Matsigenka's Mystery Morphemes

1. Introduction

Matsigenka is a language spoken at the edge of the Amazon river basin. Like most languages indigenous to the Americas, it has been studied by only a handful of linguists and much about it remains unknown. Of particular interest are two morphemes in the language, *-ak* and a null suffix (written -∅), which form a system for encoding some type of temporal information. However, the exact semantics of this system remain unknown. To date these morphemes have been analyzed as being a part of an aspectual system, but recent scholarship has shed new light suggested that these morphemes may actually form a temporal definiteness system. Thus in this paper I will attempt to address these conflicting analyses, and situate my finding within the broader scope of cross linguistic variation. The rest of this paper is organized as follows: §2 gives a brief overview of pertinent features of the language, §3 reviews methodology, §4 gives an overview of the linguistic category of aspect, and discusses aspect as an analysis for the semantics of *-ak* and -∅, §5 gives an overview of the linguistic category of temporal definiteness, and discusses temporal definiteness as an analysis for the semantics of *-ak* and -∅, and §6 closes the paper with a conclusion and future directions.

2. The Matsigenka Language

The Matsigenka language is the language of the Matsigenka people, an indigenous group located in what is now southeastern Peru. Today there are roughly 10,000 Matsigenkas living scattered about multiple river basins in the region. Their language is a member of the Kampan branch of the Arawakan language family, and much like the other members of this widespread language family, Matsigenka is typologically distinct from many of the world's more studied languages.

A prime examples of how different Matsigenka is from other languages is the the way in which it encodes time. The primary means of encoding temporality in Matsigenka is a system called reality status, which makes a distinction between events that have been realized versus those that have not (Michael 2014). However, like most languages, Matsigenka encodes more than one type of temporal information in its sentences. Another notable paradigm for encoding temporal information in Matsigenka includes the suffix *-ak* and the null suffix -∅, and is the one this paper concerns itself with. The morphological position of these suffixes is exemplified below in examples (1a) and (1b)[1]. In the form in (1a), *-ak* occupies the penultimate suffixal position. This form can be contrasted with the form given on right, which lacks *-ak*, and instead exhibits -∅ posited in the penultimate position.

(1a)	ipegaka				(1b)	ipega			
	i-	peg	-ak	-a		i-	peg	-∅	-a
	3mS-	transform	-???	-real.a		3mS-	transform	-???	-real.a

[1] The glossings '3ms' and 'real.a' refer to third person masculine subject marking and a-class reality status respectively

As stated earlier, these suffixes have been the object of previous analysis. Previous work has analyzed them as being a part of an aspectual system, with *-ak* marking the perfect aspect, and -∅ marking imperfective aspect. However, recent work on the language has suggested that this may not be the case and has suggested that they may actually form a temporal definiteness system, with *-ak* marking temporal definiteness and -∅ marking temporal indefiniteness.

3. Methodology

To investigate the semantics of *-ak* and -∅, my source of language data was a corpus of 170 translated Matsigenka texts (Michael, Beier, and O'Hagan 2013), parsed and glossed using FLEx (Fieldworks Language Explorer), the standad lexical and text corpus database used by field linguists. I utilized the corpus to first analyze *-ak* and -∅'s usages on a case by case basis, searching for near minimal pairs and seeking to identify systematic differences in there meanings by examining associated translations. While doing this I also investigated *-ak* and -∅'s distributions on a constructional level to compare with expected realizations. Once this was done, I tested the hypothesis that these two morphemes encode a perfective/imperfective contrast by attempting to apply a Klein's 1994 analysis of aspect to the data. I also tested the temporal definiteness theory by comparing the distribution of *-ak* and -∅ the distributions of temporal definiteness morphology in two other languages, Mohawk and Koro.

4. Aspect

Smith (1997) and Klein (1994) propose that aspect is a system for encoding temporal relations between the window of time through which we are viewing an event, which linguists refer to as the 'topic time' or TT, and the totality of the event we are concerning ourselves with, referred to as the 'situation time', or TSit. Let's take a closer look at this with the two aspectual categories argued to be in Matsigenka.

The first of these is the perfective viewpoint, which is typically described as showing an event as finished, closed, or in full view. Consider the following English sentence:

(2a) When I arrived, the man ate.

Here the the TT is explicitly given by the phrase *when I arrived* (this can be the exact moment of arrival or longer depending on context), and the TS (namely, the man's eating) is entirely within the view of, or fully contained in, the TT. This can be contrasted with the imperfective viewpoint, which typically depicts events as ongoing, open, or in partial view. Consider the following English sentence:

(2b) When I arrived, the man was eating.

Again, the TT is explicitly given by the same phrase, but this time the TS is only partially in view of, or only partially overlaps with, the TT. This is to say, the man is presupposed to have been eating before I arrived, and continued at least momentarily after my arrival. The statement t*he man was eating* thus expresses that the man was in the midst of eating at the time that I

arrived.

As mentioned earlier, prior analyses have posited that *-ak* encodes for perfective (complete) aspect, while null-marked forms represent imperfective (incomplete, ongoing) aspect. However, upon closer inspection of the corpus, this does not seem to always seem to be the case. There are a significant number of instances of *-ak* occurring on seemingly imperfective aspectual cases, as well as cases of -∅ occurring on seemingly perfective aspectual cases.

Due to the many examples of these apparent aspectual contradictions, it is tempting to discard the idea that *-ak* and -∅ express a perfective/imperfective contrast, but there are a few things that make me hesitant to do so. First, while sifting through the corpus, I noticed that a significant number of contradictory examples contain another morpheme, the subordinator *=ra*. Currently there is only a general understanding of *=ra*'s functions, but it is clear that it many cases it can be used to order clauses with respect to time, and thus possible that it could interact with and impose restrictions on the functions aspectual realizations in the language.

Related to this is another issue, and that is the nature of working with corpus data. A vast majority of the texts in the corpus are narratives, and ascertaining certainty regarding the TT in narrative discourse can often be challenging. Narratives rarely include utterances with TTs as neatly and/or explicitly defined as in the above examples. Thus, though it seems problematic to say that *-ak* and-∅ form a perfective/imperfective contrast, further work must be done to rule this or the possibility of a more nuanced aspectual distinction out.

5. Temporal Definiteness

Another suggested analysis for *-ak*/-∅ semantics is that they form some sort of temporal definiteness system. A temporal definiteness system can be thought of as being analogous to the English articles 'the' and 'a', but for verbs instead of nouns. For example,temporal definiteness, would correspond to unique events, as in (3a).

(3a) Yesterday, I went to the store.

On the other hand, Temporal indefiniteness would correspond to generic events, or ones that occur at multiple, unspecified times over a temporal interval, as in (3b).

(3b) I used to go to the store every Sunday.

Here, multiple events of going to the store are being referred to in a generic sense, and they are not anchored in specific moments in time.

The literature suggests that realizations of temporal definiteness systems are quite rare, and indeed few languages have been described as having them. Two good examples however, are Mohawk, a language indigenous to North America, and Koro, a language indigenous to Oceania. Mohawk displays a contrast between both definite and indefinite contrasts, marking both overtly (Baker, Travis 1997). Koro however, only marks indefinite contexts overtly, and it is unclear whether unmarked forms are truly definite, or simply non-indefinite (Cleary-Kemp 2015).

Recent work has suggested that *-ak* may be a marker for temporal definiteness, while -∅ is a marker for temporal indefiniteness. Outlined in table 1. are the distributions of *-ak* and -∅

with respect to a few key contexts in both Mohawk and Koro. Due to Koro only marking temporal indefiniteness overtly, only canonical temporally indefinite contexts have been selected.

Table 1. Comparison of Temporally Indefinite Contexts

Context	Koro	Mohawk	Matsigenka
Future	Indefinite	Indefinite	-ak
If & Whenever	Indefinite	Indefinite	mixed
Negation	Indefinite	-	-∅
Past Habitual	Indefinite	Indefinite	mixed

Assuming that *-ak* marks definite and -∅ marks indefinite, we would expect the Matsigenka forms in the right column to surface with only -∅ in each of these contexts. However, this is not the case. Additionally, due the mixed distribution of *-ak* and -∅ in these contexts, we can also rule the possibility that the reverse analysis is true, as we would expect consistent marking throughout the column. Of course it is true that not every language across the globe perfectly aligns to theoretical realizations of any grammatical category, but even accounting for this we would still expect to see a single form in most of these contexts. And so due to this, it is possible to rule out the possibility of *-ak* and -∅ being a part of a temporal definiteness system.

6. Conclusion

In this paper I considered two possible analyses for the semantics of two Matsigenka suffixes, *-ak* and -∅. Due to the distributions of *-ak* and -∅ on a constructional level, it is quite clear that they they do not express a temporal definiteness contrast. Likewise, it seems unlikely that they express simple perfective/imperfective contrast, though this analysis cannot be fully ruled out at this point. Thus my next steps will be to test this analysis further by working with a consultant to mitigate uncertainty regarding topic time, as well as to investigate further into the functions of *=ra*. Additionally, my research so far has left me with a near complete description of the distributions of *-ak* and -∅, so I will continue to search through the literature for other systems with expected realizations close to these as possible leads to follow.

Works Cited

Baker, Mark; Travis, Lisa. (1997). Mood as Verbal Definiteness in a "Tenseless" Language. Natural Language Semantics 5: 213-269.

Cleary-Kemp, Jessica. (2015). Serial Verb Constructions Revisited: A Case Study From Koro. PhD dissertation, University of California Berkeley.

Klein, Wolfgang. (1994). Time in Language. Psychology press, London.

Michael, Lev. (2014). The Nanti reality status system: Implications for the typological validity of the realis/irrealis contrast. Linguistic Typology, 18(2). doi: 10.1515/lingty-2014-0011. UC Berkeley: Retrieved from: http://escholarship.org/uc/item/9n3250df

Michael, Lev, Christine Beier, and Zachary O'Hagan (compilers), Haroldo Vargas and José Vargas (authors). (2013). Matsigenka text corpus (version June 2013; FLEx database and LaTeX interlinear output).

Smith, Carlota. (1991). The Parameter of Aspect. Springer Science and Business Media, Berlin.