all syntactic relations are case-marked, typically on the NP head. Roots can be divided into nouns (including qualitative words), verbs, and particles. Suffixes, which may have a morphological or syntactic effect, can be classified as nominal, verbal, transpositional, independent, or pragmatic suffixes. The category of independent suffixes is comprised of three suffixes which are not classifiable as members of either nominal or verbal morphology which likewise cannot be said to be pragmatic suffixes. These suffixes generally occur prior to inflectional morphology and/or the pragmatic suffixes. Pragmatic suffixes, by comparison, typically suffix to the last word of the entire sentence and/or the NP or VP. As their name suggests, their function is overall pragmatic in nature. Nominal and verbal suffixes can be subdivided into inflectional and derivational categories, but given the ease with which category-changing transpositional suffixes attach to words of any category, often multiple times, it is common to find words with several nominal, verbal, transpositional, and independent suffixes.

Detailed information can be found in (Hardman et al., 2001; Briggs, 1976; Adelaar, 2007; Cerrón-Palomino and Carvajal, 2009; Yapita and Van der Noordaa, 2008).

3 Overt Marking of Information Structure

As mentioned in the introduction, word order in Aymara is not used to express information structure. As an example, consider the following sentences in English (syntax-configurational), Russian (discourse-configurational) and Aymara (subscript **N** marks contextually new, i.e. nonpredictable information):

(1)
    (a) *Peter came*$_N$
        It's Peter$_N$ *who came* (*Came Peter*)
    (b) *Pëtr prišël*$_N$
        Prišël Pëtr$_N$
    (c) *Pedrox jutiwa*$_N$ or *Jutiw*$_N$ *Pedroxa*
        Pedrow$_N$ *jutixa* or *Jutix Pedrowa*$_N$

Due to overt morphological marking of information structure, there are discontinuous phrases in which the "gap" is not motivated by pragmatics, such as (2)[2]. The corresponding tree is given in Figure 1.

(2) *Juma-n-x*   *jiw-i-w*
     you-GEN-REF die-SMPL$_{3\to3}$-NPRED
     kimsa *ch'iyar* *phisi-ma-xa*
     three   black   cat-POSS2-REF
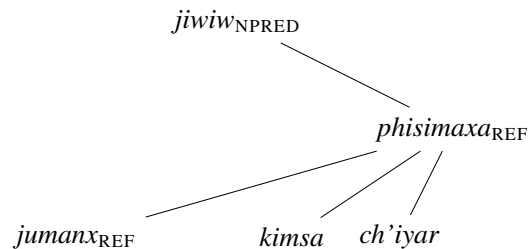   "Your three black cats died."

Figure 1: The surface syntax tree of (2) with a discontinuous noun phrase

Of course, the possessive pronoun *jumanx* could also form a continuous constituent with the rest of the NP. Or, it might be omitted since the possessiveness is already expressed with the suffix *-ma* (*phisi**ma**xa*).

Overt marking of information structure is obligatory in Aymara. A sentence without proper pragmatic suffixes (or with marking that would destroy intersentential cohesion) is considered ill-formed. It should also be noted that the marking of information structure is orthogonal to the morphosyntactic expression of evidentiality.[3]

3.1 Referentiality

Aymara has overt marking of referentiality. The notion of referentiality roughly corresponds to

---

[2]We use the following abbreviations in the glosses: ACC=accusative, AG=agentive, AGGR=aggregator, ALL=allative, BEN=benefactive, CAUS=causative, COM=comitative, CONJ=conjectural, DIR=directive, FUT=future tense, GEN=genitive, HON=honorific, HUM=human, IMPER=imperative, INCMPL=incompletive, INF=infinitive, LIM=limitative, LOC=locative, MIR=mirative, NEG=negative, NPRED=nonpredictable, PAST=past tense, PART=past participle, POSS=possessive, POSS1=1st person possessive, POSS2=2nd person possessive, POSS3=3rd person possessive, POT=potential, PROGR=progressive, QM=question marker, REF=referential, REFL=reflexive, SMPL=simple tense, TRGRDS=transgressive (different subject), TRGRSS=transgressive (same subject), VRBL=verbalizer

[3]Aymara has a three-way system of evidentiality: witnessed, knowledge through language, and inferred.

---

[1]See (Baker, 1996).

what Sgall et al. (1986) call 'contextual bound-edness'. We follow Sgall's (2009) terminology here: "[Contextually bound] items are presented by the speaker as **referring** to entities assumed to be easily accessible by the hearer(s), in the prototypical case 'given'. They refer to 'established' items, i.e. to those that were mentioned in the preceding co-text and thus still are sufficiently salient, or to the permanently established ones (indexicals, those given by the relevant culture or technical domain, etc.)."[4]

For the majority of bilingual speakers, referentiality is optionally marked with the suffix *-xa*,[5] however for monolinguals it is done so more consistently. There are also slight dialectal differences.

For example, in (3) the verb is marked as referential because the event is known from the context (the speaker supposes that the hearer knows that someone came, otherwise he would not ask).

(3) *Khiti-s   jut-i-xa?*
who-QM come-SMPL$_{3\rightarrow3}$-REF
"Who did come?"

On the other hand, in (5) the verb is not marked as referential because it expresses an event which is apparently unknown to the hearer (since he has expressed surprise in (4)).

(4) *Aka-n-ka-sk-ta-sä*
this-LOC-VRBL-PROGR-SMPL$_{2\rightarrow3}$-MIR
"Oh, you are here!"

(5) *Jïsa, wasüru-w*
yes   yesterday-NPRED
kutin-x-ta
return-COMPL-SMPL$_{1\rightarrow3}$
"Yes, I came back yesterday."

The adverbial expression *wasüruw* "yesterday" is marked as nonpredictable[6] (even though it is inherently referential) because it is more salient than the verb[7] (B is surprised to see A because

he thought that A had left or was about to leave). Nonetheless, the verb is nonpredictable, too, because B does not know what exactly happened (A may have come back earlier or may not have left at all).

Note that *-xa* marks referentiality, not known/given information, as illustrated in (6).

(6) *Qharüru-x   sara-:-wa*
tomorrow-REF go-FUT$_{1\rightarrow3}$-NPRED
"I will go tomorrow."

The adverbial expression *qharürux* "tomorrow" is marked as referential because it is inherently referential but there is no marking which would specify if it is predictable or not (the verb is already marked as nonpredictable and there can be at most one NPRED-marker per clause, as explained in the next subsection).[8]

It should be noted the the REF-marking is privative, i.e. the absence of *-xa* does not mean that the entity is not referential.

3.2 Nonpredictability

Entities are nonpredictable if they are 'new' to the hearer (cf. the usual dichotomy of information structure known/given vs. unknown/new).[9] Overt marking of nonpredictability in Aymara is much more consequent than that of referentiality. In affirmative sentences, the most salient nonpredictable entity is marked with the suffix *-wa*.[10] The placement of the NPRED-marker can be best explained by question/answer pairs, such as (7) and (8).

(7) *Khiti-taki-s   ut-x*
who-BEN-QM   house.ACC-REF
uta-ch-ta-xa
house-CAUS-SMPL$_{2\rightarrow3}$-REF
"For whom did you build the house?"

---

[4]We do not use the terminology 'topic-focus' or 'theme-rheme' because the paradigms of suffixes that express information structure do not correlate with it. For example, the two paradigms are not always mutually exclusive (two suffixes from different paradigms can mark the same word).

[5]In the literature this suffix is usually called "topic marker" or "attenuator".

[6]This grammatical feature is described in the next subsection.

[7]"More salient" means that it is more important in the current context. Informally, we could also say "less predictable". The scale of salience is important for intersentential referent identification.

[8]A reviewer suggested to compare pragmatic marking in Aymara with 'topic'-markers in Japanese and/or Korean. First of all, whereas in the mentioned languages, topic marking interferes with case marking, in Aymara they are orthogonal. Also, in Aymara the use of pragmatic markers does not affect word order. Finally, there are no free topics as in the following sentence from (Iwasaki, 2013):
Zoo-wa$_{TOP}$ *hana-ga*$_{NOM}$ *nagai*
"The elephant; its nose is long."
In Aymara, even if the word order of a NP is relaxed, all parts of the NP agree with each other.

[9]Again, we follow Sgall's (2009) terminology: "[Contextually non-bound] items are presented as **not** directly **predictable** in the given context, as 'new' information (at least as chosen from a set of available alternatives)."

[10]In the literature this suffix is often called "affirmative", "emphasizing", or "declarative".

(8) *Jilata-ja-taki-w*
    brother-POSS1-BEN-NPRED
    (uta-ch-t-xa)
    house-CAUS-SMPL$_{1\to3}$-REF
    "(I built it) for my brother."

In negative sentences, the NPRED-marker is usually attached to the negative particle *jani*, as in (9).

(9) *Jani-w   kullaka-ma-r*
    not-NPRED sister-POSS2-ALL
    uñj-k-t-ti
    see-INCMPL-SMPL$_{1\to3}$-NEG
    "I did not see your sister."

However, if an argument of the verb is more salient than the verb (i.e., it is contrastive), it attracts the NPRED-marker, as in (10).

(10) *Kullaka-ma-ru-w   jan*
     sister-POSS2-ALL-NPRED not
    uñj-k-t-ti
    see-INCMPL-SMPL$_{1\to3}$-NEG
    "It is your sister whom I did not see."

There are constructions in which NPRED-marks are disallowed, for example in imperative constructions or in those marked with the conjectural evidential suffix:

(11) *Chur-ita-ya*
     give-IMPER$_{2\to1}$-HON
     "Give it to me, please."

(12) *Jut-chi-ni*
     come-CONJ-FUT$_{3\to3}$
     "Maybe he will come."

NPRED-unmarked affirmative sentences in the future tense tend to have imperative meaning:

(13) *Sara-ñäni*
     go-FUT$_{4\to3}$
     "Let us go."

An affirmative sentence without NPRED-marking in the future tense does not alter the context of the present discourse and may have modal meaning (Hardman et al., 2001):

(14) *Nay sar-ä-xa*
     I   go-FUT$_{1\to3}$-REF
     "I will go, right? / Should I go?"

Noun phrases introduced by *mä* "one" are indefinite and therefore always nonpredictable. Noun phrases with the determiners *aka* "this", *uka* "that", *khaya* (*khä*) "yonder" and *khuri* "away yonder" are definite and therefore usually predictable unless they are explicitly NPRED-marked (in which case they are contrastive). The following example illustrates that referentiality and nonpredictability do not exclude each other.

(15) *Aka warmi-mpi-w   mä jisk'a*
     this woman-COM-NPRED one small
    marka-n   jiki-s-t-xa
    village-LOC meet-REFL-SMPL$_{1\to3}$-REF
    "It is this woman whom I met in a small village."

In verbal complexes, the main (nonfinite) verb may be NPRED-marked, as in (16) and (17):

(16) *T'ant'   ala-ñ-w*
     bread.ACC buy-INF-NPRED
    mun-ta
    want-SMPL$_{1\to3}$
    "I want to buy bread."

(17) *T'ant'   al-iri-w*
     bread.ACC buy-AG-NPRED
    sara-sk-ta
    go-PROGR-SMPL$_{1\to3}$
    "I am going to buy bread."

3.3 Deeply Embedded Nonpredictable Elements

Nonpredictability is usually marked at clause level but it can occur inside a noun phrase[11] if it is required by the discourse context, as in (18) (from (Hardman et al., 2001)).[12]

(18) *Naya-n-x   pusi-w*
     I-GEN-REF four-NPRED
    uta-ja-x   utj-itu
    house-POSS1-REF exist-SMPL$_{3\to1}$
    "As for houses, I have four."

This example shows that REF and NPRED markers need not appear at the end of a (semantic) phrase. Similarly, a nominal modifier may follow

---

[11]Hajičová et al. (1998) call this kind of information-structural marking *proxy focus*.

[12]In this example, *pusiw* is not contrastive, it is simply an answer to the question *How many houses do you have?* Of course, one could say just *Pusiwa*.

its governor if its information structure status differs from that of its head in which case both words receive pragmatic marking, as in (19) (from (Hardman et al., 2001)).

(19) *Uka-x*   *apilla-x*
    that-REF   oca-REF
    luxu-cha-ta-wa
    freeze-CAUS-PART-NPRED
    "That is FROZEN oca."

Aymara is a radically nonconfigurational language.[13] However, even languages from this class may pose word order constraints on "deeply" embedded phrases. Indeed, in Aymara noun phrases have more rigid word order than clause constituents.[14] For example, while (20) is well-formed, (21) is ill-formed because the modifier *suma* "good" does not precede its head.

(20) *Suma chuq'i-w*   *aka*
    good   potato-NPRED   this
    marka-n-x   *ach-u*
    village-LOC-REF   grow-SMPL$_{3\rightarrow3}$
    "In this village grow good potatoes."

(21) **Chuq'i-w*   *suma aka*
    potato-NPRED   good   this
    marka-n-x   *ach-u*
    village-LOC-REF   grow-SMPL$_{3\rightarrow3}$
    Intended meaning: "In this village grow good potatoes."

Nevertheless the order of a nominal modifier and its head is not restricted if both have a pragmatic marker, as in (22)–(24).

(22) *Suma-w*   *aka marka-n-x*
    good-NPRED   this   village-LOC-REF
    chuq'i-x   *ach-u*
    potato-REF   grow-SMPL$_{3\rightarrow3}$
    "GOOD potatoes grow in this village."

(23) *Uta-cha-ña-taki-x*   *qala-x*
    house-CAUS-INF-BEN-REF   stone-REF
    alluxa-w   *apthapi-ña*
    a-lot-NPRED   gather-INF
    "One must gather MANY stones to build a house."

(24) *Qullq-x*   *allux-w*
    money.ACC-REF   a-lot.ACC-NPRED
    ap-kat-ta
    put-DIR-SMPL$_{1\rightarrow3}$
    "I collected A LOT of money."

Thus Aymara, like many nonconfigurational languages, allows discontinuous constituents. For example, in (22), the noun phrase *sumaw chuq'ix* "good potatoes" is discontinuous because *aka markanx* "in this village" is not part of its surface syntax subtree.

3.4 Focalizers

"Focalizers" (i.e. focusing operators, in English adverbs such as *only, even, also, always, at least,* etc., see (Hajičová et al., 1998)) are mostly expressed by suffixes in Aymara. In many cases, the word with a focalizing suffix attracts the NPRED-marking, as in (25)–(28).

(25) *Kimsa-ni-ki-w*   *sar-i*
    three-HUM-LIM-NPRED   go-SMPL$_{3\rightarrow3}$
    "Only three went."

(26) *Wawa-pa-x*   *may-ni-ki-wa*
    child-POSS3-REF   one-HUM-LIM-NPRED
    "He has only one child."

(27) *Iki-ña-k-w*
    sleep-INF-LIM-NPRED
    mun-t-xa
    want-SMPL$_{1\rightarrow3}$-REF
    "I only want to sleep."

(28) *Nay kuna-w*   *sar-t-xa*
    I   what-NPRED   go-SMPL$_{1\rightarrow3}$-REF
    "Even I went."

Another pattern is attaching the independent aggregator *-sa* to the focalized word together with NPRED-marking of the verb, as in (29) and (30).

(29) *Naya-s*   *sara-rak-t-wa*
    I-AGGR   go-'also'-SMPL$_{1\rightarrow3}$-NPRED
    "I also went."

(30) *Juma-ki-s*   *yati-sma-wa*
    you-LIM-AGGR   know-POT$_{2\rightarrow3}$-NPRED
    "At least you should know it."

The focalizing negative particle *jani* "not" attracts the NPRED-marking in unmarked cases (but

---

[13]In terms of the generative grammar, languages with the rule S $\rightarrow \mathcal{C}^+$ at clause level.

[14]Unlike, for example, in Latin, where the word order of NPs is less rigid although not completely free.

see (10)). Two focalizers can be combined, as in the case of *jani* and *puni* "always" which gives the form *janipuni* "never" or *jani* and *raki* "also" which gives the form *janiraki* "neither". Both forms attract the NPRED-marking.

3.5 Resolving Underspecification

As was demonstrated in the above description, overt marking of information structure in Aymara is partial, as it leaves some elements underspecified. Since Aymara is a radical pro-drop language[15] (any argument of the verb may be unexpressed, not only the subject), predictable arguments are often omitted. As a consequence, direct objects are usually omitted if they can be inferred from the context, as in (31).[16]

(31) *Tata-ja-taki-w*
   father-POSS1-BEN-NPRED
   ala-sk-t-xa
   buy-PROGR-SMPL$_{1 \to 3}$-REF
   "I am buying it for my father."

It follows that if a verb is NPRED-marked and has an overt object, the latter is nonpredictable, too, unless it is REF-marked, as in (32).

(32) *Naya-w   um-x*
   I-NPRED   water.ACC-REF
   wayu-ni-:-xa
   carry-DIR-FUT$_{1 \to 3}$-REF
   "It is me who will bring the water."

The same holds for other verbal arguments. The only ambiguity arises from inherently referential expressions which appear REF-marked or unmarked (e.g., *qharürux/qharüru* "tomorrow") and this marking does not correlate with their nonpredictability.[17]

3.6 Intersentential Cohesion

Overt marking of information structure, described in the previous section, has crucial importance for intersentential cohesion and can be helpful for coreference resolution. Consider, for example, the following two sentences:

---

[15]See (Cole, 1987).

[16]The sentence is an answer to the question *Who are you buying it for?*

[17]It is possible that the category of referentiality is undergoing a reanalysis as a result of language contact. A more detailed analysis of the speech of monolingual speakers is needed.

(33) *Mä   marka-n   mä jisk'a imilla-w*
   one village-LOC   one small   girl-NPRED
   utj-i
   exist-SMPL$_{3 \to 3}$
   "There was a little girl in a village."

(34) *Jupa-x   Mariya suti-ni-wa*
   she-REF   María   name-POSS-NPRED
   "Her name was María."

In (33), both noun phrases as well as the verb are nonpredictable. The noun *imilla* "girl" is NPRED-marked (i.e., marked as the most salient part of the utterance) because in (34), it is referred to by the pronoun *jupa* "he/she".

The analysis of utterances in texts such as stories and narratives reveals that Aymara speakers consequently take intersentential cohesion into account when they decide where to place pragmatic markers.

4   Referent Identification

Surface and deep syntax as well as semantics operate on isolated sentences. Now we will discuss the formalization of pragmatics, the level of discourse context.

For the purposes of this subsection we assume that we have a discourse that consists of sentences $s_1, \ldots, s_n$ and that we have the corresponding feature structures $f_1, \ldots, f_n$. An entity we call a feature structure that represents a person, an object or an event (an event may be dynamic if described by a verb or statal if described by a nominal predicate). Every entity has a special attribute, INDEX, to represent coreferences.

The discourse context is formally a list of indices (values of the INDEX attribute) $C = \langle i_1, \ldots, i_m \rangle$. The sentences are processed one by one. At the beginning, $C = \emptyset$. For every f_i, we do the following:

1. For every entity in f_i, we try to find its referent in C (we describe below how referents are identified). If a referent was found for an entity, its index in C is moved to the beginning of the list. Otherwise, a new index is assigned to the entity and prepended to the list.

2. The index of the NPRED-marked entity is moved (or prepended) to the beginning of C.

There are various strategies of identification of referents in the preceding discourse that can combine to resolve ambiguities.

4.1 Morphosyntactic Referent Identification

Aymara has a rich system of switch-reference suffixes that help to identify referents in the discourse context. For example, the sentence

(35) *Tumasi-x$_i$*
  Thomas-REF
  klasi-n-ka-ska-:n-wa
  class-LOC-VRBL-PROGR-PAST$_{3\rightarrow3}$-NPREF
  "Thomas$_i$ was in the class room."

may be followed by the following sentence:

(36) *Yatichiri-x$_j$*   *manta-n-isin-x*
  teacher-REF   enter-DIR-TRGRSS-REF
  nuw-i-wa
  hit-SMPL$_{3\rightarrow3}$-NPRED
  "When the teacher$_j$ entered, he$_j$ hit him$_i$."

The first sentence adds the index of the entity *Tumasix$_i$* "Thomas" to C. The second sentence adds the index of the entity *Yatichirix$_j$* "the teacher" to C. Furthermore, there are two unresolved (covert) pronouns that represent the actor and the patient of *nuwiwa* "to hit". In this case the switch-reference transgressive suffix *-sina* specifies that the actor of *nuwiwa* is the actor of *mantasinx* "entered", thus the actor of *nuwiwa* has the index i. The patient is coindexed with the next entity in C with which it agrees morphologically (e.g., in terms of animacy) or semantically. On the other hand, the sentence

(37) *Yatichiri-x$_j$*   *manta-n-ipan-x*
  teacher-REF   enter-DIR-TRGRDS-REF
  nuw-i-wa
  hit-SMPL$_{3\rightarrow3}$-NPRED
  "When the teacher$_j$ entered, he$_i$ hit him$_j$."

changes the indexes of the pronouns in the matrix sentence because the switch-reference transgressive suffix *-ipana* specifies that the subject of *mantanipanx* is different from that of *nuwiwa*.

4.2 Semantic Referent Identification

If there is the sentence

(38) *Tumasi-mp$_i$*   *Marya-mp$_j$*
  Thomas-COM   Mary-COM
  uñj-t-wa
  see-SMPL$_{1\rightarrow3}$-NPRED
  "I saw Thomas$_i$ and Mary$_j$."

followed by

(39) *Jani-w*   *usuri-:-ta-p*
  not-NPRED   pregnant-VRBL-PART-POSS3
  yat-k-t-ti
  know-INCMPL-SMPL$_{1\rightarrow3}$-NEG
  "I did not now that she$_j$ was pregnant."

the referent identification of the covert pronoun which is the actor of *usurïtap* "that she was pregnant" is not morphosyntactically restricted (Aymara has one pronoun, *jupa*, for both "he" and "she") but it is semantically restricted. It is obvious that the semantic information of this kind has to come from the lexicon. Likewise, there will be lexically encoded semantic gender for words such as *tayka* "mother", *jilata* "brother", *imilla* "girl", etc.

4.3 Pragmatic Referent Identification

As described in Subsection 3.6, explicitly NPRED-marked entities are more salient than other non-predictable entities in the same sentence. If there is the sentence

(40) *Tayka-ma-w$_i$*
  mother-POSS2-NPRED
  yatichiri-ma-mp$_j$ *jik-i-si*
  teacher-POSS2-COM   meet-SMPL$_{3\rightarrow3}$-REFL
  "Your mother$_i$ met your teacher$_j$."

followed by

(41) *Usuta-:-ta-m-x*
  sick-VRBL-PART-POSS2-REF
  yat-x-i-wa
  know-COMPL-SMPL$_{3\rightarrow3}$-NPRED
  "She$_i$ already knew that you were sick."

the actor of *yatxiwa* "s/he already knew" is coindexed with *taykamaw* "your mother" because this entity is NPRED-marked in the first sentence and therefore more salient (it precedes other entities in C). If we move the NPRED-marker in the first sentence to the "teacher", the meaning of the second sentence will change:

(42) *Tayka-ma-x$_i$*
  mother-POSS2-REF
  yatichiri-ma-mpi-w$_j$
  teacher-POSS2-COM-NPRED
  jik-i-si
  meet-SMPL$_{3\rightarrow3}$-REFL
  "Your mother$_i$ met your teacher$_j$."

(43) *Usuta-:-ta-m-x*
    sick-VRBL-PART-POSS2-REF
    yat-x-i-wa
    know-COMPL-SMPL$_{3\to3}$-NPRED
    "He$_j$ already knew that you were sick."

5 Complex Predicates

A special case of pragmatic marking represent the so-called complex predicates, i.e. monoclausal predicates composed of (at least) two predicative elements. In such constructions, a (fully) semantic verb combines with a modal or auxiliary element to express complex predication such as causative, volitive, supine, etc. The concept of complex predicates, elaborated by Alsina (1996), has been applied to a number of phenomena and languages including, for example, Hindi light verbs (Mohanan, 1994) or Turkish causatives (Çetinoğlu et al., 2008). We will leave aside the rather complicated formal treatment of these constructions in unification-based grammars (Homola and Coler, 2013) and focus on the linguistic description of these form in Aymara from the pragmatic point of view here.

Sentences (16) and (17) are examples of complex predicates. The constructions contain only one NPRED-marker, i.e. they are monoclausal and should be represented by a single feature structure with a complex functor. For (16), we get:

(44) $\begin{bmatrix} \text{FUNC} & \text{'muna-'} \\ \text{ARGS} & \left\langle \text{subj}_{act}, \begin{bmatrix} \text{FUNC} & \text{'ala-'} \\ \text{ARGS} & \langle \text{subj}_{act}, \\ & \text{dobj}_{pat}\rangle \end{bmatrix}_{pat} \right\rangle \end{bmatrix}$

For (17), we get:

(45) $\begin{bmatrix} \text{FUNC} & \text{'sara-'} \\ \text{ARGS} & \left\langle \text{subj}_{act}, \begin{bmatrix} \text{FUNC} & \text{'ala-'} \\ \text{ARGS} & \langle \text{subj}_{act}, \\ & \text{dobj}_{pat}\rangle \end{bmatrix}_{pat} \right\rangle \end{bmatrix}$

In a more concise notation, (44) can be written as:

(46) $muna(\text{subj}_{act}, ala(\text{subj}_{act}, \text{dobj}_{pat})_{pat})$

Likewise, (45) can be written as:

(47) $sara(\text{subj}_{act}, ala(\text{subj}_{act}, \text{dobj}_{pat})_{pat})$

Unlike in languages with morphologically formed volitive verbal complexes, such as

Guaraní, in Aymara such constructions are formed syntactically (i.e. the complex predicate value is not created in the lexicon). As for motion verbs, such as (45), there is a morphological alternative:

(48) *Jichha-x   t'ant*
    now-REF   bread.ACC
    ala-ni-rapi-:ma-wa
    buy-DIR-BEN-FUT$_{1\to2}$-NPRED
    "I will go to buy bread for you now."

Further evidence that such predicates are monoclausal is the fact that polypersonal agreement is expressed on the auxiliary or modal verb, not the full one, as in (49) and (50).

(49) *Ch'uq   alja-ñ-w*
    potato.ACC sell-INF-NPRED
    mun-sma
    want-SMPL$_{1\to2}$
    "I want to sell potatoes to you."

(50) *Tump-iri-w   jut-sma*
    visit-AG-NPRED come-SMPL$_{1\to2}$
    "I came to visit you."

It is noteworthy that a complex predicate with a motion verb can have two complements, namely a locative phrase and a verbal complement:

(51) *Al-iri-w   Chukiaw*
    buy-AG-NPRED La.Paz.ACC
    sara-:na
    go-SMPL$_{3\to3}$
    "He went to La Paz to buy it."

Even more evidence for monoclausality can be found in sentences like (52):

(52) *Chukiaw   sara-ñ-w*
    La.Paz.ACC go-INF-NPRED
    mun-t-x   irnaqa-ña-taki-xa
    want-SMPL$_{1\to3}$-REF work-INF-BEN-REF
    "I want to go to La Paz for work."

In (52), the nominalized verb *irnaqañatakixa* "to work" depends on *sarañw* "to go" creating a long-distance dependency. Without considering the verbal complex *sarañw muntx* a monoclausal predicate, it would be linguistically counterintuitive and computationally hard to parse the sentence.

It is obvious from the mentioned examples that the NPRED-marker tends to attach to the full

verb. This fact supports the hypothesis that modal (deictic) syntactic elements should be considered synsemantic, i.e. at the level of deep syntax they should be represented as attributes of the nodes of their heads rather then autonomous nodes or feature structures. This approach, adopted by us, is consistent with (Sgall et al., 1986).

We omit from the discussion morphologically built complex predicates, such as causatives (e.g. (53)). In other languages where they are expressed syntactically, they would be treated in the same way as the constructions described above.

(53) *Jacha-y-t-wa*
cry-CAUS-SMPL$_{1\rightarrow3}$-NPRED
"I made him/her cry."

(54) $\begin{bmatrix} \text{FUNC} & \text{'caus'} \\ \text{ARGS} & \left\langle \text{subj}_{act}, \begin{bmatrix} \text{FUNC} & \text{'jacha-'} \\ \text{ARGS} & \langle \text{dobj}_{act} \rangle \end{bmatrix}^{pat} \right\rangle \end{bmatrix}$

6 Conclusions

We have described the overt marking of information structure and its pivotal role both in intersentential cohesion in Aymara as well as coreference resolution. Through an analysis of morphological information structure marking with *pragmatic suffixes* in this language, we illustrate the irrelevance of word order for information structure. While overt marking of referentiality is not always consequent, nonpredictability is marked with strict regularity. Having demonstrated that nonpredictability can be marked both at clause level and inside a NP and that focalizing suffixes tend to attract NPRED morphology, we explained how underspecification is resolved, noting, for example, that the overt object of a NPRED-marked verb is also unpredictable unless it is REF-marked. We also treated the identification of morphosyntactic, semantic and pragmatic referents through a formalization of pragmatics. The specifics of pragmatic marking in complex predicates show how the NPRED-marker typically attaches to the full verb thus supporting the hypothesis that modal syntactic constructions should be considered monoclausal.

References

Willem Adelaar. 2007. *The Languages of the Andes*. Cambridge University Press.

Alex Alsina. 1996. *The Role of Argument Structure in Grammar: Evidence from Romance*. CSLI Publications.

Mark C. Baker. 1996. *The Polysynthesis Parameter*. Oxford University Press.

Georg Bossong. 1989. Morphemic marking of topic and focus. In Michel Kefer and Johan van der Auwera, editors, *Universals of Language*, pages 27–51.

Georg Bossong. 2009. Divergence, convergence, contact. In Kurt Braunmüller and Juliane House, editors, *Convergence and divergence in language contact situations*, pages 13–40.

Lucy Therina Briggs. 1976. *Dialectal Variation in the Aymara Language of Bolivia and Peru*. Ph.D. thesis, University of Florida.

R. Cerrón-Palomino and J. Carvajal Carvajal. 2009. Aimara. In M. Crevels and P. Muysken, editors, *Lenguas de Bolivia*. Plural Editores, La Paz, Bolivia.

Özlem Çetinoğlu, Miriam Butt, and Kemal Oflazer. 2008. Mono/bi-clausality of Turkish Causatives. In *Proceedings of the 14th International Conference on Turkish Linguistics*.

Peter Cole. 1987. Null Objects in Universal Grammar. *Linguistic Inquiry*, 18(4):597–612.

Eva Hajičová, Barbara H. Partee, and Petr Sgall. 1998. *Topic-focus articulation, tripartite structures, and semantic content*. Kluwer Academic Publishers, Dordrecht.

Kenneth L. Hale. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory*, 1:5–47.

Martha Hardman, Juana Vásquez, and Juan de Dios Yapita. 2001. *Aymara. Compendio de estructura fonológica y gramatical*. Instituto de Lengua y Cultura Aymara.

Martha Hardman. 2000. *Jaqaru*. LINCOM EUROPA.

Petr Homola and Matt Coler. 2013. Causatives as Complex Predicates without the Restriction Operator. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the LFG Conference*.

Shoichi Iwasaki. 2013. *Japanese*. John Benjamins Publishing.

Katalin É. Kiss. 1994. *Discourse Configurational Languages*. Oxford University Press.

Joanna Mattissen. 2006. On the Ontology and Diachrony of Polysynthesis. In Dieter Wunderlich, editor, *Advances in the theory of the lexicon*, pages 287–354. Walter de Gruyter, Berlin.

Tara Mohanan. 1994. *Argument Structure in Hindi*. CSLI Publications.

Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. D. Reider Publishing Company.

Petr Sgall. 2009. Where to look for the fundamentals of language. *Linguistica Pragensia*, 19(1):1–35.

Juan de Dios Yapita and J.T. Van der Noordaa. 2008. *La dinámica aymara. Conjugación de verbos*. Instituto de Lengua y Cultura Aymara.