Kilu von Prince, Ana Krajinović, Manfred Krifka,
Valérie Guérin & Michael Franjieh

2019

Mapping Irreality:
Storyboards for Eliciting TAM Contexts

In

A. Gattnar, R. Hörnig, M. Störzer & S. Featherston (Eds.)

Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory

Tübingen: University of Tübingen

https://publikationen.uni-tuebingen.de/xmlui/handle/10900/87132

In A. Gattnar, R. Hörnig, M. Störzer & S. Featherston (Eds.) (2019). P*roceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory.* Tübingen: University of Tübingen. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/87132

Mapping Irreality: Storyboards for Eliciting TAM Contexts

Kilu von Prince[1], Ana Krajinović[1,3], Manfred Krifka[1,2], Valérie Guérin[4] & Michael Franjieh[5]

[1]Humboldt-Universität zu Berlin, [2]Leibniz-Zentrum Allgemeine Sprachwissenschaft, [3]University of Melbourne, [4]James Cook University, [5]University of Surrey

kilu.von.prince@hu-berlin.de

1 Introduction

In this study, we discuss temporal-modal distinctions in six Oceanic languages of Vanuatu. In particular, we ask whether the modal or the temporal dimension is dominant in each language and how much the languages differ in this respect.

We model these temporal-modal systems with a branching-time framework (going back to Thomason, 1970), which allows for more fine-grained distinctions of meanings compared to the traditional binary distinction between realis and irrealis.

Some of the domains we can thus distinguish are neither widely recognized in formal semantics nor in typology. A case in point is the counterfactual future, as expressed in *If you won the lottery tomorrow, what would you do?* This has also been referred to in the literature as *future-less-vivid* (Iatridou, 2000). This type of context is also quite rare in natural discourse so that the small corpora from language documentation that are the primary source of our research did not contain any instances.

In a branching-time model with a superimposed linear ordering of indices according to time values, the counterfactual future can be defined as indices that are not successors of the actual present i_c and are temporally later than i_c. This domain is illustrated in Figure 1.

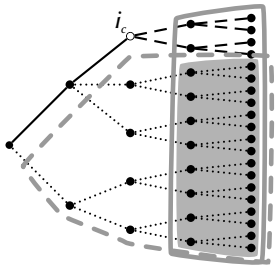

Figure 1. Shaded area: counterfactual future; solid outline: future; dashed outline: counterfactual indices; vertically stacked indices are taken to be simultaneous

Depending on whether a TAM system puts more emphasis on modal or on temporal orientation, future counterfactuals may either pattern with possible future contexts, or with past counterfactual ones – or they may receive a different marking altogether. Since the languages of our study tend to be mood prominent, we had a cautious expectation that they would group the counterfactual future with counterfactual past rather than the possible future. This expectation

was only partially borne out. At the same time, we expected significant variation between our subject languages, based on prior observations of the region. Our comparative study proves that TAM systems even in closely related languages of Vanuatu indeed show remarkable variation.

In order to answer our questions empirically, we worked with storyboard elicitations (Burton & Matthewson, 2015). Many of the existing storyboards already target specific TAM contexts so that we could use several storyboards designed by others. However, certain relevant domains have never been covered by storyboards or similar elicitation methods, to the best of our knowledge. We have therefore designed a set of storyboards that fill the gaps in previous elicitation tools. One of those gaps concerns the counterfactual future, which is the main focus of this paper. We will report our hypotheses, the storyboards we used to address them, and our results.

This study is part of a wider project that aims at better situating Oceanic languages within the debates on tense and modality, especially with respect to the distinction between realis and irrealis modalities. Our findings have implications for our understanding of tense and modality in shedding light on the intimate connection between these two dimensions. Throughout the article, we will systematically comment on how our findings bear on the following questions:

- Are our subject languages more mood-prominent or more tense-prominent?
- How important is the distinction between realis and irrealis modality in each language?
- Do our findings support a tripartite distinction of modal domains into the factual, the counterfactual and the possible?

2 Preliminary observations

The six languages of this study are Dalkalaen, Daakaka, Daakie, North Ambrym, Mavea and Nafsan (South Efate). They are all Oceanic languages of Vanuatu, with speaker populations ranging from about 30 (Mavea) to around 5000 (Nafsan). The map in Figure 2 shows the locations in which the languages are primarily spoken.

Despite many structural similarities, the subject languages differ in how they mark finite predicates for TAM. On the more analytic part of the spectrum, TAM markers are clitics or particles that occur between the subject agreement marker and the verb root. An example for this is Daakaka, as shown in Table 1.[1]

Table 1. Structure of the finite verbal complex in Daakaka

SUBJ.AGR	(=)TAM	(AUX)	(REDUP-)	**Verb**	(-RES)	(=TRANS)
na/ko/ …	*=m, …*	*du/pwer*	…	…	…	*=ne*

An example for how these elements can combine within a sentence is given in (1):

[1]Abbreviations: 1DL – first person dual; 1PL – first person plural; 1SG – first person singular; 2DL – second person dual; 2PL – second person plural; 2SG – second person singular; 3DL – third person dual; 3PL – third person plural; 3SG – third person singular; AGR – agreement; ASR – assertion; AUX – auxiliary; CL3 – possessive class 3; CL – possessive class; COMP – complementizer; COND – conditional; CONT – continuous; COP – copula; CTF – counterfactual; DIST – distal; DL – dual; DP – direct possession; EXCL – exclusive; FUT – future; GEN – general (possessive class); IMPF – imperfective; IN – inclusive; IRR – irrealis; NEG – negative; NMLZ – nominalizer; NREC – non-recent; O – object; PC – paucal; PL – plural; POSS – possessive; POS – positive; POT – potential; PSP – prospective; PST – past; REAL – realis; REC – recent; REDUP – reduplication; RES – resultative suffix; SUBJ – subject; SUB – subordinator; TAM – tense, aspect, mood; TRANS – transitivizer; V – verb.

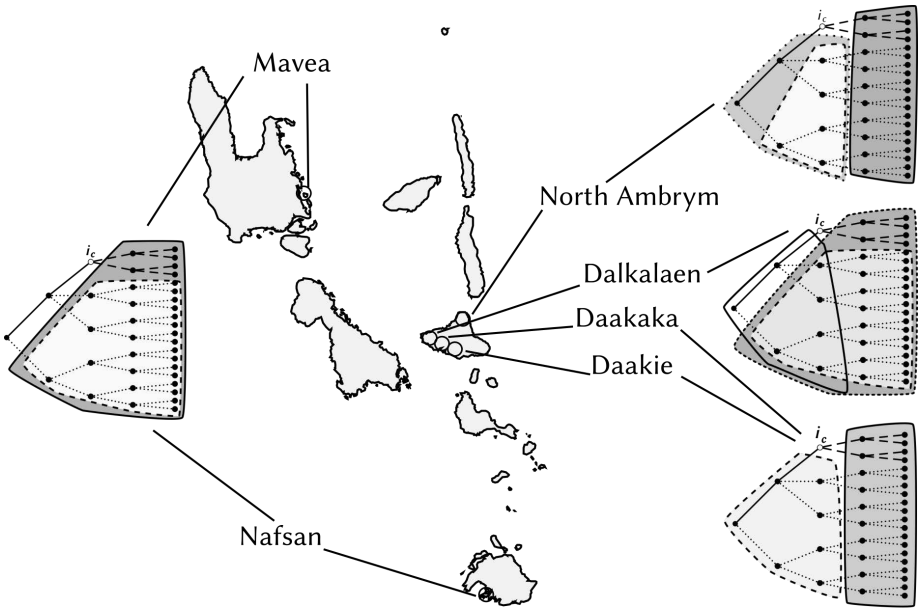

Figure 2. A map showing the section of Vanuatu in which the primary speaker populations of the subject languages are located

(1) *na=m* *yungpan=ne* *wye*
1SG=REAL thirsty=TRANS water
'I'm thirsty for water.' (von Prince, 2015: 150, ex. (244))

In some other languages under investigation, the finite TAM marking merges more closely with the subject agreement marker, yielding in some cases portmanteau subject-TAM proclitics.[2] This is illustrated by the following example from Nafsan, where the proclitic *ka* simultaneously encodes person and number features of the subject and irrealis modality.

(2) ***ka****=fan* *saof-i-r* *Ertap*
1SG.IRR=go:IRR visit-TRANS-3PL.O Eratap
'I will visit them at Eratap.' (015.004)

Portmanteau subject proclitics can also be found in Mavea and North Ambrym. Many Oceanic languages have been described as distinguishing between realis and irrealis mood. While realis expressions are restricted to the actual past and present, irrealis expressions refer to possibilities, counterfactual developments and the future. They are also often used in directives. Furthermore, the irrealis distinction interacts in complex ways with negation. The typological validity of the irrealis category has been hotly contested (Bybee et al., 1994; Bybee, 1998; Cristofaro, 2012; de Haan, 2012). But at least in the context of Oceanic languages, there is widespread agreement that irrealis is a meaningful grammatical category (Elliott, 2000; Lichtenberk, 2016).

In the project languages, too, the distinction between realis and irrealis modalities plays a central role in the organization of TAM systems. At the same time, TAM systems are usually not structured around a simple binary distinction but show a more complex situation. For example, in Nafsan there are three sets of subject proclitics – the general set, which is mostly used for realis contexts; the irrealis set mostly used for futures; and the perfect set with the corresponding aspectual information (Thieberger, 2006).

[2]A portmanteau morpheme encodes several semantically distinct functions at the same time.

Table 2. The Daakaka TAM system

	enclitic	proclitic	monosyllabic
Pos. Realis	*=m*	*mw=*	*mwe/mV*
Neg. Realis			*to*
Pos. Potential	*=p*	*w=*	*wV*
Neg. Potential	*=n*		*nV*
Distal	*=t*	*t=*	*tV*
(Open Polarity			*doo)*
(Change of State			*bwet)*

To give one more example, the Daakaka TAM paradigm has three major modal-temporal distinctions: The realis, which is responsible for the actual present and past; the potential marker, which is responsible for possible futures and epistemic possibilities of the present; and the distal marker, which refers to the actual, discontinuous past, to counterfactual developments of the past and present, and to epistemic possibilities of the past. The open-polarity marker *doo* is restricted to embedded polarity questions; the change-of-state marker *bwet* has the same temporal-modal values as the realis marker but comes with an additional aspectual interpretation. This system is shown in Table 2.

The realis marker refers to events of the actual past or present:

(3) *s-an* *naana* ***mwe*** *vyan* *yen* *too*
CL3-3SG.POSS mother REAL go in garden
'his mother went to the garden' (0274)

The potential marker refers to the possible future and to possibilities of the present. An example of the potential marker with reference to the future is given in (4):

(4) *barvinye* *swa* ***ka*** ***we*** *luk* *teve-sye* *m-ada* *em*
grass one ASR POT grow side.of-3SG.POSS CL3-1DL.IN.POSS house
'a grass will grow next to our house' (sto17:13)

The following example illustrates the distal marker with a counterfactual reference:

(5) *Nye* *na* *bwe* *dimyane* *ka* *ebya-ok* *we* *pwer* *kyun,* *na=**t*** *ka*
1SG 1SG REAL;CONT want COMP wing-3SG.POSS POT stay just 1SG=DIST fly
pini *or.*
fill place
'I wish I had wings, I would fly around everywhere.' (ess01:3)

These empirical facts have been described in more detail in von Prince (2015, 2017). The basic three-way distinction in Daakaka corresponds roughly to the three modal-temporal domains created by a branching-time frame described in the following section.

3 Mapping Irreality

The theoretical basis for our hypotheses has been fleshed out in von Prince (2017, 2019). The main ingredient for our analysis is the branching-times structure that is a well-established tool for exploring the relation between temporal and modal reference (e.g., Condoravdi, 2002; Dowty, 1977; Ippolito, 2013; Laca, 2012; Thomason, 1984). Our basic definition of the branching structure follows Thomason (1970, 1984):

(6) Definition Branching Times: A branching-times frame $\mathfrak{A}$ is a pair $\langle I, < \rangle$, where

a. I is a non-empty set of indices i; $<$ is an ordering on I such that if $i_1 < i$ and $i_2 < i$, then either $i_1 = i_2$, or $i_1 < i_2$, or $i_2 < i_1$.

b. A **branch** through i is a maximal linearly ordered subset of I containing i.

c. An index i_1 is called a **predecessor** of i_2 iff $i_1 < i_2$; it is a **successor** of i_2 iff $i_2 < i_1$

Traditionally, quantification over branching times has been restricted to those branches that are identical up to the actual present. Thus, in the toy model represented in the following figure, if i_2 is the actual present, then quantification is restricted to branches b_3, b_4.

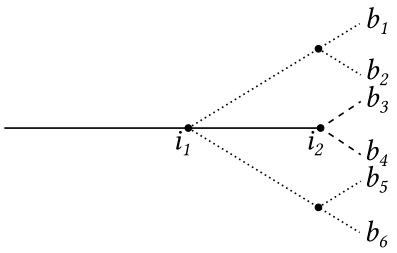

Figure 3. A branching-times structure. Relative to i_2, the solid line represents the actual past, the dashed lines the possible futures and the dotted lines counterfactual developments

It is also possible to quantify over all six branches $b_1, \ldots, b_6$, if one shifts the perspective backwards to i_1. However, it is not possible to quantify exclusively over b_1, b_2, b_5, b_6, because from i_2, they are not accessible at all, and from the perspective of i_1, the precedence relation cannot distinguish them from b_3 and b_4. The decision to restrict quantification in this way was originally well motivated, since the model was designed to define historical accessibility. We here follow von Prince (2017, 2019), however, in lifting this restriction. Instead of a two-way distinction between the actual past and present and the possible futures, we can in addition exclusively quantify over counterfactual indices as well. The precedence relation generates the following three-way distinction between modal-temporal domains relative to the contextually defined actual present i_c:

(7) a. the actual (past or present): $\{i | i \leq i_c\}$

b. the counterfactual (past, present or future): $\{i | i \not\leq i_c, i_c \not< i\}$

c. the possible (future): $\{i | i_c < i\}$

These domains will be indicated graphically as shown in Figure 4.

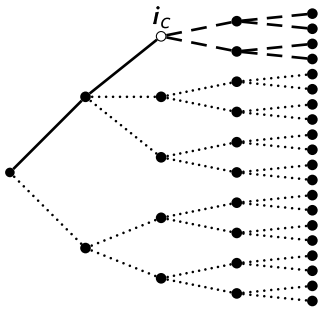

Figure 4. Solid: actual indices; dashed: the possible future; dotted: counterfactual indices

As proposed in von Prince (2017), this three-way distinction corresponds roughly to the three domains referred to by the main Daakaka TAM categories: The actual domain corresponds to the realis marker, the possible future domain to the potential mood, and the distal marker is the only one with a reference to the counterfactual past and present. The situation is different for each of the subject languages.

Prior to our storyboard elicitations, we had a preliminary picture of the main distinctions that were implemented in each TAM system. For counterfactual past contexts, however, we had only very little data. For counterfactual future contexts, we did not have any data in any of the subject languages. The state of our knowledge prior to storyboard elicitations in 2017 is depicted in Table 3.

Table 3. The state of our knowledge about distinctions between TAM contexts prior to storyboard elicitations

Language	Actual past/present	Possible future	Count. past	Count. future
Mavea	SUBJ.(REAL)	FUT, SUBJ.(IRR)	(IRR), *imte*?	(IRR), ∅?
Nafsan	∅	SUBJ.IRR	IRR?	?
Daakie	REAL/DIST	POT	DIST	?
Daakaka	REAL/DIST	POT	DIST	?
Dalkalaen	REAL/DIST	POT	?	?
North Ambrym	REC.PST/NREC.PST	IRR	CTF	?

We wanted to use the storyboards to consolidate our knowledge about counterfactual past contexts and to explore in particular counterfactual future contexts:[3] Would they pattern with the counterfactual past or with the possible future? In other words: Would they prioritize the temporal dimension or the modal one?

4 Storyboards for targeted elicitation

4.1 Methodology

Storyboards are a highly efficient tool for eliciting utterances that target a well-defined meaning, while supplying a rich discourse context and ensuring a certain degree of naturalness. The general methodology is described in Burton & Matthewson (2015). During our fieldwork, we ran each storyboard with four to ten speakers per language. We presented each speaker with the pictures and walked them through the stories before letting them retell the stories. Sometimes, speakers did a practice round. In some cases, we also used the Bislama version of the storyline to help speakers paraphrase the pictures and tell the stories. To ascertain obligatoriness of spe-

[3] In the typological literature, counterfactual future conditionals are also sometimes referred to as hypothetical conditionals (Longacre & Thompson, 1985).

cific features, we asked follow-up questions about whether alternative phrasings would also be possible in the same context.

Except for Guérin, fieldwork was carried out after a workshop in Vanuatu's capital Port Vila, which introduced the storyboards and provided instructions in the summer of 2017. Guérin had completed her fieldwork earlier the same year, after a visit to Berlin where we discussed the stimuli and methods.

For optimized searches and analysis we have used the search and visualization platform ANNIS (Krause & Zeldes, 2016), with our corpus data prepared for it using ToolboxTextModules (Druskat, 2018) for the conversion framework Pepper for linguistic data (Zipser & Romary, 2010).

4.2 Storyboards used

In order to elicit counterfactual past conditionals, we used two storyboards from the Totem Field Storyboard site: The 'Fortune Teller' storyboard (TFS Working Group, 2010) and the 'Woodchopper' storyboard (TFS Working Group, 2011b). Both storyboards also prompt speakers to utter indicative conditionals about the future and thereby provide valuable minimal pairs for the difference between counterfactual and indicative conditionals. In this context, we will focus mostly on the positive conditionals from the 'Fortune Teller'.

In TFS Working Group (2010), a woman called Mary has a hard time deciding whether to marry a (specific) tall young man or a short and fat one. She asks a fortune teller for advice. The fortune teller predicts: *If you marry the tall one, you two will have many children!* – the corresponding picture is shown in Figure 5. This first context is about the possible future.

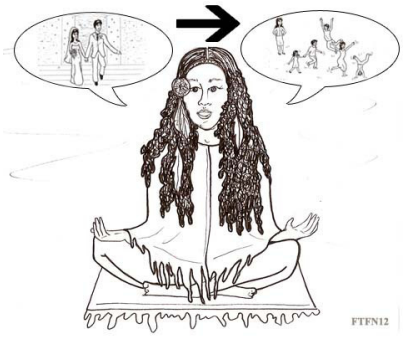

Figure 5. Possible future: 'If you marry the tall one, you two will have many children!' (TFS Working Group, 2010)

The second target context is about the counterfactual past. Many years later in the same story, Mary learns about the accidental death of her second suitor, the short man. Wondering how her life would have turned out had she chosen him, she returns to the fortune teller. The fortune teller tells her: *If you had married the short one, you two would have been rich.* The corresponding picture is shown in Figure 6.

None of the pre-existing materials, however, target the counterfactual future. In order to fill this gap, we produced the Festival storyboard (von Prince, 2018c): There are two boys, let's call them Sam and Luk. There is a three-day festival at their hometown or village, where each day comes with a different activity. There is a football game on the first day, the second day brings a concert, and on the third day, there will be a volleyball game. On the second day of the three-day event, Sam and Luk talk about the activities. Sam asks Luk whether he played football the day before. Luk says that he didn't, because it was raining. Then he says: *If I had played, I would have gotten wet.* The corresponding picture is shown in Figure 7.

Figure 6. Counterfactual past: 'If you had married the short one, you two would have been rich.'

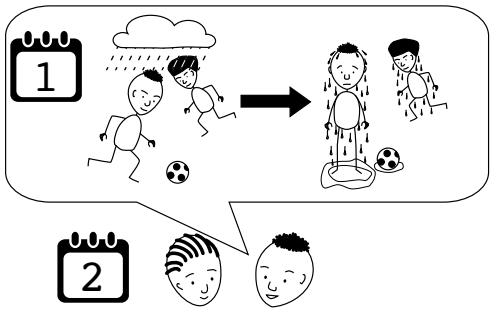

Figure 7. Counterfactual past: 'If I had played (yesterday), I would have gotten wet.'

This first target context thus produces a counterfactual conditional of the past as a closely minimal control sentence. The target context for the counterfactual future is as follows: Sam goes on to ask Luk about the volleyball game tomorrow. Is he planning to play then? Luk says, no, he is not going to play. The reason is that he has cut his finger. He says: *If I played tomorrow, my finger would bleed again.* The corresponding picture is shown in Figure 8.

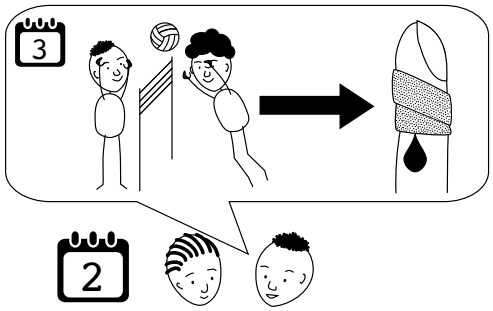

Figure 8. Counterfactual future: 'If I played (tomorrow), my finger would bleed again.'

We used a total of ten storyboards to cover also false-belief reports, intensional relative clauses, epistemic possibility, and similar (Rolka & Cable, 2010; TFS Working Group, 2010, 2011a,b; Vander Klok, 2013; von Prince, 2018a,b,c,d,e). In this article, we will primarily report the results from the two storyboards discussed here, in accordance with our focus on counterfactual future contexts. Table 4 shows by how many speakers per language each storyboard was performed. All our storyboards are freely accessible from *Zenodo*, through the links provided in the reference section.

Table 4. Numbers of speakers who performed the two storyboards in each language, and researchers responsible for elicitations in each language; Fest.: *Festival* storyboard; Fort. Tel.: *Fortune Teller* storyboard

Language	Researcher	Fest.	Fort. Tel.
Mavea	Valérie Guérin	3	4
North Ambrym	Michael Franjieh	5	5
Dalkalaen	Kilu von Prince	3	4
Daakaka	Kilu von Prince	5	4
Daakie	Manfred Krifka	10	10
Nafsan	Ana Krajinović	6	7

5 Results

In several languages we discovered previously undescribed TAM morphemes, or previously undescribed TAM-related uses for specific morphemes. Even on a merely descriptive basis, this work therefore contributes significantly to our knowledge of these languages. In this section, we will address two main topics: 1) morphology that is specific to counterfactual conditionals (of the past or otherwise); and 2) the marking of the counterfactual future: Does it pattern with the counterfactual past or with the possible future?

5.1 Nafsan

Nafsan has three paradigms of subject proclitics. Two of those paradigms are portmanteau morphemes that also encode TAM values in addition to person and number features of the subject. This system is shown in Table 5.

Table 5. Subject proclitics for singular subjects in Nafsan based on Thieberger (2006)

	General form	Irrealis	Perfect
1sg	*a=*	*ka=*	*kai=*
2sg	*ku=*	*p̃a=*	*kui=*
3sg	*i=*	*ke=*	*ki=*
		⋮	

These subject proclitics optionally combine with a TAM marker from a small paradigm including the conditional marker *fla/f*. According to Thieberger (2006), conditional clauses are always formed with this marker. In our elicited data we only found the *f* version of the conditional marker.

Alternatively, a conditional clause may be introduced by the formula *(i=)f-wel (kin)* "if it is like". These generalizations were confirmed by our findings. One example is given below, showing the use of the *if wel kin* construction:

(8) ***I=f-wel*** ***kin*** ***ku=****taulu* *John,* *akam* *rak=fo* *pitlak* *teesa*
3SG=COND-like COMP 2SG=marry John 2DL 2DL.IRR= PSP.IRR have children
ruk=fo *laap.*
3PL.IRR=PSP.IRR many
'If you marry John, you will have a lot of children.' (AK-010.24)

We found that, in the protasis of future-oriented indicatives and all counterfactual clauses, both the general and the irrealis proclitics can be found. However, the protasis and apodosis of present/past-oriented indicatives appear to be restricted to general proclitics, as illustrated by

the following example:[4]

(9) ***F-wel*** ***kin*** *npat-i-n* ***i=****miel, go Yokon m̃as kin* ***i=****paam nawi miel gaag.*
COND-like COMP teeth-V-3SG.DP 3SG=red then Yokon only COMP 3SG=eat yam red 2SG.POSS
'If her teeth are red, then Yokon is the one who ate your red yam.' (AK1-060-01.39)

The apodosis of future-oriented conditionals and all counterfactuals is always in irrealis mood. This is illustrated by example (8) for indicative conditionals above, and for future- and past-oriented counterfactuals below:

(10) ***I=f-wel*** ***kin*** ***ka=****mes "vole" matol, go nfag nen kin* ***a=****tai naru-k* ***ke=****fo mer toop.*
3SG=COND-like COMP 1SG.IRR=play volleyball tomorrow and sore that COMP 1SG=cut hand-1SG.DP 3SG.IRR=PSP again big
'If I played volleyball tomorrow, the sore that I cut on my hand would become big again.' (AK1-021-01.49)

(11) ***I=f-wel*** ***kin*** ***a=****mes futbol nanom,* ***ka=****fo lom.*
3SG=COND-like COMP 1SG=play football yesterday 1SG.IRR=PSP.IRR wet
'If I had played football yesterday, I would have gotten wet.' (AK1-021-01.39)

This means that possible-future conditionals can have the exact same form as counterfactual-future and counterfactual-past conditionals. It is however also possible to specify a conditional as counterfactual – regardless of its temporal orientation – by including the morpheme *mer* in the protasis:

(12) ***ka=f*** ***mer*** *pei p̃i "bol" nanom,* ***ka=****fo lom usrek.*
1SG.IRR=COND CTF first kick ball yesterday 1SG.IRR=PSP.IRR wet completely
'If I had played football yesterday I would have gotten soaked.' (AK1-004-01.163)

(13) ***Ka=f*** ***mer*** *mes "volibol" matol, nakni-k* ***ke=****fo mra.*
1SG.IRR=COND CTF play volleyball tomorrow finger-1SG.DP 3SG.IRR=PSP.IRR bleed
'If I played volleyball tomorrow, my finger would bleed.' (AK1-004-01.23/24)

Outside of the protasis of counterfactual clauses, *mer* means "again", as has already been observed in Thieberger (2006). Its function in counterfactual clauses has not been observed prior to this study and is a new empirical result of our fieldwork. Another new observation is that, at least in the configuration seen in (13), the conditional marker *f* can combine directly with an irrealis proclitic. Thieberger (2006) has previously stated that *f* can only combine with general proclitics. The contrast between our elicitations and the corpus data may well be an indication of diachronic change. The corpus data come from mostly older speakers in the 1990s, while the elicitations were performed primarily with younger speakers in 2017. For a detailed descriptions about possible diachronic changes, see Krajinović & Thieberger (2018).

In sum, irrealis subject proclitics are indeed restricted to irrealis contexts. They are obligatory in the apodosis of any conditional with non-actual reference. Counterfactual conditionals can optionally be distinguished from the possible future by the morpheme *mer* in the protasis. This result is illustrated in Figure 9.

[4]This is from a storyboard not discussed further for reasons of space. In it, Mary tries to find out who stole and ate her red yam and her friend suggests she look at the teeth of her prime suspect to find out (von Prince, 2018d).

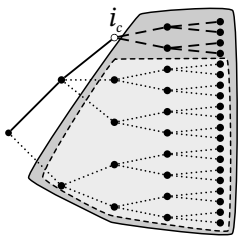

Figure 9. The irrealis domain in Nafsan. Solid outline: irrealis subject proclitics; dashed outline: optional *mer*

Our findings on Nafsan support the view that a binary distinction between realis and irrealis moods is important in the context of Oceanic languages. They also show that the domain of counterfactuality can be treated as a subdomain of irreality, irrespective of temporal reference, supporting a tripartite modal distinction. In the data presented, the modal dimension is prioritized over the temporal one.

5.2 Mavea

For Mavea, Guérin (2011) reports that conditional clauses of all kinds are marked by the conditional affix *mo*, which comes between the subject and the verb root. This was confirmed by our findings, as illustrated by (14):

(14) *ko-**mo**-l-to tuan nna me natu-mrua me i-lavoa*
2SG-COND-IMPF-stay with 3SG FUT child-1PL.EXCL.DL FUT 3SG.IRR-big
'If you stay with him, your children will be many.' (VG20171060.020)

For counterfactual contexts specifically, Guérin (2011) also reports the use of *imte* or *inte*, which is also a verb meaning *wish*. This element might have developed from a morphologically more complex structure that is shown in (15):

(15) *i-mo-te*
3SG.IRR-COND-some
'if it were' (Guérin, 2011: 234)

We can also cautiously confirm that this morpheme is specific to counterfactual contexts of the past and future. Outside of the contexts that are the focus of our discussion, there are also a few occurrences of *imte* that are ambiguous between counterfactual and indicative future contexts. In the target contexts of the present study, *imte* only appeared in counterfactual conditionals of the past, as in (16):

(16) ***imte** ka-v ka-v̋a v̋alu-na ro me […] m̋auri rarua i-isav̋ai*
if 1SG.IRR-say 1SG.IRR-go to-3SG.POSS then FUT life 3PL.DL 3SG.IRR-how
'Suppose I had stayed with him […] how would our life have been?"
(VG20171047.056-058)

The two oldest speakers used *imte* in counterfactual contexts quite consistently. The younger speakers did not use it at all. This might indicate that it is a vanishing feature of the language. This situation was also already observed by Guérin (2011: 380).

Another feature that showed up regularly in counterfactual conditionals of the past is the morpheme *me*. It has previously been described as a future marker in Guérin (2011: 217). It occurs in the apodosis of these conditionals:

(17) *ka-mo-lo-to* *tuan* ***me*** *m̋auri-ku* *i-pal* *sa*
1SG.IRR-COND-IMPF-stay with FUT life-1SG.POSS 3SG.IRR-like what
'…if i had stayed with Peri how would my life have been?' (VG20171060.031-032)

In all likelihood, though, *me* does not contribute to the absolute modal-temporal reference of the conditional but rather marks the apodosis as being temporally later than the protasis. We know from previous corpus data and description that the future reference of *me* is relative to topic time rather than utterance time, like apparently most future markers in Oceanic languages:

(18) *mo-ntao* ***me*** *ro* *i-lo-to* *aro* ***me*** *m̋arao* *i-an* *nna*
3SG-afraid FUT then 3SG.IRR-IMPF-stay here FUT eel 3SG.IRR-eat 3SG
'she was scared that she would stay there and the eel would eat her.' (06043.052)

Like the possible future, the counterfactual future and the counterfactual past are consistently marked by the irrealis version of the portmanteau subject proclitics for the first and third person singular.

Counterfactual-future contexts were equally expressed with the irrealis set of subject agreement markers:

(19) *m̋atan* *ka-v* ***ka****-mo-ple* *tuan* *varango-ku* *vutpol*
because 1SG.IRR-say 1SG.IRR-COND-kick with finger-1SG.POSS football
i*-mo-voreia* ***me*** *ro* ***i****-dae*
3SG.IRR-COND-hit FUT then 3SG.IRR-blood
'because if I play with my finger, if the ball hits it, it will bleed.' (VG20171008.051/52)

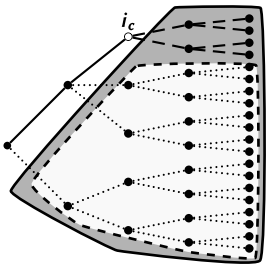

Figure 10. Our current best hypothesis about the marking of irreality in Mavea. Solid outline: irrealis subject proclitics; dashed outline: *imte*

In sum, Mavea has irrealis portmanteau subject proclitics for the first and second person singular that are used in all non-actual domains, comprising the possible future, the counterfactual past and the counterfactual future. The morpheme *imte* is used by older speakers in counterfactual context of the past, but also in a few other irrealis environments that we can not exhaustively disambiguate. The morpheme *me* probably marks relative future irrespective of modal and absolute-temporal reference. We therefore see a binary distinction between realis and irrealis modalities, with *imte* possibly being specific to counterfactual contexts. This situation is illustrated in Figure 10.

Similar to Nafsan, Mavea appears to be a mood-prominent languages. The binary distinction between realis and irrealis plays a major role in its TAM system. At the same time, the domain of counterfactuality can be marked as a sub-domain of irreality, thereby supporting a tripartite view of the modal dimension.

5.3 North Ambrym

Franjieh (2012) reports the paradigm of TAM markers that partially fuse with the subject proclitics that is summarized in Table 6.

Table 6. The core system of TAM markers in North Ambrym

Gloss	Realization	Description
REC.PST	*mV/m-/-m*	recent past, past/present for non-telic predicates.
NREC.PST	*te/te-/to-/-rr*	non-recent past
IRR	*f-/bV/b-/-∅*	irrealis
CTF	*to*	counterfactual past/present
AVE	*ne/-n*	avertive: unmet expectations about the future

The possible future is referred to by the irrealis mood in combination with the potential marker *e*:

(20) *Jon, bone* ***fō*** *ktu, lo mwen-amrō teere nyer* ***e-ve*** *lol.*
John time 2SG.IRR take then GEN.CL-2DL.POSS child PL POT-COP.IRR plenty
'If you marry John, you will have many children.' (ib1-fortune-na.35)

All five speakers produced the two target contexts for a counterfactual conditional of the past in the same way, with the counterfactual marker *to* in the protasis and the non-recent past marker *te/-rr* in the apodosis:

(21) *ō* ***to*** *yene Adam lo mwena-mrō mane* ***te*** *lam.*
2SG CTF marry Adam then POSS.CL-2DL.POSS money NREC.PST big
'If you had married Adam, you two would have been rich.' (at1-fortune-na.24)

(22) *Na* ***to*** *rrō plei bol, lo na-****rr*** *loo.*
1SG CTF CONT play ball then 1SG-NREC.PST get.wet
'If I had played football, I would have gotten wet.' (at1-lafet-na)

This confirms that the counterfactual marker *to* is specialized for counterfactual (past) contexts, while it also shows that the non-recent-past marker is not restricted to the actual past.

In some of the counterfactual past conditionals, we also find the continuous marker *rro* in the protasis. The above example is one such case. This was however rather the exception than the rule.

The target clauses for counterfactual future conditionals were realized in irrealis mood in the protasis and the apodosis. This was consistent across all five speakers. Some also used the potential marker *e* in either the protasis, the apodosis or both.

(23) *He* ***e****-na-∅ plei, lo ge rrang* ***e-b*** *gurr mōl mōn*
if POT-1SG-IRR play then SUB blood POT-IRR flow back again
'If I played then the blood would flow again.' (ib1-lafet-na.27)

In sum, we get the picture that in North Ambrym, the counterfactual marker *to* is specific to the counterfactual past. The non-recent past marker also extends to the counterfactual past in conditionals with *to* in the protasis. The counterfactual future, like the possible future, is by default referred to by the irrealis marker. This is illustrated by Figure 11.

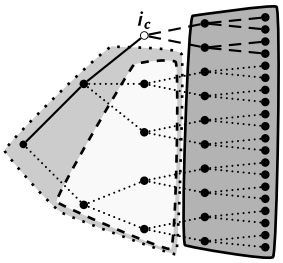

Figure 11. Our current best hypotheses about the domain of irrealis in North Ambrym. Solid outline: irrealis; dashed outline: counterfactual (past/present); dotted outline: non-recent past

North Ambrym thus presents as a language that prioritizes the temporal over the modal domain, although both dimensions are important in structuring the TAM systems. The binary divide between realis and irrealis modalities does not appear to play a major role.

5.4 Dalkalaen

Dalkalaen is the least described of the six subject languages. We have an as yet unpublished grammar sketch by von Prince based on intense fieldwork between 2009 and 2012. The language is closely related both to North Ambrym towards the north of the same island and to Daakaka towards the east. Its TAM system shows similarities with both its neighboring varieties. The core markers are shown in Table 7.

Table 7. The paradigm of core TAM markers in Dalkalaen as attested prior to this study

Gloss	Realization	Description
POS.REAL	*mV/=m/m=*	actual past or present
NEG.REAL	*to*	negative statements about the actual past/present
DIST	*tV*	non-recent past, counterfactual past, temporal clauses
IRR	*bV/-∅*	relative future
NEG.IRR	*=n*	prohibitives, negative futures

Conditionals about the possible future are formed with both the protasis and the apodosis in irrealis mood. Both parts are also accompanied by the clause-initial marker *ba*, which has a similar distribution to the North Ambrym potential marker described above. This is illustrated in (24):

(24) *Bone en ngae nga* ***ba*** *ko=∅ lene yaafu berep enti,* ***ba*** *s-amro*
when COMP COMP COMP POT 2SG=IRR marry man long this POT CL3-2DL.POSS
tejimre nye ***ba*** *ra=∅ ngor en* ***ba*** *ra=∅ ngor*
child PL POT 3PL big COMP POT 3PL big
'If you marry this tall man, you will have very many children.' (fortuneteller-am.16/17)

Counterfactual past conditionals were marked remarkably consistently across speakers: The protasis of each utterance is in potential mood, with a continuous-aspect auxiliary. The apodosis is introduced by the morpheme *bala*, or, in one out of ten cases, by *ba*; the apodosis is then marked by the distal.

(25) *a ko-∅* ***do*** *kirine yaafu mwermwer enti, ma kala lo* ***bala*** *s-amro*
COMP 2SG-IRR CONT follow man short this REAL say then CTF CL3-2DL
ver ***ti*** *fwe en ti fwe*
stone DIST many COMP DIST many
"'If you had married the short man", she said, "then you two would have been very rich."'
(fortuneteller-am.30/31)

(26) *ka na-∅* ***do*** *kirine ple futbol=an lo* ***bala*** *na* ***to*** *loo-koko*
COMP 1SG-IRR CONT follow play football=NMLZ then CTF 1SG DIST get.wet-full
'If I had joined the football game, I would have gotten soaked.' (lafet-am.17)

The morpheme *bala* has not been observed before. It was now found only in counterfactual contexts, mostly referring to the past. Two instances indicate that it might also be used in counterfactual future contexts. It did however not occur in the target contexts for the counterfactual future.

Conditionals about the counterfactual future pattern with those about the possible future: They are in irrealis mood and modified with the potential mood marker *ba*:

(27) ***ba*** *na-∅ kirine ple=an lo* ***ba*** *riy-ak* ***bo*** *rop kebu.*
POT 1SG-IRR follow play=NMLZ then POT blood.of-1SG.POSS IRR run back
'If I played [tomorrow], I would bleed again.' (lafet-am.22)

In sum, Dalkalaen uses irrealis and potential mood with reference to the possible and counterfactual future. Irrealis mood without the potential also occurs in the protasis of counterfactual-past conditionals. Distal mood is used with reference to the counterfactual past, particularly in the apodosis of conditionals. The counterfactual past is usually distinguished from other domains by the morpheme *bala*. This preliminary picture is illustrated in Figure 12. Dalkalaen also employs continuous aspect in the protasis of counterfactual conditionals.

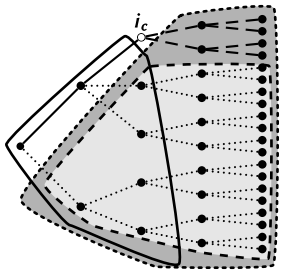

Figure 12. Our preliminary understanding of the Dalkalaen domain of irreality. Solid: distal; dotted: irrealis; dashed: *bala*

The binary distinction between realis and irrealis does appear to play a role in Dalkalaen, as well as the tripartite modal distinction between the actual, the counterfactual and the possible. The distal marker however cuts across two of these domains.

5.5 Daakaka

We have already seen an overview of the Daakaka TAM system in Section 2. Possible futures are referred to by the potential marker. The protasis of a possible-future conditional can be marked by either the potential marker or the distal:

(28) *Ko=**t*** *lene* *temeli* *man* *na* *ma* *waswas* *a* *veop,* *te* *nat-omaa*
2SG=DIST marry child male COMP REAL thin and long then child.of-2DL.POSS
nye *ka* *ye=**p*** *puo.*
PC ASR 3DL=POT many
'If you marry the skinny and tall boy, you'll have lots of children.'
(FortuneTeller_AN18/19)

More often than not, the protasis is introduced by the temporal/conditional subordinator *ka*. The apodosis is invariably marked by the homophonous assertion marker *ka* that expresses assertions about the future or possible present in combination with the potential mood.

(29) ***Ka*** *ko=**p*** *pwer* *myane* *na* *ma* *veop,* *waswas,* *te* *nat-omaa* *nyoo*
COMP 2SG=POT stay with COMP REAL long thin then child.of-2DL.POSS PL
ka *ya=**p*** *puo.*
ASR 3PL=POT many
'If you go with the tall one, the skinny one, you'll have many children'
(FortuneTeller_MT.19/20)

The counterfactual past/present is expressed by the distal marker, both in (the apodosis of) conditionals and in false belief reports. In counterfactual conditionals, the protasis is usually in potential mood, but the apodosis is always in distal mood. In many but not all cases, the apodosis is introduced by the formula *bili ka*. This is illustrated in (30):

(30) *ka* *ko=**p*** *pwer* *tevyan* *yaapu* *ente,* *te* ***bili*** ***ka*** *s-amaa* *mani* *nyoo*
COMP 2SG=POT stay with man this then time ASR CL3-3DL.POSS money PL
tu *puo.*
DIST plentiful
'If you had married this man, you two would have been rich.' (FortuneTeller_SB.038)

Counterfactual-future conditionals pattern with possible-future conditionals rather than with counterfactual-past ones, with the apodosis being in potential mood.

(31) *ka* *na=**t*** *ple* *volibol* *te* *volibol* *ka* ***we*** *me* *syute* *vy-ok*
COMP 1SG=DIST play volleyball then volleyball ASR POT come hit hand.of-1SG.POSS
te *myanok* *ente* *saka* *ne* *map.*
then wound this NEG.ASR NEG.POT heal
'If I played volleyball, the volleyball would hit my hand and then my wound wouldn't heal' (Lafet_AN.14/15)

The overall picture that we get is that the potential mood is responsible for references both to the possible and the counterfactual future. The distal marker expresses a reference to the counterfactual past and present in addition to the actual past. These values appear to be neutralized in the protasis of conditionals and we will remain agnostic at this point about the meaning of this observation. We will take the function of both markers in the apodosis as crucial for analyzing their meaning.[5] Figure 13 summarizes these conclusions.

Like in the other Ambrym languages discussed here, the temporal dimension is more dominant than the modal one in terms of obligatory distinctions. The system does support a tripartite view of modality, but does not revolve around the binary contrast between realis and irrealis.

[5] In von Prince (2017), it is assumed that the domain of the distal marker includes possible and counterfactual futures because of its occurrence in the protasis of corresponding conditionals. We do not exclude this analysis here but also consider the option that its reference may be shifted in these contexts.

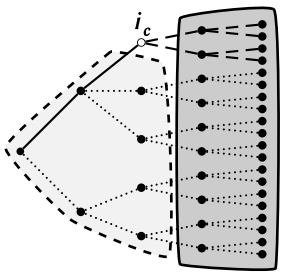

Figure 13. Our current hypotheses about the domain of irrealis in Daakaka. Solid: main domain of the potential marker; dashed: main domain of the distal marker

5.6 Daakie

The overall situation in Daakie is very similar to Daakaka. The TAM system of the language has been described in Krifka (2012, 2016). One interesting point of divergence is that, while in Daakaka the complementizer *ka* and the assertion marker *ka* are homophonous, their close Daakie counterparts are pronounced differently: The complementizer is also *ka*, but the future marker (similar in its distribution to the Daakaka assertion marker) is pronounced *a*.

(32) ***ka*** *ko=**p*** *lé-ne,* *s-amoo* *timaleh* *ngyee* ***a**=la=**p*** *pwee.*
COMP 2SG=POT marry-TRANS CL3-2DL.POSS child PL FUT=3PL=POT many
'If you marry him, you two will have many children' (Fortune_JackPaul)

The counterfactual past is marked by the distal in the apodosis and by a combination of the future marker *a* and the distal marker in the apodosis. This shows that *a* has to be understood as a relative future. It also marks another difference to Daakaka, where the assertion marker *ka* does not usually occur without the additional *bili* in connection with the distal.

(33) *ka* *ko=t* *lé-ne* *Adam,* *s-amoo* *vot* ***a-te*** *pwee.*
COMP 2SG=POT marry-TRANS Adam CL3-2DL stone FUT-3SG.DIST many
'If you had married Adam, you two would have had a lot of money.' (Fortune_JackPaul)

The counterfactual future is typically marked like the possible future, even though some speakers appear to hesitate between the potentialis and the distal in the protasis, as has also been observed in Daakaka:

(34) *ka* *na=p* *bwengbang* *ne* *volibol* *palen,* *manok* *ne* *baakon*
COMP 1SG=POT play TRANS volleyball tomorrow sore TRANS finger.of
velo-k ***a=bwe*** *top* *teteh*
hand-1SG FUT=3SG.POT break again
'If I played volleyball tomorrow, the sore on my finger would come open again.'
(Fortune_JackPaul)

6 Conclusions

We have seen remarkable variation in how our subject languages carve up the temporal modal domain. Only Mavea and Nafsan conformed to our expectation that counterfactual-future contexts would pattern with the counterfactual past rather than with the possible future. In the languages of North and West Ambrym, by contrast, they align with the possible future instead. This finding is summarized in Table 8.

In the Ambrym languages, we find that the vertical, temporal dimension is emphasized over the diagonal, modal one. They also tend to have more fine-grained distinctions, especially as we

Table 8. Our prior knowledge updated with new conclusions from the storyboard data

Language	Actual past/present	Possible future	Count. past	Count. future
Mavea	SUBJ.(REAL)	SUBJ.(IRR)	(IRR), (*imte*)	(IRR), (*imte*)
Nafsan	∅	SUBJ.IRR	IRR, (*mer*)	IRR, (*mer*)
Daakie	REAL/DIST	POT	DIST	POT
Daakaka	REAL/DIST	POT	DIST	POT
Dalkalaen	REAL/DIST	POT	DIST, *bala*	POT
North Ambrym	REC.PST/ NREC.PST	IRR	CTF	IRR

move along to coast towards the west of the island (Dalkalaen) and to the north (North Ambrym). This is illustrated in Figure 14.

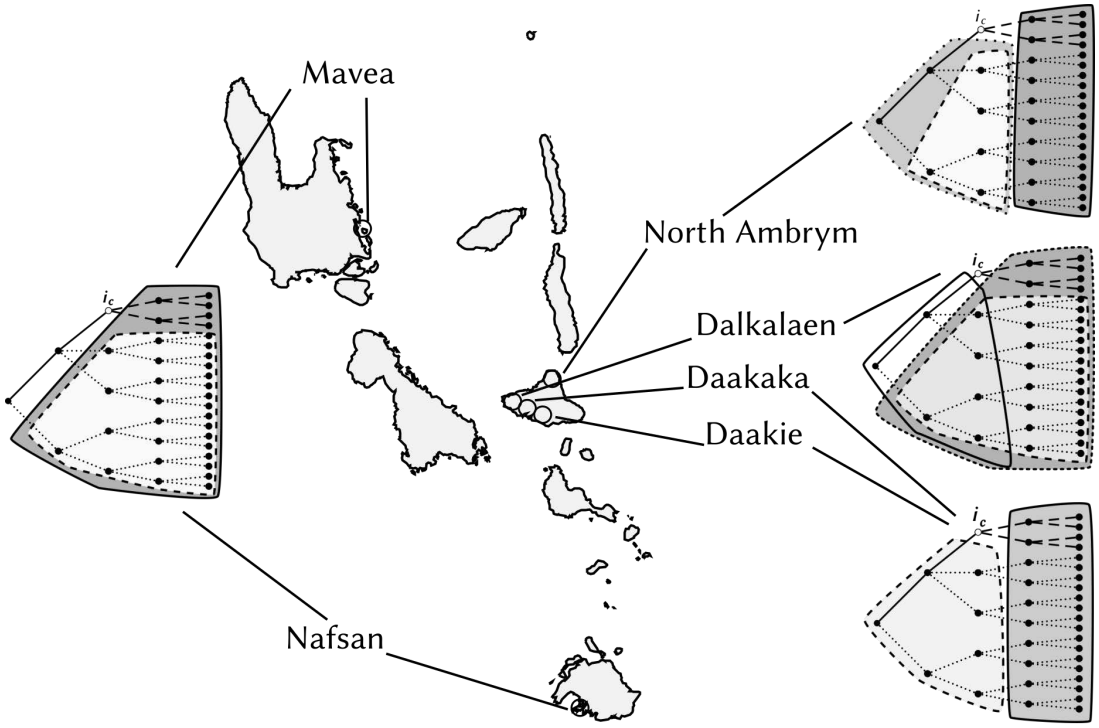

Figure 14. The spatial distribution of different systems around Vanuatu

The variation we find even between the very closely related languages from Ambrym speaks to the apparent volatility of TAM systems. This is also reflected by the observation that the distinction between counterfactual and possible contexts in Mavea and possibly Nafsan appears to be subject to diachronic change: In Mavea, the use of *imte* is restricted to older speakers. In Nafsan, the use of *mer* in counterfactual contexts is more frequent in our elicitations from young speakers than in corpus data from the 1990s to 2006.

We have seen that the distinction between realis and irrealis is an important factor for some, but not all of our subject languages. We have also seen considerable evidence for the distinction of the irrealis domain into the counterfactual and the possible. Out of our six subject languages, three use a combination of irrealis morphology and an optional counterfactuality marker, which is not restricted in terms of temporal reference. North Ambrym has a specific expression dedicated to the counterfactual past.

In contrast to generalizations by Iatridou (2000) and Ferreira (2016), we did not find imperfective aspect to play a major role in marking counterfactual meanings, except for Dalkalaen,

where it was used consistently in corresponding contexts.

Our research highlights the need to look more closely at the subdomains of irrealis mood to better understand TAM systems. It also shows that even relatively mood-prominent languages can in some subdomains prioritize temporal over modal reference. And that there are significant variation and indications of rapid diachronic change in the TAM systems of Oceanic languages that is far from being well-understood.

Acknowledgments

We would like to thank the speakers of Mavea, North Ambrym, Dalkalaen, Daakaka, Daakie and Nafsan, who have worked with us to find out more about each language. We also thank the participants of the MelaTAMP workshop in Port Vila, Vanuatu in 2017 and Robert Early for hosting this event. The reviewers for the Linguistic Evidence conference and the audience of our poster have also helped us with their comments and questions. This work has been funded by the DFG, as part of the MelaTAMP project (273640553).

References

Burton, S., & Matthewson, L. (2015). Targeted construction storyboards in semantic fieldwork. In R. Bochnak & L. Matthewson (Eds.), *Methodologies in Semantic Fieldwork* (pp. 135–156). Oxford University Press.

Bybee, J. L. (1998). "Irrealis" as a grammatical category. *Anthropological Linguistics*, *40*(2), 257–271.

Bybee, J. L., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. The University of Chicago Press.

Condoravdi, C. (2002). Temporal interpretation of modals: Modals for the present and for the past. In D. Beaver, L. Casillas, B. Clark & S. Kaufmann (Eds.), *The construction of meaning* (pp. 59–88). Stanford, CA: CSLI Publications.

Cristofaro, S. (2012). Descriptive notions vs. grammatical categories: Unrealized states of affairs and 'irrealis'. *Language Sciences*, *34*(2), 131 – 146. Papers selected from the 'What do languages encode when they encode reality status?' workshop at the 41st Annual Meeting of the Societas Linguistica Europaea, Forlì, Italy, 17-20 September 2008.

Dowty, D. (1977). Toward a semantic analysis of verb aspect and the English 'imperfective' progressive. *Linguistics and Philosophy*, *1*, 45–77.

Druskat, S. (2018). ToolboxTextModules (Version 1.1.0).

Elliott, J. (2000). Realis and irrealis: Forms and concepts of the grammaticalisation of reality. *Linguistic Typology*, *4*, 55–90.

Ferreira, M. (2016). The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. *Natural Language Semantics*, *24*(4), 353–397.

Franjieh, M. (2012). *Possessive Classifiers in North Ambrym, a Language of Vanuatu: Explorations in Semantic Classification*. PhD thesis, The School of Oriental and African Studies.

Guérin, V. (2011). *A grammar of Maʷ̌ea: An Oceanic language of Vanuatu*. Honolulu: University of Hawai'i Press.

de Haan, F. (2012). Irrealis: fact or fiction? *Language Sciences*, *34*(2), 107 – 130. Papers selected from the 'What do languages encode when they encode reality status?' workshop at

the 41st Annual Meeting of the Societas Linguistica Europaea, Forlì, Italy, 17-20 September 2008.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, *31*(2), 231–270.

Ippolito, M. (2013). *Subjunctive conditionals: a linguistic analysis*, volume 65 of *Linguistic Inquiry Monograph*. Massachussetts Institute of Technology.

Krajinović, A., & Thieberger, N. (2018). Diachrony of conditionals in nafsan. Talk given at the Vanuatu Languages Conference, Port Vila.

Krause, T., & Zeldes, A. (2016). ANNIS3: A new architecture for generic corpus query and visualization. *Digital Scholarship in the Humanities*, *31*(1), 118–139.

Krifka, M. (2012). Notes on Daakie (Ambrym, Vanuatu): Sounds and modality. In Clemens, L. E., Scontras, G., & Polinsky, M. (Eds.), *Proceedings of AFLA 18 (Austronesian Formal Linguistics Association)*, (pp. 46–65)., Cambridge, Mass. Harvard University. `http://westernlinguistics.ca/afla/proceedings/afla18/index.htm`.

Krifka, M. (2016). Realis and non-realis modalities in Daakie (Ambrym, Vanuatu). In Moroney, M., Little, C.-R., Collard, J., & Burgdorf, D. (Eds.), *Proceedings of SALT 26*, (pp. 546–565). Linguistic Society of America.

Laca, B. (2012). On modal tenses and tense modals. In C. Nishida & C. Russi (Eds.), *Building a bridge between the linguistic communities of the Old and the New World. Current research in tense, aspect, mood and modality*, number 25 in Cahiers Chronos (pp. 171–194). Amsterdam: Rodopi.

Lichtenberk, F. (2016). Modality and mood in Oceanic. In J. Nuyts & J. van der Auwera (Eds.), *The Oxford Handbook of Mood and Modality* chapter 14, (pp. 330–361). Oxford: Oxford University Press.

Longacre, R., & Thompson, S. A. (1985). Adverbial clauses. *Language typology and syntactic description*, *2*, 171–234.

von Prince, K. (2015). *A Grammar of Daakaka*. Berlin, Boston: De Gruyter Mouton.

von Prince, K. (2017). Paradigm-induced implicatures in TAM-expression: A view from the Daakaka distal. In *Proceedings of Sinn und Bedeutung 21*, Edinburgh.

von Prince, K. (2018a). Bananas (MelaTAMP storyboards). *Zenodo*. `https://doi.org/10.5281/zenodo.1230385`.

von Prince, K. (2018b). Fat pig (MelaTAMP storyboards). *Zenodo*. `https://doi.org/10.5281/zenodo.1231798`.

von Prince, K. (2018c). Festival (MelaTAMP storyboards). *Zenodo*. `https://doi.org/10.5281/zenodo.1231804`.

von Prince, K. (2018d). Red yam (MelaTAMP storyboards). *Zenodo*. `https://doi.org/10.5281/zenodo.1231808`.

von Prince, K. (2018e). Tomato and pumpkin (MelaTAMP storyboards). *Zenodo*. `https://doi.org/10.5281/zenodo.1231836`.

von Prince, K. (2019). Counterfactuality and past. *Linguistics and Philosophy*.

Rolka, M., & Cable, S. (2010). Tom and Mittens. *Totem Field Storyboards*. Retrieved from http://www.totemfieldstoryboards.org on Sep 21, 2017.

TFS Working Group (2010). The fortune teller. *Totem Field Storyboards*. Retrieved from http://www.totemfieldstoryboards.org on Sep 21, 2017.

TFS Working Group (2011a). Chore girl. *Totem Field Storyboards*. Retrieved from http://www.totemfieldstoryboards.org on Sep 21, 2017.

TFS Working Group (2011b). The woodchopper. *Totem Field Storyboards*. Retrieved from http://www.totemfieldstoryboards.org on Sep 21, 2017. Illustrated by Katie Sardinha.

Thieberger, N. (2006). *A grammar of South Efate: an Oceanic language of Vanuatu*. Number 33 in Oceanic Linguistics Special Publications. Honolulu: University of Hawai'i Press.

Thomason, R. H. (1970). Indeterminist time and truth-value gaps. *Theoria*, *36*(3), 264–281.

Thomason, R. H. (1984). Combinations of tense and modality. *Handbook of philosophical logic*, *2*, 135–165.

Vander Klok, J. (2013). Bill vs. the weather. *Totem Field Storyboards*. Retrieved from http://www.totemfieldstoryboards.org on May 31, 2018.

Zipser, F., & Romary, L. (2010). A model oriented approach to the mapping of annotation formats using standards. In *Proceedings of the Workshop on Language Resource and Language Technology Standards*.