Narrative Time in Imperfectiveless Languages: Case studies in Three Amazonian Languages*

Andrés Pablo Salanova and Adam Tallman

University of Ottawa (Canada) and Friedrich Schiller Universität (Jena, Germany)

1. A puzzle and a partial solution

It is standard to assume that aspectual viewpoint consists essentially of a relationship between a situation's runtime (t_{sit}) and a topic time (t_{top}), itself related to the utterance time (t_0) by tense. The two most commonly discussed values of viewpoint, namely perfective and imperfective, more specifically boil down to inclusion relations between topic time and situation time: $t_{sit} \subset t_{top}$ for perfective, and $t_{sit} \supset t_{top}$ for imperfective. A number of properties that are associated with perfective and imperfective are derived from this basic characterization. The imperfective's stativity (i.e., its homogeneity throughout topic time), for example, would be responsible for the fact that, contrary to the perfective, it doesn't advance narrative time. This, in turn, would be responsible for the fact that imperfective (among other viewpoints) is used to introduce background information in a narrative.

Identifying viewpoint aspect unequivocally in less-studied languages is not a trivial task. Take for instance the construction exemplified by (1) in Mẽbêngôkre, a Jê language from Central Brazil:

(1) ba i-kabẽn o= dja
1NOM 1-speak with stand.SG.V
"I'm speaking." (lit., "I'm standing with my speaking.")

The construction in question has all the morphosyntactic characteristics of a progressive construction: a stative auxiliary (in this case a postural verb) subordinating the main verb by means of an element that conveys simultaneity ("with"). The construction in fact

*We would like to thank audiences at SULA 11 and at the second Toronto-Ottawa-Montréal Workshop on the Indigenous Languages of the Americas (TOMILLA in Montréal, December 2018) for helpful discussion, as well as to SSHRC, which through Insight grant 435-2016-1375 to Ana Arregui (PI), Mar[Pleaseinsert"PrerenderUnicode–˝intopreamble]a Luisa Rivero and Andr[Pleaseinsert"PrerenderUnicode–˝intopreamble]s Pablo Salanova facilitated the development of this project.

has an ongoing meaning most of the time, and there is no question in our minds that (1) is progressive in the appropriate contexts. Puzzlingly, however, one also finds this construction in clauses that clearly advance narrative time. The following is an example:

(2) Wabi, nhỹm we kum kabẽn o= dja nhỹm . . .
go_up DS HS 3DAT 3.speak with= stand.V DS
"He went up, and the other spoke to him (lit., was speaking with him), and..."

In this case, the "progressive" clearly makes a contribution: rather than a punctual event, the clause containing *o=dja* describes something durative (i.e., "spent a little while speaking with him"); and yet the clause doesn't have the expected behavior of a progressive clause: narrative time advances such that the episode that follows is wholly after the talking ends.

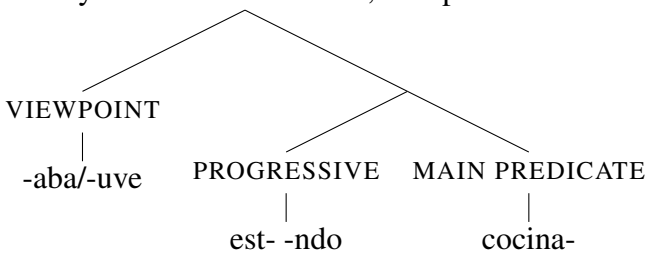
What is happening in such examples might be obvious to some readers. Progressives are generally considered a subtype of imperfectives (e.g., by Comrie 1976), but in fact their relation to imperfective viewpoint is indirect. The separation between progressivity and imperfectivity can be seen in languages that have both, such as Spanish (see Altshuler 2014, Arche 2014 for discussion):

(3) a. Estaba leyendo cuando llamaste.
be.IPF_PST.1SG reading when call.PFV_PST.2SG
"I was reading when you called."

b. Estuve leyendo (un rato), pero me aburrí.
be.PFV_PST.1SG reading a while but 1SG get_bored.PFV_PST.1SG
"I read for a little while, but I got bored."

The English aspectual auxiliary, on the other hand, may only have an imperfective reading. The Mẽbêngôkre progressive construction would illustrate a third possibility, where viewpoint is not morphologically specified but may be assigned to perfective or imperfective according to context (for a similar claim about Guaraní bare verb clauses, see Thomas 2014).

The morphology of the progressive construction in Spanish shows quite clearly that aspect (fused with tense) on the auxiliary encodes viewpoint, and that the complement of viewpoint encodes durativity and non-culmination, as represented in the following tree:



If this is the obvious solution, then what is the puzzle that remains? In fact, perfective progressives have contradictory properties. On the one hand, they are subject to the "imperfective paradox":

(4) Estuve preparando la ponencia pero me distraje y no terminé.
"I was.PFV preparing the talk but I got distracted and didn't finish."

On the other, they are clearly bounded:

(5) a. Recién estaba cocinando, debe estar en eso todavía.
"He was.IPF cooking, he should still be at it."

b. Recién estuvo cocinando, #debe estar en eso todavía.
"He was.PFV cooking, #he should still be at it."

Furthermore, they advance narrative time, as in (3b).[1]

So even if we are clear that viewpoint aspect boils down to $t_{sit} \supset t_{top}$ vs. $t_{sit} \subset t_{top}$, the distinction is, in practice, quite hard do intuit, and identifying the viewpoint of a particular construction in the field is not an trivial task, given the contradictory properties that we have examined. The puzzle then becomes a practical one: can we find an cross-linguistically translatable diagnostic trait that corresponds can be used to reliably identify (im)perfective aspect?

As tempted as we are to identify such trait as the property of narrative time advancement (which characterizes the perfective, but whose absence is characteristic not only of imperfectives but of a number of other aspectual values, such as perfects and prospectives), the approach that we will take here is less axiomatic and more exploratory: we identify a number of codable constructions or properties of clauses that we conjecture are causally related to the (non)advancement of narrative time - not in the sense that they entail a prefective or imperfective reading, but because such a reading is often implied by their presence. Speakers might infer that, *ceterus paribus*, the progressive stops narrative time advancement and the completive advances it. We would like to see what hypotheses of this general form can be motivated based on naturalistic speech.

2. Operationalizing viewpoint

Figure 1: Languages of our study.

Our database consists of clauses extracted from a collection of texts, diverse as possible as far as genre is concerned, from the following three languages: Araona (Takanan, Bolivia; ISO 639-3 code: aro), 150 speakers; Chácobo (Panoan, Bolivia; ISO 639-3 code: cao), 1200 speakers; Mẽbêngôkre (Macro-Jê, Brazil; ISO 639-3 code: txu), 13.500 speakers. Locations are shown on figure 1.

In this talk we examine viewpoint aspect in these three languages using a quantitative methodology. The key points that characterize our enquiry are the following:

[1] In Spanish, it would seem that perfective progressives advance narrative time less resolutely than a plain perfective construction, and are hence more often than not accompanied by linkers such as "y después" (and later) or adverbials such as "un rato" (for a while). Still, perfective progressives cannot be used in contexts where simultaneity with a second situation is explicitly sought. Since we do not discuss Spanish at any length here, we leave discussion of these facts for another occasion.

- The clearest way to operationalize viewpoint aspect is to consider it to be directly related to narrative time (non)advancement.
- There are languages with no dedicated markers for viewpoint aspect, but they still express the narrative time (non)advancement by default interpretations associated with certain constructions or combinations of grammatical elements.
- Typological variation in the (non)advancement of narrative time can be usefully modeled as a classification problem over such constructions treated as variables in naturalistic speech.

2.1 Viewpoint and narrative time advancement

We operationalize the perfective/imperfective distinction as one between advancing and not advancing narrative time in narrative discourse. Not advancing narrative time is a property that imperfectives share with the perfect:

IMPERFECTIVE	does not advance narrative time
PERFECTIVE	advances narrative time

Before moving on we briefly put the languages under study in typological context with regard to the expression of viewpoint aspect. Some languages have inflectional morphology that encodes whether a predicate advances narrative time or not, e.g., Spanish. In others, whether a predicate advances narrative time or not is determined by a number of factors, e.g., "jointly determined by the lexical aspect of the verb, context and Aktionsart" (Tonhauser 2011:265, n. 6). An important empirical point about the languages under study is that none of these elements (plus others, like particles) can be deterministically associated with viewpoint aspect.

As stated above, Chácobo, Mẽbêngôkre and Araona lack dedicated markers of viewpoint aspect, even though they have constructions not unlike the Mẽbêngôkre progressive discussed above. By "dedicated markers" we mean those that always encode a (or entail) a specific aspectual value.

On the other hand, the languages don't all lack dedicated markers of tense in our sense. Mẽbêngôkre displays a future vs. nonfuture distinction. Chácobo displays a nonpast versus past distinction and Araona displays a past vs. unspecified distinction. We do not apply our research strategy to tense categories in these languages, because these have dedicated markers and hence no classification problem arises. We suggest that the same research strategy applied here could be used to investigate the marking of tense relations in tenseless languages.

For each of the three languages, we created a small corpus based on approximately an hour of naturalistic speech transcribed, translated and annotated in the languages of the study. The sentences of our corpus come from different types of narrative discourse (traditional mythical narrative, personal narrative, procedural narrative): about half of this is dialogical. For each clause, we coded whether it advances narrative time or not. Each language has approximately 600 data points corresponding to 600 predicates which do or

do not advance narrative time. The data can be made available to any linguist upon request to the authors.

In order to investigate our research question we coded for a number of traits of each sentence that might influence viewpoint interpretation (see Figure 2). Unfortunately constraints of space prevent us from being able to describe all of the variables in detail. We restrict ourselves to a brief overview where we focus on making the case that none of the variables involve 'dedicated' markers of imperfectivity in our sense. Below we review those factors which we conjecture to be the most important.

2.2 Broad classes of independent variables

The list of independent variables we coded for in this study are provided in Table 2.

Figure 2: . Narrative time (non)advancement variables coded in naturalistic speech

	Type
Narrative time advancement	Binary (yes,no)
Verbal vs. nonverbal predication	Binary (yes,no)
Nominal predication	Binary (yes, no)
Quoted speech	Binary (yes, no)
Back channel	Binary (yes, no)
Dependent	Binary (yes, no)
Reduplicative pluractional	Binary (yes, no)
Posture main verb	Binary (yes, no)
Finiteness	Binary (yes, no)
Clause-type	Categorial
Conjunction before	Categorial
Conjunction after	Categorial
Temporal distance	Categorial
Associated motion	Categorial
Clause-type morpheme	Categorial
Temporal frame adverbs	Categorial
Frequency adverb	Categorial
Postural modification	Categorial

There are two types of variables: binary and categorial. A binary variable is one where the feature is either present or it is not. Categorial variables are n-valued. They contain specific morphemes or constructions, but they do not answer yes or no questions. For instance, Chácobo has a large number of 'clause-type morphemes' and the value for this variable is simply whatever clause-type morpheme is present in the relevant clause. However, whether the speech is quoted or not is a binary variable. If the predicate is embedded in a quotation in a narrative it is coded 'y', 'n' otherwise. For the purposes of this study we consider all of the variables listed above to be independent variables, except for narrative time advancement, which is dependent. Finally, an important point in the context of this study is that not

all variables are realized in all the languages. For instance, Mẽbêngôkre does not contain a set of temporal distance morphemes, while the category of verb finiteness is irrelevant for Chácobo and Araona. Our study thus mixes what might be considered comparative notions with more language specific ones. This certainly limits our ability to draw comparisons between the languages of our study at this point. Future research will work out a more cross-linguistically applicable set of typological variables.

3. Indepent variables

3.1 Spatial modifiers: motion and posture

By a spatial modifier we mean one which encodes the the spatial configuration, orientation or displacement of one of the participants of the predicate. In the context of this study, spatial modifiers include morphemes expressing posture (participant configuration) and motion (participant displacement). Araona has both, Mẽbêngôkre has only the first, and Chácobo only the second. The postural auxiliaries of Mẽbêngôkre are illustrated below.

(6) i-nhõt o=wajêt "I'm sleeping (hanging in a hammock)"
i-nhõt o=dja "I'm sleeping (standing up)"
i-nhõt o=nõ "I'm sleeping (lying down)"

Examples of associated motion in Chácobo are the following:

(7) bi=baya=ní=quë "grabbed and left"
pi=bona='ái-na "ate while going"
bi=bëria=itá=quë "grabbed and came"

Takanists refer to posture suffixes as encoding imperfective (they do not always encode posture) — however they sometimes appear in constructions that advance narrative time. This is illustrated in the following example from Araona:

(8) ewao -ti badi di-di-ani bakwa di-ta-ani
ground.LOC -IDEO moon eat-eat-sit snake eat-3A-sit
"(The moon) arrived on the earth and the moon started eating, the moon was eating [=ate] the snake."

We code in an n-valued variable which modifier is present. In the case of Araona and Mẽbêngôkre, we expect that postural modifications will tend to be associated with the non-advancement of narrative time, even if there are exceptions. For associated motion the situation is more complex, but we should expect some relationship based on the description provided by Tallman. In the context of Araona and Chácobo, Tallman argues that concurrent associated motion morphemes have a tendency to express on-going events, whereas, prior and subsequent associated motion modifiers code completive. If there is an inferential association between progressive and imperfective, and completive and perfective, then we can expect associated motion in these languages to also be associated with the (non)advancement of narrative time.

3.2 Clause type and Dependency

By clause type we refer to morphemes that code notions of grammaticalized illocutionary force such as interrogative, declarative, reportative, or imperative. By dependency we refer to whether the clause type can be subordinated under another.

In Mẽbêngôkre, verbs have distinct dependent and independent forms in the latter sense. These are illustrated below. Following prior practice, the dependent form is referred to as nonfinite and the independent form as finite.

(9)	ba omũ$_{F}$	INDEPENDENT
	ije omũnh$_{NF}$	DEPENDENT
	ije omũnh$_{NF}$ kêt	DEPENDENT
	ba omũnh$_{NF}$ o=nhỹ$_{F}$	INDEPENDENT
	ije omũnh$_{NF}$ o=i-nhỹr$_{NF}$ kêt	DEPENDENT

Note that postural auxiliaries may be finite or nonfinite, and that they govern a nonfinite form of the verb. Though there is a clear tendency for finite clauses to advance narrative time and for nonfinite clauses not to, examples may be found to show that viewpoint is independent from finiteness:

(10) ’Ỳr bôx ’ỳr ne akubyn tẽ.
up_to arrive.N almost and back go.V
“He almost got there but then returned.” (nonfinite, advancing)

(11) Kum kabẽn o=nhỹ nhỹm mã tẽ.
3DAT 3.speak with=sit.V DS away go.V
“She was talking to him and he just left.” (finite, nonadvancing)

3.3 Finiteness and clause typing: Clause-typing in Chácobo

Chácobo has a series of morphemes that are required in all clauses and that encode clause type in addition to evidentiality and tense. The following are declarative and unmarked for evidentiality:

(12) ahui ja=quí ara=ní=quë
wife 3=DAT cry=REM_PST=DEC:PST
“his wife cried for him”

(13) mi-a pí=qui yahanë
2SG-ACC eat=DEC:NONPST tick
“the ticks are eating me”

An n-valued variable encodes which clause-typing morpheme appears in the clause

Like for Mẽbêngôkre finiteness/dependence, clause-typing in Chácobo covaries with but does not determine narrative time advancement. For instance the past declarative morpheme *=quë* is typically found in verbal predicate constructions where narrative time is advanced as in the following example.

(14) toronja bo ë bo=yamé=quë toró
grapefruit pl 1SG carry_away=DISTP=DEC:PAST Toro.ERG
bëna=há=na, jaquirëqué tsi camio jo=yamé=quë
ask_for=NMLZ:P=EPEN after LNK car come=DISTP=DEC:PAST
"I took the grapefruit that Toro asked for, then the truck arrived." (*-quë* advances narrative time)

However, it is not hard to find examples where *=quë* is found where narrative time is not advanced as in the following.

(15) toa tsi cana=bá ë papí=hi i=ní=quë jarí xocobo
DEM2 LNK Cana=BEN 1SG carry=CONCUR:S ITR=REMP=DEC:P still child
xóna
a
"I was carrying (lumber, yuca etc.) for Cana, when I was still a child." [txt 1838: 59] (*=quë* does not advance narrative time)

The clause-type is coded as a categorial variable. Clause-type morphemes in Chácobo can be straightforwardly classified according to whether they are dependent or not. Depedence was coded as a binary variable.

3.4 Verbal and nonverbal predication

We coded whether the predicate was a verbal or nonverbal predicate construction — under the assumption that nonverbal predicates would be more likely to not advance narrative time

Verbal predicates are more likely overall to advance narrative time — but they can also fail to advance it. An example of a verbal predicate construction from Araona which does not advance narrative time is provided in the example below.

(16) **di-día-sha-ja=tsio** ... bajome ... kwaiña-pobea
eat-eat-DISTPAST-lie=then agouti arrive-come
"**As he was eating** the agouti arrived (to his village) and came (to him)."

An example of a verbal predicate construction which does not advance narrative time in Chácobo is provided in the example below.

(17) Anomariá quihá yotá jo=ha=na
much REP southern_wind come=NMLZ:ANT=EPEN
jo=shina=ha=na. **Matsí tsi quihá mai**
come=night=NMLZ:ANT=EPEN cold LNK REP earth
"Much southern wind came and kept coming throughout the night. **And the earth was cold**."

It is not always the case that nonverbal predicate constructions do not advance narrative time. In the following example from Chácobo, narrative time advancement is coded with an existential nonverbal predicate construction *yama* in the following example.

(18) jo-tox ë há ca ... jotsoisi ca yá yáma — bjjj — jëë
ankle-hit 1SG NMLZ:ANT =rel ankle REL COMIT NEG IDEO IDEO
rohá tsi quiha ja ní=quë
LIMIT LNK REP 3 REMP=DEC:PAST
"I hit the log with my ankle, and then there was nothing, 'bjjj' (the ghost disappeared), and then he said 'jëëë' and nothing more."

Verbal versus nonverbal predication was coded as a binary variable.

3.5 Quotation

In narrative in these three languages, it is frequent for speech to be quoted without a speech act verb being present in the clause.

(19) A Mẽbêngôkre example
Nhỹm we kũm "nhãrã", nhỹm kum "ngre" anẽ.
DS HS 3DAT where DS 3DAT egg say.V
"He [asked] him 'where are [the fledglings]?', and he said to him '[there are only] eggs [here]'."

Speech act verbs may sometimes be used, as in the second part of the example. We code in a binary value whether a clause consists of a quotation without a speech act verb. Based on our own intuition from coding we expect whether the clause is quoted speech or not to be the best predictor of narrative time non-advancement.

3.6 Connectors or Conjunctions

Conjunctions or connectors play an important part in advancing narrative time (e.g. "and then", "after that") in all three languages. However, they are not always present, and with many of them the relation with advancement or non-advancement is only a tendency. We coded for each clause which connectors are used in an n-valued variable.

In Araona *=tso*: usually advances narrative time. However, there are occasional counterexamples as in the following example.

(20) *=tso* not advancing narrative time
sada peto=ibo=tso pewe ize po tokwe pae
river cross=SIGNIF=ANT LIMIT save NMLZ EPIS REP
"(It is said and I think it's the case that) one has to cross the river to save one's self (from the undead when they chase you)" (context: talking about how when saves oneself from ghosts - a general statement) [David Washima]

(21) *=tso* advancing narrative time (19 out of 20 cases)
di=tso azio=kata ba=sha; aaa — po=ani noji
eat=ANT spicy=AUGM VIS=DISTP aaa be.ITR=sit pretend
"When he (the moon) ate the snake, it burned. He attempted or made himself cough."

3.7 Aspectual, temporal and time of day markers

Mẽbêngôkre, Araona and Chácobo all have various aspectual and temporal markers that tend to correlate with either the advancement or nonadvancement of narrative time. Both Chácobo and Araona have temporal distance markers which have a similar (though not identical) function to temporal adverbs in more famililar European languages (Tallman and Stout 2016, 2018). All three languages have a highly elaborate set of aspectual markers that encode notions like durative, punctuality, and completive. Chácobo and Araona also have sets of time of day markers, most of which are also associated with durative meanings Tallman (2020, forthcoming).

Many of these are described as markers of (im)perfectivity in the descriptive literature (Arregui et al. 2014; Tallman 2018; Guillaume 2021). The presence or absence of aspectual markers were coded as categorical variables where the values corresponded to the relevant morphemes or constructions.

=*paó* is described as a habitual or durative in Tallman (2018): it is often found in predicates where narrative time is not advanced. However, it is not deterministically associated with this function. The following example provides an example of the durative marker in Chácobo not advancing narrative time, as in the progressive in Spanish.

(22) jabi tóca pi-pi=**pao**=tsa=hiquihá jáma joni huëtsa tsi quihá jisa
surely thus eat-eat=HAB/DUR=NOW=REP but man other LNK REP look
toa mi ahuiní baquë paquë=há ca co=yáma=no
DEM2 2SG woman.GEN child fall=NMLZ:P REL wander=NEG=CONCUR
"In this way she (Ashina) **was accustomed to always** eating (children), so another man said that when your wife gives birth you cannot go about (leaving the baby behind)." (780:32)

(23) paxa ca barí jana=xo=rohá quihá pii=ca=**pao**=ni=que
raw REL run put_out=a=LIMIT REP eat=PL=HAB=REMP=DEC:PAST
"**For a long while**, they placed it (the meat) raw in the sun and they ate it." (508:364)

4. The variables that determine (non)advancement

The previous sections provide a review of morphemes and constructions that could plausibly have some association with narrative time (non)advancement in the three languages of this study. While we conjecture that each of these variables are plausibly related to the (non)advancement of narrative time in online speech, we note that we do not know what their relative importance is. We also do not know whether there is variation across the languages of our study with respect to how important one variable might be over another. We certainly should not assume that all imperfectiveless languages determine narrative time (non)advancement in an identical fashion. Consider Tonhauser's conclusion regarding temporal relations in tenseless languages.

> "[...] there may be semantic variation among tenseless languages with respect to how the temporal relation between the reference and the utterance time is constrained: by context and temporal adverbials alone, or additionally by a covert tense." (Tonhauser 2011:299)

In this section we are interested in exploring how one can characterize and describe the expression of narrative time (non)advancement in imperfectiveless languages such as Araona, Chácobo and Mẽbêngôkre in relation to constructions and morphemes which are associated with viewpoint aspect. As has been emphasized throughout this paper, the relation between the form (coded constructions and morphemes) and the function (narrative time (non)advancement) is non-deterministic or probabilistic. The relation between form and function in this context is also multivariate (Brooks 2018:2,133) meaning that narrative time (non)advancement is plausibly co-determined by many factors at once. The factors may also be associated with each other, which means that inference is not a simple matter of testing pairwise associations between a single form and function once at a time. We expect collinearity in the data set to be rampant. To illustrate the problem, consider the relationship between reduplication, clause-typing and nonverbal predication in Chácobo. Leaning on the description in Tallman (2018) the agentive nominalizer *=hai* and reduplication are plausibly associated with narrative time non-advancement.

(21) Chácobo

nëa xo ë-a mana=huë jihui chahita bara bara=hai=na iwë
here DEC 1SG-EPEN wait=IMPER log large slip slip=NMLZ:IPV=EPEN EXPL
toa tahë tsi quihá awi ja nima=ní=quë
DEM2 beside LNK REP wife 3 wait=REMP=DEC:PAST
" 'I am here, wait!' (the man said to his wife), there was a large log that slipped (that people always slipped on), you see, and he left his wife beside that (log)."

Neither of these strategies always marks the non-advancement of narrative time, but our sense is that they both play a part because their presence/absence is correlated with (im)perfectivity. However, we have problems in making inferences in this regard because they may be associated with one another. This is depicted in Figure 3.

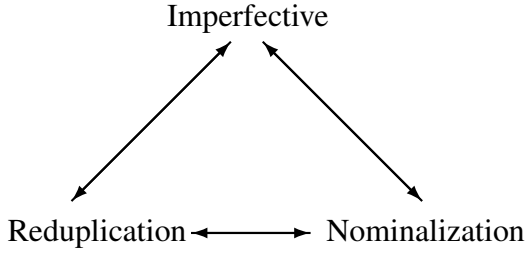
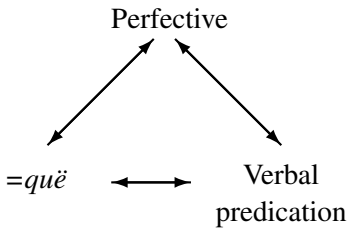


Figure 3: Collinearity problem

The problem can also be illustrated when we consider clause-typing in relation to verbal predicate constructions. We suggested that verbal predicate constructions might be as-

sociated with the advancement of narrative time. We also suggest that the declarative past marker =*quë* is associated with narrative time advancement. However, =*quë* only occurs in verbal predicate constructions. Thus simply providing bivariate correlations between narrative time (non)advancement and verbal predicate constructions and the presence of =*quë* will be misleading. The association between =*quë* and perfectivity could be spurious.



	Nonverbal predicate	Verbal predicate
No past marker	44	153
Past marker =*quë*	0	78

Figure 4: Cross-tabulation of predicate type and presence of past tense marker in Chácobo

4.1 Decision trees and Random forests

There is no perfect solution to collinearity, but some techniques deal with it better than others (Tomaschek et al. 2018, Baayen and Tagliamonte 2012). In the absence of a highly explicit causal model that allows us to tease apart different types of confounds (Pearl 2009, McElreath 2020) regressions and/or classification trees are not stable in the face of high collinearity. If we sample slightly differently from the data we could get different results. Random forests (RFs) are a strategy that attempts to rein this problem in. RFs create multiple classification trees based on resampling the data and aggregating over the results (Liaw and Wiener 2018). RFs provide more accurate results than decision trees. However, easy interpretability is sacrified in RFs. Below we, therefore, provide an overview of results from individual decision trees on the three languages and then a corroborative overview of the results of RFs with the same predictors enlisted.

The figure below provides the results of applying a decision tree to model Chácobo narrative time (non)advancement using *ctree()* in R. The following independent variables were used: verbal vs. nonverbal predication (y vs. n); quoted vs. non-quoted (y vs. n); back-channeled vs. non-back channeled (y. vs. n); temporal distance marker; connector before; connector after; clause-type morpheme; aspectual markers; associated motion markers; and independent versus dependent clauses (Tallman 2018 for an overview). The relative proportions of light grey to dark grey in the bins at the bottom of the decision tree represent relative proportions of imperfective versus perfective expressing predicates.

The model suggests that the presence of the past marker, the embedding of the clause in quoted speech, the absence of aspectual marking are important factors in marking the advancement of narrative time. Marginally, where the clause in a dependent clause and/or is embedded in an agentive nominalization is also important. A plausible reading of the model is that Chácobo speakers will infer a sentence to be narrative time advancing if, *ceteris paribus*, it contains the marker =*quë*. After the distinction between quoted and nonquoted speech, other factors appear to be relevant, but of lesser importance.

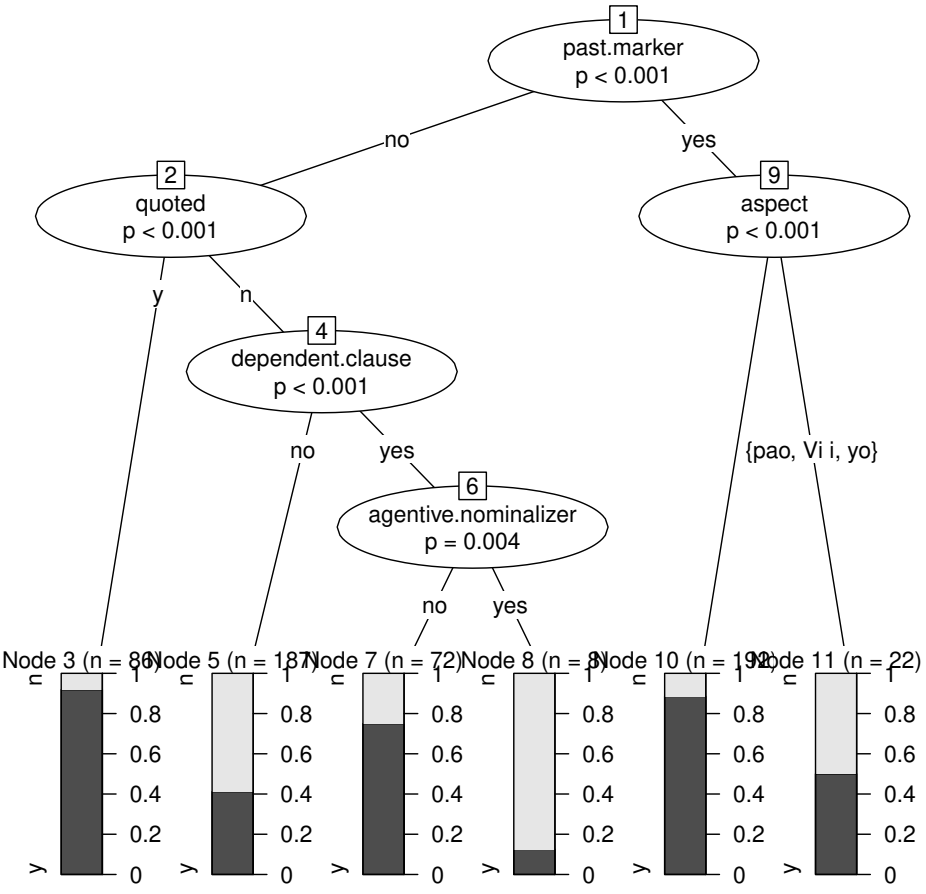


Figure 5: Decision tree for narrative time (non)advancement in Chácobo

We use the following as independent variables in a decision tree for Araona; (i) verbal versus nonverbal predication; quoted versus non-quoted speech; temporal distance markers; connectors before and after; posture modifications; aspectual markers; associated motion markers. The results are plotted below. predication.

Recall from the discussion above that Araona has no obligatory marking of tense and no obligatory clause-typing morphemes. In the absence of obligatory tense markers we might suspect that verbal predication will be associated with narrative time advancement. In fact this is what we observe from a decision tree. Narrative time nonadvancement is associated with nonverbal predication and the presence of the morpheme *=tsio* 'while, as' after the predicate. Advancement is associated with the connector morpheme *=tso* 'and then, already' and verbal predication.

We use the following as independent variables in building a decision tree for Mẽbêngôkre.; (i) nominal versus verbal predication; (ii) position versus non position main verb; (iii) quoted versus nonquoted speech; (iv) finite versus nonfinite verb; (v) connectors; (vi) particles.

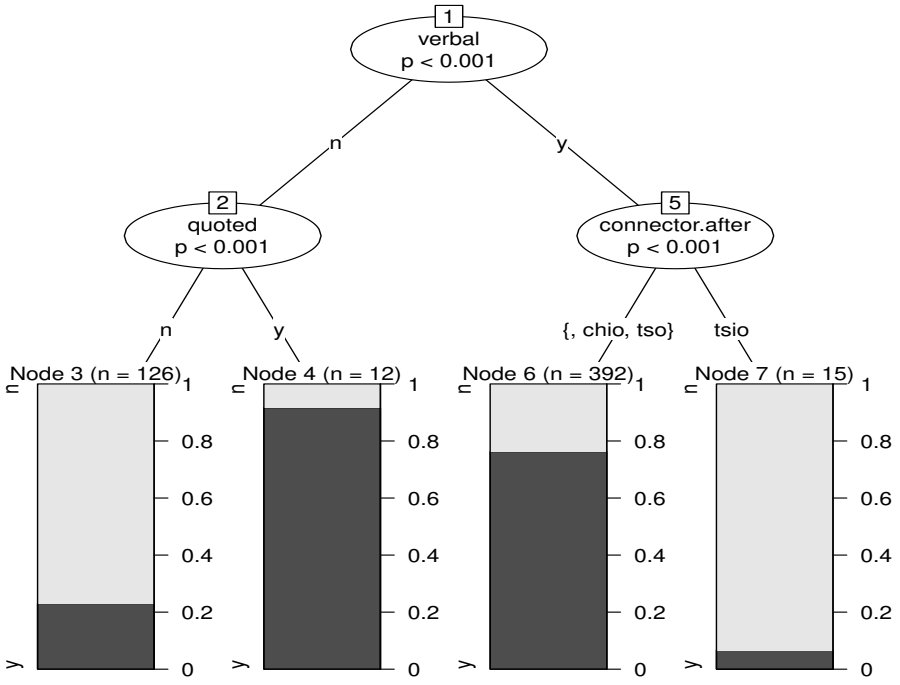


Figure 6: Decision tree for narrative time (non)advancement in Araona

The decision tree shows that, like Araona, the distinction between verbal and nonverbal predication (always nominal in the case of Mẽbêngôkre is an important predictor of narrative time (non)advancement. As with Chácobo and Araona, in Mẽbêngôkre quoted speech advances narrative time. Mẽbêngôkre's clause initial *kam* 'at that moment' might help advance narrative time in cases where the predicate in nonverbal and wherefore typically associated with non-advancement. The clause initial different subject marker *nhỹm* 'and (different subject)' appears to increase the likelihood that a sentence will advance narrative time when it involves nominal predication. Among the particles, only the reportative marker *we* 'hearsay' appears to be a good candidate for the advancement of narrative time. Position verbs are also weakly associated with non-advancement in the context of verbal predicates that are not embedded in quoted speech.

Decision trees have problems with accuracy in the presence of collinearity. The results are not always stable. From RFs we can arrive at measurements of the relative importance of different factors in the (non)advancement of narrative time by considering how much classification error is increased through their removal. We performed an RF analysis on each of the languages with the variables specified above (see Appendix for details and R code). We ran 1000 simulations with each RF analysis using the *randomForest()* package from R. The variable importance of the variables are displayed in Figure 8.

The confusion matrices for the RF models for the three languages are displayed in Figure 9. They show that all the models perform better than chance.

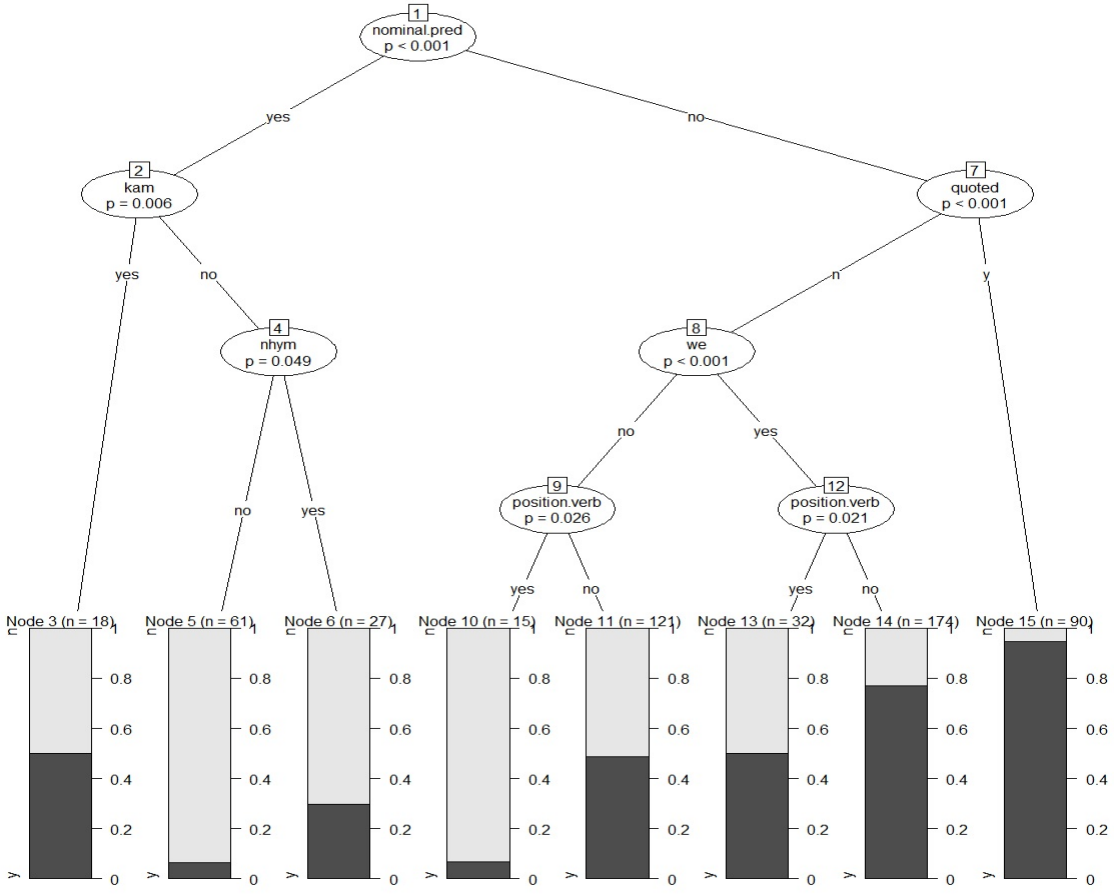


Figure 7: Decision tree for narrative time (non)advancement in Mẽbêngôkre

5. Conclusions and future directions

There are recurring patterns in the way that (non)advancement of narrative time is expressed cross-linguistically, even in languages that don't have a dedicated morpheme (i.e., verbal vs. nonverbal constructions are coopted). In cases where there is no deterministic association between a particular construction or grammatical element and narrative time advancement, random forests over naturalistic speech are a useful tool to bridge the gap between formal semantic analyses and discourse-based approaches: they can help guide more controlled semantic elicitation and show how variables interact in actual speech.

The sense that in natural speech you can't control for everything comes mostly, we suspect, from collinearity; if one uses random forests that problem can be reined in.

There are a number of theoretical issues that need to be thought through more carefully in order to home in on the best cross-linguistic operationalization of viewpoint and to identify all variables that may affect it. An important notion to examine, for instance, is that of background and foreground in narrative. Along these lines, a finer coding of narrative

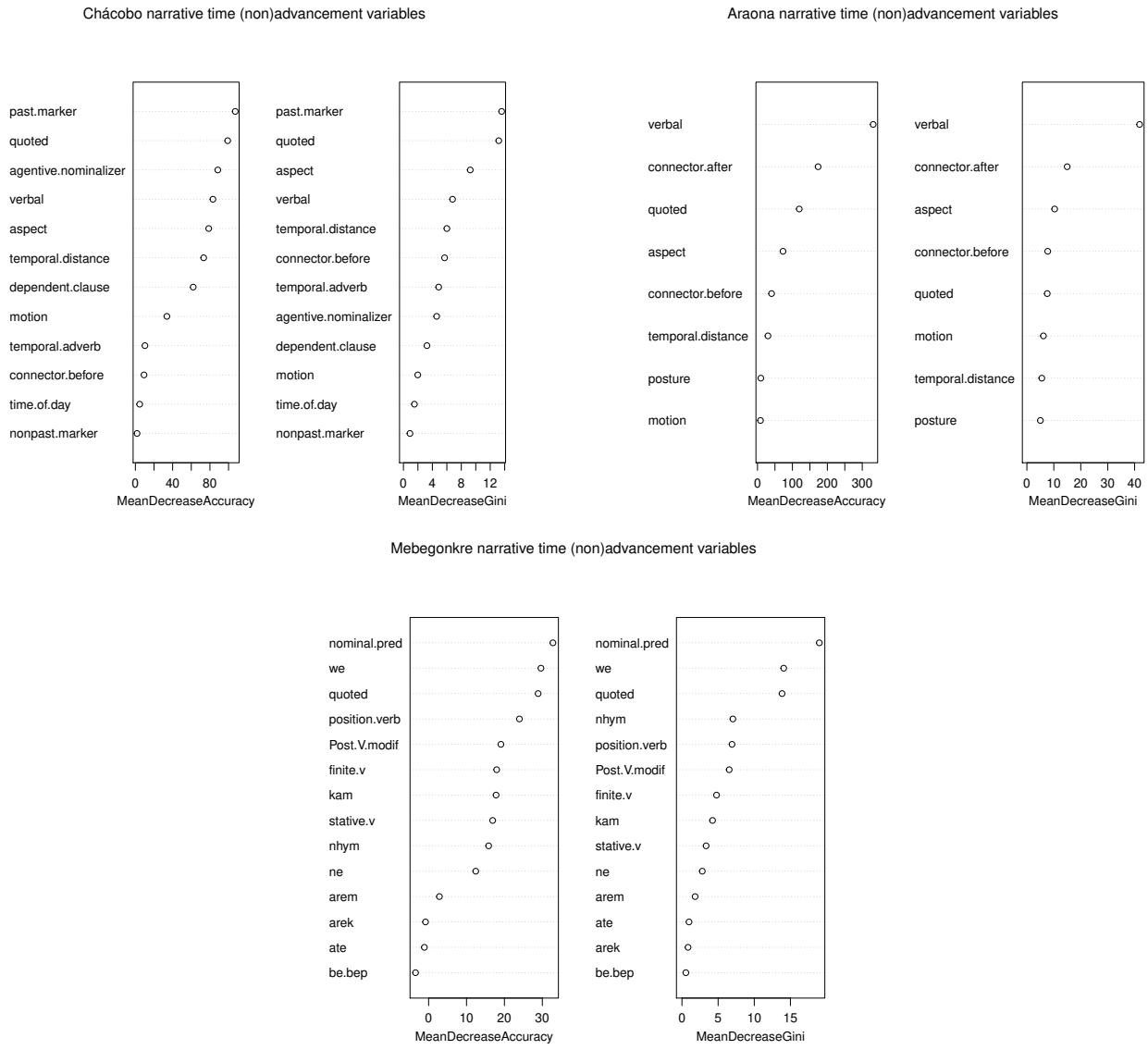


Figure 8: Relative importance of variables in the languages of the sample

structure, identifying repetition, tail-head linkage, time advancement within backgrounded episodes (i.e., an explanation that involves describing various steps in a procedure, or the description of a habit with several parts in sequence) or within quoted material, flashback, and so on. Finally, we need to determine what all of these notions mean in non-narrative text.

Regarding our operationalization of viewpoint as narrative time advancement, it would be essential to code narratives in the way we have done here in well-described languages that have explicit viewpoint morphology (i.e., Spanish, Russian, etc.) to test whether the results that are obtained are in line with traditional descriptions.

Chácobo Random Forest

	n	y	class error
n	68	107	0.61142857
y	22	370	0.05612245

Baseline 392 (69.5%) <Accuracy 438 (77.7%)

Araona Random Forest

	n	y	class error
n	110	95	0.46341463
y	34	307	0.09970674

Baseline 341 (62%) <Accuracy 417 (76.4%)

Mẽbêngôkre Random Forest

	n	y	class error
n	148	74	0.3333333
y	59	257	0.1867089

Baseline 316 (58.7%) <Accuracy 405 (75%)

Figure 9: Confusion matrices and overall Classification Error for Random Forests

References

Altshuler, Daniel. 2014. A typology of partitive aspectual operators. *Natural Language and Linguistic Theory* 32:735–775.

Arche, María J. 2014. The construction of viewpoint aspect: the imperfective revisited. *Natural Language and Linguistic Theory* 32:791–831.

Arregui, Ana, María Luisa Rivero, and Andrés Pablo Salanova. 2014. Cross-linguistic variation in imperfectivity. *Natural Language and Linguistic Theory* 32:307–362.

Baayen, Harold H., and Sali Tagliamonte. 2012. Models, forests and trees of York English: *Was*/*were* variation as a case study for statistical practice. *Language Variation and Change* 24:135–178.

Brooks, Joseph Daniels. 2018. Realis and irrealis: Chini verb morphology, clause chaining, and discourse. Doctoral dissertation, University of California Santa Barbara.

Comrie, Bernard. 1976. *Aspect*. Cambridge University Press.

Guillaume, Antoine. 2021. Takanan languages. In *Amazonian Languages. An International Handbook*, ed. by Patience Epps and Lev Michael. DeGruyter Mouton.

Liaw, Andy, and Matthew Wiener. 2018. Breiman and Cutler's random forests for classificaion and regression. Ms., CRAN repository.

McElreath, Richard. 2020. *Statistical rethinking: A bayesian course with examples in r and stan*. CRC Press.

Pearl, Judea. 2009. *Causality: Models, reasoning and inference*. Cambridge University Press.

Tallman, Adam J. R. 2018. A grammar of Chácobo, a southern Pano language of the northern Bolivian Amazon. Doctoral dissertation, University of Texas at Austin.

Tallman, Adam J. R. 2020. Associated motion and posture in Araona (Takana). Talk given at *Atelier de morphosyntaxe*, Laboratoire Dynamique du Langage, January 2020, Lyon.

Tallman, Adam J.R. forthcoming. Associated motion in chácobo. In *Associated motion*, ed. by Antoine Guillaume and Harold Koch. DeGruyter Mouton.

Thomas, Guillaume. 2014. Nominal tense and temporal implicatures: evidence from Mbyá. *Natural Language Semantics* 22:357–412.

Tomaschek, Fabian, Peter Hendrix, and Harold R. Baayen. 2018. Strategies for addressing collinearity in multivariate linguistic data. *Journal of phonetics* 71:249–267.

Tonhauser, Judith. 2011. Temporal reference in Paraguayan Guaraní, a tenseless language. *Linguistics and Philosophy* 34:257–303.