UNIVERSITY OF CALIFORNIA
Los Angeles

Aspect and evidentiality

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Linguistics

by

Margit Lia Bowler

2018

© Copyright by
Margit Lia Bowler
2018

ABSTRACT OF THE DISSERTATION

Aspect and evidentiality

by

Margit Lia Bowler
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 2018
Professor Yael Sharvit, Chair

A number of unrelated languages have portmanteau morphemes with both temporal and evidential meanings (e.g. Cherokee (Pulte 1985), Kalaallisut (Fortescue 2003), Tariana (Aikhenvald 2004), and Turkish (Slobin and Aksu 1982), among many others). These temporal and evidential meanings are strikingly consistent, cross-linguistically. Evidentiality frequently co-occurs as part of other grammatical categories (Aikhenvald 2018, 2004); however, most theories of evidentiality do not discuss its connection with other kinds of meaning. In this dissertation, I provide a formal account of the connection between temporal and evidential meanings through a case study of a set of portmanteau tense/aspect and evidential (TAE) morphemes in Tatar (Turkic). Part 1 of the dissertation provides in-depth description of the evidential and temporal interpretation of the Tatar TAE morphemes. This Tatar data was collected through original fieldwork on the language. Part 2 analyzes the Tatar TAE morphemes as having underlyingly temporal semantics. I propose that the Tatar TAE morphemes assert temporal meaning, and pragmatically implicate evidential meaning. I accomplish this by positing a tripartite event ontology (following Moens and Steedman 1988) in which events are preceded by contingently related pre-states and are followed by contingently related event post-states. I propose that the evidential meanings of the Tatar TAE morphemes arise as a "byproduct" of their ability to view events from within the runtimes of their pre- and post-states. I show that the evidential readings associated with the Tatar TAE morphemes are cancellable in some discourse contexts, and

completely absent in others. This suggests that a pragmatic mechanism is needed to account for their use. This dissertation joins and formalizes intuitions about event pre- and post-states, causality, and evidentiality that have been previously described in both the temporal and evidential literatures (e.g. Comrie 1976, Nikolaeva 1999, and Bashir 2006, among others). In doing so, I account for both the temporal and evidential contributions of the Tatar TAE morphemes.

The dissertation of Margit Lia Bowler is approved.

Roumyana Pancheva

Pam Munro

Jessica Rett

Yael Sharvit, Committee Chair

University of California, Los Angeles

2018

For my mom, dad, and Marilyn

TABLE OF CONTENTS

1 Introduction **4**

1.1 What this dissertation is about 4

1.1.1 Structure of the dissertation 7

1.2 Background on Tatar and fieldwork methodology 9

I The Tatar tense/aspect and evidential system **12**

2 Evidential contributions of the Tatar TAE suffixes **13**

2.1 Introduction 13

2.1.1 Evidential terminology 16

2.2 Evidential interpretation of the TAE suffixes in matrix clauses 18

2.2.1 Evidentiality in the Tatar past-oriented TAE suffixes 19

2.2.2 Evidentiality in the Tatar future-oriented TAE suffixes 29

2.2.3 Negation and the TAE suffixes: Lack of embedding 37

2.2.4 Recap: Evidential interpretation of TAE suffixes in matrix clauses . . 39

2.3 Evidential interpretation of semantically embedded TAE suffixes 40

2.3.1 Clausal embeddings 41

2.3.2 Evidentiality in questions: Interrogative flip 55

2.3.3 Recap: Evidential contribution of Tatar TAE suffixes in semantically embedded environments 59

2.4 Conclusion 60

3 Temporal contributions of the Tatar TAE suffixes **61**

3.1 Introduction 61

3.2 Temporal interpretation of TAE suffixes in matrix clauses 64
3.2.1 Past-oriented TAE suffixes 64
3.2.2 Present- and future-oriented TAE suffixes 70
3.2.3 Recap: Temporal interpretation of TAE suffixes in matrix clauses . . . 78
3.3 Temporal interpretation of semantically embedded TAE suffixes 78
3.3.1 Temporal interpretation of TAE suffixes in embedded clauses 79
3.3.2 Embedding TAE suffixes under *ide* 'PST' 91
3.3.3 Recap: Temporal interpretation of TAE suffixes in semantically embedded environments 97
3.4 Conclusion 97

II An aspectual proposal for evidentiality 100

4 Evaluating the Tatar data against prior theories of evidentiality 101
4.1 Introduction 101
4.1.1 Semantic variables and types 102
4.2 Preliminary discussion: Evidentials as epistemic modals 103
4.2.1 Basic epistemic modal analysis of evidentiality 103
4.2.2 Tests in the literature for modality in evidential expressions 105
4.2.3 Recap 120
4.3 Existing analyses of evidential systems that are morphosyntactically related to tense/aspect 121
4.3.1 Evidence Acquisition Time analyses of evidentiality 122
4.3.2 Recap 140
4.3.3 Trace analyses of evidentiality 141
4.3.4 Recap 147

4.4 Conclusion and foreshadowing . . . 150

5 Deriving the temporal and evidential interpretations of the Tatar TAE suffixes in matrix clauses . . . 154

5.1 Introduction . . . 154

5.2 Temporal ontology . . . 155

5.2.1 Moens & Steedman (1988) . . . 156

5.2.2 The ontology that I assume . . . 159

5.3 The link between speaker perception of pre- and post-states and evidentiality . . . 162

5.3.1 Unformalized application to *-GAn* 'RESULT' . . . 165

5.3.2 Unformalized application to *-(y)AçAK* 'PROSP' . . . 167

5.3.3 Recap . . . 169

5.4 Semantics and pragmatics of the Tatar TAE morphemes . . . 172

5.4.1 Present- and future-oriented TAE morphemes . . . 176

5.4.2 Past-oriented TAE morphemes . . . 187

5.4.3 Evaluating my theory against some proposed tests for modality in the Tatar TAE suffixes . . . 196

5.5 Conclusion . . . 202

6 Deriving the temporal and evidential interpretations of the Tatar TAE suffixes in clausal embeddings . . . 204

6.1 Introduction . . . 204

6.2 Tatar TAE suffixes in embedded CPs . . . 205

6.2.1 Updated denotations . . . 208

6.2.2 Narrow scope readings of future oriented TAE suffixes in embedded CPs . . . 211

6.2.3 Narrow scope readings of past oriented TAE suffixes in embedded CPs . . . 224
6.2.4 Wide scope readings of past oriented TAE suffixes in embedded CPs 230
6.2.5 Recap . . . 238
6.3 Tatar TAE suffixes in embedded verbal nominalizations . . . 239
6.3.1 *-(y)AçAK* in embedded verbal nominalizations . . . 242
6.3.2 *-GAn* in embedded verbal nominalizations . . . 244
6.3.3 Recap . . . 249
6.4 Conclusion . . . 250

7 Conclusion . . . 251
7.1 What this dissertation does and does not do . . . 251
7.1.1 Data this dissertation accounts for . . . 251
7.1.2 Data this dissertation does not account for . . . 253
7.2 Cross-linguistic observations and predictions . . . 256
7.2.1 Languages with tense/aspect pairs . . . 256
7.2.2 Languages without tense/aspect pairs . . . 257
7.3 Future directions . . . 259

LIST OF TABLES

1.1 Temporal and evidential meanings associated with the core set of Tatar TAE suffixes addressed in this dissertation. 7

2.1 Inventory of TAE morphemes in Tatar. 15

2.2 Tatar subject agreement paradigms and free (nominative) pronouns. 15

2.3 Core set of Tatar TAE suffixes addressed in this dissertation. 19

2.4 Core set of Tatar TAE morphemes addressed in this dissertation (repeated from Table 2.3). 40

2.5 Semantic interpretations of embedded clause types in Tatar. 42

2.6 Available evidential interpretations of the Tatar TAE suffixes in embedded CPs. 45

2.7 Grammaticality and evidential interpretation of the Tatar TAE suffixes in embedded verbal nominalizations. 52

2.8 Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings. 60

3.1 Temporal contributions of the Tatar TAE suffixes. 62

3.2 Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings. 80

3.3 Grammaticality of the Tatar TAE suffixes in combination with *ide* 'PST.' 92

3.4 Comparison of my description of the Tatar TAE suffixes to Greed (2014), Tatevosov (2007), and Poppe (1961). 98

4.1 Semantic variables and types used in this dissertation. 102

4.2 Summary of semantic components of Smirnova (2013)'s analysis of Bulgarian evidentiality. 132

4.3 Summary of semantic components of Koev (2011)'s analysis of Bulgarian evidentiality. 137

4.4 Components of Koev (2017)'s trace analysis of evidentiality in Bulgarian (Koev 2017, 17). 145

4.5 Temporal meanings of the Tatar TAE suffixes and their associated evidential interpretations, as motivated in Chapter 5. 151

5.1 Interpretation of the Tatar TAE suffixes in semantically unembedded environments. 173

6.1 Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings, where "shift" refers to evidential shift. 204

6.2 Grammaticality and evidential interpretation of the Tatar TAE suffixes in embedded verbal nominalizations (repeated from Chapter 2). 240

ACKNOWLEDGMENTS

I'm lucky to have so many wonderful people to thank. Still, I'll try to keep this short.

I am honored to have Yael, Jessica, Pam, and Roumi on my dissertation committee. Thank you, Yael, for believing in me and putting up with all of my questions, for meeting with me many (many!) times over Skype, often on short notice and across many time zones, and for being very understanding of my struggles at maintaining a work/life/fieldwork balance throughout grad school. Thank you to Jessica for thorough, helpful critiques of my drafts and handouts (of this project, and of several others), and for being such a supportive member of the Semantics Tea. Thank you to Pam for introducing me to all kinds of new and wonderful data in AIS, for fieldwork advice, and for always pressuring me to be as clear as possible (in this dissertation, and in many past AIS talks). Thank you to Roumi for the encouragement to pursue this topic; it means a lot, coming from you! My committee's feedback on earlier versions of this project have guided and improved it immensely. Thank you, thank you!

I could not have written this dissertation if it were not for the many hours spent working with my primary Tatar consultant, Sofia Mazgarova. (Sorry we talked about Michael Phelps so much.) Sofia is one of the smartest and most interesting people I know, and she has taught me so much about Tatar (and Russian) language and culture—I can't thank her enough! I also worked with a number of other wonderful Tatar consultants in LA and in Burlingame; huge, huge thanks to Alina Salikhova Bailey, Guzel Ahmadullina, Zuhra Salahutdinova, Farida Sulikova, Rukiye Safa, Deyyan Safa, Aygol Gimranova Lyon, Alfiya Safiullina, and Ahmet Mazgarov. Räxmät! The Tatar community overall is so welcoming and friendly; I feel so lucky to be involved in it.

My favorite thing about the UCLA Linguistics department is how supportive the grad students are of one another. There are many, many people in the department who have helped me in one way or another during my time at UCLA, and I can't thank them enough. First, thank you to my psycholinguistics lab officemates over the years; I've loved working alongside all of you! Thanks to Deborah Wong, Sözen Özkan, Meaghan Fowlie, John

Gluckman, Iara Mantenuto, Marju Kaps, Nikos Angelopoulos, Rachel Jacobson, and Victoria Mateu. Thanks to Maayan Abenina-Adar and Jos Tellings for being faithful Semantics Tea attendees; I've learned a lot from you both. Thanks to Adam Royer and Scott Gaines for being great pals. Thank you to Lauren Winans for helping me get settled in to the department/LA when I first arrived, and to the other UCLA Linguistics alums who were role models for me in my early years here: Laura McPherson, Laura Kalin, Natasha Korotkova, Yu Tanaka, and Vanya Kapitonov. Thanks also to my friends in the Indo-European department for Sunday dinners (and Mr. Jingles): Chiara Bozzone, Ryan Sandell, and Tony and Sam Yates.

I would like to give extra-special thanks to Travis Major for introducing me to Tatar in the first place! Travis and I co-taught a summer sessions field methods course in 2016, and he suggested that we work on Tatar. This dissertation wouldn't exist if it weren't for him! Thanks also to Maura O'Leary for Tatar data discussion, and to Ani Alaberkyan for helping me collect and organize data.

I would like to especially thank Jesse Zymet for being an endless source of laughter, warmth, and groundedness, especially during the first few years at UCLA. I absolutely could not have gotten through this PhD without him.

Thanks to my parents for supporting my weird interest in languages for so many years now, even when that interest takes me far away from home. I'm looking forward to bringing back more stories from Australia—and now, the UK—in the future. Thank you to my friends and family in Oregon for being there for me throughout grad school: Maren, Maura, Dan, Colleen, and the Swearingens.

Even though this dissertation is not about Warlpiri, my most formative fieldwork trips were spent working on Warlpiri in central Australia. I miss my Australian friends very much (Cecilia, Gloria, Jono, Fiona, Otto, and too many others), and I hope to get back to Australia as soon as I can.

My graduate studies at UCLA were supported in part by a NSF Graduate Research Fellowship, a Charles E. & Sue K. Young Graduate Student Award, and a Dissertation Year

Fellowship. I am also very grateful to the UCLA Linguistics department for a Ladefoged Scholarship that funded one of my fieldwork trips to Australia.

Finally, above all, I'd like to thank John Gluckman, for being a great partner in all things—in fieldwork, and in life. You're simply the best.

VITA

2011 B.A. in Linguistics, Reed College

2012 Fulbright Scholar, Australian National University

2014 M.A. in Linguistics, University of California, Los Angeles

2016 Ph.D. candidate in Linguistics, University of California, Los Angeles

PUBLICATIONS

Bowler, Margit & Ivan Kapitonov. To appear. "Quantification in Australian languages." In Claire Bowern (ed.), *Oxford Handbook of Australian Languages*. Oxford University Press.

Bowler, Margit & John Gluckman. 2018. "Intensifying ideophones in three Luhya languages." In Elizabeth Bogal-Allbritten and Elizabeth Coppock (eds.), *Proceedings of The Semantics of African, Asian and Austronesian Languages (AAA) 4*: 31-47.

Bowler, Margit & Sözen Özkan. 2018. "Turkish *-mIş* does not contribute (English-type) present perfect." In Céleste Guillemot, et al. (eds.), *Proceedings of the 13th Workshop on Altaic Formal Linguistics (WAFL 13)*: 115-126. MIT Working Papers in Linguistics.

Bowler, Margit. 2017. "Quantification in Warlpiri." In Edward Keenan and Denis Paperno (eds.), *Handbook of Quantifiers in Natural Language, Volume II*: 963-994. Springer.

Gluckman, John & Margit Bowler. 2016. "Expletive agreement, evidentiality, and modality

in Logooli." In Mary Moroney, et al. (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 26*: 1063–1082.

Gluckman, John & Margit Bowler. 2016. "The anticausative alternation in Luragooli." In Doris Payne, et al. (eds.), *Diversity in African Languages: Selected Papers from the 46th Annual Conference on African Linguistics (ACAL)*: 271–289.

Bowler, Margit. 2016. "The status of degrees in Warlpiri." In Mira Grubic & Anne Mucha (eds.), *Proceedings of The Semantics of African, Asian and Austronesian Languages (AAA) 2*: 1-17.

Bowler, Margit. 2015. "'Might be something': Information protection in central Australia." In Ksenia Ershova, et al. (eds.), *Proceedings of Chicago Linguistic Society (CLS) 51*: 33-45.

Bowler, Margit. 2015. "Evaluative morphology in Warlpiri." In N. Grandi and L. Körtvélyessy (eds.), *The Handbook of Evaluative Morphology*: 438-447. Edinburgh University Press.

Bowler, Margit. 2014. "Conjunction and disjunction in a language without 'and.'" In Todd Snider, et al. (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 24*: 137-155.

Rouhizadeh, Masoud, Margit Bowler, Richard Sproat & Bob Coyne. 2011. "Collecting semantic data from Amazon's Mechanical Turk for a lexical knowledge resource in a text to picture generating system." *9th International Conference on Computational Semantics*: 380-384.

Rouhizadeh, Masoud, Margit Bowler, Richard Sproat, & Bob Coyne. 2010. "Data collection and normalization for the scenario–based lexical knowledge resource of a text-to-scene conversion system." *2010 Fifth International Workshop on Semantic Media Adaptation and Personalization*: 25-30.

ORTHOGRAPHY

The Tatar language has historically been written in a number of different orthographies: Arabic (until the 1920s), Latin (1920s-1930s), and Cyrillic (1930s-present). Tatar speakers in Tatarstan currently write Tatar using Cyrillic, while emigrant communities outside of Tatarstan generally use Latin scripts. (See Sebba 2006 for a history of the orthographies used for Turkic languages in the former Soviet Union.)

All of the Tatar data in this dissertation is written in a Latin orthography that roughly follows Turkish spelling conventions, with the addition of a small number of characters representing Tatar phonemes that are not found in Turkish. In Figure 0.1, I provide a list of the Latin characters that I use with their Cyrillic equivalents (I note that the Tatar Cyrillic script does not distinguish between all of the phonemes of the language). I omit some sound sequences that are written using single characters in Cyrillic but are represented using digraphs in Latin (e.g. Cyrillic ю, Latin *yu* [ju]; Cyrillic ё, Latin *yo* [jo]). In the Tatar Cyrillic script, these characters are only used in loanwords from Russian.

I also provide IPA transcription in Figure 0.1 of the phonemes that these symbols roughly represent. Tatar has significant featural harmony, and allophones are not represented in this orthography. There are also dialectal differences with respect to the pronunciation of some characters; for instance, the sound that I indicate using *ç* is pronounced as [tɕ], rather than [ɕ], by speakers of Mishar Tatar. The phonemic inventory that I give is based on the dialect of Kazan Tatar spoken by my primary consultant as well as on Comrie (1997). See Comrie (1997) and Poppe (1961) for much more information on Tatar phonology.

Finally, I follow Turkicist convention in this dissertation by capitalizing underspecified segments that are subject to featural harmony. For example, I write the underlying form of one of the suffixes that I discuss as "*-GAn*." This is because this suffix can take a number of different phonological forms in which its first two segments differ, depending on the phonological properties of the verb root it combines with: *-qan, -kän, -gän*, and so on.

Latin uppercase	Latin lowercase	Cyrillic uppercase	Cyrillic lowercase	IPA
A	a	А	а	[ɑ]
Ä	ä	Ә	ә	[æ]
B	b	Б	б	[b]
V	v	В	в	[v]
G	g	Г	г	[g]
Ğ	ğ	Г	г	[ɣ]
D	d	Д	д	[d]
E	e	Е	е	[e]
J	j	Ж	ж	[ʒ]
C	c	Җ	җ	[ʑ]
Z	z	З	з	[z]
İ	i	И	и	[i]
Y	y	Й	й	[j]
K	k	К	к	[k]
Q	q	К	к	[q]
L	l	Л	л	[l]
M	m	М	м	[m]
N	n	Н	н	[n]
Ñ	ñ	Ң	ң	[ŋ]
O	o	О	о	[o]
Ö	ö	Ө	ө	[ø]
P	p	П	п	[p]
R	r	Р	р	[r]
S	s	С	с	[s]
T	t	Т	т	[t]
U	u	У	у	[u]
Ü	ü	Ү	ү	[y]
F	f	Ф	ф	[f]
X	x	Х	х	[x]
H	h	Һ	һ	[h]
Ç	ç	Ч	ч	[ɕ]
Ş	ş	Ш	ш	[ʃ]
I	ı	Ы	ы	[ɤ]

Figure 0.1: Orthographic conventions used for the Tatar data in this dissertation.

ABBREVIATIONS

I use the following abbreviations in this dissertation. This list does not include abbreviations that are used by other authors in examples that I cite; I refer the reader to the original sources for information on these abbreviations. For more information on Tatar grammar, see Poppe (1961).

1	first person
2	second person
3	third person
ABL	ablative
ACC	accusative
ADD	additive
COMP	complementizer
DAT	dative
EVID	evidential
EXIST	existential
FUT	future
GEN	genitive
INF	inferential
INFIN	infinitive
LOC	locative
NEG	negation
NOM	nominative
NPST	nonpast
PST	past
PL	plural
POSS	possessive
PRES	present
PROSP	prospective aspect
Q	question
RESULT	resultative aspect
SG	singular

CHAPTER 1

Introduction

1.1 What this dissertation is about

In the most general terms, this dissertation is about how the way that we locate ourselves in time relative to an event can determine what we know about that event. In doing so, this dissertation links two not obviously related components of grammar: **tense/aspect** and **evidentiality**. I examine these topics through the lens of portmanteau tense/aspect and evidential (TAE) morphemes in **Tatar** [ISO: tat], a Turkic language spoken primarily in Russia by approximately 5 million people (Simons and Fennig 2018). (For more information on Tatar, see §1.2.)

Tenses and aspects let us locate events in time. For the purpose of this discussion, we can assume Comrie (1985, 1976)'s descriptive definitions of tense and aspect. Comrie describes tenses as locating the event time relative to some other time, typically the utterance time. Aspects relate the event time to a reference time from which it is "viewed"; this reference time can be located e.g. within the runtime of the event, or in a pre- or post-state of the event, as I propose in this dissertation.[1] There is a long descriptive and theoretical tradition of the study of how languages grammatically encode tense and aspect (Reichenbach 1947; Comrie 1976, 1985; Klein 1994; Dahl 1985, 2000; among many others).

Evidentials indicate a speaker's "way of knowing" the proposition that they are reporting. The most basic evidential distinction is between direct and indirect evidence; that is, whether the speaker has eyewitness evidence for the truth of the proposition they

[1] I use the term "aspect" in this dissertation to refer to what other authors have termed grammatical aspect or viewpoint aspect, as opposed to lexical aspect or Aktionsart (in the sense of e.g. Vendler 1957). I provide a formal definition of event pre- and post-states in Chapter 5.

are reporting ("direct evidence"), or whether they are inferring its truth indirectly ("indirect evidence"). Evidential meanings are noted for their cross-linguistic tendency to co-occur as part of other grammatical categories, including complementizers (Noonan 1985), tense/aspect systems (Izvorski 1997), and noun class marking (Gluckman and Bowler 2016), among others (Aikhenvald 2018, 2004).

I give an example of a portmanteau tense/aspect and evidential Tatar morpheme in (1).[2] In the expression in (1), the verbal suffix *-GAn* appears to descriptively contribute two different kinds of meaning.[3] First, it indicates that the event of Mansur going to Moscow happened in the past. This component of meaning belongs to the domain of tense/aspect. Second, it indicates that the speaker has indirect evidence for this event (i.e., they did not witness the event of Mansur going to Moscow). This component of meaning belongs to the domain of evidentiality.

(1) **Context:** You found a train ticket to Moscow in Mansur's desk. You infer from this that Mansur traveled to Moscow. You say:

Mansur Mäskäü-gä bar-**ğan**-∅.
Mansur Moscow-DAT go-GAN-3SG
'[I have indirect evidence that] Mansur went to Moscow.'

At first glance, this connection between tense/aspect and evidentiality in Tatar may appear to be random. However, portmanteau TAE morphemes are documented in a range of unrelated languages worldwide. What's more, these languages all display the **same kinds** of related temporal and evidential meanings.[4] For instance, past-oriented morphemes are typically described as encoding a distinction between direct and indirect evidentiality. I provide examples of more portmanteau TAE morphemes in Kalaallisut (also called West

[2]Here and elsewhere in the dissertation, I indicate the morpheme(s) that I am discussing in bold.

[3]I follow Turkicist convention by capitalizing underspecified segments that are subject to featural harmony.

[4]This empirical fact is puzzling, since evidentials should in theory be able encode any kind of way of knowing. For example, we might expect to find evidentials that indicate "evidence obtained through the telephone," or "evidence obtained through psychic premonition," and so on. However, the set of cross-linguistic evidential meanings that we find is restricted to four or five categories, at most (Aikhenvald 2004; see also Kalsang et al. 2013 for discussion on this topic).

Greenlandic) and Cherokee in (2)-(3). In both examples, the relevant morphemes also mark (i) that the event happened in the past, and (ii) that the speaker has indirect evidence for it. In this sense, they are analogous to Tatar *-GAn* in (1).

(2) KALAALLISUT (ESKIMO-ALEUT) (modified from Fortescue 2003, 293)

Context: The speaker goes outside and sees a pool of water.

siallir-**sima**-vuq.
rain-SIMA-3SG.INDIC
'[I have indirect evidence that] It rained.'

(3) CHEROKEE (IROQUOIAN) (modified from Aikhenvald 2004, 26-27)

Context: The speaker wakes up, looks out, and sees puddles of water.

u-gahnan-**eʔi**.
it-rain-EʔI
'[I have indirect evidence that] It rained.'

In addition to languages with portmanteau TAE morphemes, a significant number of other languages have morphologically complex expressions (e.g. perfects) that express both temporal and evidential meanings (e.g. Bulgarian (Izvorski 1997), Korean (Lee 2013)). Strictly speaking, these languages do not have portmanteau morphemes conveying both temporal and evidential meanings; however, their tense/aspect and evidential systems are clearly related.

Languages with related tense/aspect and evidential systems (including both those with portmanteau TAE morphemes and with morphologically complex TAE expressions) include (but are not limited to) Aymara (Klose 2014), Azerbaijani (Öztopçu 2000), Bashkir (Poppe 1964), Bulgarian (Izvorski 1997, Sauerland and Schenner 2007, among others), Cherokee (Pulte 1985), Cuzco Quechua (Faller 2004), English (Winans 2016), Gagauz (Pokrovskaya 1964), Kalaallisut (Fortescue 2003), Kalasha (Bashir 2006), Karachay (Seegmiller 1996), Karaim (Mysaev 1964), Khanty (Nikolaeva 1999), Kumyk (Johanson and Csató 1998), Kyrgyz (Abduldaev and Zakharova 1987), Macedonian (Friedman 1986), Matses (Fleck 2007), Norwegian (Izvorski 1997), Persian (Simeonova and Zareikar 2015), Salar (Dwyer 2000), Slovenian (Rivero and Sheppard 2015), Spanish (Rivero 2014), Tariana (Aikhenvald 2004), Tofa (Rassadin 1978), Tsez (Comrie and Polinsky 2007), Turkish (Şener 2011,

Slobin and Aksu 1982, among others), Tuvan (Harrison 2000), Uzbek (Straughn 2011), and Wakhi (Bashir 2006).

This cross-linguistic prevalence of morphologically linked tense/aspect and evidential meanings suggests that the Tatar data in (1) is not simply a quirk of the language. To understand data like (1), we need an explanation as to how temporal and evidential meanings can be related. This dissertation is an attempt at addressing this topic.

1.1.1 Structure of the dissertation

This dissertation is composed of two parts; one is descriptive and one is theoretical.

Part I (Chapters 2 and 3) provides a descriptive overview of the Tatar TAE system.[5] Chapter 2 describes the evidential contribution of the TAE suffixes, while Chapter 3 describes the temporal contribution of the TAE suffixes. I list the core set of Tatar TAE suffixes addressed in this dissertation, and their temporal and evidential contributions, in Table 1.1. (These are merely descriptive labels for how they are interpreted in matrix clauses; I ultimately propose in Chapter 5 that the Tatar TAE suffixes have no evidential component in their semantics.)

TAE suffix	**Temporal meaning**	**Evidential meaning**
-DI	past tense	speaker has direct evidence
-GAn	resultative aspect	speaker has indirect evidence
-(y)AçAK	prospective aspect	speaker has "specific" evidence
-(y)Er	future tense	speaker has "non-specific" evidence

Table 1.1: Temporal and evidential meanings associated with the core set of Tatar TAE suffixes addressed in this dissertation.

There are a number of important takeaways from the descriptive portion of the dissertation. First, I show that the Tatar TAE suffixes appear to encode an evidential distinction

[5]There are a small number of other English language publications on Tatar: Greed (2014, 2009), Davliyeva (2011), Tatevosov (2007), Şahan (2002), Comrie (1997), and Poppe (1961). Greed (2014, 2009) also discusses evidentiality and modality in Tatar, but does not give a formal analysis of the data. To my knowledge, no other English language publications explicitly address the Tatar tense/aspect system, either descriptively or formally.

in the future as well as in the past; this is typologically unusual (Aikhenvald 2004). Second, I show that the evidential readings associated with the TAE suffixes are absent in some morphosyntactic environments (namely, verbal nominalizations). Third, I show that the evidential readings of some of the TAE suffixes are cancellable in certain contexts. I take these latter two points to support treating the evidential readings of the Tatar TAE suffixes as arising pragmatically.

Chapters 2 and 3 are largely theoretically "neutral," and are intended to be accessible to all linguists who are interested in Tatar and how the TAE suffixes are interpreted. However, this linguistic description is theoretically informed. I use technical terminology, when necessary, with the goal of steering the reader towards my eventual analysis of the data.

Part II (Chapters 4, 5, 6, and 7) lays out my theoretical semantic and pragmatic analysis of the Tatar TAE data. Chapter 4 reviews some analyses of languages with related evidential and temporal systems (Koev 2017, 2011 and Smirnova 2013, 2011 for Bulgarian; Lee 2013, 2011 for Korean; Faller 2004 for Cuzco Quechua). I show that these theories, while compelling, cannot account for the observed Tatar data. In particular, I show that these theories are not suited for a TAE system like Tatar in which evidential distinctions are apparently encoded in the future as well as the past.

In Chapter 5, I introduce my semantic and pragmatic proposal for the Tatar TAE suffixes in matrix clauses. Thus far, most formal analyses of evidentiality have focused solely on modeling evidential meaning (Izvorski 1997, Faller 2002, Murray 2010, Korotkova 2016, among many others).[6] These proposals do not address the relationship between evidentiality and other kinds of meanings. However, as noted by Aikhenvald (2004), evidentiality is often "parasitic" on some other grammatical category. Chapters 5 and 6 are an attempt to formalize the relationship between evidential meaning and temporal meaning in Tatar.

In a nutshell, I propose that the Tatar TAE suffixes in Table 1.1 have underlyingly temporal semantics. I argue that they assert temporal meaning, and pragmatically implicate

[6]To be clear, this is a **huge** task; my observation is not intended to be a criticism of these authors. Furthermore, in some of the languages that these analyses address, evidentiality does not appear to be morphosyntactically related to tense/aspect (or any other kind of meaning).

evidential meaning. That is, I argue that there is nothing in the semantics of the TAE suffixes themselves that explicitly encodes evidentiality; rather, it arises as a byproduct of their temporal semantics. This differs from several prior theories of evidentiality, which explicitly encode evidential meaning as e.g. presuppositions (Izvorski 1997) or as part of the sincerity conditions of an utterance (Faller 2002).

My analysis formalizes and expands upon two previously described intuitions from the evidential and tense/aspect literature. The first intuition is that aspects (e.g. perfects, resultatives) can relate ongoing states to previous events. These states and events can stand in a causal relationship to each other, as proposed by e.g. Comrie (1976). The second intuition is that indirect evidentials—like these aspectual morphemes—indicate the speaker's observation of the results of some previous event (e.g. Nikolaeva 1999, Bashir 2006).

I couch my analysis in terms of a tripartite ontology of events in which events are preceded by causally related pre-states, and are followed by causally related post-states.[7] In brief, I propose that Tatar speakers use the TAE suffixes in Table 1.1 to express whether or not they locate themselves in a causally related pre- or post-state of the event that their utterance describes. The aspectual suffixes in Table 1.1 assert the existence of an event pre- and post-state, whereas the tense suffixes do not. When a Tatar speaker utters a matrix clause using an aspectual TAE suffix, they therefore assert the existence of an event pre- or post-state and thereby pragmatically commit themselves to having perceived that event pre- or post-state. Since pre- and post-states stand in causal relationships to their associated event, this implicates the evidential readings described in Table 1.1. (I expand upon the details of this analysis at length in Chapter 5.)

In Chapter 6, I extend this temporal proposal (with some tweaking) to uses of the Tatar TAE suffixes in embedded CPs and verbal nominalizations. I show that my temporal proposal explains the distribution of the TAE suffixes in verbal nominalizations: aspects are grammatical in Tatar verbal nominalizations, while tenses are not. My pragmatic pro-

[7]This ontology is based in part on a proposal by Moens and Steedman (1988). I give a formal definition of event pre- and post-states in Chapter 5.

posal also accounts for why the evidential readings of the TAE suffixes disappear in some morphosyntactic environments.

Chapter 7 summarizes the puzzles that this dissertation answers, and provides future directions for work on this topic.

1.2 Background on Tatar and fieldwork methodology

Tatar belongs to the Kipchak branch of the Turkic language family (Hammarström et al. 2018). As such, it has a number of grammatical features that are common to Turkic languages. Tatar is SOV and has grammatical and semantic case marking, including locative and ablative cases. It also displays extensive vowel and consonant harmony. For a descriptive overview of Tatar grammar, see Poppe (1961).

Tatar is primarily spoken in the republics of Tatarstan and Bashkortostan in western Russia, although there are sizable diaspora populations of Tatar speakers in Finland, Kazakhstan, Turkey, and the US, among other countries (Simons and Fennig 2018). Tatars are predominantly Sunni Muslim, and are an ethnic and linguistic minority within Russia. As such, they have historically experienced pressure to assimilate into Russian culture by e.g. adopting the Russian language and the Cyrillic script. Following the breakup of the Soviet Union, the government in Tatarstan enacted a policy of "de-Russification" promoting the use of Tatar and making it an official language of the republic (Wertheim 2002). Nonetheless, all of the Tatar speakers I worked with speak Tatar in addition to another first language, usually Russian.

All of the data in this dissertation was collected through original fieldwork with Tatar speakers in the US. My primary Tatar consultant, Sofia Mazgarova, is in her late 30s. She was born in Kazan, the capital of Tatarstan, and emigrated to the US at age 19. Sofia travels back to Tatarstan roughly every summer, and uses Tatar regularly when communicating with her family and friends. She is bilingual in Tatar and Russian, fluent in English, and additionally speaks Turkish, Persian, and some Arabic. Due to her extensive language back-

ground, she has a very high degree of etymological knowledge of Tatar words. Elicitation with Sofia was conducted in Los Angeles.

I conducted secondary fieldwork on Tatar with a number of other speakers in California; the purpose of this fieldwork was primarily to double-check the data I collected with Sofia. On two occasions, I traveled to Burlingame, CA to meet with Tatar speakers there. There is a significant Tatar population in Burlingame, which is home to the American Turko-Tatar Association. The speakers I worked with in Burlingame formed two distinct sociolinguistic groups. One group (n = 4) consisted of speakers in their 20s-30s who were born in Tatarstan and emigrated to the US within the last decade. The second group (n = 2) were in their 80s and were born in Japan; their parents had emigrated from Tatarstan to Japan following the Russian revolution. The former group of speakers identified strongly as being Russian as well as Tatar, while the latter group had no connections to Russian identity at all. I noted some small differences with respect to the use of the TAE suffixes by these groups of speakers (see e.g. my discussion in Chapter 2 on the availability of mirative readings). However, on the whole, all of the speakers I worked with provided the same evidential and temporal interpretations of the TAE suffixes. In addition to the speakers in Burlingame, I worked with both of Sofia's parents while they visited Los Angeles.

Elicitation was primarily conducted using English as a metalanguage. I occasionally used Russian as a metalanguage when e.g. eliciting the Sequence of Tense data discussed in Chapter 3, so as not to bias speakers with English tenses. Elicitation generally took the form of providing a discourse context to the speaker; I then either asked how they would say a target sentence in that context, or how they would judge the felicity of a provided Tatar expression in that context (i.e., a felicity judgment task). My initial elicitations were guided by Aikhenvald (2004, 385-390)'s guide to collecting evidential data, Dahl (2000, 789-818)'s tense and aspect questionnaire, and Vander Klok (2014)'s modality questionnaire. Later elicitations involved more targeted, Tatar-specific lines of inquiry.

Part I

The Tatar tense/aspect and evidential system

CHAPTER 2

Evidential contributions of the Tatar TAE suffixes

2.1 Introduction

Tatar, like many other Turkic languages, has a number of portmanteau morphemes that are interpreted as marking both tense or aspect and evidentiality. Henceforth, I refer to these as **TAE (tense/aspect/evidential) morphemes**. While TAE morphemes are a common feature of Turkic languages, they are not a uniquely Turkic phenomenon; TAE morphemes are documented in a number of unrelated languages around the world (e.g. Cherokee (Pulte 1985), Kalaallisut (Fortescue 2003), Khanty (Nikolaeva 1999), Tariana (Aikhenvald 2004), and Tsez (Comrie and Polinsky 2007), among many others).

Other Turkic TAE suffixes have been described as encoding evidential distinctions effectively only in the past (e.g. Meriçli 2016 and Slobin and Aksu 1982 for Turkish; Dwyer 2000 for Salar; Boeschoten 1998 for Uzbek; among others).[1] However, as I show in this chapter, Tatar is of particular interest due to the fact that it also encodes an apparently evidential distinction in the **future**: descriptively speaking, the use of both the past and future Tatar TAE morphemes appears to hinge on the evidence that is available to the speaker. This is typologically unusual; Aikhenvald (2004) notes that languages rarely make evidential distinctions in the future. The Tatar TAE data is therefore interesting from both descriptive and theoretical perspectives.

[1] Many of these sources describe the future Turkic suffixes as having "modal" interpretations; the readings that they describe are similar to the readings of the Tatar future-oriented TAE morphemes that I document in this chapter. However, I ultimately show that it is more fruitful to think about the contrast as something evidential, rather than something modal. The close relationship between evidentiality and modality is has been discussed at length (e.g. de Haan 1999, among many others).

The Tatar TAE system has been described previously by Greed (2014, 2009), Poppe (1961), and Tatevosov (2007) (writing on Mishar Tatar). These are strictly descriptive sources; this dissertation is the first attempt to give a formal analysis of the Tatar TAE data. The Tatar data in this and the following chapter is therefore presented in such a way as to steer the reader towards my eventual theoretical conclusions. As a result, my descriptive labels for the Tatar TAE morphemes (and their semantic contribution) sometimes differ from the labels proposed by other authors.[2] I compare my description of the Tatar TAE system to Greed (2014), Tatevosov (2007), and Poppe (1961) at the end of Chapter 3.

This chapter lays out the evidential interpretations of the Tatar TAE morphemes that the formal portion of this dissertation addresses, while the following chapter describes their temporal interpretations. I refrain from using theoretical (semantic) terminology in these chapters, since they are intended to be a primarily descriptive resource on the Tatar TAE system. I list the TAE morphemes in Table 2.1, along with the glosses that I will use. I note that these glosses only indicate the tense/aspect contribution of these morphemes and not their evidential contribution. This is due to the fact that, as I eventually argue in Chapter 5, the semantics of these morphemes do not explicitly encode any evidential meaning; i.e., their evidential reading is a byproduct of their temporal semantics. For the purpose of this chapter, however, I will nonetheless refer to the "evidential" readings of these morphemes, since this term provides the simplest description of the kinds of contexts that license their use.

The TAE morphemes all occur as verbal suffixes with the exception of *ide*, which occurs as a free morpheme directly following the predicate. All finite verbs must host one of the TAE morphemes; tense/aspect is obligatorily marked in Tatar.

Subject agreement in Tatar is marked as a verbal suffix following TAE marking. The TAE morphemes differ in the subject agreement paradigms that they occur with. The past

[2]There is a tradition in the descriptive evidential literature to refer to TAE systems like Tatar as having "definite" and "indefinite" tenses (e.g. Davliyeva 2011, Straughn 2011, Comrie and Polinsky 2007, among many others). My primary Tatar consultant also uses these labels to refer to the Tatar TAE suffixes. According to these labels, *-DI* and *-(y)AçAK* are "definite" tenses, whereas *-GAn* and *-(y)Er* are "indefinite" tenses. In Chapter 5, I propose a classification of the Tatar TAE suffixes that cross-cuts these descriptive labels.

Morpheme	Gloss	Abbreviated gloss
ide	past tense	PST
-DI	past tense	PST
-GAn	resultative aspect	RESULT
-A	present tense	PRES
-(y)AçAK	prospective aspect	PROSP
-(y)Ar	future tense	FUT

Table 2.1: Inventory of TAE morphemes in Tatar.

tense morphemes *-DI* and *ide* occur with the past tense subject agreement markers in the first column in Table 2.2; all of the other TAE morphemes occur with the "elsewhere" agreement marking in the second column. I provide the set of free nominative pronouns in the third column for comparison with the agreement suffixes.

Gloss	Past tense agreement	Elsewhere agreement	Nominative free pronouns
1SG	*-m*	*-mIn/-m*	*min*
1PL	*-K*	*-bIz*	*bez*
2SG	*-n*	*-sIn*	*sin*
2PL	*-GIz*	*-sIz*	*sez*
3SG	*-∅*	*-∅*	*ul*
3PL	*-lAr/-∅*	*-lAr/-nAr/-∅*	*alar*

Table 2.2: Tatar subject agreement paradigms and free (nominative) pronouns.

(4)-(5) show examples of past tense agreement versus elsewhere agreement.

(4) **Past tense agreement**

a. bez yeger-de-**k**.
1PL.NOM run-PST-1PL
'We ran.'

b. bez student ide-**k**.
1PL.NOM student PST-1PL
'We were students.'

(5) **Elsewhere agreement**

a. bez Mäskäü-gä bar-a-**bız.**
1PL.NOM Moscow-DAT go-PRES-1PL
'We are going to Moscow.'

b. bez Mäskäü-gä bar-ğan-**bız.**
1PL.NOM Moscow-DAT go-RESULT-1PL
'We went to Moscow.'[3]

With the exception of 3SG agreement, which is phonologically null in both paradigms, the "elsewhere" agreement paradigm closely matches the set of free nominative pronouns in Tatar. This suggests that they originated historically as pronominal clitics. Unlike the free pronouns, the Tatar elsewhere agreement markers are subject to some phonological conditions, chiefly vowel harmony.[4]

2.1.1 Evidential terminology

In unembedded contexts, evidential morphemes mark the speaker's evidence source for the proposition that they are reporting (henceforth, the **scope proposition**; I will abbreviate the scope proposition as *p* elsewhere in this chapter). Common cross-linguistic evidential categories are indicated in the tree in Figure 2.1 (Willett 1988, Aikhenvald 2004).

Direct evidentiality typically refers to events that occurred within the speaker's perceptual field. Conversely, indirect evidentiality typically refers to events that occurred outside of the speaker's perceptual field. As indicated in Figure 2.1, indirect evidentiality can include both inferential and reportative evidence. Inferential evidence generally refers to

[3]As I eventually discuss in §3.2.1.2, this expression is infelicitous unless it occurs in a context in which the speaker is somehow unaware of their trip to Moscow (e.g. they were drunk or extremely sick when they went). I include this example here primarily to show the contrast in 1PL past tense and elsewhere agreement.

[4]In combination with the TAE suffix *-GAn*, the 3PL agreement marker *-lAr* is subject to a rule that assimilates the suffix-initial /l/ to [n] when following another [n] at (some) morpheme boundaries, as in (6). This rule is also documented in a number of other Turkic languages (Schönig 1999, 76).

(6) alar Mäskäü-gä {bar-ğan-**nar** / *bar-ğan-**lar**}.
3PL.NOM Moscow-DAT go-RESULT-3PL / go-RESULT-3PL
'[I have indirect evidence that] they went to Moscow.'

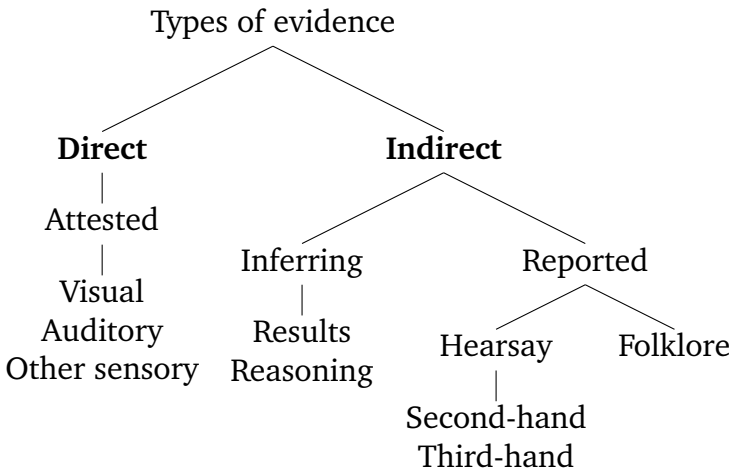


Figure 2.1: Types of evidence that can be grammaticalized in evidential systems (adapted from Willett 1988, 57).

(visual or aural) evidence that leads the speaker to conclude that the scope proposition is true, whereas reportative evidence refers to reports that the speaker heard from another individual.

Evidential categories are noted to be subject to significant contextual fuzziness; that is, what counts as "direct" or "indirect" evidence can in part be determined contextually. An example of this is described by Faller (2002) for Cuzco Quechua. In most contexts, the Cuzco Quechua direct evidential *-mi* requires having visual evidence for the scope proposition *p*, as in (7). However, *-mi* is also felicitous in contexts in which the speaker has not witnessed *p* but is nonetheless highly confident in their belief that *p* is true, as in the expression in (8).[5]

(7) CUZCO QUECHUA (QUECHUAN) (adapted from Faller 2002, 131-133)

Context: The speaker saw Mary eat soup.

Marya-qa lawa-ta-**n** mikhu-raq-n.
Marya-TOP soup-ACC-MI eat-PST.1-3
'Marya ate soup.'

[5] Faller (2002) ultimately uses this data to argue that *-mi* should therefore be termed a "best possible grounds" evidential, as opposed to a direct evidential.

(8) **Context**: The speaker learned in school that there are elephants in Africa.

Africa-pi-**n** elefante-kuna-qa ka-n.
Africa-LOC-MI elephant-PL-TOP be-3
'In Africa, there are elephants.'

Data like (7)-(8) suggests that the terms "direct" and "indirect" can serve as a useful shorthand for the stereotypical properties of what we call direct or indirect evidence (e.g. having "direct evidence" requires that the speaker be an eyewitness to the event described by the scope proposition, whereas having "indirect evidence" requires that the speaker not have witnessed the event). However, the availability of (7)-(8) shows that these terms do not distinguish between all the possible evidence types as neatly as we might like.[6]

I will ultimately argue in Chapter 5 that the semantics of the Tatar TAE morphemes do not in fact have any evidential component, and that a temporal account of the Tatar TAE data can avoid some of the fuzziness associated with the evidential terminology. My use of evidential labels in this chapter should be taken simply as descriptive generalizations, rather than reflecting any theory of the Tatar TAE system.

2.2 Evidential interpretation of the TAE suffixes in matrix clauses

In the following sections, I describe the kinds of evidence that license the Tatar TAE suffixes in matrix clauses. I primarily focus on the set of four core TAE suffixes shown in Table 2.3. My analysis of the Tatar data ultimately only addresses these four suffixes; I set aside the free past tense morpheme *ide* due to its different morphosyntactic distribution, and the present tense morpheme *-A* due to the fact that it lacks any evidential interpretation. (I briefly discuss *-A* in §2.2.2.3 to show that it lacks any evidential meaning; I discuss it only because it occupies the same morphosyntactic "slot" on the Tatar verb as the other TAE suffixes.)

I show that the two past-oriented TAE suffixes roughly distinguish between requiring direct (*-DI*) and indirect (*-GAn*) evidence for the scope proposition; i.e., they make the ba-

[6] Similar contextual effects are also observed in Tatar; I discuss this data in §3.2.1.1.

	Morpheme	Gloss	Kind of evidence required
Past-oriented	*-DI*	'PST'	speaker has direct evidence
	-GAn	'RESULT'	speaker has indirect evidence
Future-oriented	*-(y)AçAK*	'PROSP'	speaker has "specific" evidence
	-(y)Er	'FUT'	speaker has "non-specific" evidence

Table 2.3: Core set of Tatar TAE suffixes addressed in this dissertation.

sic evidential distinction described in Figure 2.1. The two future-oriented TAE suffixes are also sensitive to a difference in the evidence that the speaker has for the scope proposition. Roughly speaking, this contrast is one of having "specific" (*-(y)AçAK*) versus "non-specific" (*-(y)Er*) evidence for the scope proposition. I note that these are not canonically accepted types of evidence; for instance, the chart in Figure 2.1 does not include specific and non-specific evidential categories. However, I find these to be the most useful descriptive labels for the kinds of evidence that license the use of these morphemes.

2.2.1 Evidentiality in the Tatar past-oriented TAE suffixes

We can describe the evidential contrast between the two past-oriented Tatar TAE suffixes *-DI* and *-GAn* as one of direct versus indirect evidentiality. When a speaker asserts a proposition *p* using *-DI*, it is understood that they have direct evidence for *p*. "Direct evidence" can include both visual observation of *p*, as in (9a), as well as auditory observation of *p* in some cases, as in (9b). (Here and elsewhere in the dissertation, I set off the evidential component of an expression in square brackets [].)

(9) ***-DI*: Past-oriented, direct evidence**

a. **Visual context:** You saw Mansur get on a train to Moscow.

Mansur Mäskäü-gä {bar-**dı**-∅ / #bar-**ğan**-∅}.
Mansur Moscow-DAT go-PST-3SG / go-RESULT-3SG
'[I have direct evidence that] Mansur went to Moscow.'

b. **Auditory context:** You are at home and hear thunder. The next day, you say:

kiçä tön-ne kük {kükrä-**de**-∅ / #kükrä-**gen**-∅}.
yesterday night-LOC sky thunder-PST-3SG / thunder-RESULT-3SG
'[I have direct evidence that] it thundered last night.'

(Lit. 'The sky thundered last night.')

Conversely, when a speaker asserts a proposition using *-GAn*, it is understood that they have indirect evidence for *p*. "Indirect evidence" can include having inferential evidence for *p*, as in the context in (10a), in which the speaker infers that *p* is true based on evidence for *p*. It can also include reportative evidence, as in the context in (10b).

(10) ***-GAn*: Past-oriented, indirect evidence**

a. **Inferential context:** You found a train ticket in Mansur's desk that is from Moscow.

b. **Reportative context:** Wäğıyz told you that Mansur went to Moscow.

Mansur Mäskäü-gä {bar-**ğan**-∅ / #bar-**dı**-∅}.
Mansur Moscow-DAT go-RESULT-3SG / go-PST-3SG
'[I have indirect evidence that] Mansur went to Moscow.'

The same evidential contrast has been described for cognate verbal suffixes in a number of other Turkic languages, including Bashkir *-de* and *-gan* (Poppe 1964), Karachay *-dI* and *-gEn* (Seegmiller 1996), Karaim *-dI* and *-gEn* (Mysaev 1964), Kyrgyz *-de* and *-gan* (Abduldaev and Zakharova 1987), Salar *-(d)ʒi* and *-GAn* (Dwyer 2000), and Tuvan *-di* and *-gan* (Harrison 2000, Krueger 1977).[7]

In the following sections, I give an in-depth discussion of the evidential contribution of the two past-oriented Tatar TAE suffixes in turn.

[7]Another set of Turkic languages exhibit a similar contrast between past-oriented verbal suffixes roughly of the form *-DI* and *-mIş*. These include Gagauz (Pokrovskaya 1964) and Turkish (Kornfilt 1997, Slobin and Aksu 1982), among several others.

2.2.1.1 Direct evidence: *-DI* 'PST'

The Tatar TAE suffix *-DI* is interpreted as marking direct evidentiality. If the event described by *p* occurred within the speaker's perceptual field (including both visual and auditory perception), they must use *-DI* to report *p*. This ultimately amounts to the observation that speakers must make the strongest claim possible, given their available evidence.[8] Basic examples of contexts that license *-DI* are given in (9).

In addition to these contexts, there are a number of additional contexts in which it is felicitous for speakers to use *-DI* even if the event described by *p* did not occur within their perceptual field. These contexts correlate with a high degree of speaker certainty that *p* occurred. These contexts include reporting facts that the speaker obtained from a source that they consider to be highly reliable with respect to *p* (11), facts about which the speaker has "privileged" knowledge (12), reliably scheduled events (13), and widely known facts (14). In the age of social media, my primary Tatar consultant reports that it is also felicitous to use *-DI* in contexts in which a sufficient amount of evidence is available (through Facebook, Instagram, etc.) to assert confidently that *p* occurred (15).[9]

[8] I use "strong" here in a non-technical sense, referring rather to the idea (previously observed by Davis et al. 2007, Faller 2012, and AnderBois 2014, among others) that direct evidence is the strongest kind of evidence that a speaker can have for a proposition.

[9] In Tatar, "privileged" knowledge typically refers to the knowledge that close family members have of one another; this knowledge is considered to be especially reliable. This is in line with Aikhenvald (2004)'s observation that the use of direct evidentials to report other peoples' activities is often reserved, but can be licensed in contexts in which the speaker is considered sufficiently "close" to the other individual. This can, but need not, be licensed by standing in a kinship relation to one another (Aikhenvald 2004, 350). For instance, in the case of Tatar, the simple past in (12) is no longer felicitous if the speaker is reporting on where someone else's grandmother was born, since they are no longer assumed to have privileged knowledge of her life.

(11) **Highly reliable source context:**[10] Läylä's friend Wäğıyz works for a news channel as a sports broadcaster. Wäğıyz told her that Michael Phelps won a race in the Rio Olympics. Läylä did not watch the race herself, but she considers Wäğıyz as a good source on events in sports. She can tell another person:

Michael Phelps ciñ**-de**-∅.
Michael Phelps win-PST-3SG
'[I have direct evidence that] Michael Phelps won.'

(12) **"Privileged" knowledge context:** You are describing your grandmother's life.

äbi-em Qazan-da tu**-de**-∅.
grandmother-1SG.POSS Kazan-LOC born-PST-3SG
'[I have direct evidence that] My grandmother was born in Kazan.'

(13) **Reliably scheduled event context:** You are discussing the TV schedule.

kiçä Titanik-nı tilivisor-da kürsät-te-lär, läkin min anı
yesterday Titanic-ACC television-LOC show-PST-3PL but 1SG.NOM 3SG
qara-ma**-dı**-m.
watch-NEG-PST-1SG
'[I have direct evidence that] They showed Titanic on TV yesterday, but I didn't watch it.'

(Speaker's comment: "Television is a pre-programmed thing, so you know that they showed it. They have a schedule.")

(14) **Widely known facts context:** You are a history professor lecturing your class on WWII.

Germaniya Polşa-nı bas**-tı**-∅.
Germany Poland-ACC invade-PST-3SG
'[I have direct evidence that] Germany invaded Poland.'

[10] Speakers can consider sources to be reliable with respect to some topics, but not others. For instance, in (11), Läylä considers Wäğıyz to be an authority with respect to sports, because his job involves reporting facts about sporting competitions. However, this does not mean that she considers Wäğıyz to be an authority about other topics, e.g. who won a local election, or what band won a local "battle of the bands" competition.

(15) **Social media context:** You are Facebook friends with your friend Aygöl. You saw her post pictures of herself traveling in Kenya. She also "checked in" online to the Nairobi airport, Nairobi national park, and so on. While talking about her with some other friends, you say:

Aygöl Kenya-ğa bar-ıp qayt-**tı**-∅.
Aygöl Kenya-DAT go-IP travel-PST-3SG
'[I have direct evidence that] Aygöl traveled to Kenya.'

Both *-DI* and *-GAn* are licensed in all of the above contexts. That is, *-GAn* is always licensed in contexts in which the speaker has indirect evidence for the proposition. However, there are a small set of indirect evidential contexts in which speakers can also make stronger claims using *-DI*.

The use of *-DI* is not licensed by having "best possible grounds" for the proposition.[11] This is shown by the context in (16) below. In (16), there is no possible way that the speaker could have witnessed the event of the child catching a fish, since it happened before they were born. However, (16) could be considered an instance of best possible evidence for the child having caught a fish. Nonetheless, *-DI* is unavailable.

(16) **Thrift store context:** You are at a thrift store that sells old photographs. You are looking at a photo from the 1800s. It is of a child holding up a fish on a fishing line. You say:

bu bala balıq {tot-**qan**-∅ / #tot-**tı**-∅}.
this child fish catch-RESULT-3SG / catch-PST-3SG
'[I have indirect/#direct evidence that] This child caught a fish.'

Finally, Tatar speakers use *-DI* to report dreams and hallucinations, as in (17)-(18). Tatar speakers conceptualize experiencing dreams or hallucinations as constituting direct evidence for the events in the dream/hallucination. The choice of using direct versus indirect evidentials to report dreams and hallucinations is a noted point of variation across languages (Aikhenvald 2004, 344-347).

[11] A "best possible grounds" account was proposed by Faller (2002) to account for the distribution of the Cuzco Quechua evidential *-mi*.

(17) **Dream context:** You had a dream that your cat, Didi, talked to your husband, Ali. You say:

töş-em-dä, Didi Ali belän {söyläş-**te**-∅ / #söyläş-**kän**-∅}.
dream-1SG.POSS-LOC Didi Ali with talk-PST-3SG / talk-RESULT-3SG
'[I have direct/#indirect evidence that] In my dream, Didi talked with Ali.'

(18) **Hallucination context:** You are at home sick with a high fever. You are really not feeling well and you think you're having strange hallucinations. At one point, your cat talked to you. However, you know cats can't talk. You say:

mäçe minem belän {söyläş-**te**-∅ / #söyläş-**kän**-∅}.
cat 1SG.GEN with talk-PST-3SG / talk-RESULT-3SG
'[I have direct/#indirect evidence that] My cat talked with me.'

(Speaker's comment: "You use *-DI* because it was your experience.")

The example in (18) suggests that the use of *-DI* is not governed solely by speaker certainty that *p* occurred, as might be suggested by the data in (11)-(15). That is, it suggests that *-DI* does not make a strictly modal contribution. In (18), the speaker is highly certain that her cat did not actually speak with her; nonetheless, she must use *-DI* to report her hallucination, because she perceived it occur.

2.2.1.2 Indirect evidence: *-GAn* 'RESULT'

The Tatar TAE suffix *-GAn* is interpreted as marking indirect evidentiality. This evidential category includes both reportative and inferential evidentiality, as shown in Figure 2.1. I demonstrated in (10) that this suffix can be used in both basic inferential and reportative evidential contexts. *-GAn* can be used to make very "strong" claims as well as "weak" claims: that is, a speaker can use *-GAn* to report propositions about which they are certain are true, as well as propositions about which they are uncertain. (*-GAn* is infelicitous if the speaker is certain that the scope proposition is false.)

The weaker use of this morpheme is typically associated with reportative contexts, in line with the extensive cross-linguistic data in AnderBois (2014). The reportative context in (20) shows that a speaker can use *-GAn* to express a proposition that they heard from

an individual that they do not trust; here it is not the case that the content of the scope proposition itself makes them believe that it is unlikely, but rather they do not trust the source of the report. The reportative context in (21) shows that the speaker can also use *-GAn* to report a proposition that they believe itself is unlikely due to their knowledge of facts about the world.[12] In both contexts, the speaker has reportative evidence for the scope proposition.

(20) **Unreliable source context #1:** Wäğıyz told you that Läylä went to Kazan, but you think Wäğıyz is a compulsive liar and you don't really trust what he says. Later, you tell Mansur:

Läylä Qazan-ğa {kit-**kän**-∅ / #kit-**te**-∅}.
Läylä Kazan-DAT leave-RESULT-3SG / leave-PST-3SG
'[I have indirect/#direct evidence that] Läylä left for Kazan.'

(21) **Unreliable source context #2** (adapted from Smirnova 2011, 277): You visited your friend Alsu in a psychiatric clinic. Alsu was hospitalized because of severe hallucinations and thinks that she was abducted by aliens. When a friend asks you what is going on with Alsu, you say:

başqa planeta-da yäşäüçe-lär Alsu-ne url-ap al-ıp {kit-**kän**-när /
other planet-LOC resident-PL Alsu-ACC steal-IP take-IP leave-RESULT-3PL /
#kit-**te**-lar}.
leave-PST-3PL
'[I have indirect/#direct evidence that] Aliens abducted Alsu.'[13]

[12]The clause-final particle *imeş* is felicitous in contexts in which the speaker doubts that or is uncertain whether the scope proposition is true, but is not required. However, *imeş* is required if the speaker is **certain** that the scope proposition is false.

Tatevosov (2007) (writing on Mishar Tatar) describes *imeş* as a marker of uncertainty and indirect evidentiality. *imeş* is one of a small set of clause-final evidential or modal particles in Tatar, which I do not address in this dissertation. It is typically used in reportative contexts like (20)-(21), and expressions containing *imeş* are often translated into English by my Tatar consultants as "It is rumored that..." My primary Tatar consultant reports that the expression in (19) sounds sarcastic or snide.

(19) başqa planeta-da yäşäüçe-lär Alsu-ne url-ap al-ıp kit-kän-när **imeş**.
other planet-LOC resident-PL Alsu-ACC steal-IP take-IP leave-RESULT-3PL IMEŞ
'[I have indirect evidence that] Aliens abducted Alsu [and I don't think it happened].'

[13]I note that reporting (21) without *imeş* requires that the speaker at least believe that it is possible that aliens exist. (21) is infelicitous if the speaker does not believe that aliens exist. I return to the issue of speaker commitment to the truth of propositions embedded under the TAE suffixes in Chapters 4 and 5.

-GAn can also be used to make strong claims. In the context in (22), the speaker can be highly certain that the scope proposition "Germany invaded Poland" is true, since it is a fact that they have been taught in their history class. However, since they are not a scholar of WWII history and were not alive to witness the invasion of Poland, they use *-GAn* to report the proposition in (22). The availability of *-GAn* in all of the contexts in (20)-(22) therefore shows that its use is independent of speaker certainty.

(22) **Context**: You are a student studying for your history exam on WWII. You say:

Germaniya Polşa-nı bas-**qan**-∅.
Germany Poland-ACC invade-RESULT-3SG
'[I have indirect evidence that] Germany invaded Poland.'

If a speaker uses *-GAn* in combination with a first person subject, the resulting interpretation is typically that the speaker is reporting an event that they witnessed or participated in while impaired in some way.[14] I refer to these readings as "blackout" readings.[15] Blackout readings typically arise when recounting events that occurred when the speaker was under the influence of drugs or alcohol, was sick, was asleep, or was too young to consciously remember. Such readings are common cross-linguistic interpretations of indirect evidentials in combination with first person subjects (Aikhenvald 2004, 219-223). I give two examples of blackout readings of *-GAn* in the following examples.

[14]This appears to be a point of variation across Turkic. Straughn (2011, 64) reports that in Kazakh and Uzbek, the cognate morphemes *-GAn* and *-gan* may co-occur with first person subjects without any associated blackout readings. (Straughn does not indicate whether the choice of *-GAn/-gan* versus *-DI/-di* changes the meaning of the utterances or if the use of *-GAn/-gan* is licensed only by particular contexts.)

Similarly, Göksel and Kerslake (2005, 310) report that in Turkish, it is possible for speakers to use *-mIş* (the Turkish analog of Tatar *-GAn*) with a first person subject to report other individuals' opinions of the speaker (e.g. '[I have indirect evidence that] I annoyed my sister'). This use is unavailable in Tatar.

[15]Greed (2014) discusses similar Tatar data, but labels the associated readings as mirative due to their meaning of "unprepared mind." I believe that the examples in (23)-(24) are not mirative uses of *-GAn* due to the fact that they do not necessarily convey speaker surprise. I argue in this section that *-GAn* does not synchronically have a mirative usage.

(23) **Drunk context:** You got drunk last night and smashed a plate. When you wake up in the morning, you have no memory of doing this; however, you find the broken plate in your kitchen and remember being drunk. You say:

tälinkä-ne {wat-**qan**-mın / #wat-**tı**-m}.
plate-ACC break-RESULT-1SG / break-PST-1SG
'[I have indirect evidence that] I broke the plate.'

(24) **Sleeping context:** You fall asleep with a glass of water by your bed. When you wake up, the glass of water has been tipped over. You don't have any pets, and no one else was in the house with you. You infer that you knocked over the water in your sleep. You say:

min su-nı bär-ep {töşer-**gän**-men / #töşer-**de**-m}.
1SG.NOM water-ACC hit-IP drop-RESULT-1SG / drop-PST-1SG
'[I have indirect evidence that] I knocked over the water.'

I finally note that Tatar *-GAn* differs from similar Turkic morphemes (e.g. Turkish *-mIş* (Slobin and Aksu 1982)) in that it does not express **mirativity**.[16] Mirativity is defined as an expression of speaker surprise with respect to the scope proposition. Several authors have noted a cross-linguistic link between mirativity and indirect evidentiality; this is observed across a number of genetically unrelated languages (Rett and Murray 2013, Aikhenvald 2004, 195-215). In these languages, so-called "mirative evidential" morphemes mark indirect evidentiality in some contexts and mirativity in others. In (25), I show examples of the Turkish TAE morpheme *-mIş* in both its evidential and mirative uses.

[16]The Turkic languages can be divided into those that have *-GAn* as a past-oriented verbal suffix (Kazakh, Tatar, Uzbek, among others), and those that have *-mIş* (Azerbaijani, Gagauz, Turkish, among others). The exact origins of *-GAn* and *-mIş* are unknown (Johanson 2003, 287). The use of *-GAn* is a hallmark of central Asian Turkic languages, and is more common than *-mIş*, which is primarily a feature of the southwestern Turkic languages (Johanson 2000, Schönig 1999).

(25) TURKISH (TURKIC) (Sözen Özkan, p.c.)

a. **Indirect evidential *-mIş* context:** You saw a picture of Ayşe posing in front of the Blue Mosque in Istanbul.

Ayşe Istanbul-a git-**miş**-∅.
Ayşe Istanbul-DAT go-MIS-3SG
'[I have indirect evidence that] Ayşe went to Istanbul.'

b. **Mirative *-mIş* context:** For all you know, your friend Seren is out of town. However, you just saw her walk past your office. You say:

Seren dön-**müs**-∅!
Seren turn-MIS-3SG
'[I am surprised that] Seren came back!'

Tatar *-GAn* differs from Turkish *-mIş* in that has no mirative usage.[17] When asked if an expression containing *-GAn* is felicitous in a mirative context, my consultants respond that it is not and "fix" the expression by adding the free clause-final particle *ikän*.[18]

(27) **Mirative context:** You were not expecting your friend Läylä to return from vacation for another week, but you just saw her walk past your office. You say:

Läylä qayt-**qan**-∅ #(**ikän**)!
Läylä return-RESULT-3SG IKÄN
'[I am surprised that] Läylä came back!'

[17] Fieldwork with older Tatar speakers in their 70s and 80s suggests that *-GAn* could have had a mirative interpretation in the past. These older consultants judged *-GAn* to be felicitous in the following mirative context, and rejected the free particle *ikän* due to its indirect evidential requirement.

(26) Mirative context: You were not expecting your friend Läylä to return from vacation for another week, but you just saw her walk past your office. You say:
Läylä qayt-**qan**-∅ (#**ikän**)!
Läylä return-RESULT-3SG IKÄN
'[I am surprised that] Läylä came back!'
(Speaker's comment about the use of *ikän*: "This is like I heard it from somebody, like my sister called me up and told me.")

[18] *Ikän* belongs to a small set of clause-final particles that convey evidential or modal meanings, and which I do not address in this dissertation. (These particles are also discussed in footnote 12.) *Ikän* has been described previously by Tatevosov (2007) (writing on Mishar Tatar) as a marker of mirativity, and by Greed (2009) as a marker of indirect (non-reportative) evidentiality. My consultants frequently translate *ikän* expressions into English using "It turns out that..." or "Apparently..."

2.2.2 Evidentiality in the Tatar future-oriented TAE suffixes

As a first pass, the descriptive difference between the two future-oriented Tatar TAE suffixes appears to be one of modality: that is, high versus low (or neutral) speaker certainty that the event described by the scope proposition will occur. This descriptive generalization is supported by my consultants' English translations of expressions containing the future-oriented suffixes; expressions containing *-(y)AçAK* are typically translated with *definitely*, whereas expressions containing *-(y)Er* are typically translated with *maybe*.

I propose, however, that it is more productive to conceptualize the Tatar future-oriented TAE suffixes as differing with respect to evidentiality. This evidential contrast is one of "specific" versus "non-specific" evidence for the scope proposition. When a speaker asserts a proposition using *-(y)AçAK*, it is understood that they have "specific" evidence for the scope proposition, as in (28).

(28) ***-(y)AçAK*: Future-oriented, specific evidence**

Specific evidence context: You are planning a party, and you have assigned your friends different things to bring. Your friend Güzäl is assigned to bring cookies to the party (i.e., you have specific evidence that Güzäl will bring cookies to the party).

Güzäl peçeniye al-ıp {kil-**äçäk**-∅ / #kil-**er**-∅}.
Güzäl cookie take-IP come-PROSP-3SG / come-FUT-3SG
'[I have specific evidence that] Güzäl will bring cookies.'

Conversely, when a speaker asserts a proposition using *-(y)Er*, it is understood that they have "non-specific" evidence for the scope proposition, as in (29).

(29) ***-(y)Er*: Future-oriented, non-specific evidence**

No specific evidence context: You are planning a party, and you've asked all of your friends to bring things to contribute. Your friend Güzäl has a delicious chocolate chip cookie recipe that she usually brings to parties. (You haven't specifically asked her to bring the cookies, nor has she told you that she will bring them; i.e., you have no specific evidence that she will bring the cookies to the party.)

Güzäl peçeniye al-ıp {kil-**er**-∅ / #kil-**äçäk**-∅}.
Güzäl cookie take-IP come-FUT-3SG / come-PROSP-3SG
'[I have non-specific evidence that] Güzäl will bring cookies.'

A similar semantic contrast in the future-oriented verbal suffixes has been described in a number of other Turkic languages. As noted previously, many authors describe this contrast as one of modality (i.e., speaker certainty), rather than evidentiality.[19] Languages with this contrast include Azerbaijani *-(y)acaq* and *-(y)Ar* (Öztopçu 2000), Bashkir *-(y)asak* and *-(e)r* (Poppe 1964), Kumyk *-(A)zhAK* and *Ar* (Johanson and Csató 1998), Tofa *-azhek* and *-Ar* (Rassadin 1978), Turkish *-(y)AcAk* and *-(A)r* (Kornfilt 1997), and Uzbek *-(y)ajak* and *-(a)r* (Straughn 2011).

As I will show in the following sections, the apparently modal readings of the future-oriented TAE suffixes are in fact due to the available evidence that the speaker has that the event will occur. Conceptualizing the future-oriented suffixes as differing in evidentiality brings the future-oriented suffixes in line with the past-oriented suffixes; i.e., both sets make a fundamentally evidential distinction. The future-oriented Tatar TAE data is therefore of particular interest due to the fact that evidential readings tend to be much rarer cross-linguistically in future-oriented expressions than in non-future-oriented expressions (Aikhenvald 2004, 261).

[19] In fact, Greed (2014, 72) states that "Tatar evidentiality expressed with verb suffixes pertains to the past tense only."

2.2.2.1 "Specific" evidence: *-(y)AçAK* 'PROSP'

Conceptually speaking, if a speaker has specific evidence that the event described by *p* will occur, then they must use *-(y)AçAK* to report *p*. This results in a reading of high speaker certainty: if a speaker has specific evidence that *p* will occur, they are therefore typically also highly certain that *p* will occur.[20]

With respect to an expression with an agentive subject, "specific" evidence typically requires speaker knowledge of the relevant agent's plans to carry out the event described by *p*. Furthermore, the speaker should believe that the agent has the ability to carry out those plans. In the context in (30) below, the use of *-(y)AçAK* is licensed by the fact that the speaker knows that Mansur has a plan to go to Kazan (as shown by his purchasing tickets), and also that he has the ability to carry out that plan (since he can afford to go).[21]

(30) **Context:** You know that your friend Mansur bought plane tickets to go to Kazan. He can afford the trip and will be able to take time off of work to go. You report:

Mansur Qazan-ğa bar-**açaq**-∅.
Mansur Kazan-DAT go-PROSP-3SG
'[I have specific evidence that] Mansur will go to Kazan.'

-(y)AçAK is therefore preferred in expressions with first person subjects, since speakers are assumed to be authorities with respect to knowledge of their own plans and their ability to carry out those plans. However, it is not strictly required; *-(y)Er* is felicitous in expressions with first person subjects if the speaker has not made concrete plans yet. I provide an example of *-(y)AçAK* in a first person expression in (31).

[20] I argue in Chapter 5 that *-(y)AçAK* should be analyzed as a marker of prospective aspect. Cross-linguistically, prospective aspects frequently contribute readings of increased speaker certainty with respect to the reported event. Descriptions of these morphemes often explicitly link the reading of increased certainty to the presence of (specific) evidence that the event will occur. Such meanings have been described for prospective aspects in Syrian Arabic (Jarad 2014), Plains Cree (Wolvengrey 2006), and some Neo-Aramaic dialects from the Mosul Plain (Coghill 2010).

[21] My primary Tatar consultant notes that when Vladimir Putin (the president of Russia)'s speeches are translated from Russian into Tatar, his future tense utterances are always translated using *-(y)AçAK*. This is because Putin is assumed to have concrete plans, as well as the ability to carry out those plans, since he is a powerful political leader.

(31) **Context:** You booked tickets to go to Moscow in ten days. You are reporting your plans. You say:

un kön-nän min Mäskäü-gä {bar-**açaq**-min / #bar-**ır**-mın}.
ten day-ABL 1SG.NOM Moscow-DAT go-PROSP-1SG / go-FUT-1SG
'[I have specific/#non-specific evidence that] I will go to Moscow in ten days.'

(Speaker's comment on the use of *-(y)Er*: "Are you going or not? Make up your mind.")

-(y)AçAK is also licensed in expressions without agentive subjects. These include utterances regarding events that will necessarily occur, given what we know about the scientific facts of the world, as in (32). (This sentence was said in Los Angeles, where the sun rises and sets every day of the year.)

(32) qoyaş irtägä {bat-**açaq**-∅ / #bat-**ar**-∅}.
sun tomorrow set-PROSP-3SG / set-FUT-3SG
'[I have specific/#non-specific evidence that] The sun will set tomorrow.'

Speakers can also use *-(y)AçAK* to make predictions about the future. *-(y)AçAK* is licensed in predictions given (i) specific evidence that the event will occur, and (ii) the speaker's ability to authoritatively interpret that evidence. In the following examples, I give contrasting contexts to show that these criteria must be satisfied for *-(y)AçAK* to be felicitous. (I use a # to indicate that the context is incompatible with the utterance; in these incompatible contexts, the speaker must use *-(y)Er* instead of *-(y)AçAK*.)

In (33a), the speaker's scientific expertise enables them to interpret the weather reports and claim authoritatively that it will rain next week. If they lack this expertise, as in (33b), then they cannot felicitously use *-(y)AçAK*. In (34a), the speaker's medical expertise as a doctor enables them to assess the child's height and weight and assert authoritatively that they will grow up to be big and strong. If the speaker lacks this medical expertise, as in (34b), then they cannot use *-(y)AçAK*.

(33) a. **Meteorologist context:** You are a meteorologist. You see on your scans that a cold front is coming. You know that, given the climate of the area and the time of the year, the cold front will cause it to rain next week.

b. # **Speculation context:** You have no training in meteorology. While talking with your friends, you speculate that it will rain next week.

kiläse atna yañgır yaw-**açaq**-∅.
next week rain rain-PROSP-3SG
'[I have specific evidence that] It will rain next week.'
(Lit. 'Rain rains next week.')

(34) a. **Doctor's appointment context:** You take your son to a doctor's appointment. The doctor notes that your son has grown three inches since his last appointment, and that his weight is also steadily increasing. The doctor says to him:

b. # **Grandmother context:** Your grandmother is talking to your son and speculating about his bright future. She says to him:

sin däü häm köçle yeget bul-ıp üs-**äçäk**-sen.
2SG.NOM big and strong guy become-IP grow-PROSP-2SG
'[I have specific evidence that] You will grow up to be a big and strong guy.'

The expressions in (33) and (34) are acceptable in the contexts in (33b) and (34b), respectively, if the speaker has some way to exert control over future events. Perhaps the speaker has a magical ability to control the weather, and the grandmother possesses special medicines to ensure that her grandson will grow up big and strong. In these cases, the felicity of (33b) and (34b) would rely on the speaker's plans and ability to carry them out.[22]

Finally, *-(y)AçAK* is licensed when discussing the plots of films and video games. It is licensed in these contexts despite the fact that the speaker has no ability to control the outcome of the events.

[22]Alternately, in a context like (33b), uttering (33) would be felicitous if the speaker is intentionally flouting conversational norms in an attempt to sound authoritative.

(35) **Titanic context:** You are watching *Titanic* with your friend. You have seen it many times before, but she hasn't. You decide to spoil the ending for her. You say:

Jack {ül-**äçäk**-∅ / #ül-**är**-∅}.
Jack die-PROSP-3SG / die-FUT-3SG
'[I have specific/#non-specific evidence that] Jack will die.'

2.2.2.2 "Non-specific" evidence: *-(y)Er* 'FUT'

The Tatar TAE suffix *-(y)Er* is used in contexts in which the speaker lacks specific evidence that the event described by the scope proposition will occur. That is, *-(y)Er* is used in contexts in which speakers do not have the kind of evidence described in §2.2.2.1 that licenses the use of *-(y)AçAK*. I therefore treat *-(y)Er* as an "elsewhere" morpheme with respect to describing events located in the future of the utterance time.

My consultants report that Tatar expressions including *-(y)Er* convey that the speaker has a low or neutral degree of certainty that the event described by the scope proposition will occur. As a result, my consultants often translate these expressions using the English modals *possibly* or *maybe*. I give some examples of basic expressions with *-(y)Er* in (36)-(37). I note that in both of the following contexts, the speaker lacks any piece of specific evidence that the event described by the scope proposition will occur.

(36) **Playground context:** You are at a playground with Timur and his mother. Timur is playing on a platform high off the ground. You warn his mother that he might fall. You say:

Timur yeğıl-**ır**-∅.
Timur fall-FUT-3SG
'[I have non-specific evidence that] Timur will fall.'

Speaker's translation: 'Timur might fall.'

(37) **Work trip context**: A group of co-workers are discussing who will be sent on an upcoming work trip to Paris. Mansur thinks that Alsu is a pretty hard worker, so he speculates that she will go to Paris. He says:

Alsu Parij-ğa bar-**ır**-∅.
Alsu Paris-DAT go-FUT-3SG
'[I have non-specific evidence that] Alsu will go to Paris.'

Speaker's translation: 'Alsu might go to Paris.'

-(y)Er is also used in contexts in which the speaker is discussing events that occur repeatedly. In the context in (38), the speaker's knowledge that Mansur typically arrives at 1 o'clock every day is not enough to license the use of *-(y)AçAK*. The speaker would need to have some additional piece of knowledge (e.g. that Mansur rented a car specifically to come for lunch that day) to assert the expression using *-(y)AçAK*.

(38) **Context:** Mansur comes to your house at 1 o'clock every day for lunch. It's almost 1 o'clock now. You say:

Mansur kit-**er**-∅.
Mansur come-FUT-3SG
'[I have non-specific evidence that] Mansur will come.'

I note crucially that while my Tatar consultants frequently translate *-(y)Er* expressions using English epistemic possibility modals like *maybe*, *-(y)Er* expressions make stronger claims than Tatar expressions of epistemic possibility. I contrast *-(y)Er* and the epistemic possibility modal *bälki* 'maybe' in (39)-(40).[23] In the context in (39), the speaker generally knows that Nadir goes to school every week day. They therefore cannot use *bälki* 'maybe' in (39), since it would make too weak of a claim regarding the scope proposition.

[23] *Bälki* 'maybe' can only co-occur with *(y)Er*, as in (39)-(40); it is ungrammatical in combination with *-(y)AçAK*. The possible combinations of the Tatar modals and TAE morphemes is a major topic for future research.

(39) **Context:** Your next door neighbors have a child named Nadir. Since he is a child, you know that he generally goes to school every week day. Someone asks you what Nadir will do tomorrow. You say:

Nadir (#**bälki**) irtägä mäktäbkä bar-**ır**-∅.
Nadir maybe tomorrow school go-FUT-3SG
'[I have non-specific evidence that] Nadir will go to school tomorrow.'

Conversely, in the context in (40), omitting *bälki* would make too strong of a claim, since the speaker is uncertain that Nadir will go to school tomorrow. This data shows that *-(y)Er* expressions and *bälki* expressions are not truth-conditionally equivalent; there are some contexts in which one expression is felicitous and the other is not.

(40) **Sick child context**: You are Nadir's mother. He has been sick recently, and is still sick. You are telling his teacher that he might come to school tomorrow. You say:

Nadir #(**bälki**) irtägä mäktäbkä bar-**ır**-∅.
Nadir maybe tomorrow school go-FUT-3SG
'Nadir might go to school tomorrow.'

Finally, my Tatar consultants report that speakers overall use *-(y)Er* much more frequently than *-(y)AçAK* when reporting future events. This is unsurprising, given the "specific" evidence that is required to license the use of *-(y)AçAK*. This observation is also in line with my claim that *-(y)Er* is an "elsewhere" morpheme. If a speaker has strong enough evidence to license the use of *-(y)AçAK*, then they will use *-(y)AçAK*. If not, then they will use *-(y)Er*.

2.2.2.3 Evidentially neutral: *-A* 'PRES'

The Tatar TAE suffix *-A* is neutral with respect to evidentiality; I review it briefly in this section primarily to show that it lacks any evidential reading. Expressions containing *-A* are felicitous in both of the contexts in (41a)-(41b), regardless of whether the speaker has direct or indirect evidence for the scope proposition.[24]

[24]Since I showed in §2.2.2.1 and §2.2.2.2 that the future oriented Tatar TAE morphemes seem to contrast with respect to having "specific" versus "non-specific" evidence, as opposed to direct versus indirect evidence,

(41) a. **Direct evidentiality context:** You see Nima jumping. You say:

b. **Indirect evidentiality context:** Your cat runs and hides under your bed whenever Nima jumps. You are in your bedroom when your cat suddenly runs in and hides under the bed. You say:[25]

Nima siker-**e**-∅.
Nima jump-PRES-3SG
'Nima is jumping.'

I show in Chapter 3 that *-A* has a highly productive futurate use; futurates are expressions with no obvious means of future time reference that nonetheless locate the time of the described event in the future of the utterance time. Copley (2009, 2008) notes that in English (among other languages), futurate uses of present tense morphemes like *-A* have readings that could be described as modal or evidential; that is, they convey that the described event is "planned, scheduled, or otherwise determined" (Copley 2008, 261). Tatar *-A* patterns similarly to English in this respect. However, these modal or evidential readings have been documented for futurates cross-linguistically. These readings of *-A* therefore seem to involve a phenomenon that is separate from the consistently evidential readings of the other TAE suffixes in matrix clauses. I therefore do not focus on the semantics of *-A* in this dissertation.

2.2.3 Negation and the TAE suffixes: Lack of embedding

The evidential contribution of the Tatar TAE suffixes is interpreted above clausal negation; that is, in a expression with sentential negation, the evidential reading of the TAE suffixes "escapes" the negation. This is the case for both past oriented and future oriented TAE suffixes.

---

we might expect that *-A* might also be sensitive to some other sort of non-standard evidential distinction. However, I do not find any evidence for this.

[25]The speaker can optionally include the free particle *ikän* (previously discussed in footnote 18) in (41b) to emphasize their indirect evidence source, or to express surprise (i.e., mirativity). However, this particle is not required.

In the context in (42), the speaker has direct evidence that the scope proposition is false; they witnessed that Läylä did not perform in the concert. They can therefore only report the proposition "Läylä did not perform in the concert" using *-DI*, rather than *-GAn*.

(42) **Direct evidence that the scope proposition is false**

Concert context: You attended a concert in which a number of musicians performed. Aygöl says later that your mutual friend Läylä performed in it, but you were there, and you didn't see her play. You say:

Läylä kontsert-ta {qatnaş-ma-**dı**-∅ / #qatnaş-ma-**ğan**-∅}.
Läylä concert-DAT perform-NEG-PST-3SG / perform-NEG-RESULT-3SG
'[I have direct/#indirect evidence that] Läylä didn't perform in the concert.'

(Speaker's comment: "You can't say *qatnaşmağan* because you were at the concert.")

In the context in (43), the speaker has indirect evidence that the scope proposition is false. They can therefore only report the proposition "Mansur didn't go to Moscow" using *-GAn*, and not *-DI*.

(43) **Indirect evidence that the scope proposition is false**

Context: You thought Mansur was going to go to Moscow, but then you went past his house and saw his car in the driveway. You say:

Mansur Mäskäü-gä {bar-ma-**ğan**-∅ (ikän) / #bar-ma-**de**-∅}.
Mansur Moscow-DAT go-NEG-RESULT-3SG IKÄN / go-NEG-PST-3SG
'[I have indirect/#direct evidence that] Mansur didn't go to Moscow.'[26]

A similar pattern holds with respect to the negation of future oriented expressions. In the context in (44), the speaker has specific evidence that the scope proposition will not occur; that is, Aygöl told them that she plans **not** to read *The Idiot*. They therefore can only report the proposition using *-(y)AçAK*, and not *-(y)Er*.[27]

[26]My primary Tatar consultant prefers including the free particle *ikän* in (43), because the context is such that the statement contradicts a previously held belief that she had. That is, the proposition "Mansur didn't go to Moscow" contradicts her previously held belief that Mansur went to Moscow. This data can be explained by the use of *ikän* as a mirative morpheme, as discussed briefly in footnote 18.

[27]When *-(y)Er* co-occurs with the negative suffix *-m(A)*, it surfaces as the allomorph *-As*.

(44) **Specific evidence that the scope proposition is false**

Book club schedule context: You and your friend Aygöl are in a book club together, and are going over the schedule of upcoming assigned books. *The Idiot* is one of the upcoming books. Aygöl tells you that she is not going to read it. Later, you tell the other members of your book club:

Aygöl Idiot-nı {uqı-mı-**yaçaq**-∅ / #uqı-m-**as**-∅}.
Aygöl Idiot-ACC read-NEG-PROSP-3SG / read-NEG-FUT-3SG
'[I have specific/#non-specific evidence that] Aygöl will not read *The Idiot*.'

(Speaker's comment: "[The use of *uqımas* is bad] because that would give the impression that you haven't talked to her about her plans.")

Finally, in the context in (45), the speaker lacks any specific evidence for the scope proposition "Alsu will not read The Idiot." As a result, they can only report the proposition using *-(y)Er*, rather than *-(y)AçAK*.

(45) **Non-specific evidence that the scope proposition is false**

Book shopping context: You and your partner are going shopping for a present for your mutual friend Alsu. He suggests buying her *The Idiot*. However, you know that she doesn't like long Russian novels. You say:

Alsu Idiot-nı {uqı-m-**as**-∅ / #uqı-mı-**yaçaq**-∅}.
Alsu Idiot-ACC read-NEG-FUT-3SG / read-NEG-PROSP-3SG
'[I have non-specific/#specific evidence that] Alsu will not read *The Idiot*.'

The ability of evidentials to scope above negation is well-documented (e.g. Johanson 2003, 286 for other Turkic languages, Faller 2002, 227 for Cuzco Quechua, Izvorski 1997 for Bulgarian, among many others). The Tatar data is therefore unexceptional in this respect.

2.2.4 Recap: Evidential interpretation of TAE suffixes in matrix clauses

I have shown that in matrix clauses, the set of four core Tatar TAE suffixes addressed in this dissertation contribute both evidential and temporal meanings. I summarize the evidential requirements of the four core suffixes in Table 2.4, repeated from Table 2.3.

	Morpheme	Gloss	Evidential requirement
Past-oriented	*-DI*	'PST'	speaker has direct evidence
	-GAn	'RESULT'	speaker has indirect evidence
Future-oriented	*-(y)AçAK*	'PROSP'	speaker has "specific" evidence
	-(y)Er	'FUT'	speaker has "non-specific" evidence

Table 2.4: Core set of Tatar TAE morphemes addressed in this dissertation (repeated from Table 2.3).

I have shown that in matrix clauses, the past-oriented TAE suffixes *-DI* and *-GAn* require that the speaker has either direct or indirect evidence for the scope proposition, respectively. The distribution and use of these suffixes is similar to previously described direct/indirect evidential systems. However, the future-oriented TAE suffixes *-(y)AçAK* and *-(y)Er* appear to make an evidential distinction that has not been described in prior evidential typologies, e.g. Aikhenvald (2018, 2004) and Willett (1988). I will ultimately argue in Chapter 5 that this apparently "specific"/"non-specific" evidential requirement can be accounted for under an analysis of these suffixes as temporal morphemes. That is, the future-oriented Tatar TAE data presented in this chapter ultimately does not require reformulating our existing evidential typologies.

2.3 Evidential interpretation of semantically embedded TAE suffixes

In the following section, I describe the distribution and interpretations of the Tatar TAE suffixes in some semantically embedded environments, focusing specifically on their evidential readings. I take **semantically embedded environments** to include instances of clausal embedding as well as expressions in which propositions are embedded under operators, e.g. question operators. This section sets my descriptive work apart from Greed (2014, 2009), Tatevosov (2007), and Poppe (1961); these sources do not address the distribution and use of the TAE suffixes in semantically embedded environments.

I showed in Table 2.4 that the four core Tatar TAE suffixes appear to contribute evidential meanings in matrix clauses/semantically unembedded environments. In this section,

I show that their distribution and interpretation in semantically embedded environments raises questions with respect to their description as evidential morphemes. For instance, the evidential readings of the TAE suffixes disappear in some semantically embedded environments. This suggests that evidentiality may not in fact be a core part of their semantics.

What's more, the TAE suffixes also differ with respect to their morphosyntactic distribution in some embedded environments; specifically, they differ in embedded verbal nominalizations. The ways that the TAE suffixes pattern in these verbal nominalizations also suggests that evidentiality may not be a core part of their meaning. I describe the relevant data in §2.3.1.2.

Finally, I note that I do not ultimately provide a full-fledged analysis for all of the data that is presented in the following section. In Chapter 6, I sketch analyses of the embedded CP data discussed in §2.3.1.1 and the embedded verbal nominalization data in §2.3.1.2.

2.3.1 Clausal embeddings

Tatar has two primary strategies for forming embedded clauses: (i) embedding full CPs, as in (46a), and (ii) embedding verbal nominalizations (diagnosed by the presence of accusative case marking on the embedded nominalization), as in (46b). The choice of embedding strategy is itself semantically nontrivial. This is reflected by the different English translations for (46a) and (46b); both of the examples in (46) have the same matrix verb, *belergä*, but their interpretation hinges on the choice of embedding strategy. For the purpose of this discussion, I do not provide a gloss for *belergä*, since its interpretation varies based on the kind of clause it embeds. In the following examples, I simply gloss it as BEL.

(46) a. **Embedded CP**

Läylä [Michael Phelps ciñ-gän-∅ dip] bel-ä-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] BEL-PRES-3SG
'Läylä **thinks** that Michael Phelps won.'

b. **Embedded nominalization**[28]

Läylä [Michael Phelps ciñ-gän-e-n] bel-ä-∅.
Läylä [Michael Phelps win-RESULT-3SG.POSS-ACC] BEL-PRES-3SG
'Läylä **knows** that Michael Phelps won.'

If a speaker embeds a full CP under *belergä*, as in (46a), the resulting reading is that they (the speaker) do not necessarily believe that the embedded proposition is true. A possible context for (46a) would be one in which Läylä believes that Michael Phelps won a race, but the speaker doubts that he did. A natural English translation for (46a) would therefore be "Läylä thinks that Michael Phelps won," rather than "Läylä knows that Michael Phelps won."

Conversely, if the speaker embeds a verbal nominalization under *belergä*, as in (46b), there is no associated reading of speaker doubt that the embedded proposition is true. A possible context for (46b) would be one in which both Läylä and the speaker believe that Michael Phelps won a race. A natural English translation would therefore be "Läylä knows that Michael Phelps won."[29] I summarize the readings typically associated with the different clausal embedding strategies in Table 2.5.

Embedding strategy	**Typical interpretation**
Full CP	No speaker belief in truth of embedded proposition
Verbal nominalization	Speaker belief in truth of embedded proposition

Table 2.5: Semantic interpretations of embedded clause types in Tatar.

Tatar embedding verbs differ in whether they can combine with full CPs, verbal nominalizations, or both. *uylarğa* 'to think' can only embed full CPs, whereas *kürergä* 'to see'

[28]The right bracket marking the embedded clause in (46b) could technically be placed between the possessive and accusative morphemes on the nominalized verb: *ciñ-gän-e]-n*. I choose to place them after the accusative marking for readability, keeping in mind that the accusative case is marking the entire possessive construction.

[29]Özyıldız (2016) describes this semantic contrast between embedded CPs and verbal nominalizations in Turkish as a factivity alternation.

can only embed verbal nominalizations. *äytergä* 'to say' can combine with both embedded clause types.

My consultants report that the interpretation of expressions with *äytergä* as a matrix verb are **generally** consistent regardless of the type of clausal embedding. That is, *äytergä* seems to be less semantically sensitive to the type of embedded clause than other embedding verbs like *belergä* in (46). However, I show in §2.3.1.1 that *äytergä* is not completely immune to the semantic effects of embedded clause type. In fact, I propose that clause type influences the felicity of some embedding expressions.

For simplicity, I primarily discuss and analyze examples of clausal embeddings under *äytergä*, in this section and elsewhere in the dissertation.[30]

2.3.1.1 Evidentiality in embedded CPs

All of the Tatar TAE suffixes in Table 2.3 can occur in embedded CPs. In embedded CPs, the TAE suffixes maintain the same evidential contributions as in matrix clauses. However, the suffixes appear to differ with respect to **evidential shift**.

Evidential shift refers to the ability of the evidential to "shift" from its default interpretation reflecting the **speaker's** evidence for the embedded proposition; shifted evidentials instead reflect the **matrix subject's** evidence for the embedded proposition. Evidential shift is a noted point of variation across evidential systems (Korotkova 2016). Evidential shift is unavailable in some languages (Georgian), obligatory in some languages (Korean, Japanese, Zazaki), and optional in others (German, Turkish). I provide an example of non-shifted and shifted readings of the Turkish TAE suffix *-mIş* in (47). I underline the individual whose evidence *-mIş* reflects.

[30]When discussing embedded clauses, I only use examples of *äytergä* combined with the TAE suffix *-DI* (i.e. *äytte*). This is for simplicity. As I described in §2.2.1, *-DI* contributes a reading of direct evidentiality, whereas *-GAn* contributes a reading of indirect evidentiality. Testing the evidential readings of embedded propositions is challenging enough without tinkering with the evidential status of the matrix clause.

(47) TURKISH (TURKIC) (adapted from Korotkova 2016, 182)

a. **Non-shifted context**

Speaker: Indirect evidence

Matrix subject: Direct evidence

I was told by Mary, Anna's roommate, that Anna got a dog. Jay visited them recently and has seen the dog himself.

Jay [Anna bir köpek al-**mış**] di-yor.
Jay [Anna INDEF puppy get-MIŞ] say-PST
'Jay said that—and I've heard it—Anna got a puppy.'

b. **Shifted context**

Speaker: Direct evidence

Matrix subject: Indirect evidence

I recently visited Anna and found out that she finally got a dog. Jay hasn't visited yet, but she called him to share the news.

Jay [Anna bir köpek al-**mış**] di-yor.
Jay [Anna INDEF puppy get-MIŞ] say-PST
'Jay said that, as he has heard, Anna got a puppy.'

Before I introduce the Tatar data, I would like to discuss some of the issues with respect to collecting evidential shift data in Tatar. As I noted in §2.3.1, Tatar embedding verbs differ with respect to whether or not they can combine with both embedded clause types (i.e., both embedded CPs and verbal nominalizations). I choose to use *äytergä* 'to say' as a matrix verb due to the fact that it can combine with both embedded clause types, therefore enabling me to compare minimal (or near-minimal) pairs of expressions.

In addition to this morphosyntactic property, *äytergä* is compatible with both the speaker and the matrix subject having either direct or indirect evidence for the embedded proposition. This contrasts with other Tatar embedding verbs that can combine with both embedded clause types, like *işeterğa* 'to hear.' *işeterğa* can embed both clause types; however, its meaning is such that the matrix subject necessarily has indirect evidence for the embedded proposition. The combination of morphosyntactic and semantic properties of *äytergä*

therefore make it a useful verb with which to test evidential shift.

However, the meaning of *äytergä* is such that its use in contexts in which the speaker has direct evidence for the embedded proposition are pragmatically odd. These are contexts like the Turkish example in (47b). When eliciting Tatar examples analogous to (47b), my primary consultant typically remarks that "If you witnessed [the event described by the embedded proposition], you wouldn't be saying the sentence to begin with," i.e., if the speaker had direct evidence for the embedded proposition, they would relay their own observation of it rather than someone else's report of it. My attempts at eliciting examples like (47b) were therefore frequently unsuccessful, since the overall expressions were judged as pragmatically odd in the contexts.

With these caveats in mind, I can nonetheless lay out some generalizations regarding the availability of evidential shift in embedded CPs in Tatar. The past oriented and future oriented TAE suffixes pattern differently with respect to evidential shift. The future oriented TAE suffixes **obligatorily** shift: that is, they can only reflect the evidence of the matrix subject. However, I tentatively conclude that the past oriented TAE suffixes pattern like Turkish *-mIş* in that they can **optionally** shift: that is, they can reflect the evidence of either the matrix subject or the speaker. I summarize these facts in Table 2.6. ✓ indicates that the evidential interpretation is available, whereas ✗ indicates that it is unavailable.

TAE suffix	**Can reflect speaker's evidence for embedded proposition**	**Can reflect matrix subject's evidence for embedded proposition**
-DI 'PST'	✓(?)	✓
-GAn 'RESULT'	✓	✓(?)
-(y)AçAK 'PROSP'	✗	✓
-(y)Er 'FUT'	✗	✓

Table 2.6: Available evidential interpretations of the Tatar TAE suffixes in embedded CPs.

I begin by discussing the evidential readings of the future oriented TAE suffixes in embedded CPs, since my consultants' judgements are much crisper with respect to this data.

Future oriented TAE suffixes in embedded CPs

I give examples of embedded future oriented TAE morphemes in (48)-(49), and show that the evidential reading of the morphemes obligatorily shifts. That is, future oriented TAE morphemes in embedded CPs obligatorily reflect the matrix subject's evidence for the embedded proposition.

In (48), the matrix subject has specific evidence for the embedded proposition "Michael Phelps will win." However, the speaker does not have specific evidence for this proposition. The speaker must report (48a) with *-(y)AçAK* in the embedded clause; this use of *-(y)AçAK* reflects the matrix subject's evidence, not the speaker's. It is infelicitous in this context for the speaker to use *-(y)Er* in the embedded clause, as in (48b). *-(y)Er* is infelicitous in this context because the matrix subject has specific evidence for the embedded proposition.

(48) **Shifted context: Swimming fan context #1**

Speaker: Non-specific evidence

Matrix subject: Specific evidence

Läylä is watching an Olympic swimming event. She is a huge fan of competitive swimming, and knows the statistics about all of the swimmers. She believes that Michael Phelps has trained the best out of everyone competing at the Olympics this year: his diet, practice regimen, etc. are all impeccable. She tells you he will definitely win. However, you have no strong opinion about his ability to win. Later, you tell someone else:

a. Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG
'Läylä said that [and has specific evidence that] Michael Phelps will win.'

b. # Läylä [Michael Phelps ciñ-**er**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-FUT-3SG COMP] say-PST-3SG
'Läylä said that [and has non-specific evidence that] Michael Phelps will win.'

(Speaker's comment: "This means that [Läylä] is not 100% sure that he will win.")

Conversely, in (49), the speaker has specific evidence for the embedded proposition "Michael Phelps will win." However, the matrix subject does not. The speaker therefore must report (49a) with *-(y)Er* in the embedded clause. This use of *-(y)Er* reflects the fact that the matrix subject does not have specific evidence for the embedded proposition. *-(y)AçAK* is infelicitous in this context (49b), since the matrix subject does not have specific evidence.

(49) **Shifted context: Swimming fan context #2**

Speaker: Specific evidence

Matrix subject: Non-specific evidence

You are watching an Olympic swimming event with your friend Läylä. You are a huge fan of competitive swimming, and know the statistics about all of the swimmers. You believe that Michael Phelps has trained the best out of everyone competing at the Olympics this year: his diet, practice regimen, etc. are all impeccable. However, Läylä has no strong opinion about Michael Phelps's ability to win. Before the race starts, Läylä speculates that he will win. You tell someone else:

a. Läylä [Michael Phelps ciñ-**er**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-FUT-3SG COMP] say-PST-3SG
'Läylä said that [and has non-specific evidence that] Michael Phelps will win.'

b. # Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG
'Läylä said that [and has specific evidence that] Michael Phelps will win.'

It is possible for Tatar speakers to explicitly assert that their evidence for the embedded proposition differs from the matrix subject's. In (50), the speaker reports that the matrix subject has specific evidence that the event described by the embedded proposition will occur; this is conveyed via their use of *-(y)AçAK* in the embedded clause. They can then follow this by asserting explicitly that they (i.e., the speaker) lack specific evidence for the embedded proposition, or have evidence to the contrary. This shows that the use of *-(y)AçAK* in the embedded clause in (50) does not reflect the speaker's evidence.

(50) Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅, läkin min
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG but 1SG.NOM
uzim anı turında kup bel-m-i-m.
self.1SG 3SG.ACC about much BEL-NEG-PRES-1SG
'Läylä says that [and has specific evidence that] Michael Phelps will win, but I myself don't know much about him.'

The opposite scenario holds in (51). In (51), the speaker reports that the matrix subject lacks specific evidence regarding the embedded proposition "Michael Phelps will win." This is indicated by their use of *-(y)Er* in the embedded clause. They can then follow this expression by explicitly asserting that they have specific evidence that the event described by the embedded proposition will occur, or that they believe that it will occur. This shows that the use of *-(y)Er* in the embedded clause in (51) does not reflect the speaker's evidence.

(51) Läylä [Michael Phelps ciñ-**er**-∅ dip] äyt-te-∅, min-de
Läylä [Michael Phelps win-FUT-3SG COMP] say-PST-3SG 1SG.NOM-ADD
[anın ciñ-eyeçeg-u-n] bel-ä-m.
[3SG.GEN win-PROSP-3SG.POSS-ACC] BEL-PRES-1SG
'Läylä says that [and has non-specific evidence that] Michael Phelps will win, and I know that he will win.'

Past oriented TAE suffixes in embedded CPs

Speakers' judgments regarding the evidential interpretation of past oriented TAE suffixes in embedded CPs are much murkier. Unlike the future oriented TAE suffixes in (48)-(49), both of the past oriented TAE suffixes are consistently felicitous in embedded CPs, regardless of the context. Since there is no contrast in acceptability depending on context, I tentatively conclude that the past oriented TAE suffixes differ from the future oriented TAE suffixes in that they can optionally undergo evidential shift.

I begin by discussing contexts in which the speaker has indirect evidence for the embedded proposition. Speakers report that it is natural to use *äytergä* as a matrix verb in these contexts, since the speaker is expressing someone else's report of an event that they

themselves did not directly perceive. I show examples of such contexts in (52)-(53).

In (52), the context is such that the speaker has indirect (reportative) evidence for the embedded proposition "Michael Phelps won," and the matrix subject has direct evidence for the same proposition. The speaker's use of *-GAn* in the embedded clause therefore reflects their indirect evidence for the embedded proposition; i.e., in (52), *-GAn* remains speaker oriented and does not undergo evidential shift.

(52) **Non-shifted context: Olympics attendee context**

Speaker: Indirect evidence

Matrix subject: Direct evidence

Läylä went to the Olympics and saw Michael Phelps compete. She calls you and tells you he won; you did not see the race yourself. You then tell someone else:

Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] say-PST-3SG
'Läylä said that [and I have indirect evidence that] Michael Phelps won.'

(Speaker's comment: "It's possible to say (52) even if Läylä saw it happen, because you are the speaker reporting about Läylä.")

The context in (53) is again such that the speaker has indirect (reportative) evidence for the embedded proposition "Michael Phelps won," and the matrix subject has direct evidence for the same proposition. In this example, the use of *-DI* in the embedded clause reflects the matrix subject's direct evidence for the embedded proposition; i.e., *-DI* in this example undergoes evidential shift.[31]

[31]Recall from §3.2.1.1 that it is possible for Tatar speakers to use *-DI* to report events that they did not witness as long as they heard about the event from someone who they believe to be a reliable source (e.g. (11)). The availability of examples like (11) complicates our understanding of (53). Theoretically, it could be the case that *-DI* in (53) has remained unshifted and is reflecting the speaker's assessment of Läylä as a reliable source regarding Michael Phelps winning.

(53) **Shifted context: Olympics attendee context**

Speaker: Indirect evidence

Matrix subject: Direct evidence

Läylä went to the Olympics and saw Michael Phelps compete. She calls you and tells you he won; you did not see the race yourself. You then tell someone else:

Läylä [Michael Phelps ciñ-**de**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and has direct evidence that] Michael Phelps won.'

As I noted previously, contexts in which the speaker has direct evidence for the embedded proposition are judged to be somewhat pragmatically odd. I propose that this stems from two sources. This stems primarily from the fact that it is odd for a speaker to report a third party's report of an event that they themselves witnessed. Secondarily, it stems from the semantic contribution of CP embeddings. As discussed in §2.3.1, CP embeddings convey that the speaker does not know whether the embedded proposition is true. It is therefore pragmatically odd for a speaker to indicate that they do not know whether an event that they witnessed is true. Given these major caveats, I will nonetheless report the relevant Tatar CP embedding data; the marginality of these expressions stems from the semantics of the matrix verb and embedding strategy, rather than the embedded TAE suffixes themselves.

In (54), the context is such that the speaker has direct evidence for the embedded proposition "Michael Phelps won," and the matrix subject has indirect evidence for the same proposition. The speaker's use of *-DI* in the embedded clause can reflect their direct evidence for the embedded proposition; i.e., *-DI* can remain unshifted.

(54) **Non-shifted context: Neighboring sports fan context**

Speaker: Direct evidence

Matrix subject: Indirect evidence

You are attending the Rio Olympics; you see Michael Phelps compete and win. Your friend Aygöl did not attend the Olympics; however, her next door neighbor is a huge fan of Michael Phelps. She knows that the Olympics are going on right now, and she hears her neighbor cheering through the wall. She calls you and says that she heard that Michael Phelps won. You then tell someone else:

?Aygöl [Michael Phelps ciñ**-de**-∅ dip] äyt-te-∅.
Aygöl [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Aygöl said that [and I have direct evidence that] Michael Phelps won.'

Conversely, in (55), the speaker has direct evidence for the embedded proposition "Michael Phelps won," whereas the matrix subject has indirect evidence for the same proposition. The use of *-DI* in the embedded clause in (55) can therefore reflect the speaker's direct evidence for the embedded proposition; i.e., it can undergo evidential shift.

(55) **Shifted context: Neighboring sports fan context**

Speaker: Direct evidence

Matrix subject: Indirect evidence

You are attending the Rio Olympics; you see Michael Phelps compete and win. Your friend Aygöl did not attend the Olympics; however, her next door neighbor is a huge fan of Michael Phelps. She knows that the Olympics are going on right now, and she hears her neighbor cheering through the wall. She calls you and says that she inferred that Michael Phelps won. You then tell someone else:

?Aygöl [Michael Phelps ciñ**-gän**-∅ dip] äyt-te-∅.
Aygöl [Michael Phelps win-RESULT-3SG COMP] say-PST-3SG
'Aygöl said that [and has indirect evidence that] Michael Phelps won.'

In sum, I conclude that the Tatar TAE suffixes can optionally shift in embedded CPs, primarily due to the availability of data like (52)-(53).

2.3.1.2 Evidentiality in embedded verbal nominalizations

Tatar embedded verbal nominalizations differ from embedded CPs in two major ways. First, only a subset of the TAE suffixes can occur in them, as shown in Table 2.7. Second, the TAE suffixes make no evidential contribution in embedded nominalizations; their contribution is primarily temporal.[32]

TAE suffix	Grammaticality in verbal nominalizations
-DI 'PST'	*
-GAn 'RESULT'	✓ (no evidential reading)
-(y)AçAK 'PROSP'	✓ (no evidential reading)
-(y)Er 'FUT'	*

Table 2.7: Grammaticality and evidential interpretation of the Tatar TAE suffixes in embedded verbal nominalizations.

I begin by addressing the first point of variation. Only the TAE suffixes *-GAn* and *-(y)AçAK* can occur in verbal nominalizations; *-DI* and *-(y)Er* are ungrammatical. I demonstrate the ungrammaticality of *-DI* and *-(y)Er* in embedded verbal nominalizations in (56)-(57). (The ungrammaticality of these examples does not stem from e.g. selectional restrictions on the part of the matrix verb; *-DI* and *-(y)Er* are robustly ungrammatical in all embedded verbal nominalizations.)

(56) * Mansur [Läylä-nen Qazan-ğa bar-**dı**-ı-n] bel-ä-∅.
Mansur [Läylä-GEN Kazan-DAT go-PST-3SG.POSS-ACC] BEL-PRES-3SG
Intended: 'Mansur knows Läylä went to Kazan.'

(57) * Mansur [Läylä-nen Qazan-ğa bar-**ır**-ı-n] bel-ä-∅.
Mansur [Läylä-GEN Kazan-DAT go-FUT-3SG.POSS-ACC] BEL-PRES-3SG
Intended: 'Mansur knows Läylä will go to Kazan.'

I take this lack of morphosyntactic cohesion within the set of Tatar TAE suffixes to suggest that we should treat *-DI/-(y)Er* and *-GAn/-(y)AçAK* as forming two different natural

[32] Straughn (2011, 56) notes that embedded verbal nominalizations in Kazakh and Uzbek also lack any evidential meaning.

classes, despite their very different evidential interpretations in matrix clauses. In §2.2, I described the TAE suffixes *-DI* and *-(y)AçAK* as making "strong" claims (in the sense of e.g. Faller 2002) with respect to the speaker's evidence for the scope proposition. Conversely, the TAE suffixes *-GAn* and *-(y)Er* make "weaker" claims with respect to the speaker's evidence. Descriptively speaking, these pairs of suffixes appear to be past- and future-oriented counterparts of the other with respect to their evidential strength. If the TAE suffixes have underlyingly evidential meanings, we might therefore expect that the two sets would pattern together morphosyntactically (i.e., *-DI* and *-(y)AçAK* would pattern together and *-GAn* and *-(y)Er* would pattern together). However, data like (56)-(57) shows that this expectation is not borne out: *-(y)AçAK* and *-GAn* pattern together morphosyntactically to the exclusion of *-DI* and *-(y)Er*. (I ultimately use this data in Chapters 5 and 6 to argue that the TAE suffixes do not have evidential semantics, and that they in fact differ with respect to their temporal meanings.)

The second point of variation in embedded verbal nominalizations is with respect to the availability of evidential readings. I propose that the TAE suffixes in embedded verbal nominalizations contribute **no** evidential reading with respect to either the speaker's or matrix subject's evidence for the event described by the embedded clause. For simplicity, I only show the felicity of *-GAn* and *-(y)AçAK* in verbal nominalizations relative to the matrix subject's evidence for the embedded proposition. I showed in §2.3.1.1 that this shifted reading is always available in embedded CPs (and, as previously discussed, embeddings under *äytergä* 'to say' are pragmatically odd if the speaker has direct evidence for the embedded proposition).

The embedded *-GAn* expression in (58) is compatible with the matrix subject having either direct or indirect evidence for the embedded proposition "Läylä went to Kazan yesterday," as shown by the contexts in (150a)-(150b). I take this to show that in an embedded verbal nominalization, *-GAn* makes no evidential contribution.[33]

[33]Anecdotally, my Tatar consultants also describe instances of *-GAn* in embedded verbal nominalization as having "nothing to do with" evidential readings.

(58) a. **Matrix subject has direct evidence for embedded proposition:**
Mansur tells you that he accompanied Läylä to the train station yesterday and watched her get on a train to go to Kazan. You express Mansur's report of what Läylä did yesterday.

b. **Matrix subject has indirect evidence for embedded proposition:**
Mansur finds a receipt for a train ticket to Kazan in Läylä's desk. He tells you that he infers that Läylä went to Kazan yesterday. You express Mansur's report of what Läylä did.

Mansur [Läylä-nen Qazan-ğa kit-ep bar-**ğan**-ı-n] äyt-te-∅.
Mansur [Läylä-GEN Kazan-DAT leave-IP go-RESULT-3SG.POSS-ACC] say-PST-3SG
'Mansur said Läylä left for Kazan.'

The interpretation of embedded *-(y)AçAK* expressions like (59) differs slightly from embedded *-GAn* expressions. (59) is compatible with the matrix subject having specific evidence for the embedded proposition "Läylä will go to Kazan tomorrow," as in the context in (59a). However, my primary Tatar consultant reports that she is slightly less likely to utter (59) in a context like (59b), in which the matrix subject lacks specific evidence for the embedded proposition.

(59) a. **Specific evidence context:** Mansur is a close friend of Läylä's. He knows that Läylä has bought tickets to go to Kazan next week, and has taken time off work specifically to go on vacation there. He tells you about her plans. You express Mansur's report of what Läylä will do.

b. ?**Non-specific evidence context:** Mansur is a colleague of Läylä's. He overhears another colleague saying that Läylä will go to Kazan next week, and tells you this. He does not know about any particular plans of hers, and he has no reason to think this colleague is an authority on Läylä's actions. You express Mansur's report of what Läylä will do.

Mansur [Läylä-nen Qazan-ğa bar-**açağ**-ı-n] äyt-te-∅.
Mansur [Läylä-GEN Kazan-DAT go-PROSP-3SG.POSS-ACC] say-PST-3SG
'Mansur said Läylä will go to Kazan.'

I propose that this weak contrast is due to the semantic contribution of the embedded verbal nominalization itself, rather than from any evidential contribution of the embedded *-(y)AçAK*. Recall from §2.3.1 that embedded verbal nominalizations generally convey that the speaker believes that the embedded proposition is true (in Özyıldız 2016's terminology, embedded verbal nominalizations are factive). This makes uttering (59) in the context in (59a) quite natural, since the speaker knows that Mansur has strong, specific evidence that the event described by the embedded proposition will happen. However, in the context in (59b), the speaker knows that Mansur is not an authority on Läylä's actions, since he is "only" a colleague and not a close friend of hers; furthermore, his knowledge of her going to Kazan is through hearsay. It is therefore less likely that the speaker would believe that Mansur's report in (59b) is true. This in turn would make it less likely that the speaker would talk about Mansur's report using an embedded verbal nominalization.

I ultimately conclude from this data that Tatar TAE suffixes do not contribute any evidential meaning in embedded verbal nominalizations.[34]

2.3.2 Evidentiality in questions: Interrogative flip

Interrogative flip concerns the behavior of evidentials in questions. Aikhenvald (2004) and Aikhenvald and Dixon (2003) note that evidentials tend to be restricted in questions cross-linguistically; in some languages, only a subset of the language's evidentials can occur in questions (Tariana, Tucano), whereas in others, evidentials are completely ungrammatical in questions (Abkhaz, Baniwa, among others) (Aikhenvald 2004, 242-243). All of the Tatar TAE suffixes are grammatical in questions.

[34]Relative clauses in Tatar also involve verbal nominalizations, as in (60). (Like the embedded verbal nominalizations, only *-GAn* and *-(y)AçAK* are grammatical in relative clauses.)

(60) Timur [Aygöl peşer-gän tort-nı] urla-dı-∅.
Timur [Aygöl make-RESULT cake-ACC] steal-PST-3SG
'Timur stole the cake that Aygöl made.'

The TAE suffixes in relative clauses pattern similarly to the embedded verbal nominalization data with respect to their evidential contribution. For instance, *-GAn* is licensed in relative clauses regardless of whether the speaker has direct or indirect evidence for the event described by the relative clause. Similarly, *-(y)AçAK* appears to contribute a weak evidential reading in relative clauses. There is a slight preference for using *-(y)AçAK* when the speaker believes that the event described by the relative clause occurred.

Many of the languages that permit evidentials to occur in questions are subject to **interrogative flip**. Interrogative flip describes a phenomenon in which evidentials in questions mark the kind of evidence that the speaker assumes that the addressee has for the answer to the question (San Roque et al. 2015). This differs from declarative clauses, in which the evidentials reflect the speaker's evidence source. That is, in instances of interrogative flip, the interpretation of the evidential "flips" from the subject to the addressee. I give a basic schema for this in Figure 2.2.

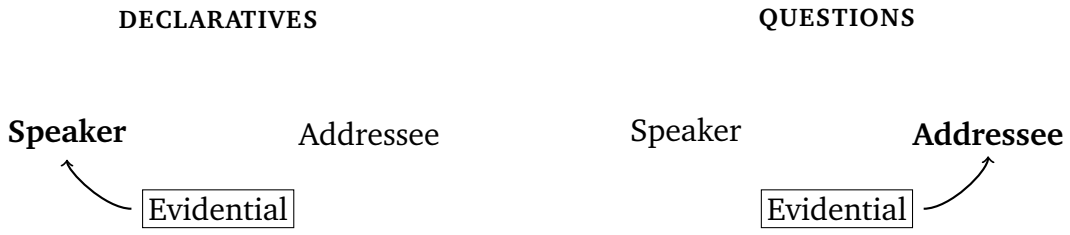


Figure 2.2: Interrogative flip: The interpretation of evidentials in declarative versus interrogative expressions.

Interrogative flip occurs in Tatar.[35] In the context in (62), the mother must use *-DI* when asking the babysitter whether Yasmin ate the yogurt. This is because she assumes that the babysitter has direct evidence for the proposition "Yasmin ate the yogurt." *-GAn* is infelicitous in this context, despite the fact that the mother lacks direct evidence for this proposition. This shows that the evidential meaning of *-DI* reflects the babysitter's evidence, and not the mother's.[36]

[35]Interrogative flip also occurs in embedded questions in Tatar, as in (61).

(61) Context: A babysitter has been watching a child named Yasmin for the evening. Yasmin is still young enough that she must be sat in a highchair and fed by someone else in order to eat. Her mother, Aygöl, returned home and asked the babysitter if Yasmin ate her yogurt.
Aygöl [Yasmin yogurt-nı {aşa-**dı**-∅ / #aşa-**ğan**-∅} mı dip] sara-dı-∅.
Aygöl [Yasmin yogurt-ACC eat-PST-3SG / eat-RESULT-3SG Q COMP] ask-PST-3SG
'Aygöl asked whether [the babysitter had direct evidence that] Yasmin ate the yogurt.'
(Speaker's comment: "[If you use *-GAn*], it's like she asked someone who wasn't there [whether Yasmin ate the yogurt].")

[36]If the speaker uses an infelicitous evidential in their question, the addressee will respond with the appropriate evidential. The addressee will not match their evidential use to accommodate that of the speaker's question. An example of this would be if the mother asked (62) using *-GAn*, which would suggest that the

(62) **Addressee: Direct evidence**

Babysitter context (adapted from Meriçli 2016, 10): A babysitter has been watching a child named Yasmin for the evening. Yasmin is still young enough that she must be sat in a highchair and fed by someone else in order to eat. (Consequently, at any meal, the person feeding her sees everything that Yasmin eats.) Her mother, Aygöl, returns home and asks the babysitter if Yasmin ate her yogurt.

Yasmin yogurt-nı {aşa-**dı**-∅ / #aşa-**ğan**-∅} mı?
Yasmin yogurt-ACC eat-PST-3SG / eat-RESULT-3SG Q
'[Do you have direct evidence that] Did Yasmin eat the yogurt?'

Conversely, if the speaker assumes that the addressee has indirect evidence for their answer, they must use *-GAn* in their question.[37] In the context in (64), Yasmin's father assumes that her mother has indirect evidence for the proposition "Yasmin ate the yogurt." He therefore asks the question using *-GAn*. *-DI* is infelicitous in this context, since he should know that her mother was also away for the evening and did not see Yasmin eat her dinner.

---

babysitter had indirect evidence for whether Yasmin ate the yogurt. If the babysitter in fact saw Yasmin eat the yogurt, she would simply provide her answer using *-DI*.

[37] *-GAn* is also required in so-called conjectural questions, in which interrogative flip does not occur. Conjectural questions are defined by Littell et al. (2010) as questions in which the speaker assumes that the addressee does not know the answer. In conjectural questions like (63), the indirect evidential marker is used to indicate that the speaker is not expecting an informative response, and that they are simply "wondering" about the answer. (Similar expressions are also described in Turkish (Meriçli 2016) and Salishan languages (Littell et al. 2010).)

(63) Teddy bear context: You find a teddy bear in your sons' room that you don't remember giving them. They are away at school right now, so you can't ask them who gave it to them. Talking aloud to yourself, you say:
bu qayan {kil-**gän**-∅ / #kil-**de**-∅} tağın?
this where come-RESULT-3SG / come-PST-3SG more?
'Where did this come from?'
(Speaker's comment: "[You can't say *kilde*] because you obviously don't know where it came from. You're wondering.")

(64) **Addressee: Indirect evidence**

Updated babysitter context (adapted from Meriçli 2016, 10): Yasmin's father comes home and asks his wife Aygöl whether Yasmin ate her yogurt. He knows that Yasmin was with a babysitter and that Aygöl did not witness Yasmin eating the yogurt; he assumes that she heard from the babysitter whether or not she ate it.

Yasmin yogurt-nı {aşa-**ğan**-∅ / #aşa-**dı**-∅} mı?
Yasmin yogurt-ACC eat-RESULT-3SG / eat-PST-3SG Q
'[Do you have indirect evidence that] Did Yasmin eat the yogurt?'

The future-oriented TAE morphemes are also felicitous in questions, and are also subject to interrogative flip. To my knowledge, there are no examples in the interrogative flip literature of the interpretation of evidentials in future-oriented questions. The Tatar data on this topic is therefore especially interesting.

In contexts in which the speaker assumes that the addressee has specific evidence for the answer to their question, the speaker must ask their question using *-(y)AçAK*. I show an example of this in (65). *-(y)Er* is infelicitous in this context, since the teacher should have already decided on the students' grades.

(65) **Addressee: Specific evidence**

Teaching assistant context: You are a teaching assistant at a university. It is the end of the quarter and you are discussing final grades with the professor, who ultimately will assign students their grades. The professor has already decided on the students' grades. You are curious if your favorite student, Alsu, will get an A. You ask:

Alsu kurs-ta A {al-**açaq**-∅ / #al-**er**-∅} mı?
Alsu course-LOC A take-PROSP-3SG / take-FUT-3SG Q
'[Do you have specific evidence that] Will Alsu get an A in the class?'

Conversely, in the context in (66), the speaker assumes that the addressee lacks any specific evidence for their answer. Since the addressee is another student, they have no access to Timur's academic records; furthermore, since Timur has not even enrolled in the

course yet, the addressee has no evidence for how well he would perform. The speaker therefore must use *-(y)Er* in their question; *-(y)AçAK* is infelicitous in this context.

(66) **Addressee: Non-specific evidence**

"Easy A" context: Timur is not a great student and needs to take an easy class to boost his GPA. You're speculating with another student about whether or not he could take a certain class in order to get an easy A in it. Neither of you have any control over the class grades, or have any specific idea about how he would perform in the class. You ask them:

Timur kurs-ta A {al-**er**-∅ / #al-**açaq**-∅} mı?
Timur course-LOC A take-FUT-3SG / take-PROSP-3SG Q
'[Do you have non-specific evidence that] Will Timur get an A in the class?'

2.3.3 Recap: Evidential contribution of Tatar TAE suffixes in semantically embedded environments

In this section, I described the evidential interpretations of the Tatar TAE suffixes in some semantically embedded environments. I showed that in full CP embeddings, the TAE suffixes contribute the same evidential meanings as in matrix clauses (modulo the ability/requirement to undergo evidential shift).

However, I also showed that the distribution and interpretation of the TAE suffixes in some embedded environments differs from their use in matrix clauses in two major ways. First, only *-GAn* and *-(y)AçAK* are grammatical in embedded environments involving verbal nominalizations. Second, the evidential readings of these morphemes effectively disappears in the embedded verbal nominalizations. I summarize these facts in Table 2.8.

In sum, the semantic embedding data discussed in this section complicates our understanding of the Tatar TAE suffixes as underlyingly evidential morphemes.

TAE suffix	Grammaticality in embedded CPs	Grammaticality in embedded verbal nominalizations
-DI 'PST'	✓	*
-GAn 'RESULT'	✓	✓ (no evidential reading)
-A 'PRES'	✓	*
-(y)AçAK 'PROSP'	✓	✓ (no evidential reading)
-(y)Er 'FUT'	✓	*

Table 2.8: Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings.

2.4 Conclusion

This chapter provided a descriptive overview of the evidential readings that are associated with the Tatar TAE morphemes in matrix clauses and some semantically embedded environments. I focused particularly on the set of four core TAE suffixes for which I provide a semantic and pragmatic analysis in Chapter 5 (shown in Table 2.3).

The Tatar TAE system behaves generally as how we would expect for a Turkic language (cf. Johanson 2003, among many others). My description of the Tatar TAE data differs from other authors' in that I propose to treat both past and future oriented TAE suffixes as making an evidential distinction, descriptively speaking. I propose that the unconventional "specific"/"non-specific" evidential distinction that I make with respect to the future oriented TAE suffixes will ultimately help us come to a unified theoretical understanding of the semantics of the Tatar TAE system.

In the following chapter, I describe the temporal contributions of the Tatar TAE suffixes. I will ultimately argue in Chapter 5 that the TAE suffixes have underlyingly temporal, not evidential, meanings: I propose that *-DI*, *-A*, and *-(y)Er* should be analyzed as tenses, whereas *-GAn* and *-(y)AçAK* should be analyzed as aspects. The evidential meanings that I described in this chapter are therefore byproducts of their underlyingly temporal semantics. My theoretical analysis of these suffixes must therefore be able to derive the evidential meanings described in this chapter from their temporal meanings.

CHAPTER 3

Temporal contributions of the Tatar TAE suffixes

3.1 Introduction

This chapter continues the descriptive overview of the portmanteau Tatar tense/aspect/evidential (TAE) system. In Chapter 2, I described the evidential meanings that the Tatar TAE suffixes contribute. In this chapter, I lay out the **temporal** contributions of the same morphemes; i.e., how they locate the event described by the scope proposition in time. I will ultimately propose in Chapter 5 that these morphemes are fundamentally tense and aspect markers. The goal of this chapter is therefore to describe their distribution and use as tense/aspect markers such that my eventual temporal analysis is well-motivated.

The temporal contributions of the Tatar TAE suffixes are similar to those of other Turkic tense/aspect systems. Cognate past-oriented *-DI* and *-GAn* morphemes are widespread across Turkic languages; Johanson (1998, 44) asserts that "all Turkic languages have [simple past] items (...) of the type *-DI*."[1] Cognates of *-GAn* are described as perfects in a number of languages (e.g. Johanson 2000 for Karachay, Kumyk, and Tuvan; Csató and Karakoç 1998 for Noghay); however, I discuss in §3.2.1.2 why I do not analyze Tatar *-GAn* as a (present) perfect.[2] Cognates of Tatar future-oriented *-(y)AçAK* and *-(y)Er* are also

[1] Johanson (2003, 2000, 1998, among others) refers to the temporal contributions of the Turkic TAE suffixes using the labels **intraterminal**, **postterminal**, and **terminal**. Roughly speaking, intraterminal morphemes view an event from within its runtime (e.g. Tatar *-A*). Postterminal morphemes envisage the event "after its relevant limit, i.e., typically after it has been carried out" (Johanson 1998, 44) (e.g. Tatar *-GAn*). Terminal morphemes "view the event as a whole" and take the form of *-DI*. While I do not adopt Johanson's terminology in this dissertation, the conceptual distinctions that he makes are similar to certain aspects of the analysis that I propose in Chapter 5.

[2] As I noted previously, the Turkic languages can be divided into those that have *-GAn* as a past-oriented verbal suffix, and those that have *-mIş*.

documented in a number of other Turkic languages. However, Johanson (1998) notes that "genuine future" items are less common than past items in Turkic. Nonetheless, *-(y)AçAK* and *-(y)Er* cognates contributing roughly future or aorist meanings are documented in e.g. Azerbaijani (Öztopçu 2000), Bashkir (Berta 1998a; Poppe 1964), and Tofa (Rassadin 1978), among other languages.[3]

TAE morpheme	**Temporal contribution**
-DI	past tense
-GAn	resultative aspect
-A	present tense
-(y)AçAK	prospective aspect
-(y)Er	future tense

Table 3.1: Temporal contributions of the Tatar TAE suffixes.

For the purpose of this chapter, I do not yet assume a fully fleshed out semantic theory of tenses and aspects. As in Chapter 2, I largely refrain from using theoretical (semantic) terminology, since this chapter is intended to be a descriptive resource on the language. However, I assume the descriptive definitions of tense and aspect given by Comrie (1985, 1976). I also use some technical terminology to give a more precise description of the temporal contribution of these morphemes, and how they interact.

Generally speaking, **tenses** relate the time of the event described by the utterance to some other time, typically the utterance time.[4] We can further distinguish between **absolute** and **relative** tenses. Absolute tenses relate the time of the event to the utterance time, whereas relative tenses relate the time of the event to some other time. I show an example of English past tense being interpreted as absolute versus relative in (67). In (67a), the past tense is interpreted as placing the time of Vince meeting Howard prior to

[3]The term "aorist" is used frequently in the Turkic literature; it describes something like habitual or continuous aspect.

[4]Later theories of tense (e.g. Klein 1994) propose that the utterance time is related to a "topic time," or a time about which the speaker wants to say something. This topic time is then temporally related to the time of the event. This can result in the same temporal interpretations as in theories in which tenses directly relate the event time and utterance time, but with an additional "mediated" step involving the topic time.

the utterance time; i.e., it is an absolute tense. In (67b), the past tense in the second clause places the time of Vince meeting Howard prior to some other time, i.e., the time of Vince walking down the road. This is a relative tense.

(67) a. Vince met Howard.

b. When walking down the road, Vince met Howard.

I use the term "aspect" in this dissertation to refer to what other authors have termed **grammatical aspect** or **viewpoint aspect**, as opposed to lexical aspect or Aktionsart (in the sense of e.g. Vendler 1957). Comrie (1976) defines aspect as a grammatical strategy encoding how the internal constituency of a described event is "viewed"; that is, aspects relate the runtime of the event to a reference time from which it is viewed. This reference time can, but need not be, the utterance time. I give an example of the English progressive aspect in (68). In this sentence, the singing event in the second clause is viewed from the reference time established in the first clause (the time at which Kirk arrived at the recital). The use of the progressive aspect indicates that this reference time (i.e., the time of Kirk's arrival) is contained within the runtime of the event of Susan singing.

(68) Kirk arrived at the recital while Susan was singing.

With these descriptive labels in mind, I will turn to the task of describing the temporal contributions of the Tatar TAE suffixes. In §3.2, I show that in matrix clauses, the Tatar TAE suffixes seem to simply mark past, present, or future meaning: that is, they seem to simply indicate the relation between the utterance time and the time of the event described by the scope proposition (henceforth, the "described event"). In §3.3, however, I show that the interpretation of some of the morphemes differs in some semantically embedded environments, particularly in embedded verbal nominalizations. The temporal meanings that I ultimately propose for the TAE suffixes are summarized in Table 3.1. I conclude the chapter by comparing my descriptive labels of the Tatar TAE suffixes to the descriptions previously proposed by Greed (2014, 2009), Tatevosov (2007), and Poppe (1961).

3.2 Temporal interpretation of TAE suffixes in matrix clauses

I begin by giving a descriptive overview of the TAE morphemes in matrix clauses. I use temporal adverbials to demonstrate the relation that they mark between the utterance time and the time of the described event. I show that in matrix clauses, the morphemes appear to simply mark past, present, or future meaning. These readings arise despite the fact that (as I argue) two of the morphemes denote aspects rather than tenses.

3.2.1 Past-oriented TAE suffixes

In matrix clauses, the Tatar TAE suffixes *-DI* and *-GAn* both place the time of the described event prior to the utterance time. Descriptively speaking, they appear to have identical temporal contributions in matrix clauses; the two suffixes differ only in their evidential interpretation, as described in Chapter 2.

3.2.1.1 Past tense: *-DI* 'PST'

The TAE suffix *-DI* marks past tense: it locates the time of the described event prior to the utterance time.[5] *-DI* is grammatical in combination with temporal adverbs that pick out times prior to the utterance time (henceforth, **past adverbials**), like *kiçä* 'yesterday' in (70). However, these adverbials are not required for a past-oriented interpretation of *-DI*; it is interpreted as placing the event time prior to the utterance time even in the absence

[5] *-DI* is diachronically related to the free past tense morpheme *ide*. Historically, *ide* could have consisted of a free copula **i* (from Old Turkic **(y)Vr*) in combination with the past tense suffix *-DI* (Erdal 2004). In Tatar as it is currently spoken, *ide* directly follows the predicate and hosts the past tense subject agreement paradigm. It is now generally restricted to combining with stative predicates, i.e., nominal or adjectival predicates, although in §3.3.2 I discuss its interpretation in combination with finite verbs.

In copular expressions like (69), *ide* communicates a cessation implicature akin to what Altshuler and Schwarzschild (2013) describe for English stative past tense expressions. That is, when a Tatar speaker says an expression with *ide*, they imply that the proposition described by the expression is no longer true at the utterance time; i.e., by saying the first clause in (69), the speaker implies that Läylä is no longer a student. However, this cessation implicature can be cancelled in Tatar, as in English.

(69) Läylä student **ide**-∅ (xäzer dä student).
Läylä student PST-3SG now ADD student
'Läylä was a student (and in fact, she still is a student).'

of a past adverbial.

(70) ***-DI* and past adverbial:**

min (**kiçä**) yeger-**de**-m.
1SG.NOM yesterday run-PST-1SG
'[I have direct evidence that] I ran (yesterday).'

As expected for a past tense morpheme, *-DI* is robustly ungrammatical with temporal adverbs picking out times following the utterance time (henceforth, **future adverbials**). I show an ungrammatical example of *-DI* in combination with *irtägä* 'tomorrow' in (71).

(71) ***-DI* and future adverbial:**

* min **irtägä** yeger-**de**-m.
1SG.NOM tomorrow run-PST-1SG

-DI is also largely ungrammatical with temporal adverbs picking out times overlapping with the utterance time (henceforth, **present adverbials**), like *xäzer* 'now.' Again, this is as expected given its past tense meaning. I show an ungrammatical example of *-DI* combining with *xäzer* in (72).

(72) ***-DI* and present adverbial:**

* Aygöl **xäzer** yeger-**de**-∅.
1SG.NOM now run-PST-3SG

Combining *DI* and *xäzer* appears to be marginally acceptable with some punctual verbs (e.g. *yegerep kitergä* 'to run away'), as in (74).[6] However, even in these marginally grammatical expressions, *-DI* still contributes past tense semantics; the gloss of (74) is such that the time of the event described by the verb precedes the utterance time.

[6] Altshuler (2016, 2010) notes that the English temporal adverb *now* is also compatible with a range of tenses; for instance, it can be used in past tense expressions like (73).

(73) Pilate raised his martyred eyes to the prisoner and saw how high the sun **now stood** above the hippodrome. (*The Master and Margarita*, M. Bulgakov)

Given this English data, the marginal availability of Tatar expressions like (74) is unsurprising to me; further investigation of expressions containing *xäzer* 'now' could reveal data similar to what Altshuler describes for English *now*.

(74) ? Aygöl **käzer** yeger-ep kit-**te**-∅.
Aygöl now run-IP leave-PST-3SG
Marginally possible reading: '[I have direct evidence that] Aygöl just ran away.'

As noted previously, Johanson (2000) argues that all Turkic languages have a past tense marker that is cognate with *-DI*. Other Turkic languages with cognate past-oriented TAE morphemes include Bashkir (*-DĚ*; Berta 1998a), Kyrgyz (*-de*; Abduldaev and Zakharova 1987), Turkish (*-DI*; Kornfilt 1997), and Tuvan (*-di*; Harrison 2000), among many others. These cognate morphemes are all generally described as contributing past tense semantics. Tatar *-DI* is also described as a past tense morpheme by Greed (2009) and Poppe (1961).

3.2.1.2 Resultative aspect: *-GAn* 'RESULT'

I propose that the Tatar TAE suffix *-GAn* contributes **resultative** aspect. Nedjalkov and Jaxontov (1988, 6) conceptually define resultative expressions as conveying "a state implying a previous event"; that is, they express "both a state and the preceding action [the state] has resulted from." (I do not otherwise follow Nedjalkov and Jaxontov 1988's diagnostics for what counts as a resultative construction, although I find their basic definition to be useful.) For now, I will define resultative aspect as a variety of viewpoint aspect in which the described event is viewed from a reference time—in matrix clauses, the utterance time—in the event's "post-state." Conceptually speaking, event post-states can be thought of as periods of time in which the event has concluded and its result(s) are still perceptible.[7]

In matrix clauses, *-GAn* is interpreted identically to *-DI* with respect to its temporal contribution; that is, it expresses that the time of the described event precedes the utterance time. This temporal meaning is more restricted than we might expect for an aspectual morpheme, since aspects are typically compatible with events occurring in the past, present, or future of the utterance time.[8] This restricted temporal meaning of *-GAn* can be accounted

[7] I discuss the properties of event post-states at length, and provide a formal definition, in Chapter 5.

[8] For example, the English progressive aspect views events from times within their runtime; this aspect is compatible with events being located in the past, present, or future of the utterance time:

for if, in matrix clauses, the utterance time (i.e., the reference time from which the event is viewed) is necessarily located in the post-state of the event. (I expand on this at length in my analysis in Chapter 5.)

Like *-DI*, *-GAn* is grammatical in matrix clauses with past adverbials, as in (76). However, past adverbials are not required in order to locate the event time prior to the utterance time, as shown in (77).

(76) **-*GAn* and past adverbial:**

alar **kiçä** kür-**gän**-när.
3PL.NOM yesterday swim-RESULT-3PL
'[I have indirect evidence that] They swam yesterday.'

(77) Läylä yal-ğa kit-**kän**-∅.
Läylä rest-DAT leave-RESULT-3SG
'[I have indirect evidence that] Läylä left on vacation.'

-GAn is ungrammatical with future adverbials, as in (78).

(78) **-*GAn* and future adverbial:**

* Aygöl **irtägä** yeger-**gän**-∅.
Aygöl tomorrow run-RESULT-3SG

-GAn is also ungrammatical with present adverbials in matrix clauses, as in (79). When asked to compare the acceptability of *-GAn* and *-DI* expressions with present adverbials (e.g., (72) and (79)), my consultants report that *-GAn* expressions are uniformly worse.

(79) **-*GAn* and present adverbial:**

* Aygöl **xäzer** yeger-**gän**-∅.
Aygöl now run-RESULT-3SG

---

(75) a. Past progressive: I **was swimming**.
b. Present progressive: I **am swimming**.
c. Future progressive: I **will be swimming**.

Tatar *-GAn* (and cognate morphemes across Turkic) have a long history of being described as (present) perfects. For instance, Tatevosov (2007, 408) (writing on Mishar Tatar) describes *-GAn* as a perfect. Greed (2009) describes Tatar *-GAn* as a "resultative past tense"; however, she glosses all examples with *-GAn* using the English present perfect, and later describes it as a resultative/perfect in Greed (2014).[9] Cognate morphemes in other Turkic languages have been described as perfects or "perfect-like" by Johanson (2000). Johanson (2000, 70) cites specific examples of cognate perfects from Chaghatay, Crimean Tatar, Karachay, Kumyk, and Tuvan, and glosses them using the English present perfect.

Despite this long history of treating Tatar *-GAn* (and cognate morphemes) as (present) perfects, I do not find that analyzing *-GAn* as an English present perfect is well-motivated. I will briefly review why I do not treat *-GAn* as a present perfect. I focus primarily on comparison of *-GAn* expressions with English present perfect expressions, and set aside the issues that are raised by evaluating *-GAn* against theories of the English present perfect.

Semantic theories of the present perfect are typically concerned with accounting for a number of well-described properties of English present perfect expressions (as described by Klein 1992, McCawley 1981, McCoard 1978, and Comrie 1976, among others). Two central properties of the English present perfect are (i) the general inability of the English present perfect to co-occur with temporal adjuncts (Klein 1992's "present perfect puzzle"), as in (80); and (ii) the requirement that the result state of a present perfect expression be true at the utterance time, as in (81).[10] ("Result state" is used here to refer strictly to results that are linked to the lexical semantics of the verb; for example, the result state of an event of someone losing their keys is that their keys are lost; the result state of an event of someone breaking a plate is that the plate is broken, and so on.)

[9] In English, present perfect expressions are formed using a present tense auxiliary *have* in combination with a past participial form of the main verb, e.g. *Susan **has eaten** the satsuma*.

[10] English present perfects are compatible with a small set of temporal adverbials like *just* and *recently*, e.g. *Leroy has **just** left for Albuquerque* (Klein 1992). Klein (1992) proposes that this can be explained by positing a restriction on the ability of the present perfect to combine with temporal adverbials that pick out "definite" times (e.g. *last year, yesterday*).

(80) **Ungrammaticality with (most) temporal adjuncts:**

Leroy has left for Albuquerque (*yesterday).

(81) **Result state must be true at utterance time:**

Sofia has lost her keys (#but now she found them).

Neither of these properties apply to Tatar expressions with *-GAn*. As I showed in (76), *-GAn* expressions are grammatical in combination with temporal adjuncts like *kiçä* 'yesterday,' unlike English present perfects. Furthermore, *-GAn* expressions are not subject to the requirement that the result state of the described event be true at the utterance time. I show this in (82). In this example, the speaker first uses *-GAn* to assert that (they have indirect evidence that) Ali has lost his keys. They can then felicitously follow this assertion by stating that Ali later found his keys, i.e., the keys are no longer lost and the result state of the first expression no longer holds. *-GAn* is licensed in both clauses, since the speaker has indirect evidence for both propositions. The felicity of this continuation in Tatar constrasts with the English present perfect example in (81).

(82) **Perfect of result context:** Ali calls you and tells you that he can't find his keys. Later, he calls you back and says he found his keys. Your friend asks what happened. You say:

Ali açqıç-ın yuğalt-**qan**-∅ (läkin beraz-dan soñ tap-**qan**-∅).
Ali key-ACC lose-RESULT-3SG but short.time-ABL after find-RESULT-3SG
'[I have indirect evidence that] Ali lost his keys (but after a little while he found them).'

Tatar *-GAn* expressions therefore pattern differently from English present perfect expressions with respect to two major properties.[11] Since our semantic theories of the present

[11]There are a number of additional ways in which Tatar *-GAn* expressions pattern differently from English present perfects, which I discuss in Bowler and Özkan (2017). These include the fact that the described event need not be salient at the utterance time; a lack of lifetime effects (i.e., the requirement that individuals in present perfect expressions be alive at the utterance time; Chomsky 1970), and a lack of the repeatability requirement (i.e., the requirement that events described by present perfect expressions can be repeated; Katz 2003).

perfect are primarily based on English data, these contrasts between Tatar and English suggest to me that we should not analyze Tatar *-GAn* as a present perfect.[12]

In sum, my decision to label *-GAn* a resultative aspect is driven by two main considerations. First, I find that the term "resultative" impressionistically captures the fact that the reference time picked out by *-GAn* follows the inception of the event.[13] Second, there are no major formal semantic analyses currently associated with resultative aspect, unlike the (English) present perfect. Since there are no theoretical assumptions attached to this term, I need not commit to additional (potentially unwanted) theoretical assumptions when using it.

3.2.2 Present- and future-oriented TAE suffixes

Tatar has three TAE suffixes that, in matrix clauses, can place the time of the described event after the utterance time. Two of these suffixes, *-(y)Er* and *-(y)AçAK*, strictly place the time of the described event after the utterance time. I propose that *-(y)Er* marks simple future tense, whereas *-(y)AçAK* marks prospective aspect. Descriptively speaking, these suffixes appear to have identical temporal contributions in matrix clauses; they differ only in their evidential interpretation.

The third TAE suffix, *-A*, can place the event at a time that overlaps with the utterance time. I propose that this morpheme is underlyingly a present tense. However, it also permits a "futurate" reading in which it places the time of the event after the utterance time. (As I noted in Chapter 2, I do not address *-A* in the theoretical component of this dissertation, since it is not associated with any evidential meaning.)

---

[12] Tatar **does** have an expression that appears to pattern more like the English present perfect; this construction includes a verbal nominalization with genitive case on the subject, and an existential particle *bar*:

(83) minem monda berniçä märtäbä yöz-gän-em bar.
1SG.GEN here few time swim-RESULT-1SG.POSS EXIST
Speaker's translation: 'I have swum here a few times.'

[13] As I will show in §3.3.1.2, this can sometimes include the reference time being located within the runtime of the event itself.

3.2.2.1 Future tense: *-(y)Er* 'FUT'

The TAE suffix *-(y)Er* marks future tense: it places the time of the described event after the utterance time. This suffix is grammatical in combination with future adverbials, as in (84). However, future oriented adverbials are not required for it to convey future tense meaning, as shown in (85).

(84) xat **irtägä** kil-ep cit-**er**-∅.
letter tomorrow come-IP arrive-FUT-3SG
'[I have non-specific evidence that] The letter will arrive tomorrow.'

(85) sin söyläş-**er**-sen.
2SG.NOM talk-FUT-2SG
'[I have non-specific evidence that] You (sg.) will talk.'

-(y)Er is ungrammatical in combination with past adverbials, as in (86).

(86) * xat **kiçä** kil-ep cit-**er**-∅.
letter yesterday come-IP arrive-FUT-3SG

-(y)Er is grammatical in combination with present adverbials like *xäzer* 'now,' as in (87). However, it still contributes a future tense reading in such expressions. The interpretation of (87) is such that the letter-arriving event described by the expression will occur immediately after the utterance time.[14]

(87) **Online tracking context:** You are expecting a letter. You are tracking this letter online; the tracking service says that it will be delivered by 3PM today. You look at the clock and see that it is almost 3PM. You say:

xat **xäzer** kil-ep cit-**er**-∅.
letter now come-IP arrive-FUT-3SG
'[I have non-specific evidence that] The letter will arrive right now.'

A number of Turkic languages have verbal suffixes that are cognate with Tatar *-(y)Er*. However, cognates of *-(y)Er* are typically described as aorists (marking something like

[14]This patterns similarly to the English future tense, which is also felicitous in combination with *now* and gives rise to a similar reading (e.g. *I **will now** introduce you to the gentleman who is in charge of your transportation*) (Altshuler 2016).

habitual or continuous aspect) or present tenses, rather than future tenses. Languages with present-oriented *-(y)Er* cognates include Turkish (*-(A/I)r*; Göksel and Kerslake 2005), Turkmen (*-yAr*; Schönig 1998), and Yakut (*-Ar/Ir*; Stachowski and Menz 1998), among others.[15] Nonetheless, as I showed in this section, Tatar *-(y)Er* has a synchronically future-oriented meaning. A future-oriented *-(y)Er* cognate is also described in Noghay (*-(A)r*; Karakoç 2017).

3.2.2.2 Prospective aspect: *-(y)AçAK* 'PROSP'

The TAE suffix *-(y)AçAK* contributes **prospective** aspect. Conceptually speaking, I assume the definition of prospective aspect given by Comrie (1976, 64-65). Comrie defines prospective aspect as relating a (present) state to some subsequent, future situation; he describes prospective aspectual expressions as alluding to "the already present seeds of some future situation." Comrie (1976) cites expressions containing *going to* as examples of prospective meaning in English, e.g. *Bill is going to throw himself off the cliff*.

I propose that prospective aspect can be thought of as the "mirror image" of resultative aspect, described informally in §3.2.1.2. That is, it is a variety of viewpoint aspect in which the described event is viewed from a reference time—in matrix clauses, the utterance time—in the event's "pre-state." Conceptually speaking, event pre-states can be thought of as periods of time preceding the beginning of an event in which plans or causes of the event are perceptible.[16]

In matrix clauses, *-(y)AçAK* appears to simply contribute future meaning. Its temporal interpretation in matrix clauses is identical to *-(y)Er*; that is, it locates the time of the event described by the verb after the utterance time. Like *-GAn*, this temporal meaning is again more restricted than we might expect for an aspectual morpheme. However, this restricted meaning can be accounted for if, in matrix clauses, the utterance time (i.e., the reference

[15]These roughly present-oriented meanings could historically stem from the Old Turkic aorist/copula **(y)Vr* (Erdal 2004, 240-242). This suggests that Tatar *-(y)Er* has undergone semantic change to become future-oriented.

[16]I provide a formal definition of event pre-states in Chapter 5.

time from which the event is viewed) is necessarily located in the pre-state of the event. If the utterance time is located in the pre-state of the event, it must precede the event. (I expand on this at length in my analysis in Chapter 5.)

I illustrate the basic temporal contribution of *-(y)AçAK* in (88)-(89). This suffix is grammatical in combination with future adverbials, as in (88). However, future adverbials are not required for *-(y)AçAK* to contribute future meaning, as in (89).

(88) **un kön-nän** min Mäskäü-gä bar-**açaq**-mın.
ten day-ABL 1SG.NOM Moscow-DAT go-PROSP-1SG
'[I have specific evidence that] I will go to Moscow in ten days.'

(89) sin söyläş-**äçäk**-sen.
2SG.NOM talk-PROSP-2SG
'[I have specific evidence that] You (sg.) will talk.'

-(y)AçAK is ungrammatical in combination with past adverbials, as in (90).

(90) * xat **kiçä** kil-ep cit-**äçäk**-∅.
letter yesterday come-IP arrive-PROSP-3SG

In combination with present adverbials, *-(y)AçAK* patterns identically to *-(y)Er* in that it still maintains its future-oriented contribution; that is, it still locates the event described by the expression after the utterance time.[17] Like (87), the interpretation of (92) is such that the letter-arriving event described by the expression will occur immediately after the utterance time.

[17]Tatar *-(y)AçAK* (and *-(y)Er*) are both unable to express present inference. That is, these suffixes cannot be used in expressions like the English example in (91a), as shown in (91b). In (91a), the future tense marker *will* does not mark future tense, but rather indicates that the speaker is making an inference about an ongoing event. (See Winans 2016 for more information on the use of future tenses to express present inference.)

(91) Context: Your neighbors barbecue every Friday night. You are out of town, and you realize that it's Friday night. You say:
a. The neighbors **will** be barbecuing (right now).
b. # kürşe-lär şaşlıq {yas-**ar**-lar / yas-**ıyaçaq**-∅}.
neighbor-PL barbecue make-FUT-3PL / make-PROSP-3PL
'[I have non-specific/specific evidence that] the neighbors will barbecue.'
(Speaker's comment: "[These sentences are] something that would be said earlier in the day"; i.e., they must make a claim about a future event)

Tatar *-(y)AçAK* differs from Turkish *-(y)AcAK* in this sense; Turkish *-(y)AcAK* is grammatical in expressions of present inference (Winans 2016, 60).

(92) **Mailman context:** You are expecting a letter from a friend. You look out the window and see the mailman approaching with a letter. You say:

xat **xäzer** kil-ep cit-**äçäk**-∅.
letter now come-IP arrive-PROSP-3SG
'[I have specific evidence that] The letter will arrive right now.'

I note that overall, there is a slight preference for *-(y)AçAK* to co-occur with temporal adverbials. Anecdotally, my consultants report that this preference is linked to the reading of increased speaker certainty regarding the scope proposition that is conveyed by *-(y)AçAK* and which I described previously in Chapter 2. My consultants report that if a speaker has specific evidence that (i.e., is relatively certain that) an event will occur, then they should also know when it will happen. However, this is not a strict preference; *-(y)AçAK* expressions are felicitous without temporal adverbials, given the correct context. I show such a context in (93). In this context, the speaker uses *-(y)AçAK* to convey that they have specific evidence that the described event will occur, but explicitly states that they do not know when it will happen.

(93) **Moscow trip context**: You know Mansur has bought tickets to visit Moscow. He loves Moscow and you are certain that he will go on the trip. However, you don't know what dates the tickets are for. Aygöl asks you whether Mansur will visit Moscow. You say:

Mansur Mäskäü-gä bar-**açaq**-∅ (läkin qayçan bel-m-im).
Mansur Moscow-DAT go-PROSP-3SG but when know-NEG-1SG
'[I have specific evidence that] Mansur will go to Moscow (but I don't know when).'

Some typologists have termed cognate morphemes in Turkic "proximatives" or "immediate/imminent futures" (following Heine 1994; see Nevskaya 2018, 2005 for discussion on terminology). These terms suggest that the described event will occur in the near future of the utterance time; i.e., they imply that the utterance time and event time are temporally close. This is not the case with Tatar *-(y)AçAK*. *-(y)AçAK* can be used in expressions describing events that are in the near future of the utterance time, as in (94); the remote

future, as in (95); and even when describing events that will occur at some indefinite future time, as in (96).

(94) Mansur Mäskäü-gä **säğat unike-dä** bar-**açaq**-∅.
Mansur Moscow-DAT hour twelve-LOC go-PROSP-3SG
'[I have specific evidence that] Mansur will go to Moscow at 12 o'clock.'

(95) Mansur Mäskäü-gä **biş yel-dan** bar-**açaq**-∅.
Mansur Moscow-DAT five year-ABL go-PROSP-3SG
'[I have specific evidence that] Mansur will go to Moscow in five years.'

(96) **ber kön** Mansur Mäskäü-gä bar-**açaq**-∅.
one day Mansur Moscow-DAT go-PROSP-3SG
'[I have specific evidence that] Mansur will go to Moscow one day.'

I note that, overall, if an expression contains a near future adverbial (e.g. *irtägä* 'tomorrow,' *säğat unikedä* 'at twelve o'clock'), speakers prefer to use *-(y)AçAK* rather than *-(y)Er*. Expressions in which *-(y)Er* co-occurs with a near future adverbial are typically judged as marked in some way if they are presented without an accompanying context. My consultants report that this is due to the intuition that if a speaker knows "enough" about the state of the world to be able to assert that the described event will happen in the near future of the utterance time, then they should also be certain that (i.e., have specific evidence that) the event will occur. The speaker should therefore report the future event using *-(y)AçAK*.

I demonstrate this preference for *-(y)AçAK* when describing events close to the utterance time in (97). This preference for *-(y)AçAK* decreases when referring to more distant times, as in (98). When the temporal adjunct picks out a distant or indefinite future time, both *-(y)AçAK* and *-(y)Er* are again judged as fully felicitous, as in (99).

(97) Mansur Mäskäü-gä **säğat unike-dä** {bar-**açaq**-∅ / ??bar-**ır**-∅}.
Mansur Moscow-DAT hour twelve-LOC go-PROSP-3SG / go-FUT-3SG
'[I have specific/??non-specific evidence that] Mansur will go to Moscow at 12 o'clock.'

(98) **kiläse atna** Mansur Mäskäü-gä {bar-**açaq**-∅ / ?bar-**ır**-∅}.
next week Mansur Moscow-DAT go-PROSP-3SG / go-FUT-3SG
'[I have specific/?non-specific evidence that] Mansur will go to Moscow next week.'

(99) **ber kön** Mansur Mäskäü-gä {bar-**açaq**-∅ / bar-**ır**-∅}.
one day Mansur Moscow-DAT go-PROSP-3SG / go-FUT-3SG
'[I have specific/non-specific evidence that] Mansur will go to Moscow one day.'

I propose that these distributional facts about *-(y)AçAK* and *-(y)Er* do not need to be hardwired into the grammar. As I showed in (93), *-(y)AçAK* is felicitous in contexts without near future adverbials as long as its evidential requirement is satisfied; i.e., as long as the speaker has specific evidence that the event will occur. Furthermore, it is natural to use *-(y)Er* in combination with near future adverbials, given the correct evidential context. I show such a context in (100), in which the speaker lacks specific evidence that the described event will occur in the near future of the utterance time and therefore is able to assert the proposition using *-(y)Er*.

(100) **Karaoke context**: You are at karaoke with your friends. You know that Läylä is on the list to sing, and you think that her turn will come up in a few minutes. You say:

Läylä **berniçä minut-tan** cırl-**ar**-∅.
Läylä few minute-ABL sing-FUT-3SG
'[I have non-specific evidence that] Läylä will sing in a few minutes.'

Several other Turkic languages have morphemes that are cognate with *-(y)AçAK* and that have also been described as marking prospective aspect. These languages include Bashkir (*-(y)AsAK*; Berta 1998a), Gagauz (*-(y)eǰek*; Menz 2000), Turkish (*-(y)AcAK*; Key and Schreiner 2014), and Turkmen (*-ǰAK*; Schönig 1998), among others.

3.2.2.3 Present: *-A* 'PRES'

The TAE suffix *-A* contributes present tense. In matrix clauses, *-A* can locate the event described by the verb at a time that overlaps with the utterance time. This leads to expressions that are glossed with the English present tense or progressive aspect, as in (101)-

(102). Tatar *-A* is also described as a present tense by Tatevosov (2007, 408) (writing on Mishar Tatar) and Poppe (1961, 70).

(101) min yeger-**ä**-m.
1SG.NOM run-PRES-1SG
'I run/am running.'

(102) Mansur kit-**ä**-∅.
Mansur leave-PRES-3SG
'Mansur is leaving.'

-A is also compatible with habitual and generic readings. In (103), the speaker uses *-A* to express that Alsu habitually bites.

(103) **Context:** You are introducing someone to your young daughter, Alsu. You warn them that Alsu likes to bite people. You say:

Alsu teşl-**i**-∅.
Alsu bite-PRES-3SG
'Alsu bites.'

This suffix also has a highly productive **futurate** use. Copley (2009, 2008) defines futurates as expressions with no obvious means of future time reference that nonetheless locate the time of the described event in the future of the utterance time. Futurates additionally convey that the described event is "planned, scheduled, or otherwise determined," giving the expressions a somewhat modal or evidential flavor (Copley 2008, 261).[18] Tatar *-A* is compatible with futurate readings even in the absence of future adverbials. However, future adverbials can be used to force a futurate reading of *-A*, as in (104)-(105).

(104) Mansur **irtägä** kit-**ä**-∅.
Mansur tomorrow leave-PRES-3SG
'Mansur leaves/is leaving tomorrow.'

(105) **un kön-nän** min Mäskäü-gä bar-**a**-m.
ten day-ABL 1SG.NOM Moscow-DAT go-PRES-1SG
'I go/am going to Moscow in ten days.'

[18]Roughly speaking, Copley uses the presence of this additional "planned"/"scheduled" meaning to differentiate between futurates and nonpast tenses; nonpast tenses (as in e.g. German) lack this "planned" reading.

Other Turkic languages with cognate present tense morphemes include Bashkir (*-A*; Berta 1998a); Crimean Tatar, Karachay-Balkar, Karaim, and Kumyk (*-A*; Berta 1998b), Kyrgyz (*-A*; Kirchner 1998), and Noghay (*-A*; Csató and Karakoç 1998), among others.

3.2.3 Recap: Temporal interpretation of TAE suffixes in matrix clauses

In matrix clauses, the Tatar TAE suffixes are all relatively restricted with respect to where they can locate the event time relative to the utterance time. I demonstrated this by testing the ability of the TAE suffixes to co-occur with past, present, and future-oriented temporal adverbials. The distribution of the TAE suffixes *-(y)AçAK* and *-GAn* are somewhat more restricted than what we might expect for aspectual morphemes. However, I propose that their restricted temporal meanings can be understood if, in matrix clauses, we treat these morphemes as locating the utterance time in a pre- or post-state of the described event. I provide a formal account of this proposal in Chapter 5.

The temporal interpretations of the Tatar TAE suffixes in matrix clauses are generally similar to cognate verbal suffixes in other Turkic languages. One exception to this is the future-oriented interpretation of Tatar *-(y)Er*, which is present-oriented in a number of other Turkic languages. I also argued that the Tatar TAE suffix *-GAn* should not be treated as a (present) perfect. This differs from descriptions of cognate morphemes in other Turkic languages as well as prior descriptions of Tatar (Greed 2014, Tatevosov 2007).

3.3 Temporal interpretation of semantically embedded TAE suffixes

Our picture of the Tatar TAE system becomes more complex when considering expressions in which the morphemes are semantically embedded. This section sets my descriptive work apart from Greed (2014, 2009), Tatevosov (2007), and Poppe (1961); these sources do not describe the TAE suffixes in semantically embedded environments. In this section, I show data on TAE suffixes embedded under attitude predicates, in both full CPs and verbal nominalizations (§3.3.1). I then discuss how the TAE suffixes are interpreted when

embedded under the free past tense morpheme *ide* (§3.3.2). In my theoretical analysis of this data in Chapter 5, I ultimately only give a compositional account of embedding under attitude predicates; I do not give a full compositional analysis for the embedding data under *ide*. Furthermore, my analysis of the embedding data in Chapter 6 primarily concentrates on accounting for the **evidential** readings of the embedded TAE suffixes, rather than all of their possible (and impossible) temporal interpretations.

I noted in Chapter 2 that the evidential interpretation of the Tatar TAE suffixes in embedded environments is significantly more complex than in matrix clauses. For instance, the evidential contribution of the TAE suffixes seems to always be interpreted relative to the speaker's evidence in matrix clauses; however, the evidential contribution of the past-oriented TAE suffixes appears to be able to be interpreted relative to either the speaker or the matrix subject when it is embedded. For simplicity, I therefore do not gloss the evidential readings of the Tatar TAE suffixes in this section, since there may be more than one evidential reading available. Furthermore, since this dissertation is primarily concerned with the semantics and pragmatics of the verbal TAE suffixes, I do not discuss the distribution and use of the free past tense morpheme *ide* in embedded environments.

3.3.1 Temporal interpretation of TAE suffixes in embedded clauses

As I described at length in Chapter 2, Tatar has two strategies for forming embedded clauses. Embedded clauses can take the form of full CPs with the complementizer *dip*, as in (106a), or verbal nominalizations with accusative case marking, as in (106b). (I refer the reader to Chapter 2 for a full discussion of the semantic differences between the embedded CPs and embedded verbal nominalizations.)

(106) a. **Embedded CP**

Läylä [Michael Phelps ciñ-gän-∅ dip] bel-ä-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] BEL-PRES-3SG
'Läylä thinks that Michael Phelps won.'

b. **Embedded nominalization**

Läylä [Michael Phelps ciñ-gän-e-n] bel-ä-∅.
Läylä [Michael Phelps win-RESULT-3SG.POSS-ACC] BEL-PRES-3SG
'Läylä knows that Michael Phelps won.'

The TAE suffixes differ with respect to their grammaticality in embedded environments. In Table 3.2, I summarize the availability and interpretation of the TAE suffixes in embedded CPs versus embedded verbal nominalizations. I note that under an analysis of the TAE suffixes as tenses versus aspects, their availability in embedded verbal nominalizations corresponds to whether they are tenses or aspects: the Tatar aspectual suffixes are permitted to occur in embedded verbal nominalizations, whereas the tense suffixes are not. I take this as morphosyntactic evidence in support of my analysis as the suffixes as tenses versus aspects. I assume that the Tatar tense suffixes are simply morphologically blocked from occurring in verbal nominalizations (I address why this is in Chapter 6).

TAE suffix	**Meaning**	**Grammaticality in embedded CPs**	**Grammaticality in embedded verbal nominalizations**
-DI	past tense	✓	*
-GAn	resultative aspect	✓	✓ (no evid. reading)
-A	present tense	✓	*
-(y)AçAK	prospective aspect	✓	✓ (no evid. reading)
-(y)Er	future tense	✓	*

Table 3.2: Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings.

Like in the preceding chapter, in the interest of simplicity, I primarily provide examples of CPs and verbal nominalizations embedded under past tense tokens of the matrix verb *äytergä* 'to say.'

3.3.1.1 Temporal interpretation of the TAE suffixes in embedded CPs

The Tatar TAE suffixes are relatively restricted in the temporal readings that they permit in embedded CPs. In this sense, they pattern similarly to their use in matrix clauses.

Temporal interpretation of past-oriented TAE suffixes in embedded CPs

The past-oriented TAE suffixes *-GAn* and *-DI* pattern identically in embedded CPs. When these suffixes occur in CPs embedded under past tense matrix verbs, they are strictly interpreted as placing the time of the embedded event in the past of the time of the matrix event, which is itself in the past of the utterance time (i.e., a "backshifted" reading). (I will refer to these as "past-under-past" expressions, even though *-GAn* is not, strictly speaking, a past tense.) I show this backshifted reading in (107) through the use of optional temporal adverbs.

(107) (**kiçä**) Timur [Aygöl (**ütkän atna**) {cırla-**de**-∅ / cırla-**ğan**-∅} dip]
yesterday Timur [Aygöl past week sing-PST-3SG / sing-RESULT-3SG COMP]
äyt-te-∅.
say-PST-3SG
'Timur said (yesterday) that Aygöl sang (last week).'

(108) shows that these past-under-past CP embeddings require that the time of the embedded event be in the past of the event described by the matrix clause, rather than simply in the past of the utterance time. The intended reading of the highly marginal sentence in (108) is one in which the time of the embedded event is in the past of the utterance time, but in the future of the time of the matrix event (i.e., a "future-in-the-past" reading). This reading is impossible; (108) is unacceptable.

(108) ?? **uzğan atna** Timur [**kiçä** Parij-da yañgır {yaw-**dı**-∅ /
past week Timur [yesterday Paris-LOC rain fall-PST-3SG /
yaw-**ğan**-∅}] dip äyt-te-∅.
fall-RESULT-3SG] COMP say-PST-3SG
Intended meaning: 'Timur said last week that it would rain in Paris yesterday.'

Temporal interpretation of present- and future-oriented TAE suffixes in embedded CPs

The future-oriented TAE suffixes *-(y)AçAK* and *-(y)Er* also pattern identically in embedded CPs. When these suffixes are embedded under past tense matrix verbs, they place the time

of the embedded event in the future of the utterance time. The time of the embedded event is also necessarily in the future of the matrix event. I show this through the use of optional temporal adverbs in (109).

(109) (**kiçä**) Timur [Alsu (**irtägä**) {cırl-**ar**-∅ / cırlı-**yaçaq**-∅} dip]
yesterday Timur [Alsu tomorrow sing-FUT-3SG / sing-PROSP-3SG COMP]
äyt-te-∅.
say-PST-3SG
'Timur said (yesterday) that Alsu will sing (tomorrow).'

The present tense suffix *-A* has two possible readings when it is embedded under a past tense matrix verb. The first reading is one in which *-A* is interpreted (like *-(y)AçAK* and *-(y)Er*) as placing the time of the embedded event in the future of both the utterance time and the matrix event. This is in accordance with its futurate use in matrix clauses, as described in §3.2.2.3. I show an example of this in (110).

(110) **kiçä** Mansur [Timur **irtägä** Qazan-ğa bar-**a**-∅ dip] äyt-te-∅.
yesterday Mansur [Timur tomorrow Kazan-DAT go-PRES-3SG COMP] say-PST-3SG
'Mansur said yesterday that Timur goes to Kazan tomorrow.'
(Speaker's comment: "Yesterday Mansur said that Timur would go to Kazan in two days.")

The second reading of embedded *-A* is one in which it indicates that the runtime of the embedded event includes the utterance time; this is an instance of a so-called "double access" reading. I show an example of a double access reading of *-A* in (111). This interpretation parallels the use of *-A* in matrix clauses as marking the overlap of the event time and utterance time (i.e., its use as a simple present tense).

(111) **Context:** You are at a concert, and you stepped outside for some fresh air. Your friend Timur pokes his head out the door and tells you that Aygöl, who you have been waiting to hear, is singing right now. You tell your friends:

Timur (**äle genä**) [Aygöl (**xäzer**) cırl-**i**-∅ dip] äyt-te-∅.
Timur now only [Aygöl right.now sing-PRES-3SG COMP] sing-PST-3SG
'Timur (just) said that Aygöl is singing (right now).'

Data like (112) demonstrate that when embedded *-A* is interpreted as a simple present tense (i.e., not as a futurate), the double access reading is required. In (112), the embedded predicate is *bala kötergä* 'to be pregnant' (literally, 'to expect a child'). Our knowledge of the world is such that pregnancies typically only last nine months. The expression in (112) is therefore infelicitous, since the use of *-A* in the embedded clause combined with the temporal adverbial *biş yel elek* 'five years ago' in the matrix clause suggests that Alsu was pregnant five years ago, and is also still pregnant at the utterance time.

(112) # Mansur **biş yel elek** [Alsu bala köt-**ä**-∅ dip] äyt-te-∅.
Mansur five year ago [Alsu child expect-PRES-3SG COMP] say-PST-3SG
'Five years ago, Mansur said that Alsu is pregnant.'
(Lit. 'Five years ago, Mansur said that Alsu is expecting a child.')
(Speaker's comment: "It sounds like she's still pregnant right now, and that's why it doesn't make any sense.")

Finally, in embedded CPs, these three morphemes cannot be used to place the time of the embedded event in the past of the utterance time but the future of the time of the matrix event. That is, future-in-the-past readings are unavailable for all present- and future-oriented TAE suffixes in embedded CPs. I illustrate this through the highly marginal example in (113).

(113) ?? **uzğan atna** Timur [**kiçä** Parij-da yañgır {yaw-**a**-∅ / yaw-**er**-∅
past week Timur [yesterday Paris-LOC rain fall-PRES-3SG / fall-FUT-3SG
/ yaw-**açaq**-∅}] dip äyt-te-∅.
/ fall-PROSP-3SG] COMP say-PST-3SG
Intended meaning: 'Timur said last week that it was going to rain in Paris yesterday.'

3.3.1.2 Temporal interpretation of the TAE suffixes in embedded verbal nominalizations

As I noted previously in Table 3.2, only the TAE suffixes *-GAn* and *-(y)AçAK* are grammatical in embedded verbal nominalizations. As I described in Chapter 2, when these suffixes

occur in embedded verbal nominalizations, they no longer make any evidential contribution. In verbal nominalizations, they permit a wider range of temporal interpretations than in either matrix clauses or embedded CPs.

Temporal interpretation of *-GAn* 'RESULT' in embedded verbal nominalizations

When *-GAn* occurs in an embedded verbal nominalization, it is compatible with two different temporal interpretations. The first is effectively the same temporal contribution that it makes in embedded CPs: it can express that the time of the embedded event occurred in the past of the time of the matrix event, which is itself in the past of the utterance time. In (114), temporal adverbials indicate that the time of the embedded event is in the past of the matrix event; i.e., it is a backshifted reading of the embedded clause.

(114) **bügen irtä belän** Timur [Alsu-nın **kiçä** çäkçäk-ne
today morning with Timur [Alsu-GEN yesterday çäkçäk-ACC
aşı-**ğan**-ı-n] äyt-te-∅.
eat-RESULT-3SG.POSS-ACC] say-PST-3SG
'Timur said this morning that Alsu ate the çäkçäk yesterday.'

-GAn is incompatible with future-in-the-past readings in embedded verbal nominalizations; this again parallels its behavior in embedded CPs. I demonstrate this incompatibility through the highly marginal example in (115). In this example, the intended reading is that the time of the embedded event is in the past of the utterance time, and in the future of the matrix event.

(115) ?? **uzğan atna** Timur [Parij-da **kiçä** yañgır yaw-**ğan**-e-n]
past week Timur [Paris-LOC yesterday rain fall-RESULT-3SG.POSS-ACC]
äyt-te-∅.
say-PST-3SG
Intended meaning: 'Timur said last week that it would rain in Paris yesterday.'

The second reading of *-GAn* in embedded verbal nominalizations is one in which the runtime of the embedded event overlaps with the utterance time; that is, it can express something akin to the present tense use of *-A* in embedded CPs, described in §3.3.1.1. I show an example of this in (116). In this example, the runtime of the event described by

the embedded verb contains the utterance time, as indicated by the felicity of the temporal adverbial *xäzer* 'right now' in the embedded clause. The availability of this temporal interpretation of *-GAn* is puzzling, since in matrix clauses it is solely used to refer to times in the past of the utterance time.[19]

(116) Timur [Alsu-nın (**xäzer**) çäkçäk-ne aşı-**ğan**-ı-n] äyt-te-∅.
Timur [Alsu-GEN right.now çäkçäk-ACC eat-RESULT-3SG.POSS-ACC] say-PST-3SG
'Timur said that Alsu is eating the çäkçäk (right now).'

Embedded *-GAn* contrasts with *-A* in that double access effects generally do not arise, since *-GAn* is also compatible with a backshifted reading of the embedded event. I demonstrate this in (117). In this expression, the embedded predicate is *bala kötergä* 'to expect a child.' This embedded predicate (marked with *-GAn*) is felicitous in combination with the matrix temporal adverbial *biş yel elek* 'five years ago,' since the time of Alsu's pregnancy is not necessarily interpreted as overlapping with the utterance time. This contrasts with the infelicitous example in (112), in which double access effects arise due to the use of *-A* in the embedded CP. (I return to data like (117) in §3.3.1.3.)

(117) **biş yel elek** Mansur [Alsu-nın bala köt-**kän**-e-n]
five year ago Mansur [Alsu-GEN child expect-RESULT-3SG.POSS-ACC]
söylä-**de**-∅.
talk-PST-3SG
'Five years ago, Mansur said Alsu was pregnant.'
(Lit. 'Five years ago, Mansur said Alsu was expecting a child.')
(Speaker's comment: "[Alsu]'s not pregnant anymore.")

Finally, I note that in embedded verbal nominalizations, *-GAn* is (generally) not able to place the time of the embedded event in the future of the utterance time.[20] This shows

[19]In the absence of any temporal adverbials, an expression like (116) is ambiguous between a past- and present-oriented reading of embedded *-GAn*. However, my primary Tatar consultant reports that in embedded verbal nominalizations, a past-oriented reading of embedded *-GAn* is the most natural.

[20]In an embedded verbal nominalization, it is possible to use *-GAn* to place the time of the embedded event in the future of the utterance time only with a very limited set of matrix verbs, perhaps only *kötergä* 'to expect'/'to wait for' (this observation was made previously by Şahan 2002, 204). I assume that the future-oriented reading of the embedded nominalization in this case is required by the semantics of the matrix verb

that *-GAn* is not completely "atemporal" in these verbal nominalizations. I demonstrate this in the ungrammatical example in (119). ((119) was said in 2018; i.e., 2020 is in the future of the utterance time.)

(119) * **biş yel elek**, Timur [Läylä-nen **ike men yegermençe yel-da** manda
five year ago Timur [Läylä-GEN two thousand twentieth year-LOC here
tor-ğan-ı-n] äyt-te-∅.
live-RESULT-3SG.POSS-ACC] say-PST-3SG
Intended meaning: 'Timur said 5 years ago that Läylä will live here in 2020.'

Based on the data in (114)-(119), I propose that we can descriptively think of *-GAn* in embedded verbal nominalizations as contributing "non-future" meaning.

Temporal interpretation of *-(y)AçAK* 'PROSP' in embedded verbal nominalizations

When *-(y)AçAK* occurs in an embedded verbal nominalization, it places the time of the embedded event in the future of the matrix event. There are two options for the placement of this embedded event time with respect to the utterance time; it can be (i) in the future of both the matrix event time and the utterance time; or (ii) in the future of the matrix event time and the past of the utterance time (i.e., a future-in-the-past).

I give an example of *-(y)AçAK* placing the embedded event time in the future of both the matrix event time and the utterance time in (120). This use of *-(y)AçAK* is analogous to its interpretation in embedded CPs. My primary Tatar consultant reports that, absent any temporal adverbials, this is the most natural interpretation of *-(y)AçAK* in an embedded verbal nominalization.

---

and is not being contributed by *-GAn*.

(118) Mansur [Timur-nın kil-ep cit-**kän**-e-n] köt-ä-∅.
Mansur [Timur-GEN come-IP arrive-RESULT-3SG.POSS-ACC] wait.for-PRES-3SG
'Mansur is waiting for Timur's arrival.'

(120) **uzğan atna** [Timur **irtägä** Parij-da yañgır yaw-**açağ**-ı-n]
past week [Timur tomorrow Paris-LOC rain fall-PROSP-3SG.POSS-ACC]
äyt-te-∅.
say-PST-3SG
'Timur said last week that it would rain in Paris tomorrow.'

I show an example of *-(y)AçAK* being used to express future-in-the-past in (121). In this example, the temporal adverbials force a reading in which the time of the embedded event is in the future of the matrix event but in the past of the utterance time.

(121) **uzğan atna** Timur [**kiçä** Parij-da yañgır yaw-**açağ**-ı-n]
past week Timur [yesterday Paris-LOC rain fall-PROSP-3SG.POSS-ACC]
äyt-te-∅.
say-PST-3SG
'Timur said last week that it would rain in Paris yesterday.'

Expressions like (121) are the only (non-periphrastic) strategy that Tatar speakers can use to express future-in-the-past when embedding under attitude verbs.[21] As I showed previously in this section, no other combinations of embedding strategy and TAE suffix are compatible with future-in-the-past readings of the embedded event relative to the matrix event.

3.3.1.3 Sequence of tense (SOT)

In some languages, past tenses in embedded clauses can make essentially no temporal contribution. This phenomenon, termed **sequence of tense** (SOT), can occur in expressions in which a past tense is embedded under another past tense attitude verb. In an SOT expression, the time of the event in the embedded clause is evaluated as being the same as the time of the event in the matrix clause (i.e., the embedded past tense is semantically "null").

English is a language that permits SOT. I show an example of SOT in English in (122); the SOT interpretation of the expression in (122) is given in (122a). In (122a), the time

[21] *-(y)AçAK* can also be used to express future-in-the-past when it is embedded under *ide* 'PST,' as shown in §3.3.2.4.

of Leroy loving Howard (the embedded event) is at the same time as Howard's belief (the matrix event). Another possible paraphrase for this reading is "Howard believed: 'Leroy loves me.'" The SOT reading in (122a) contrasts with the non-SOT/backshifted reading in (122b), in which the time of the embedded event precedes the time of the matrix event.

(122) Howard believed that Leroy **loved** him.

a. ✓ SOT: Howard believed that Leroy loved him at his belief time.

b. ✓ No SOT/Backshifted: Howard believed that Leroy loved him in the past, at a time prior to his belief time.

SOT is a noted point of variation across languages. SOT is available in English, but is unavailable in Japanese, Polish, and Russian, among other languages (Kubota et al. 2009, Arregui and Kusumoto 1998, among others). The English data in (122) therefore contrasts with the Polish data in (123); Polish, unlike English, lacks SOT.

When a past tense verb is embedded under a past tense attitude verb in Polish, as in (123), only a backshifted reading of the embedded event is available, as in (123a). This reading is one in which the time of Marcin being sick (the embedded event) precedes the time of Ania reporting it (the matrix event). An SOT reading of (123) is unavailable, as shown in (123b); that is, the time of the embedded event and the matrix event cannot be the same.

(123) POLISH (SLAVIC) (Arregui and Kusumoto 1998, 6)

Ania **powiedziała** ze Marcin **był** chory.
Ania say.PRF.**PST** that Marcin be.**PST** sick

a. ✓ No SOT/Backshifted: Ania said that Marcin was sick in the past, at some time prior to her speaking time.

b. ✗ SOT: Ania said that Marcin was sick at her speaking time.

For a Polish speaker to express a reading akin to the English SOT expression in (122a), they must embed a present tense verb under a past tense attitude verb, as in (124).

(124) Ania **powiedziała** ze Marcin **jest** chory.
Ania say.PRF.**PST** that Marcin be.**PRES** sick
'Ania said that Marcin was sick [at her speaking time].'

(Alternate gloss: 'Ania said: 'Marcin is sick.'')

With this background in mind, we can now address the question of whether or not Tatar permits SOT readings of the embedded past-oriented TAE suffixes.

Lack of SOT in Tatar embedded CPs

SOT readings are unavailable when past-oriented Tatar TAE suffixes are embedded in CPs under past tense attitude verbs. I illustrate this through the example in (125). The Tatar expression in (125) is only compatible with the backshifted interpretation in (125a), in which the time of Timur loving his mother (the embedded event) precedes the time at which he said that he would tell her (the matrix event). This expression is incompatible with an SOT reading in which the times of these two events overlap, as in (125b).

(125) **Context:** Timur has had a difficult relationship with his mother. However, last week, he decided that he loved her and in ten days, he would tell her this.

English prompt: 'Last week Timur said that he would tell his mother he loved her in ten days.'

uzğan atna Timur [äni-se-nä un kön-nän [anı
past week Timur [mother-3SG.POSS-ACC ten day-ABL [3SG.ACC
{yarat-**te**-∅ / yarat-**qan**-∅} dip] äyt-äçäg-e-n] äyt-**te**-∅.
love-**PST**-3SG / love-**RESULT**-3SG COMP] say-PROSP-3SG.POSS-ACC] say-**PST**-3SG

a. ✓ No SOT/Backshifted: Last week Timur said that in ten days, he would tell his mother that he loved her in the past, at some time prior to his speaking time.

(Speaker's comment: "That would be like he used to love her, but it's all over now.")[22]

[22]My primary Tatar consultant additionally notes that the use of *-GAn* in the embedded CP is somewhat semantically odd in this context. This is because the most natural evidential reading of embedded *-GAn* in (125) is one in which it reflects the matrix subject's (i.e., Timur's) evidence. Its pragmatic oddness stems

b. ✗ SOT: Last week Timur said that in ten days, he would tell his mother that he loved her at his speaking time last week.

If a Tatar speaker wants to express a meaning akin to the SOT reading of the English prompt in (125), they must embed a present tense verb under a past tense matrix attitude verb. I show this in (126).

(126) uzğan atna Timur [äni-se-nä un kön-nän [anı yarat-**a**-∅
past week Timur [mother-3SG.POSS-ACC ten day-ABL [3SG.ACC love-**PRES**-3SG
dip] äyt-äçäg-e-n] äyt-**te**-∅.
COMP] say-PROSP-3SG.POSS-ACC] say-**PST**-3SG
'Last week Timur said that he would tell his mother he loved her [at his speaking time last week] in ten days.'

Additional, anecdotal evidence for a lack of SOT in Tatar embedded CPs comes from the translational equivalents that my primary Tatar consultant provides when prompted with SOT expressions in English. My primary Tatar consultant is fluent in both Tatar and English. When I prompt her with an English sentence in which the SOT reading is the most natural interpretation of the embedded tense, she consistently provides a Tatar translational equivalent using a present-under-past expression, as in (127). This parallels the Polish data in (124).

(127) Mansur [Alsu-ne yarat-**a**-∅ dip] äyt-**te**-∅.
Mansur [Alsu-ACC love-**PRES**-3SG COMP] say-**PST**-3SG
English prompt: 'Mansur said that he loved Alsu.'

I conclude from this data that SOT does not apply in embedded CPs in Tatar.

SOT effects in Tatar embedded verbal nominalizations

Recall from Table 3.2 that of the past-oriented Tatar TAE suffixes, only the resultative aspect suffix *-GAn* is grammatical in embedded verbal nominalizations. As I described in

from the fact that Timur should have direct evidence for his feelings for his mother. My consultant reports that this evidential "mismatch" therefore suggests a reading in which the subject of the embedded clauses is not Timur, but rather some third party who Timur is reporting about.

§3.3.1.2, *-GAn* is compatible with both past- and present-oriented temporal interpretations in embedded verbal nominalizations; that is, it effectively marks "non-future." As a result, we predict that SOT readings should be available for embedded verbal nominalizations with *-GAn*, since it encompasses both past- and present-oriented meanings.

This prediction is generally borne out. As shown in (117), embedded *-GAn* can be used to express that the time of the embedded event overlaps with the time of the matrix event. In (117), this indicates that the time of Alsu's pregnancy overlaps with the time of Mansur's saying event, both of which are in the past of the speaker's utterance time.

I show another example of SOT effects with embedded *-GAn* in (128). In this mini narrative, the speaker first uses an embedded CP (lacking an SOT reading) to state that Mansur loved Alsu (at his saying time in the past). The speaker then continues the narrative by stating that Alsu believed that Mansur loved her. In the second sentence, the speaker uses an embedded verbal nominalization containing *-GAn*. This is compatible with an SOT interpretation, in which the time of the event of Mansur loving Alsu overlaps with the time of Alsu's belief.

(128) English prompt: 'Mansur said that he loved Alsu. At that moment, Alsu believed that he loved her.'

Mansur [Alsu-ne yarat-a-∅ dip] äyt-te-∅. Alsu [Mansur-nın
Mansur [Alsu-ACC love-PRES-3SG COMP] say-PST-3SG Alsu [Mansur-GEN
anı yarat-**qan**-ı-na] ışan-**dı**-∅.
3SG.ACC love-**RESULT**-3SG.POSS-ACC] believe-**PST**-3SG
'Mansur said that he loved Alsu. Alsu believed Mansur loved her.'

3.3.2 Embedding TAE suffixes under *ide* 'PST'

In this section, I briefly describe the distribution and interpretations of the Tatar TAE suffixes embedded under the free simple past tense morpheme *ide*. Combinations of TAE suffixes with *ide* result in complex tense and aspectual readings; English translational equivalents of these expressions include e.g. past perfects and counterfactuals. I ultimately do not give a compositional analysis of these expressions in Chapter 6; I set aside their

compositional semantics for future work.

I begin by reiterating that only one TAE suffix can occur on a given Tatar verb; that is, there appears to be only one tense/aspect/evidential "slot" available on the Tatar verb. Tatar differs in this respect from other Turkic languages like Turkish, which permits more than one TAE suffix to occur on the same verb. In the Turkish example in (129), the TAE suffixes *-mIş* (the Turkish counterpart of Tatar *-GAn*) and *-DI* co-occur and result in a reading that is glossed in English as past perfect. Analogous Tatar expressions are ungrammatical, regardless of the order of TAE suffixes, as shown in (130).[23]

(129) TURKISH (Sözen Özkan, p.c.)

Seren sene-ler önce Ankara-ya gel-**miş-ti**-∅.
Seren year-PL ago Ankara-DAT come-MIŞ-DI-3SG
'Seren had come to Ankara years ago.'

(130) a. * sin yeger-**de-gen**-sen.
2SG.NOM run-PST-RESULT-2SG

b. * sin yeger-**gen-de**-n.
2SG.NOM run-RESULT-PST-2SG

Despite this morphological restriction on the Tatar verb, all of the Tatar TAE suffixes (with the exception of the past tense marker *-DI*) can co-occur with the free past tense morpheme *ide*. I summarize this in Table 3.3.

TAE suffix	**Grammaticality in combination with *ide* 'PST'**
-DI 'PST'	*
-GAn 'RESULT'	✓
-A 'PRES'	✓
-(y)AçAK 'PROSP'	✓
-(y)Er 'FUT'	✓

Table 3.3: Grammaticality of the Tatar TAE suffixes in combination with *ide* 'PST.'

[23]Past perfect expressions are generally more natural in combination with temporal adverbials, as in the Turkish example in (129). However, Tatar expressions with "stacked" TAE suffixes like (130) are robustly ungrammatical, regardless of whether or not they contain temporal adverbials.

3.3.2.1 *ide* and *-GAn* 'RESULT'

Tatar expressions in which the resultative aspect suffix *-GAn* is embedded under *ide* result in a range of different English glosses. I propose that these glosses reflect two underlying meanings that are available for the combination of *-GAn* and *ide*. In both cases, the beginning of the described event is located in the past of some past reference time.

The first reading is akin to the English past perfect (e.g. *Maren **had left** when I arrived*). In a past perfect expression, the time of the event described by the verb is interpreted as occurring in the past relative to a reference time; this reference time is itself in the past of the utterance time. I give an example of a past perfect interpretation of a *-GAn ide* expression in (131). In (131), the reference time is established by the temporal adverbial clause 'when the police arrived'; the event of the addressee escaping is then located in the past relative to this past time.

(131) **Escaped thieves context:** You work at a police station. You collected a statement from someone who was arrested. You want to make sure you understand the timeline of events, so you retell their statement to them.

politsia kil-ep cit-kän-dä, sez kit-ep bar-**ğan** **ide**-gez inde.
police come-IP arrive-RESULT-LOC 2PL.NOM leave-IP go-RESULT PST-2PL already
'When the police arrived, you (pl.) had already left.'

Another possible interpretation of a *-GAn ide* expression is one in which the time of the event described by the verb overlaps with a past reference time. This meaning is conceptually very similar to the English past progressive (e.g. *Maren **was leaving** when I arrived*).[24] I give an example of a past progressive interpretation of a *-GAn ide* expression in (132). In this example, the past reference time is established by the adverbial clause 'when Aygöl arrived.' The singing event described by the main clause overlaps with this reference time.

[24]Combining the present tense marker *-A* and *ide* also results in a past progressive reading; I discuss this in §3.3.2.2. At present, I do not address the differences between the use of *-A ide* and *-GAn ide* expressions when expressing past progressive meaning.

(132) **Party context:** You had a party that your friend Mansur was going to sing at. Aygöl hates Mansur's singing, so she was planning arrive late enough to not hear it. However, when Aygöl arrived, Mansur was still singing.

Aygöl kil-ep cit-kän-dä, Mansur haman cırla-**ğan ide**-∅.
Aygöl come-IP arrive-RESULT-LOC Mansur still sing-RESULT PST-3SG
'When Aygöl arrived, Mansur was still singing.'

My primary consultant reports that it is possible that the events described by the main clauses in the above expressions could still be occurring at the utterance time; i.e., double access readings are available. That is, with respect to (132), it could be the case that Mansur is still singing when the speaker says this expression. The use of *-GAn ide* therefore does not place the entire event in the past of the utterance time, but rather locates the beginning of the event in the past of some past reference time.

To summarize: Tatar *-GAn ide* expressions can be interpreted as describing an event time that completely precedes a past reference time (past perfect), or as describing an event time that overlaps with a past reference time (past progressive). In both cases, (at least) the beginning of the event is located in the past of some past reference time. These interpretations of *-GAn* embedded under *ide* parallel its temporal interpretations in embedded verbal nominalizations (§3.3.1.2).

3.3.2.2 *ide* and *-A* 'PRES'

When the present tense suffix *-A* is embedded under *ide*, the resulting expression is glossed in English as a past progressive. It conveys that the event described by the verb was in progress at some reference time in the past of the utterance time. This can be broken down conceptually into the past tense contribution of *ide* in combination with the present/progressive contribution of *-A*.

Like the English past progressive, Tatar *-A ide* expressions generally require that the past reference time be established overtly in the utterance. In (133), this past reference time is picked out by the adverbial clause 'when JFK was shot and killed.'

(133) **JFK atıl-ıp üterel-gän-dä**, min Mäskäü-dä uqıt-**a** **ide**-m.
JFK shoot-IP kill-RESULT-LOC 1SG.NOM Moscow-LOC teach-PRES PST-1SG
'When JFK was shot and killed, I was teaching in Moscow.'

3.3.2.3 *ide* and *-(y)Er* 'FUT'

When the future tense suffix *-(y)Er* is embedded under *ide*, the resulting expression is interpreted as a counterfactual. As such, (134) is felicitous only if the speaker believes that the described event (here, the event of Timur going to Korea) did not occur. I illustrate this through the continuation in (134), in which the speaker explains why Timur did not go.

(134) Timur Koreya-ğa bar-**ır** **ide**-∅ (läkin viza bir-mä-de-lär).
Timur Korea-DAT go-FUT PST-3SG but visa give-NEG-PST-3PL
'Timur would have gone to Korea (but they didn't give him a visa).'

Counterfactual expressions are infelicitous if the speaker does not know whether the described event occurred or not. This is shown by the infelicity of the continuation in (135), in which the speaker asserts that they are unsure whether or not the event of Timur going to Korea actually occurred.

(135) Timur Koreya-ğa bar-**ır** **ide**-∅ (#läkin çınnap-ta bar-ğan-ı-n
Timur Korea-DAT go-FUT PST-3SG but really-ADD go-RESULT-3SG-ACC
bel-m-i-m).
BEL-NEG-PRES-1SG
'Timur would have gone to Korea (#but I don't know if he really went).'
(Speaker's comment: "You have to follow [the first clause] with an explanation of why it didn't happen.")

While I do not provide a compositional analysis of the counterfactual Tatar data in this dissertation, I note that the occurrence of past tense morphology in counterfactual expressions is well documented cross-linguistically (e.g. Iatridou 2000). Tatar therefore patterns like many other languages in this respect.

3.3.2.4 *ide* and *-(y)AçAK* 'PROSP'

When the prospective aspect suffix *-(y)AçAK* is embedded under *ide*, the expression is interpreted as locating an event in the future of some past time ("future-in-the-past"). These *-(y)AçAK ide* expressions are the only (non-periphrastic) strategy that Tatar speakers can use to express future-in-the-past in matrix clauses.

In future-in-the-past expressions, the time of the described event may or may not be in the future of the utterance time. (136) gives an example of a future-in-the-past time that is in the past of the utterance time, whereas (137) gives an example of a future-in-the-past time that is in the future of the utterance time.

(136) **kiçä** Parij-da yañgır yaw-**açaq ide**-∅, läkin cil bolıt-lar-nı kuw-ıp
yesterday Paris-LOC rain rain-PROSP PST-3SG but wind cloud-PL-ACC chase-IP
cibär-de-∅.
send-PST-3SG
'It was going to rain in Paris yesterday, but the wind blew away the clouds.'

(137) Timur **irtägä** Koreya-ğa bar-**açaq ide**-∅, läkin viza bir-mä-de-lär.
Timur tomorrow Korea-DAT go-PROSP PST-3SG but visa give-NEG-PST-3PL
'Timur was going to go to Korea tomorrow, but they didn't give him a visa.'

Future-in-the-past expressions can have a counterfactual flavor. However, Tatar future-in-the-past expressions differ from genuine counterfactual expressions in that when an individual says a future-in-the-past expression, they can be unsure whether or not the described event actually happened. I show that this uncertainty is possible through the felicitous continuation in (138), in which the speaker explicitly states that they are not sure whether or not the described event occurred.

(138) Timur Koreya-ğa bar-**açaq ide**-∅ (läkin çınnap-ta bar-ğan-ı-n
Timur Korea-DAT go-PROSP PST-3SG but really-ADD go-RESULT-3SG.POSS-ACC
bel-m-i-m).
BEL-NEG-PRES-1SG
'Timur was going to go to Korea (but I don't know if he really went).'

This future-in-the-past data contrasts with the counterfactual data in (135); I showed in (135) that counterfactual expressions are felicitous only if the speaker believes that the

event did not happen. While Tatar *-(y)Er ide* and *-(y)AçAK ide* expressions have similar counterfactual flavors, they are not truth-conditionally equivalent.

3.3.3 Recap: Temporal interpretation of TAE suffixes in semantically embedded environments

In semantically embedded environments, the Tatar TAE suffixes are compatible with a wider range of temporal interpretations than in unembedded environments. The available range of interpretations differs based on whether they occur in embedded CPs or verbal nominalizations, with verbal nominalizations permitting the greatest flexibility. For instance, I showed in §3.3.1.3 that SOT readings are available only in embedded verbal nominalizations, and are blocked from embedded CPs. This stems from the fact that in verbal nominalizations, *-GAn* is compatible with both past- and present-oriented readings of the embedded event.

When the TAE suffixes are embedded under the free past tense morpheme *ide*, the resulting readings are temporally complex. My consultants gloss these expressions in English using past perfects, past progressives, and so on. These combinations of finite verbs and *ide* are the most productive strategies by which Tatar speakers form complex tense/aspect expressions, since multiple TAE suffixes cannot co-occur on a single Tatar verb (i.e., the TAE suffixes cannot "stack").

3.4 Conclusion

This chapter concludes my descriptive overview of the Tatar TAE system. In Chapter 2, I described the evidential interpretations of the TAE suffixes. In this chapter, I described the temporal interpretations of the TAE suffixes. In my analysis in Chapter 5, I will argue that these suffixes have underlying temporal semantics; i.e., their evidential meanings are a pragmatic byproduct of their temporal meanings.

As I noted previously, my glosses of the TAE suffixes differ somewhat from previous

descriptive accounts of the Tatar data. I compare my descriptive terminology to the terminology used by Greed (2014), Tatevosov (2007), and Poppe (1961) in Table 3.4. Like these three authors, I also propose that the underlying meanings of these suffixes are temporal, and not evidential.

Suffix	**My description**	**Greed (2014)**	**Tatevosov (2007)**	**Poppe (1961)**
-DI	past tense	past tense	preterite	past tense
-GAn	resultative aspect	perfect	perfect	past participle
-A	present tense	present tense	present tense	present tense
-(y)AçAK	prospective aspect	future tense	future tense	future participle
-(y)Er	future tense	future tense	future tense	future participle

Table 3.4: Comparison of my description of the Tatar TAE suffixes to Greed (2014), Tatevosov (2007), and Poppe (1961).

I am generally in agreement with these authors with respect to the glosses proposed for the past tense suffix *-DI*, present tense suffix *-A*, and future tense suffix *-(y)Er*. However, my account differs from theirs with respect to my glosses for the resultative aspect suffix *-GAn* and the prospective aspect suffix *-(y)AçAK*. As I showed in §3.2.1.2, *-GAn* patterns very differently from English (present) perfect expressions. Since analyses of the present perfect are primarily based on English data, I argue therefore that *-GAn* should not be termed a (present) perfect.

I theoretically motivate my aspectual gloss of *-(y)AçAK* in Chapter 5. Descriptively speaking, I note that readings of high speaker certainty (like those associated with *-(y)AçAK*) are described for prospective aspects cross-linguistically, as in e.g. Syrian Arabic (Jarad 2014), Plains Cree (Wolvengrey 2006), and other Turkic languages (Korn and Nevskaya 2017). This suggests that *-(y)AçAK* has prospective, rather than plain future tense, semantics. As I will eventually argue in Chapter 5, the evidential interpretations of the Tatar TAE suffixes can be linked to the aspectual semantics of *-GAn* and *-(y)AçAK*.

With this descriptive data in hand, I will now turn to the task of providing a theoretical semantic and pragmatic analysis. I outline prior analyses of similar TAE data from other languages in Chapter 4, and argue that these analyses cannot satisfactorily account for all

of the observed Tatar data. I provide my analysis of the Tatar TAE suffixes in matrix clauses in Chapter 5; in Chapter 6, I sketch an analysis of their interpretation in some semantically embedded environments.

Part II

An aspectual proposal for evidentiality

CHAPTER 4

Evaluating the Tatar data against prior theories of evidentiality

4.1 Introduction

In Part 1 of this dissertation, I presented the Tatar data that this dissertation addresses. I showed that Tatar has a set of four verbal suffixes that, in semantically unembedded environments, are descriptively interpreted as contributing a combination of evidential information as well as temporal information about the scope proposition.

In this chapter, I evaluate some existing theories of evidentiality against how well they can explain the following questions:

1. **Why are tense/aspect and evidentiality fused in Tatar?**

2. **How can we account for the future-oriented Tatar TAE suffixes?**

There are no existing theoretical accounts of evidentiality and tense/aspect in Tatar; prior linguistic literature on the Tatar evidentiality and tense/aspect system is descriptive (Greed 2014, 2009; Tatevosov 2007; Poppe 1961).[1] Furthermore, of the current theories of evidentiality available to us, only a subset of these are concerned with addressing Question 1 above, i.e., how it is that evidentiality and tense/aspect are linked in their language of study. I therefore do not discuss theories formulated to account for data in which evidentiality is realized completely separately from the language's tense/aspect system (e.g.

[1] I have discovered references to one paper that appears to provide an analysis of the future oriented Tatar TAE suffixes (Tatevosov 2017); however, to the best of my knowledge, this paper is only available in Russian and I have been unable to locate the full text. I therefore do not discuss it in this chapter.

Faller 2002's illocutionary operator analysis of Cuzco Quechua evidentials; Murray 2010's Not-At-Issue analysis of Cheyenne evidentials; among others).

This chapter contains three sections. In §4.2, I discuss why some previously proposed tests for modality in the semantics of evidentials do not accurately diagnose its presence, and suggest that the Tatar TAE morphemes do not include a modal component in their semantics. In §4.3, I evaluate my Tatar data against some prior theories in which evidentiality is treated as a byproduct of temporal meaning. I refer to these as Evidence Acquisition Time analyses (§4.3.1; Lee 2013, 2011; Smirnova 2013, 2011; Koev 2011) and trace analyses (§4.3.3; Koev 2017, Faller 2004). I ultimately show that these theories, while compelling, cannot account for all of the observed Tatar data. In §4.4, I sketch the analysis that I will eventually propose for my data, and acknowledge the intuitions that my analysis borrows from prior EAT and trace theories of evidentiality.

4.1.1 Semantic variables and types

Before presenting any formal theories of evidentiality, I spell out some of the details of the semantic formalisms that I will use in this dissertation. In Table 4.1, I list the variables and types that I use to refer to different semantic objects.

Object	**Variable**	**Type**
Truth values	—	t
Individuals	x, y	e
Events	e, e'	v
Times	i, i'	i
Worlds	w, w'	s

Table 4.1: Semantic variables and types used in this dissertation.

In this chapter, I adapt other authors' formalisms to use the semantic variables and types in Table 4.1, where necessary. This is intended only to provide uniform variables and types across the dissertation, and is not intended to alter the content of any of the theories that I present. I also standardize some of the interlinear glosses used in the following chapter (e.g. I gloss all past tense morphemes using PST); again, this is solely for presentation, and

is not intended to suggest that I am altering other authors' analyses.

4.2 Preliminary discussion: Evidentials as epistemic modals

There is an established tradition in the evidential literature, beginning with Izvorski (1997), of analyzing evidentials as including an (epistemic) modal component. This analysis has been adopted by a range of authors (Lee 2013, 2011; Smirnova 2013, 2011; Matthewson 2011, 2012b; McCready and Ogata 2007; among others). Several of the Evidence Acquisition Time analyses of evidentiality reviewed in §4.3.1 themselves include epistemic modality in their semantics.

In this section, I briefly review a basic analysis of evidentiality as epistemic modality (Izvorski 1997). I then discuss a number of tests that have been used in the literature to motivate modal analyses of evidentiality. I argue (in part following observations by Murray 2017; Korotkova 2016; Matthewson 2011, 2012a) that the majority of these tests do not necessarily diagnose modality in the semantics of evidentials. This discussion is relevant to my evaluation of the (modal) Evidence Acquisition Time analyses in §4.3.1; I choose not to adopt these theories to account for the Tatar data in part because I find that including modality in the semantics of the Tatar TAE suffixes is not motivated.

4.2.1 Basic epistemic modal analysis of evidentiality

Izvorski (1997) addresses evidential data from Bulgarian, Turkish, and Norwegian. Evidential expressions in these languages are linked either historically or synchronically with the present perfect (Izvorski terms this connection the "perfect of evidentiality"). I show in (139)-(140) that present perfect expressions in these languages also have indirect evidential readings. However, accounting for this connection is not a central component of Izvorski's analysis. She is concerned primarily with giving a semantics for the evidential component of these morphemes.[2]

[2]In addition to her modal account of indirect evidentiality, Izvorski (1997, 13) provides an unformalized

(139) BULGARIAN (SLAVIC) (Izvorski 1997, 1)

Az sâm došâl.
I be.1SG.PRES come.P.PART

a. 'I have come.'

b. 'I apparently came.'

Alternate translation: '[I have indirect evidence that] I came.'

(140) NORWEGIAN (GERMANIC) (Izvorski 1997, 1)

Jeg har kommet.
I have.1SG.PRES come.P.PART

a. 'I have come.'

b. 'I apparently came.'

Alternate translation: '[I have indirect evidence that] I came.'

Izvorski (1997) proposes that the evidential component of the expressions in (139)-(140) should be analyzed as a universal epistemic modal (in the Kratzerian framework, as a universal quantifier over possible worlds). She supplements this modal with a presupposition that the speaker has indirect evidence for the scope proposition, as in the basic schema in (141).[3]

(141) **The interpretation of EVID(*p*):**

a. **Assertion**: $\Box p$ in view of the speaker's knowledge state

b. **Presupposition**: Speaker has indirect evidence for p

This proposal can be formalized in more detail as follows. In (142), the modal base is defined as the set of propositions that the speaker considers indirect evidence in w.

account of the link between indirect evidentiality and the present perfect. For the purpose of this chapter, I designate papers as addressing the theoretical link between tense/aspect and evidentiality only if they provide a formal account of the relationship between the categories.

[3] Izvorski (1997) notes that the strength of evidential assertions (in these languages) can vary depending on whether the evidential is being interpreted relative to reportative or inferential evidence. In reportative contexts, the modal component of the evidential can range from existential to universal force; however, in inferential contexts, the modal consistently has universal force. She therefore proposes that this variable force is derived contextually, and presents a uniform analysis of the indirect evidential as a universal quantifier.

This modal base is restricted by the ordering source in (143), which orders worlds *w′*, *w″* according to whether there are more propositions in *g(w)* which are true in *w″* than in *w′*. Finally, the denotation of an evidential expression in (144) can be represented identically to an expression containing a necessity modal. (All formalisms below are slightly adapted from Izvorski 1997, 9.)

(142) **Modal base**

f(w) = {p: speaker considers p indirect evidence in w}

(143) **Ordering source**

g(w) = {p: speaker believes p with respect to the indirect evidence in w}

For all worlds w′,w″ ∈ W, w′ $<_{g(w)}$ w″ iff:

{p: p ∈ g(w) & w″∈ p} ⊂ {p: p ∈ g(w) & w′ ∈ p}

(144) **Denotation**

$[\![\text{EVID}(p)]\!]^{w,f,g}$ = ∀w′[w′ ∈ $\text{BEST}_{g(w)}\bigcap$f(w) → p(w′) = 1]

Izvorski (1997) motivates this presuppositional modal analysis primarily through the observation that the indirect evidential contribution of the morphemes cannot be targeted by negation, as in the Bulgarian example in (145). Presuppositions are noted for their ability to project above sentential negation (Karttunen 1973, Strawson 1950, among many others). This interaction with negation is also described for the Tatar TAE morphemes, as discussed in Chapter 2.

(145) Ivan ne izkaral izpita.
Ivan not passed.PERF.EVID the.exam

a. = 'Ivan didn't pass the exam [it is said/I infer].'

b. ≠ 'It is not the case that [it is said/I infer] that Ivan passed the exam.'

In the following section, I discuss some tests in the literature that have been used to diagnose modality, and briefly evaluate Tatar against some of these tests.

4.2.2 Tests in the literature for modality in evidential expressions

A number of tests have been proposed in the literature to diagnose modality in an evidential expression. These include (i) the ability to be semantically embedded (e.g. occur under an attitude verb); (ii) the ability to participate in modal subordination; (iii) the inability of the evidential contribution to scope under negation; (iv) the inability of the speaker to assert EVID(p) if they know that p is false; and (v) the inability of the speaker to assert EVID(p) if they know that p is true. I review each of these tests with respect to their ability to accurately diagnose modality in a given evidential expression, and show how the Tatar data patterns with respect to each of them. I also show that these tests themselves do not definitively diagnose modality in a given expression.

4.2.2.1 Modal test #1: Ability to be semantically embedded

The ability of an evidential to be semantically embedded (under e.g. attitude verbs or in the antecedents of conditionals) is typically taken in the literature to definitively diagnose modality in an evidential expression. Authors that use embeddability as a test for modality include Lee (2013) for Korean; Matthewson (2012a) and Matthewson et al. (2007) for St'át'imcets; and McCready and Ogata (2007) for Japanese. However, as noted previously by Korotkova (2016), the ability of a given expression to be semantically embedded merely suggests that it forms part of the propositional content of the expression. That is, embeddability does not necessarily diagnose modality, as many other linguistic expressions also have the ability to be semantically embedded.[4]

In (146)-(147), I give two examples of semantically embedded evidentials that have been used to argue for modal analyses of the given data. In the Korean example in (146),

[4]The origin of this test as a diagnosis for modality stems from what Korotkova (2016) terms the "dichotomy view" of evidential analyses; that is, that evidentials can either be modals that form part of the propositional content of an expression (following Izvorski 1997) or illocutionary operators that combine at the speech act level (following Faller 2002). Current research (Korotkova 2016; Koev 2011, 2017; Murray 2010; among others) suggests that this dichotomy view is far too simple. Furthermore, Korotkova (2016) notes that modals can be blocked from embedding, suggesting that a lack of embeddability does not necessarily prove that an evidential expression is non-modal.

the embedded proposition "It was raining yesterday" is marked with the direct evidential *-te*.[5] This use of *-te* is anchored to the evidence possessed by the the matrix subject, Chelswu, and not to the speaker. This is shown by the ability of the speaker to deny the embedded proposition in the following sentence.

(146) KOREAN (KOREANIC) (Lee 2013, 22)

Chelswu-nun pi-ka ecey o-∅-**te**-la-ko malha-yess-e.
Chelswu-TOP rain-NOM yesterday fall-PRES-TE-DECL-COMP say-PST-DECL
Kulentay ecey pi-ka an-o-ass-e.
but yesterday rain-NOM NEG-fall-PST-DECL
'Chelswu said that [he made a sensory observation that] it was raining yesterday. But it did not rain yesterday.'

In the St'át'imcets example in (147), the embedded proposition "Maria hit her younger brother" is marked with the inferential evidential *k'a*. The context is such that this evidential particle can only reflect the evidence of the matrix subject, Lémya7, and not the speaker (who has reportative evidence for the embedded proposition).

(147) ST'ÁT'IMCETS (SALISHAN) (Matthewson et al. 2007, 47)

Context: Lémya7 was babysitting your nephew and niece and she noticed at one point that the boy had a red mark on his face and his sister was looking guilty. She tells you when you get home what she noticed. Then you tell the mother of the kids.

tsuts-Lémya7 kw s-tup-un'-ás **k'a** s-Maria ta
say.NOM-Lémya7 DET NOM-punch-DIR-3.ERG INFER NOM-Maria DET
sésq'wez'-s-a
younger.sibling-3.POSS-DET
'Lémya7 said that Maria must have hit her younger brother.'

I showed in Chapter 2 that the evidential contribution of the past oriented Tatar TAE suffixes—unlike in Korean and St'át'imcets—are not always interpreted as being embedded. That is, the evidential contribution of embedded past oriented TAE suffixes can be

[5]The interpretation of *-te* as a direct or indirect evidential varies based on the tense morpheme that it co-occurs with; (146) could therefore be more accurately described as a direct evidential use of *-te*.

interpreted as reflecting the speaker's evidence, rather than the matrix subject's. I repeat an example of a speaker oriented embedded TAE suffix in (149). In this example, the embedded TAE suffix *-GAn* reflects that the speaker, not the matrix subject, has indirect evidence for the embedded proposition.[6]

(149) **Olympics attendee context**

Speaker: Indirect evidence

Matrix subject: Direct evidence

Läylä went to the Olympics and saw Michael Phelps compete. She calls you and tells you he won; you did not see the race yourself. You then tell someone else:

Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] say-PST-3SG
'Läylä said that [and I have indirect evidence that] Michael Phelps won.'

(Speaker's comment: "It's possible to say (149) even if Läylä saw it happen, because you are the speaker reporting about Läylä.")

If we were to assume that the Tatar TAE suffixes have underlyingly modal semantics—i.e., that their evidential meaning is hardwired into their semantics—we also could not account for the fact that the evidential readings of the TAE suffixes disappear in some semantically embedded environments. For instance, I showed in Chapter 2 that the evidential readings of the TAE suffixes are absent in embedded verbal nominalizations.

[6] Izvorski (1997)'s modal semantics in (144) are such that the evidential is only predicted to be speaker-oriented, as in the data in (149). However, as I showed in Chapter 2, embedded Tatar TAE suffixes can also reflect the matrix subject's evidence, as repeated in (148). The availability of this data is not predicted by the semantics in (144).

(148) Swimming fan context
Speaker: Non-specific evidence
Matrix subject: Specific evidence
Läylä is watching an Olympic swimming event. She is a huge fan of competitive swimming, and knows the statistics about all of the swimmers. She believes that Michael Phelps has trained the best out of everyone competing at the Olympics this year: his diet, practice regimen, etc. are all impeccable. She tells you he will definitely win. However, you have no strong opinion about his ability to win. Later, you tell someone else:
Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG
'Läylä said that [and has specific evidence that] Michael Phelps will win.'

I repeat relevant Tatar embedding data in (150). In this example, the use of *-GAn* in the embedded clause is felicitous regardless of whether the matrix subject has direct or indirect evidence for the embedded proposition.

(150) a. **Direct evidence context:** Mansur tells you that he accompanied Läylä to the train station yesterday and watched her get on a train to go to Kazan. You express Mansur's report of what Läylä did yesterday.

b. **Indirect evidence context:** Mansur finds a receipt for a train ticket to Kazan in Läylä's desk. He tells you that he infers that Läylä went to Kazan. You express Mansur's report of what Läylä did yesterday.

Mansur [Läylä-nen Qazan-ğa kit-ep bar-**ğan**-ı-n] äyt-te-∅.
Mansur [Läylä-GEN Kazan-DAT leave-IP go-RESULT-3SG.POSS-ACC] say-PST-3SG
'Mansur said Läylä went to Kazan.'

According to a modal analysis of evidentiality, the evidential/modal contribution of the TAE suffix *-GAn* should be hardwired into its semantics, along the lines of Izvorski (1997)'s denotation in (144). If so, we cannot explain why its evidential reading disappears entirely in some semantically embedded contexts. Data like (150) therefore provides a strike against a modal analysis of the Tatar TAE suffixes.

4.2.2.2 Modal test #2: Ability to participate in modal subordination

Modals can be interpreted as being semantically subordinate to previous modals in the discourse (Roberts 1989); this phenomenon is termed **modal subordination**. I give an example of modal subordination in English in (151). In this example, the indefinite noun *a thief* is introduced within the scope of the epistemic necessity modal *must*. Since the existence of a thief is not entailed by this modalized expression, the speaker cannot anaphorically refer back to this individual using the non-modalized expression in (151a). Instead, the speaker must embed the co-referential pronoun *it* under another modal, as in (151b). The modal *must* in (151b) therefore restricts the domain of the original modal expression in (151).

(151) **Tall thief context** (modified from Lee 2011, 299): When you got home yesterday, you found that your belongings were scattered all over your apartment. You infer that someone broke into your house while you were gone. You see that a shelf very high up has had things taken off of it. You say:

A thief$_i$ must have broken into my house.

a. # ... He$_i$ is/was tall.

b. ✓ ... He$_i$ must be/must have been tall.

A number of authors argue that the evidential expressions they analyze can participate in modal subordination, supporting a modal analysis of their evidential data. These authors include Smirnova (2011, 2013) for Bulgarian; Lee (2011, 2013) for Korean; and McCready and Ogata (2007) for Japanese.[7]

I give a Bulgarian example from Smirnova (2013) in (152). Smirnova proposes that the pronoun *toj* 'he' in the unmodalized expression in (152a) cannot be anaphorically linked to the nominal *kradec* 'thief' in the indirect evidential expression in (152). An anaphoric dependency between *toj* 'he' and *kradec* 'thief' is available if the pronoun is embedded under the epistemic necessity modal *trjabva* 'must,' as in (152b). Smirnova (2013) argues that this data shows that the Bulgarian indirect evidential must include a modal component in its semantics, making the Bulgarian data analogous to the English data in (151).

(152) BULGARIAN (Smirnova 2013, 506-507)

Inferential context: You just came home and discovered that your house has been broken into. Besides your laptop, a tray of baklava is missing. There are baklava crumbs all over the floor. You inferred that the thief ate the baklava. Later you tell Maria:

Vkăšti **vleznal** kradec$_i$.
in.house enter.PERF.PST.PLE thief.
'A thief$_i$ broke into the house [I inferred].'

[7]However, Koev (2017, 25) gives Bulgarian data that he argues demonstrates the availability of anaphora across evidential expressions—i.e., a lack of modal subordination effects—in Bulgarian.

a. # Toj$_i$ beše gladen.
he be.3SG.PST hungry.
'... He$_i$ was hungry.'

b. ✓ Toj$_i$ **trjabva** da e bil gladen.
he must.PRES SUBJ be be.PLE hungry.
'... He$_i$ must have been hungry.'

I show analogous Tatar data in (153). The genitive pronoun *anın* 'his' in the unmodalized expression in (153a) cannot be anaphorically linked to the noun *qaraq* 'thief' in the preceding expression, which is marked with the TAE suffix *-GAn*. However, an anaphoric dependency between *anın* 'his' and *qaraq* 'thief' is available if the pronoun is embedded under the epistemic necessity modal *tiyeş* 'must,' as in (153b).

(153) **Tall thief context** (modified from Lee 2011, 299): When you got home yesterday, you found that your belongings were scattered all over your apartment. You infer that someone broke into your house while you were gone. You see that a shelf very high up has had things taken off of it. You say:

öye-bız-gä qaraq$_i$ ker-**gän**-∅.
house-1PL.POSS-DAT thief enter-RESULT-3SG
'[I have indirect evidence that] a thief$_i$ entered our house.'

a. # anın$_i$ buye ozın.
3SG.GEN height tall
'... He$_i$ is tall.'

(Lit. 'His$_i$ height is tall.')

b. ✓ anın$_i$ buye ozın bul-ırğa **tiyeş**.
3SG.GEN height tall be-INFIN must
'... He$_i$ must be tall.'

(Lit. 'His$_i$ height must be tall.')

However, the Tatar data in (153) does not necessarily require a modal explanation. When speakers report evidentially unmarked expressions as in (153a), the implicature is that they have direct (i.e., sensory) evidence for the proposition that they are expressing. However, in the context in (153), the speaker does not have sensory evidence either for the proposition "A thief entered our house" or the continuation "The thief is tall." Uttering the

"plain" proposition in (153a) is therefore infelicitous in this context, since it would make a stronger evidential claim than is licensed by the speaker's available evidence. (Koev 2017, 24-25 raises similar objections to Smirnova 2013's Bulgarian data in (152).)

The speaker can make their continuation felicitous by including the epistemic necessity modal *tiyeş*, as in (153b). This could also fall out from evidential requirements; von Fintel and Gillies (2010) argued that English epistemic modals like *must* require that the speaker has indirect evidence for the scope proposition. At present, I have no reasons to propose that Tatar *tiyeş* patterns differently semantically or pragmatically from English *must*. The use of *tiyeş* in (153b) could therefore be required to satisfy the indirect evidential requirement given by the context, rather than to enforce modal subordination. Regardless of the correct analysis of the Bulgarian data in (152), the Tatar data in (153) could be explained away by evidential requirements, and does not necessarily require positing a modal component in the denotations of the TAE suffixes.

4.2.2.3 Modal test #3: Inability to scope under negation

Izvorski (1997) argued that the inability of the Bulgarian indirect evidential to scope under sentential negation supports an underlyingly presuppositional (i.e., modal) analysis of the evidential. Izvorski encodes the indirect evidence requirement as a presupposition on the modal. As described previously, presuppositions are noted for their ability to project above negation (Karttunen 1973, Strawson 1950, among many others). I give the relevant Bulgarian data in (154), repeated from (145).

(154) BULGARIAN (Izvorski 1997, 7)

Ivan **ne izkaral** izpita.
Ivan not passed.PERF.EVID the.exam

a. = 'Ivan didn't pass the exam [it is said/I infer].'

b. ≠ 'It is not the case that [it is said/I infer] that Ivan passed the exam.'

Tatar patterns identically to Bulgarian with respect to this test, as in (155). The Tatar

example in (155) is incompatible with a reading in which the speaker denies having indirect evidence for Mansur going to Moscow.

(155) Mansur Mäskäü-gä bar-**ma-ğan**-∅.
Mansur Moscow-DAT go-NEG-RESULT-3SG
a. = '[I have indirect evidence that] Mansur didn't go to Moscow.'
b. ≠ '[It is not the case that I have indirect evidence that] Mansur went to Moscow.'

However, this inability to scope under sentential negation could also be accounted for by assuming that evidentials operate at the speech act level, i.e., at a level higher than sentential negation. Faller (2002) proposes such an explanation for the Cuzco Quechua example in (156) with the direct evidential *-mi*.[8]

(156) CUZCO QUECHUA (QUECHUAN) (Faller 2002, 227)
Ines-qa **mana-n** qaynunchaw ñaña-n-ta-**chu** watuku-rqa-n.
Inés-TOP not-MI yesterday sister-3-ACC-CHU visit-PST.1-3
a. = '[I have direct evidence that] Inés did not visit her sister yesterday.'
b. ≠ '[It is not the case that I have direct evidence that] Inés visited her sister yesterday.'

The interpretation of the evidential above negation could also be accounted for under an analysis of the evidential as contributing Not-At-Issue content, along the lines of the proposals in Koev (2017) for Bulgarian and Murray (2010) for Cheyenne. This test therefore does not definitively diagnose modality in a given evidential expression.

4.2.2.4 Modal test #4: Inability of the speaker to assert EVID(*p*) if they believe that *p* is false

Modal statements allow speakers to discuss a proposition without asserting whether it is true or false in the actual world. In fact, the received wisdom is that if a speaker asserts

[8] Technically Faller (2002) terms *-mi* a "best possible grounds" evidential. *-mi* surfaces as *-n* in the example in (156). Negation in Cuzco Quechua is expressed through a bipartite construction including both *mana* and *-chu*.

an epistemic expression MODAL(p), they must **not** be certain whether p is true or false in the actual world.[9]

I show data in (157) suggesting that the English epistemic possibility modal *might* is infelicitous if the speaker believes that the scope proposition is false (modified from Faller 2002, 193). It is infelicitous for a speaker to report *might*(p) and then immediately deny the truth of p, as in (157a). (I assume for the purpose of this example that speakers believe the propositions that they report.) It is similarly infelicitous for a speaker to assert *might*(p) and then give a continuation in which they assert that they believe $\neg p$, as in (157b).

(157) a. It might be raining (#and it's not raining).

b. It might be raining (#and I don't believe it's raining).

I set aside the issue of how precisely to analyze the data in (157), since data along these lines has been discussed at length by a number of other authors (e.g. Yalcin 2007). Instead, I simply note that several authors have observed that evidential expressions seem to pattern similarly to the epistemic modals in these examples (Matthewson et al. 2007, Smirnova 2013, Lee 2011, among others).[10] This contradicts what we might expect if evidentials solely marked the speaker's evidence source for the scope proposition p; if an evidential solely marked the speaker's evidence for p, and not the speaker's epistemic state regarding p, it should be possible for the speaker to felicitously assert EVID(p) while believing that p is false. This falls out from the fact that reporting a speaker's evidence source for p is theoretically independent from reporting their epistemic stance regarding the truth of p.

Authors give a range of continuations to demonstrate speaker belief that p is false. I define two main classes. In the first, speakers follow EVID(p) expressions with plain

[9] I discuss an exception to this from von Fintel and Gillies (2010) in §4.2.2.5.

[10] Reportative evidentials are a noteworthy exception to this generalization. AnderBois (2014, 238) gives multiple examples of REPORTATIVE.EVID(p) expressions that are felicitously followed by DIRECT.EVID($\neg p$) and plain $\neg p$ continuations. AnderBois ultimately argues that the availability of these continuations is pragmatically derived; the reportative evidential facilitates perspective shift from the speaker to the source of the reported proposition.

$\neg p$ assertions, as in (157a). These expressions are of the form EVID(p), $\neg p$. In the second, speakers follow EVID(p) assertions with negated expressions containing another evidential, typically a direct evidential. These expressions are of the form EVID(p), EVID($\neg p$). Since the TAE suffixes obligatorily occur in past- and future-oriented finite Tatar clauses, Tatar can only be evaluated with respect to the second class of continuations. I review both classes of continuations below, and argue that they do not necessarily require modal semantics.

EVID(p), $\neg p$ expressions

The following infelicitous St'át'imcets examples are of the form EVID(p), $\neg p$. Matthewson et al. (2007) test two St'át'imcets indirect evidentials in the initial expression: the inferential evidential *k'a* (158a) and the conjectural evidential *an'* (158b). In both examples in (158), St'át'imcets expressions containing one of these evidentials cannot be followed by a denial of the original expression. (Other authors that cite this test in support of a modal analysis of evidentiality include Lee 2011 for Korean and Smirnova 2013 for Bulgarian; Smirnova 2013, 507-509 notes that while this test can support a modal analysis, it does not definitively diagnose modality.)

(158) ST'ÁT'IMCETS (Matthewson et al. 2007, 27-28)

a. * wa7 **k'a** kwis, t'u7 **aoz** t'u7 k-wa-s kwis.
IMPF INFER rain but NEG just DET-IMPF-3POSS rain
'[I have indirect evidence that] it's raining, but it's not raining.'

(Matthewson et al.'s original gloss: 'It may/must be raining, but it's not raining.')

b. * wá7-as-**an'** kwis, t'u7 **aoz** t'u7 k-wa-s.
IMPF-3CONJ-PERC.EVID rain but NEG just DET-IMPF-3POSS rain
'It's apparently raining, but it's not raining.'

The data in (158) is compatible with a modal analysis of evidentiality. However, this data could also be accounted for under e.g. an illocutionary operator analysis of evidentiality (Faller 2002). If evidentials operate at the speech-act level and EVID(p) expressions nonetheless assert p, as according to Faller (2002)'s theory, examples like (158) would be

infelicitous due to being contradictory. Faller uses an argument along these lines to account for the infelicity of the Cuzco Quechua illocutionary direct evidential *-mi* in a similar expression (159). (I note that (159) differs marginally from the St'át'imcets examples in (158) in that the Cuzco Quechua speaker asserts that they do not believe p, rather than explicitly stating $\neg p$.)

(159) CUZCO QUECHUA (Faller 2002, 163)

#Para-sha-n-**mi**, ichaqa mana crei-ni-chu.
rain-PROG-3-MI but not believe-1-NEG
'It is raining, but I don't believe it.'

Since the infelicity of EVID(p), $\neg p$ expressions can be accounted for by other theories of evidentiality (namely, speech act theories), the infelicity of these expressions in a language therefore does not necessarily diagnose modality in the evidential.

EVID(p), EVID($\neg p$) expressions

In the infelicitous Cuzco Quechua example in (160), the initial EVID(p) assertion contains the conjectural evidential morpheme *-cha*. This is followed by a negated expression including the direct (best possible grounds) evidential *-mi*. This example has the form EVID(p), $\neg$EVID(p). Faller (2002) cites the infelicity of this data in support of a modal analysis of *-cha*.

(160) CUZCO QUECHUA (Faller 2002, 178)

#Llave-qa muchila-y-pi-**chá** ka-sha-n, ichaqa **mana-n** aqhay-pi-**chu**.
key-TOP backpack-1-LOC-CHA be-PROG-3 but not-MI there-LOC-NEG
'The keys {may be/are possibly/probably} in my backpack, but they are not there.'

The Tatar TAE morphemes pattern identically to Cuzco Quechua *-cha* in this respect. In (161), the speaker uses *-GAn* to indicate that they have indirect evidence for the initial scope proposition "The dove exploded." They cannot felicitously follow this with a direct evidential assertion denying this initial proposition.[11]

[11] Following the initial *-GAn*(p) assertion with a continuation of the form *-GAn*($\neg p$) is also infelicitous. My consultants report that it sounds contradictory.

(161) **Magic show context:** You went to a magic show in Las Vegas. As part of the act, the magician went behind a curtain with a dove. You saw a puff of feathers come up above the curtain, and you inferred that he made the dove explode. But later in the act, he took the same dove out of his pocket and showed that it was completely fine. You say:

kügärçen şartla-**ğan**-∅ (#läkin şartla-**ma-tı**-∅).
dove explode-RESULT-3SG but explode-NEG-PST-3SG
'[I have indirect evidence that] the dove exploded (but [I have direct evidence that] it didn't explode).'

However, (161) does not require a modal analysis of the evidential. Data like (161) could also be explained under an illocutionary account in which uttering EVID(p) asserts p. If the speaker asserts p in the initial expression, it would be contradictory to then assert $\neg p$ in the continuation. Furthermore, analogous data holds of evidentials that have not been argued to include modality in their semantics, e.g. in Cheyenne (Murray 2010, 54) and in other Cuzco Quechua evidentials (Faller 2002). While this data is compatible with a modal analysis of evidentiality, it does not necessarily diagnose it.

4.2.2.5 Modal test #5: Inability of the speaker to assert EVID(p) if they believe that p is true

Received wisdom states that cooperative speakers cannot felicitously assert a proposition embedded under an epistemic modal if they believe that this scope proposition is true. I illustrate this with English examples in (162). It is infelicitous for a speaker to report *might*(p) and then immediately assert p, as in (162a). It is also infelicitous for a speaker to report *might*(p) and then assert that they believe that p, as in (162b). I again set aside an analysis of this data, since it is outside the scope of this dissertation.

(162) a. It might be raining (#and it's raining).

b. It might be raining (#and I believe it's raining).

Matthewson et al. (2007, 20) propose that, under an analysis of evidentials as epistemic

modals, it should be similarly infelicitous for a speaker to assert EVID(p) if they are certain that p is true, i.e., if they have direct (eyewitness) evidence for p. Matthewson et al. (2007) test such expressions in St'át'imcets using the reportative evidential *ku7* and the inferential evidential *k'a* in (163). In both examples in (163), the speaker reports propositions for which they have direct (eyewitness) evidence using non-direct evidentials. Both examples are infelicitous.

(163) ST'ÁT'IMCETS (Matthewson et al. 2007, 30)

a. **Context:** You were invited to Ted's wedding and you went there and watched him get married. Marilyn (Ted's sister) didn't see you at the wedding and didn't know you had been invited. She told you 'Ted got married.' Later, you see me and you tell me:

#melyíh **ku7** kw s-Ted
marry REPORT DET NOM-Ted
'[I heard] Ted got married.'

b. ts'um'-qs-án'-as **k'a** kw s-Lémya7 kw s-Roger
lick-nose-DIR-3.ERG INFER DET NOM-Lémya7 DET NOM-Roger
(#ats'x-en-lhkán wi7 zam').
see-DIR-1SG.SUBJ EMPH after.all
'Lémya7 must have kissed Roger (actually, I saw it).'

The Tatar TAE suffix *-GAn* patterns similarly to the St'át'imcets data in (163); i.e., *-GAn(p)* expressions are infelicitous in contexts in which the speaker has direct evidence for p.

(164) **Fishing trip context**: You go fishing with your family. You see your son, Nima, catch a fish. Later, you tell your mother that Nima caught a fish. You say:

Nima balıq tot-**qan**-∅.
Nima fish catch-RESULT-3SG
'[I have indirect evidence that] Nima caught a fish.'

However, this inability to assert EVID(p) in contexts in which the speaker is certain that p is true appears to generally not be the case for expressions containing direct or

best possible grounds evidentials. In the Cuzco Quechua example in (165), the context establishes that the speaker has eyewitness evidence for the scope proposition "He is living with my niece." Since the speaker presumably believes that what they see is true, we can assume that the speaker is certain that the prejacent in (165) is true. Nonetheless, the use of the direct (best possible grounds) evidential *=mi* is permitted (in fact, required).

(165) CUZCO QUECHUA (Faller 2011, 663)

Context: The speaker has seen that he is living with her niece.

Subrina-y-wan**=mi** tiya-sha-n.
niece-1-COM=MI live-PROG-3
'He is living with my niece.'

Analogous data occurs in Tatar with respect to the TAE suffix *-DI*. I show in (166) that *-DI* is required if the speaker has direct (eyewitness) evidence for the scope proposition "Nima caught a fish." In the context in (166), we can reasonably assume that the speaker believes in the truth of the proposition "Nima caught a fish," since she witnessed the event occur. Nonetheless, the speaker must use *-DI* to report the proposition.[12] This would be unexpected if, as Matthewson et al. 2007 claim, evidential expressions (as modals) are incompatible with speaker belief in the truth of the scope proposition.

(166) **Fishing trip context**: You go fishing with your family. You see your son, Nima, catch a fish. Later, you tell your mother that Nima caught a fish. You say:

Nima balıq tot-**tı**-∅.
Nima fish catch-PST-3SG
'[I have direct evidence that] Nima caught a fish.'

Furthermore, the use of the Tatar indirect evidential suffix *-GAn* does not correlate with speaker certainty that *p*. *-GAn* is infelicitous if the speaker has direct evidence for *p*, as shown in (164). However, *-GAn* is felicitous even in contexts in which the speaker is certain that *p* is true, as long as they do not have direct (i.e., sensory) evidence for *p*. I

[12] I attempted to elicit Tatar expressions that were equivalents of the English expressions in (162) (modulo substituting the TAE suffixes for the English epistemic modals). However, my primary consultant judged these expressions as pragmatically odd, since the continuation sounded redundant.

provide a context in (167) in which the speaker is certain that *p* is true, but nonetheless uses *-GAn* to express it.

(167) Germaniya Polşa-nı bas-**qan**-∅.
Germany Poland-ACC invade-RESULT-3SG
'[I have indirect evidence that] Germany invaded Poland.'

The speaker of (167) did not witness the event described by the scope proposition "Germany invaded Poland," since the invasion occurred before they were born. Nonetheless, since this is a historical fact, they can be certain that it occurred. However, since they did not witness *p* occur, they can still felicitously assert this proposition using *-GAn*.[13] The use of *-GAn* therefore does not correlate with speaker certainty that *p*.

Finally, (as noted also by Matthewson 2011, 352), von Fintel and Gillies (2010, 362) give the following example as a counterargument to the claim that speakers cannot assert a proposition under an epistemic modal if they are certain that it is true. In (168), the speaker is able to felicitously state *must*(*p*) despite the fact that she can be certain that *p* is true.[14]

(168) **Context:** Chris has lost her ball, but she knows with full certainty that it is in either Box A or Box B or Box C. She says:
The ball is in A or B or C. It is not in A... it is not in B... so, it **must** be in C.

Given these observations, I propose that this test not does motivate including modality in the semantics of the Tatar TAE suffixes.

[13] Izvorski (1997, 5) similarly notes that the modal component of the Bulgarian indirect evidential appears to have variable force. In its reportative evidential use, it can have a possibility reading; in its inferential use, it makes a much stronger claim. This correlates with cross-linguistic observations made by AnderBois (2014).

[14] von Fintel and Gillies 2010 use the data in (168) to argue for an analysis of *must* in which it is infelicitous if the speaker's evidence for *p* is direct, i.e., the use of *must* is governed by evidence type and is independent from the speaker's certainty.

4.2.3 Recap

I reviewed a number of tests in the evidential literature that have been proposed to diagnose or support analyses of evidentials as epistemic modals. The Tatar TAE suffixes explicitly fail two of these tests. As discussed in §4.2.2.1, the evidential contributions of the (past oriented) Tatar TAE suffixes need not be embedded; furthermore, the evidential readings of the TAE suffixes disappear when they occur in embedded verbal nominalizations. This would be unexpected if their evidential meaning was part of their propositional content. Furthermore, as discussed in §4.2.2.5, speakers can felicitously use the TAE suffixes to report propositions that they believe are true.

I proposed (following prior observations made by e.g. Murray 2017, Korotkova 2016, among others) that the other tests reviewed in this section do not necessarily diagnose modality in a given evidential expression; rather, they are simply **compatible** with a modal analysis. Recent work on this topic may reflect an ongoing shift away from the "dichotomy view" of evidentiality in which evidentials are assumed to be (only) either epistemic modals or speech act operators. In the following section, I review a number of relatively recent theories of evidentiality that also depart from this "dichotomy" view and derive evidentiality as a byproduct of temporal meaning.

I propose, based on the tests reviewed in this section, that the Tatar TAE suffixes do not require an epistemic modal analysis. Furthermore, a "plain" epistemic modal analysis (along the lines of Izvorski 1997) cannot account for why the Tatar TAE suffixes also contribute temporal meaning. In the following section, I evaluate the Tatar data against some theories that account for data in which evidentiality and tense/aspect are morphosyntactically linked.

4.3 Existing analyses of evidential systems that are morphosyntactically related to tense/aspect

In the following sections, I review theories of evidentiality addressing data in which evidentiality and tense/aspect are morphosyntactically linked. These theories all share the property of not explicitly encoding evidentiality as part of semantic meaning. Instead, they propose that evidential readings fall out as a byproduct of the relationships between certain times or events. In this sense, these theories all theoretically link evidentiality and tense/aspect.

I define two main classes of these theories; although their formal implementation differs, the two classes are conceptually similar. I review the first class, which I term **Evidence Acquisition Time** analyses of evidentiality, in §4.3.1 (Lee 2013, 2011; Smirnova 2013, 2011; Koev 2011). These theories posit a novel Evidence Acquisition Time (or Learning Event) that evidentials can manipulate. I review the second, smaller, class, which I term **trace** analyses of evidentiality, in §4.3.3 (Faller 2004, Koev 2017).

I show that while both classes of theories are theoretically compelling, they cannot account for the observed Tatar data. More specifically, they cannot (i) account for the evidential readings of the future oriented TAE suffixes; and (ii) explain why tense/aspect and evidentiality are fused in Tatar; i.e., why these meanings do not occur as separate elements in the syntax. That is, these theories do not answer Questions 1 and 2 from §4.1. Furthermore, I show in §4.3.1.3 that Tatar fails the primary test used by Koev (2011, 2017) to demonstrate sensitivity to Learning Events in the grammar. These issues are not shortcomings for the EAT and trace theories in and of themselves; the purpose of this section is simply to show that they are not extendable to Tatar.

4.3.1 Evidence Acquisition Time analyses of evidentiality

Evidence Acquisition Time analyses of evidentiality are based on the intuitively compelling observation that the relationship between the time of the event described by a proposition

p and the time that a speaker acquired evidence that p can affect the type of evidence that the speaker has for p (following intuitions from Woodbury 1986, Nikolaeva 1999, and Fleck 2007, among others).[15] More specifically, if a speaker acquired evidence that p or learned that p after the time of the event described by p, then their evidence for p must be indirect. Conversely, if the time that they acquired evidence that p or learned evidence that p overlaps with the time of the event described by p, then they could potentially have direct evidence for p. This can be represented graphically as in Figure 4.1 and Figure 4.2, respectively.[16]

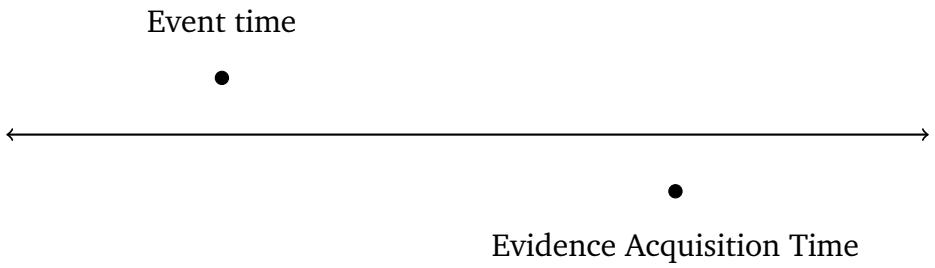


Figure 4.1: No overlap of the Evidence Acquisition Time and the event time; speaker must have indirect evidence for the event.

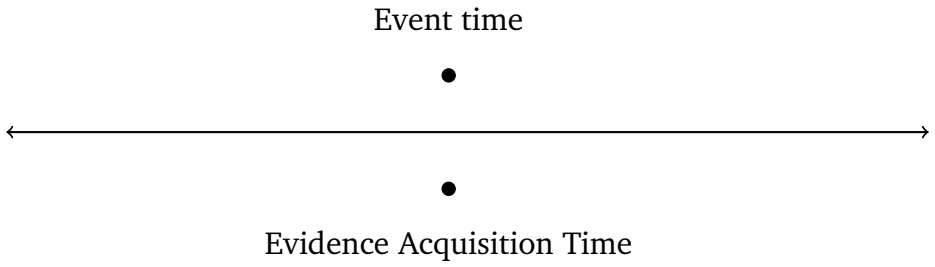


Figure 4.2: Overlap of the Evidence Acquisition Time and the event time; speaker may have direct evidence for the event.

[15]There are a number of descriptive papers that informally account for evidential data using concepts in line with EAT theories, including Woodbury (1986) for Sherpa (Tibeto-Burman) and Fleck (2007) for Matses (Panoan). Since these sources do not provide formal semantic proposals for their data, I do not discuss them here. I acknowledge, however, that formal EAT theories of evidentiality are indebted to these preliminary descriptive observations.

[16]For the purpose of these simplified diagrams, I represent event times as points.

These analyses typically involve manipulating the theoretical ingredients in (169); terminology varies slightly between authors.

(169) a. ET = Event Time

b. RT = Reference Time

c. UT = Utterance Time

d. EAT = Evidence Acquisition Time; i.e., the time that the speaker acquired evidence that *p* (Lee 2013, 2011; Smirnova 2013, 2011)

e. LE = Learning Event; i.e., the event in which the speaker learned that *p* (Koev 2017, 2011)

I will now evaluate the Tatar TAE data against each of these theories in turn.

4.3.1.1 Lee (2013, 2011): Korean

Lee (2011) bases her EAT analysis of evidentiality on previously unformalized intuitions by Woodbury (1986) and Nikolaeva (1999), among others. Her analysis addresses the Korean data in (170).

Evidentiality is optionally marked in Korean using the verbal suffix *-te*. In these examples, the Korean evidential suffix *-te* co-occurs with past, present, and future tense morphemes. When *-te* co-occurs with the (phonologically null) present tense, it requires that the speaker have direct (i.e., sensory) evidence for the scope proposition, as in (170a). However, when *-te* co-occurs with the past or future tenses, it requires that the speaker have indirect evidence for the scope proposition, as in (170b)-(170c).

(170) KOREAN (KOREANIC) (Lee 2013, 2)

a. **Context**: Yenghi saw it raining yesterday. Now, she says:

Pi-ka o-∅-**te**-la.
rain-NOM fall-PRES-TE-DECL
'[I made a sensory observation that] it was raining.'

b. **Context**: Yenghi saw yesterday that the ground was wet. Now, she says:

Pi-ka o-**ass-te**-la.
rain-NOM fall-PST-TE-DECL
'[I inferred (from the acquired sensory evidence) that] it had rained.'

c. **Context**: Yenghi saw the overcast sky yesterday. Now, she says:

Pi-ka o-**kyess-te**-la.
rain-NOM fall-FUT-TE-DECL
'[I inferred (from the acquired sensory evidence) that] it would rain.'

Lee (2011, 2013) argues that all *-te* expressions encode that the speaker has some sensory evidence for the scope proposition. She argues against explicitly encoding the evidence type that *-te* requires (i.e., direct versus indirect) in its semantics, contrary to other analyses of evidentials in which their evidence type is hardwired into their meaning (e.g. Izvorski 1997, Faller 2002). Rather, the evidential reading of *-te* falls out from its compositional interaction with the tense morphemes.

Lee posits a modal analysis of *-te*, which she motivates through a series of tests that she argues support treating *-te* as a necessity modal. These tests include (i) the infelicity of *-te*(p), $\neg p$ expressions; (ii) the embeddability of *-te* under verbs of saying; and (iii) modal subordination of *-te* expressions. I discuss in §4.2.2 why I find that these tests do not necessarily diagnose modality in a given evidential expression. Although the behavior of *-te* with respect to these tests does not argue against a modal analysis, (some of) its distribution could also be explained by other accounts of evidentiality, e.g. Faller (2002)'s illocutionary analysis of evidentials.[17]

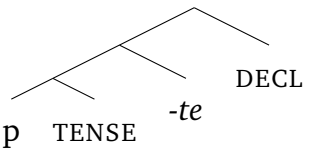
Lee proposes that *-te* combines above tense (past, present, or future) and below DECL, a

[17]An additional test that Lee (2011, 2013) gives (and which I do not discuss in §4.2.2) is what she terms the "non-equi subject constraint," i.e., the pragmatic restriction that evidential expressions are typically infelicitous if the subject is the speaker. (I discuss the relevant Tatar data in Chapter 2, and show how such expressions are only licensed in blackout contexts.)

While this data is compatible with a modal analysis of evidentiality, it again does not necessarily diagnose it. For instance, under Faller (2002)'s illocutionary analysis of evidentiality, evidentials can weaken the illocutionary force of expressions from e.g. ASSERT to PRESENT. It seems to me that the infelicity of first person subjects in combination with evidential expressions could also be derived from an illocutionary account along these lines, assuming that speakers (i) are authorities on their actions and (ii) make the strongest claims possible.

declarative morpheme that shifts the world of evaluation to the actual world and the time of evaluation to the utterance time. I give a basic tree according to Lee (2011, 2013) in (171) (Korean has a left-branching syntax).

(171)



Lee (2013, 29) gives a denotation for *-te* as in (172). This denotation has two components, one temporal and one modal. The temporal component of the denotation asserts that the EAT i'' precedes i; in unembedded contexts, i is the utterance time (valued by DECL). The modal component of the denotation asserts that in all the worlds in the speaker's sensory modal base (SO), as ordered by a stereotypical/doxastic ordering source (ST/DX), p is true at the EAT in those worlds.

(172) $[\![\text{-}te]\!] = \lambda p_{<s,<i,t>>} \lambda w \lambda i.\ [i''<i\ \&\ \forall w'[w' \in \text{BEST}(SO, ST/DX, w, i'') \rightarrow p(w')(i'')]]$

According to this analysis, the Korean tense morphemes and *-te* share the same semantic type: $<<s,<i,t>>,<s,<i,t>>>$. However, *-te* is an absolute tense (placing the EAT relative to the utterance time), whereas (in expressions with *-te*) the other Korean tense morphemes are relative tenses (placing the event time relative to the EAT). I summarize the temporal contribution of the various tense morphemes in Lee's analysis in (173)-(174).

(173) **Evidential**

a. *-te*: EAT $<$ UT

(174) **Tenses** (when co-occurring with *-te*)

a. PAST: ET $<$ EAT

b. PRESENT: EAT $\leq$ ET

c. FUTURE: EAT $<$ ET

I give the compositional semantics of two *-te* expressions in (175)-(176), omitting the declarative morpheme (which values the open world and time arguments). In (175), *-te*

co-occurs with the phonologically null present tense morpheme. As a result, the EAT overlaps with the ET. This permits the speaker to have direct evidence for the scope proposition. In (176), *-te* co-occurs with the past tense morpheme *-ass*. This places the ET prior to the EAT. The speaker therefore cannot have direct sensory evidence for the event described by the scope proposition. I indicate the temporal contribution of the tense morphemes in blue and the temporal contribution of *-te* in orange.[18] In these denotations, i'' denotes the EAT, i' denotes the ET, and i (eventually valued by DECL) denotes the UT.

(175) a. pi-ka o-∅-te
rain-NOM fall-PRES-TE
'[I made a sensory observation that] it was raining.'

b. $[\![$*pi-ka o-∅-te*$]\!] =$

$\lambda w \lambda i.\ \mathbf{i''} < \mathbf{i}\ \&\ \forall w'[w' \in \text{BEST}(SO, ST/DX, w, i'') \rightarrow \mathbf{i''} \leq \mathbf{i'}\ \&\ \text{AT}(i', w', \text{rain})]$

(176) a. pi-ka o-ass-te
rain-NOM fall-PST-TE
'[I inferred that] it had rained.'

b. $[\![$*pi-ka o-ass-te*$]\!] =$

$\lambda w \lambda i.\ \mathbf{i''} < \mathbf{i}\ \&\ \forall w'[w' \in \text{BEST}(SO, ST/DX, w, i'') \rightarrow \mathbf{i'} < \mathbf{i''}\ \&\ \text{AT}(i', w', \text{rain})]$

I proposed in Chapter 2 that, descriptively speaking, both past and future Tatar TAE suffixes make evidential contributions. In future-oriented Tatar expressions, the contrast between *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' is linked to the speaker's evidence for the scope proposition p. If the speaker has specific evidence that p will occur—i.e., if they have evidence that that they are in a pre-state of p—then they will use *-(y)AçAK*. Otherwise, speakers will use *-(y)Er*. Lee (2013)'s proposal for *-te* in combination with the future tense can not easily account for this observed Tatar future-oriented data.

Lee (2013) provides Korean data in which *-te* co-occurs with *-kyess* 'FUT.' These examples appear to be interpreted similarly to Tatar *-(y)AçAK(p)* expressions. That is, the kinds

[18] The AT predicate (adopted from Condoravdi 2002) takes a property of eventualities (of type $<s,<i,t>>$), a world argument, and a time argument. Informally speaking, AT*(t, w, P)* means that the property of eventualities P is instantiated in w at t (Lee 2013, 25).

of contexts that Lee describes as licensing Korean *-kyess-te* expressions are analogous to the "specific" evidential contexts that I described as licensing Tatar *-(y)AçAK(p)* expressions in Chapter 2.[19]

(177) a. **Context**: Yenghi found curry powder with sliced vegetables and meat in Chelswu's kitchen yesterday. Now, she says: (Lee 2013, 8)

Chelswu-ka khaley-lul mantul-**kyess-te**-la.
Chelswu-NOM curry-ACC make-FUT-TE-DECL
'[I inferred (from the acquired sensory evidence) that] Chelswu would make curry.'

b. **Context**: The speaker knows that Yenghi's favorite food is curry, and she usually cooks it for parties. The speaker was invited to Yenghi's party tomorrow. Now, he says: (Lee 2013, 6)

#Yenghi-ka nayil phathi-ey khaley-lul mantul-**kyess-te**-la.
Yenghi-NOM tomorrow party-at curry-ACC make-FUT-TE-DECL
Intended: '[I inferred (from reasoning based on previous experience) that] Yenghi would make curry for the party tomorrow.'

Lee (2011, 2013) argues that future tense expressions with *-te* necessarily indicate indirect evidence for the scope proposition, since the EAT and the ET cannot overlap. However, this requires being able to distinguish between contexts like (177a) (in which *-te* is felicitous) and (177b) (in which *-te* is infelicitous). Lee proposes to distinguish between these contexts by stipulating that *-te* requires sensory evidence for the scope proposition, which she argues the speaker has in (177a) and lacks in (177b). I don't find this argument compelling, since the speaker could be argued to have sensory evidence for the scope proposition in both contexts. In (177b), this sensory evidence could include things like "I see Yenghi eat curry often" and "I see Yenghi bring curry to parties." (Lee 2013, 5, footnote

[19] Lee (2013, 2011) consistently translates Korean *-kyess-te* 'FUT-TE' expressions with English *would*, as opposed to *will* (e.g. (170c)). She uses *will* to translate Korean future tense marked expressions without *-te* (e.g. Lee 2013, 21). It is unclear to me if the choice of *would* is intended to suggest a counterfactual reading of *-kyess-te* expressions or otherwise indicate a difference in modality between the two expressions.

5 also acknowledges the pitfalls to this approach.)[20]

Finally, Lee (2011, 2013)'s Korean evidential data differs descriptively from the Tatar evidential data in Chapter 2 in two major ways. First, the Korean evidential morpheme *-te* is optional in any given utterance, and does not form part of any grammatical paradigm. This means that some Korean utterances are arguably evidentially "neutral." However, since tense/aspect is obligatorily marked in Tatar, Tatar past- and future-oriented expressions necessarily receive an evidential interpretation. Second, evidentiality and tense are marked by two separate morphemes in Korean (and, as Lee 2011, 2013 argues, two separate heads in the syntax).[21] However, evidentiality and tense/aspect are marked by a single morpheme in Tatar. I therefore have no morphosyntactic motivation to decompose the Tatar TAE morphemes into separate syntactic tense/aspect and evidential heads, unlike the Korean data addressed by Lee (2013, 2011).

4.3.1.2 Smirnova (2013, 2011): Bulgarian

Smirnova (2013) is concerned with the basic Bulgarian data in (179). She argues that participial verbs in Bulgarian (henceforth, "evidential verb forms") are compatible with a range of evidential meanings, as shown in (179).

In (179a), the evidential verb form expresses that the speaker has either reportative or inferential evidence for the scope proposition. In (179b), the evidential verb form expresses that the speaker has direct evidence for the scope proposition. The only difference

[20] In Chapter 5, I propose to distinguish between contexts like (177a) and (177b) on the basis of whether or not the speaker locates themselves in a causally related preparatory state of the event described by the scope proposition. (I formally define these event pre-states in Chapter 5.) However, my analysis as it currently stands is only intended to apply to the Tatar data, not the Korean data from Lee (2013, 2011).

[21] The evidential interpretation of *-te* also appears to be sensitive to other compositional factors; for instance, if it is both preceded and followed by a declarative morpheme, then it requires a reportative evidential context (Lee 2013, 3):

(178) Context: The speaker talked to Chelswu on the phone yesterday. Now, she says:
(Chelswu-ka) pi-ka o-**ass-ta-te-la**.
Chelswu-NOM rain-NOM fall-PST-DECL-TE-DECL
'[I was told (from Chelswu) that] it had rained.'

There is no way to force a reportative reading of a Tatar expression through functional morphology.

between the expressions in (179a) and (179b) is that the latter is necessarily pronounced with exclamative intonation. (Other authors treat expressions like (179b) as exceptional mirative uses of the Bulgarian evidential verb form, e.g. Koev 2017. Since (179b) necessarily conveys speaker surprise, I am also of the opinion that (179b) should be treated as a mirative.)

(179) BULGARIAN (SLAVIC) (Smirnova 2013, 480-481)

a. i. **Reportative evidence context**: You and your sister were out of touch for a couple of years. Today she calls you on the phone to catch up. She tells you that her daughter Maria plays the piano. Later, you tell your husband:

ii. **Inferential evidence context**: You and your sister were out of touch for a couple of years. Today you visit her for the first time. As she shows you around her apartment, you see that there is a piano in her daughter Maria's room. You infer that Maria plays the piano. Later, you tell your husband:

Maria **svirela** na piano.
Maria play.IMPERF.$_{PRES}$.PLE on piano
'Maria plays the piano, [I inferred].'

b. **Direct evidence context:** Your niece Maria stays with you over the summer. She usually spends most of the time reading in her room. While passing by Maria's room today, you see her playing the piano. You say to yourself:

Maria **svirela** na piano!
Maria play.IMPERF.$_{PRES}$.PLE on piano
'Maria plays the piano, [I see].'

Like Lee (2011, 2013), Smirnova (2011, 2013) analyzes the Bulgarian evidential as a epistemic necessity modal. She motivates a modal analysis of the evidential through a number of tests for modality; these are (i) modal subordination of evidential expressions, (ii) inability to assert EVID(p) if the speaker believes that p is true, and (iii) inability to assert EVID(p) if the speaker believes that p is false. (Smirnova 2013, 508 notes that the latter test, while compatible with modal analyses of evidentiality, does not definitively

diagnose modality.)[22] Again, I discussed in §4.2.2 why I do not believe that these tests adequately motivate a modal analysis of evidentiality.

Furthermore, the following Bulgarian data is problematic for a modal analysis. The context in (181) (and in (179b)) is such that the speaker has direct (sensory) evidence for the scope proposition *p*. In English, epistemic necessity modals are infelicitous if the speaker has direct evidence for *p* (von Fintel and Gillies 2010), as shown in (181b). Nonetheless, the Bulgarian evidential verb form is felicitous in a context in which the speaker has direct evidence for *p*. (This is a mirative use of the evidential.)

(181) **Direct evidence context**: You visit your brother, who works on a church restoration in a remote Bulgarian village. At midday, you hear the church bell ring. You say: (Smirnova 2013, 486)

a. Tja kombanata **zvănjala**!
it bell ring.IMPERF$_{PRES}$.PLE
'The bell is ringing, [I hear].'

b. # The bell must be ringing!

Setting aside these issues with the modal component of Smirnova's proposal, I will outline her analysis. Smirnova (2013) provides an analysis of Bulgarian evidential expressions that derives their meaning from a number of components: aspect, tense, an evidential operator, and an exclamative/declarative operator. The basic syntax is as in (182).

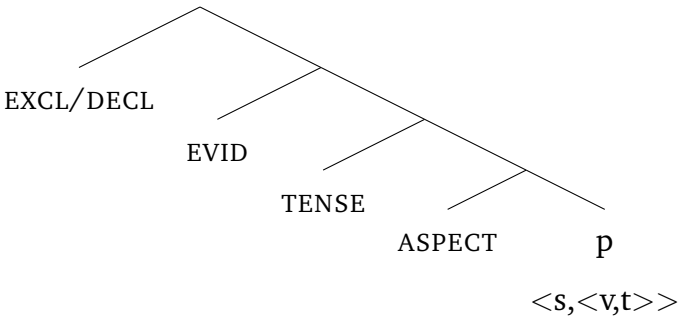
---

[22] In addition to these tests, Smirnova (2011, 279-280) provides the data in (180) as further support for a modal analysis. She argues that (180) shows that the speaker's belief worlds, and not the actual world, are relevant for the interpretation of the scope proposition. This is shown by the fact that Ivan's report is felicitous despite the fact that Maria is not writing a book in the actual world.

(180) Inferential evidence context: After your aunt Maria died, Ivan found a first chapter of an unauthored manuscript in her apartment. He inferred that Maria was writing a book. You know that it was Maria's sister who was writing the book, not Maria. When someone asks Ivan what Maria was doing before she died, he says:
Maria **pisala** kniga.
Maria write.IMPERF$_{PAST}$.PLE book
'Maria was writing a book, [I inferred].'

I find this data unconvincing with respect to a modal analysis since for all Ivan knows, he **is** reporting about the actual world. (180) could just be a mistaken report about the actual world as Ivan believes that it is. This does not require a modal analysis of evidentiality.

(182)



Each of these components encodes a relation between times. I summarize the contributions of each of these components, and their semantic types, in Table 4.2.

Component	**Temporal relation encoded**	**Semantic type**
EXCL	EAT = UT	$\langle\langle s,\langle i,\langle i,t\rangle\rangle\rangle,\langle s,\langle i,\langle i,t\rangle\rangle\rangle\rangle$
DECL	EAT $<$ UT	
EVID	EAT $\leq$ UT	$\langle\langle s,\langle i,t\rangle\rangle,\langle s,\langle i,\langle i,t\rangle\rangle\rangle\rangle$
Tense: PST	RT $<$ EAT	$\langle\langle s,\langle i,t\rangle\rangle,\langle s,\langle i,t\rangle\rangle\rangle$
PRES	RT = EAT	
FUT	EAT $<$ RT	
Aspect: IMPF	RT $\subseteq$ ET	$\langle\langle s,\langle v,t\rangle\rangle,\langle s\langle i,t\rangle\rangle\rangle$
PERF	ET $\subset$ RT	

Table 4.2: Summary of semantic components of Smirnova (2013)'s analysis of Bulgarian evidentiality.

Smirnova (2013, 516)'s evidential operator in (183) is similar in spirit to Lee (2011, 2013)'s semantics for Korean *-te* in that it has both a temporal and a modal component in its semantics. The temporal component of this evidential asserts that the EAT (i'') precedes or overlaps with the UT (i). The modal component is interpreted relative to an epistemic agent α; in direct or inferential evidence contexts, α is the is the speaker, whereas in reportative contexts, α is the reporter.[23] The evidential asserts that in all of α's doxastically accessible world-time pairs $\langle w',i'''\rangle$ (where i''' is α's "now"), the scope proposition Q is true in w' at i'''. The evidential is additionally subject to a definedness condition that the speaker have "external evidence" for p; according to Smirnova, "external evidence" contains reports, evidence-based inferences, and direct perception of p.

[23] Smirnova (2013) uses this component of the semantics to account for the fact that Bulgarian EVID(p), $\neg p$ assertions are felicitous in reportative evidential contexts.

(183) $[\![\text{EVID}]\!] = \lambda Q_{<s,<i,t>>} \lambda w \lambda i \lambda i''.\ i'' \leq i\ \&\ \forall <w',i'''>[<w',i'''> \in MB_{DOX\alpha}(w,i'') \rightarrow Q(w')(i''')]$

Where i = UT, i'' = EAT, and i''' = α's "now"

Like Lee (2011, 2013), Smirnova (2011, 2013) links the availability of direct evidence for the scope proposition p to overlap of the EAT and the ET. If the EAT and ET do not overlap, then the speaker cannot have direct evidence for p. To illustrate this, I give Smirnova's denotation for an inferential evidential Bulgarian expression in (184).

In (184), the speaker has indirect evidence for the scope proposition "Maria was writing a book." Smirnova (2013) derives this evidential reading through a combination of the components in Table 4.2. I label the temporal contribution of aspect in purple, tense in blue, the evidential operator in orange, and the declarative operator in red.

(184) a. **Inferential evidence context**: You are writing a biography of your late aunt Maria. After finding an unauthored manuscript in her apartment, you inferred that Maria was writing a book. When one of the relatives starts to wonder what Maria did before she died, you say: (Smirnova 2013, 519)

Maria **pisala** kniga.
Maria write.IMPERF$_{PST}$.PLE book
'Maria was writing a book, [I inferred].'

b. $[\![\text{DECL } \textit{Maria pisala kniga}]\!] =$

$\lambda w \lambda i \lambda i''.$ $\mathbf{i''} \leq \mathbf{i}$ & $\forall <w',i'''>[<w',i'''> \in MB_{DOX\alpha}(w,i'') \rightarrow$

$\exists i'[\exists e[\text{write.book}(w')(e)(\text{Maria})$ & $\mathbf{i'} \subseteq \tau\mathbf{(e)(w')}$ & $\mathbf{i'} < \mathbf{i'''}]]$ & $\mathbf{i''} < \mathbf{i}]$

The evidential operator in (184b) contributes that the EAT (i'') precedes or overlaps with the UT (i); this is further specified by the declarative operator, which asserts that the EAT strictly precedes the UT. The past tense places the RT (i') in the past with respect to the speaker's "now" (i'''), which contributes the overall past tense reading of the expression. Finally, imperfective aspect asserts that the RT is contained within the runtime (i.e., the temporal duration) of the event.[24]

[24] Smirnova (2013, 519-520) proposes that the open world and time arguments in (184b) are eventually

To account for the fact that the event of Maria writing a book in (184) is located in the past of the UT, Smirnova (2013, 520) crucially assumes that the speaker's "now" (i''') is structurally equivalent to the EAT (i''). (This assumption does not permit the speaker to be mistaken about the time of their "now.") If we assume that these times are equivalent, we can derive that $i' < i''$ and $i''' < i$, where $i'' = i'''$. That is, the ET $<$ EAT, and the speaker's "now" $<$ UT, where the EAT and the speaker's "now" are the same. Since the ET and EAT do not overlap, the speaker must have indirect evidence for the scope proposition.

The Tatar data differs from Smirnova (2011, 2013)'s Bulgarian data in a number of ways. First, as noted in Chapter 2, the Tatar resultative aspect suffix *-GAn* is (synchronically) infelicitous in mirative contexts. This differs from Bulgarian evidential verb forms, which Smirnova (2011, 2013) shows are felicitous in mirative contexts like (181). Smirnova uses this mirative data to motivate her denotations of the Bulgarian exclamative and declarative operators; however, since no such mirative use exists synchronically in Tatar, this analysis is not applicable to the Tatar data.

Furthermore, Smirnova (2011, 2013)'s proposal is not applicable to the future-oriented Tatar TAE morphemes. In Smirnova (2011, 277-278), she argues that in Bulgarian, future tense evidential verb forms are compatible only with reportative evidence for the scope proposition, and not inferential evidence. I show relevant Bulgarian data in (185)-(186).[25]

(185) **Reportative future tense context**: In the morning Ivan told you that Maria will spend the evening writing a portion of her book. When in the afternoon your friend asks you what Maria will do in the evening, you say:

Maria **štjala** da piše kniga.
Maria FUT.PLE SUBJ write.IMPERF.3SG.PRES book
'Maria will be writing a book, [I heard].'

valued by the utterance time (i) and existential closure (w, i'').

[25] Smirnova (2011, 278) states that this contrast is attested "not only in Bulgarian but in all Balkan languages that grammatically express evidentiality."

(186) **Inferential future tense context**: By looking at your roommate Maria's schedule this morning you inferred that she plans to spend the evening writing a portion of her book. In the afternoon, when your friend asks you what Maria will do in the evening, you say:

#Maria **štjala** da piše kniga.
Maria FUT.PLE SUBJ write.IMPERF.3SG.PRES book
Intended: 'Maria will be writing a book, [I inferred].'

Smirnova (2011, 288-290) argues that the infelicity of examples like (186) is due to pragmatics. She proposes that when a speaker says (186), they assert that they believed at some past time, prior to the UT, that the event of Maria writing a book would occur at some future time. This assertion of a past belief regarding *p* is weaker than an assertion of a current belief regarding *p*. As a result, (186) is infelicitous due to violating the Maxim of Quantity. However, in reportative contexts, the modal is interpreted relative to the beliefs of the reporter, and not the speaker. Asserting a third party's past belief regarding *p* no longer violates the Maxim of Quantity; the reportative use of the evidential in (185) is therefore felicitous.

I showed in Chapter 2 that the future-oriented Tatar TAE suffixes are both compatible with reportative as well as inferential evidence for the scope proposition. In contexts in which the speaker has reportative evidence, their choice of *-(y)Er* 'FUT' versus *-(y)AçAK* 'PROSP' is dictated by whether or not they take the report as evidence that they are in a pre-state of the described event. This provides additional evidence that I do not need to incorporate Smirnova (2013, 2011)'s modal analysis into my proposal for the Tatar TAE data; namely, I do not need to incorporate any shifting of the epistemic agent into my semantics.

Finally, as noted similarly for Lee (2013, 2011)'s analysis in §4.3.1.1, I see no morphosyntactic motivation to decompose the Tatar TAE morphemes into separate tense, aspect, evidential, and declarative/exclamative heads. Tatar patterns differently from Bulgarian in that tense/aspect and evidentiality are not morphosyntactically decomposable into separate morphemes. I therefore do not adopt Smirnova (2013, 2011)'s analysis for

Tatar.

4.3.1.3 Koev (2011): Bulgarian

Koev (2011) provides a non-modal analysis of roughly the same Bulgarian evidential data as Smirnova (2011).[26] I show the basic data that Koev is concerned with in (187); he links the evidential contrast between the two expressions with the use of a phonologically null direct evidential suffix in (187a) and the reduced present perfect suffix *-l* (glossed IND) in (187b).

(187) BULGARIAN (Koev 2011, 115-116)

a. Ivan celu-n-a-∅ Maria.
Ivan kiss-PFV-3SG.PST-DIR Maria
'[I saw that] Ivan kissed Maria.'

b. Ivan celu-n-a-**l** Maria.
Ivan kiss-PFV-3SG.PST-IND Maria
'[I have indirect evidence that] Ivan kissed Maria.'

For the purpose of this chapter, I will not recreate Koev's motivation for a non-modal analysis of the Bulgarian data, since I already suggested in §4.2 that the Tatar TAE suffixes need not be analyzed as modal. Instead, I focus on the temporal component of his analysis.

Koev (2011) argues that the Bulgarian evidential morphemes are underlyingly tense morphemes that denote a relationship between the Learning Event (LE) (i.e., the time at which the speaker learned the scope proposition) and RT. He proposes (in this spirit of Lee 2011's analysis and intuitions from Nikolaeva 1999 and Fleck 2007, among others) that the indirect evidential suffix marks that the speaker learned the scope proposition after the RT. Conversely, the direct evidential marks that the speaker learned the scope proposition during or before the RT.[27]

[26] Klose (2014) assumes Koev (2011)'s analysis to account for a portmanteau past tense and indirect evidential morpheme in Aymara. For the purpose of this section, I solely describe Koev's analysis and acknowledge that it is also used by Klose to account for similar Aymara data.

[27] Koev (2011) motivates his proposal that the direct evidential is compatible with an EAT prior to the RT

With respect to the data in (187): when discussing the events of a party last night, the use of the direct evidential in (187a) signals that the speaker learned that Ivan kissed Maria during the party (and could therefore have direct evidence for the proposition). Conversely, the use of the indirect evidential in (187b) signals that the speaker learned that Ivan kissed Maria after the party was over (and therefore must have indirect evidence for the proposition). These secondary tense morphemes co-occur with (primary) tense marking, which indicates the relationship between the RT and UT. In both expressions in (187), the primary tense marking is past tense. I summarize the semantic components of Koev (2011)'s analysis in Table 4.3.

Koev (2011)'s formal account is couched in Dynamic Predicate Logic (AnderBois et al. 2010, Groenendijk and Stokhof 1991), in which the (At-Issue) assertion is a proposal to update the context set of propositions shared by speaker and addressee, whereas backgrounded content (e.g. appositives) is directly imposed on the context set. Koev argues that in Bulgarian expressions like (187), the contribution of the evidential (i.e., the relation between RT and EAT) is directly added to the context set as backgrounded content.

Component		**Temporal relation encoded**
Evidential:	IND	RT $<$ EAT
	DIR	EAT $\leq$ RT
Tense:	PST	RT $<$ UT
	PRES	RT $=$ UT
	FUT	UT $<$ RT
Assumption		EAT $<$ UT

Table 4.3: Summary of semantic components of Koev (2011)'s analysis of Bulgarian evidentiality.

with the example in (188). He argues that in this example, the EAT precedes the RT for the scope proposition and the direct evidential reading is nonetheless licensed.

(188) Context: Two days ago your colleague Ivan calls you and tells you that he is very sick and will skip work the next day. As expected, yesterday he does not show up for work. Today you say: (Koev 2011, 126)

Včera Ivan {beš-e-∅ / #bi-**l**} bolen.
yesterday Ivan be.PST-3SG-DIR / be.PST-IND sick
'Yesterday Ivan was sick [as I learned before that].'

I find that Tatar patterns very differently from Bulgarian with respect to some of the predictions that Koev (2011)'s analysis makes. For example, Koev provides the following "delayed learning time" example in (189) in support of his proposal to encode a LE in the semantics.

(189) **Delayed learning time context:** One of Nixon's aides vividly recalls walking into the President's office and seeing the President erase some tapes. A few months later she learns about the Watergate scandal from the newspapers and makes sense of what she has seen. When asked what happened on that day, she says: (Koev 2011, 125)

a. Kogato vljaz-ox-∅, Niskyn tri-e-še-∅ njakav-i zapis-i.
when enter-1SG.PST-DIR Nixon erase-3SG-PST-DIR some-PL tape-PL
'When I walked in, [I saw] Nixon erase some tapes.'

b. A toj {#zalič-ava-še-∅ / zalič-ava-**l**} ulik-i-te.
but he remove-3SG-PST-DIR / remove-3SG-IND clue-PL-DEF.PL
'He was covering up the clues [as I learned later].'

In the context in (189), the time at which the speaker learned the proposition "He was covering up the clues" follows the time at which the event described by this proposition occurred; although the speaker witnessed *p* occur, they did not realize what was happening at the time. The speaker then uses the indirect evidential to report this proposition in (189b). Koev (2011) argues that this falls out from the fact that the speaker's LE does not precede or overlap with the ET; as a result, the speaker must use the indirect evidential.

However, analogous data is not available in Tatar. In the modified delayed learning time context in (190), my primary Tatar consultant reports that the use of the TAE suffix *-GAn* is infelicitous, since the speaker nonetheless had eyewitness evidence for the scope proposition. This suggests that inclusion of Koev (2011)'s LE in the semantics, while appropriate for Bulgarian, is not necessarily motivated for the Tatar data.

(190) **Delayed learning time context** (adapted from Koev 2011, 125): One of Nixon's aides vividly recalls walking into the President's office and seeing the President messing around with the tape recorder, although she didn't know what he was doing at the time. A few months later she learns about the Watergate scandal from the newspapers and makes sense of what she has seen. She realizes that she saw Nixon erase some tapes. When asked later by a reporter what happened that day, she says:

Nikson magnitafon tasma-lar-ın {boz-**dı**-∅ / #boz-**ğan**-∅}.
Nixon tape.player tape-PL-ACC destroy-PST-3SG / destroy-RESULT-3SG
'[I have direct/#indirect evidence that] Nixon destroyed some tapes.'

(Speaker's comment: "[*bozğan* is unacceptable] because she witnessed it. Even though she didn't realize it at the time, it's still OK for her to use *bozdı*.")

The future oriented Tatar TAE morphemes also fall outside the scope of Koev (2011)'s analysis. I noted in Chapter 2 that the future oriented Tatar TAE suffixes *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' appear to contrast with respect to the speaker's evidence for the scope proposition. Descriptively speaking, if the speaker has "specific" evidence that *p* will occur, then they will use *-(y)AçAK*; otherwise, they will use *-(y)Er*. In both cases, the speaker has acquired some evidence that *p* will occur in the future: that is, their LE for *p* precedes the RT.

According to Koev's analysis, if the LE for a given proposition precedes the RT, only a direct evidential interpretation should be available (as shown in Table 4.3). This is not the case for the Tatar data; indeed, it is unclear what it would mean for a speaker to have "direct" evidence for an event that has not yet occurred. Koev (2011)'s proposal is therefore unable to account for the "specific"/"non-specific" evidential distinction between the two future-oriented Tatar TAE suffixes.[28] Furthermore (as noted previously with respect to

[28]Koev (2011, 129-131) acknowledges this issue with his analysis, and notes that the Bulgarian indirect evidential *-l* is also available in future tense expressions. He tentatively proposes to revise his analysis such that Bulgarian evidentials can only relate LEs to past RTs; when the event described by the scope proposition *p* is non-past, the RT that the evidential targets would then be located within some salient stage of the pre-state of the event described by *p*. In my analysis in Chapter 5, I rely similarly on the inclusion of preparatory processes or states in the ontology of events.

both Lee 2013, 2011 and Smirnova 2013, 2011), I see no morphosyntactic motivation in Tatar to decompose the Tatar TAE suffixes into separate evidential and tense/aspect heads.

4.3.2 Recap

The EAT theories represent an interesting step forward with respect to the range of possible analyses of evidential data. The intuitive link between the time of the described event, the time at which the speaker acquired evidence for the event, and the kinds of evidence that the speaker can then have for the event is compelling. My analysis is indebted to EAT analyses of evidentiality in that I, too, propose that evidential readings can arise as a byproduct of temporal semantics. However, as I showed in this section, the EAT analyses proposed by Lee (2013, 2011), Smirnova (2013, 2011), and Koev (2011) are not applicable to the Tatar data.

I additionally find that introducing a separate EAT/LE into the semantics is perhaps not linguistically well-motivated. For instance, we do not find any functional morphology that picks out the EAT/LE.[29] I do not assume an EAT analysis of Tatar primarily due to the empirical issues involved with applying these analyses to the Tatar data, as discussed previously in this section. However, I am also interested in pursuing an analysis of the Tatar TAE data that does not require introducing additional theoretical machinery along the lines of the Evidence Acquisition Time. To my knowledge, EATs have only been posited to account for evidential data by the authors discussed in this section. In my analysis in Chapter 5, I explore the ability to account for the Tatar TAE data using some theoretical tools (namely, event pre- and post-states) that have been previously independently motivated to account for very different data in the literature. These tools bring together the evidential and tense/aspect literatures.

[29] Koev (2011, 126) discusses some English data which he argues could involve linguistic reference to a LE. For instance, he proposes that *was* in (191) targets the time at which the speaker learned that John's blood is A-positive, rather than a past reference time at which John's blood was A-positive.

(191) John's blood was drawn and tested a few weeks ago and his blood type **was** A-positive.

I am not convinced by this argument, since *was* could refer to the past time at which John's blood was drawn and tested. Since testing produces a result, the past time at which the test was performed is also the time at which the speaker learned that John's blood is A-positive.

4.3.3 Trace analyses of evidentiality

Trace analyses of evidentiality (Koev 2017, Matthewson 2011, Chung 2007, Faller 2004) are conceptually very similar to EAT analyses of evidentiality. These analyses rely on the intuition (initially from Nikolaeva 1999) that the "situation" (informally speaking) in which the speaker gained evidence for the scope proposition *p* must overlap with the situation of *p* itself in order for the speaker to have direct evidence for *p*.[30] Like the EAT analyses, trace analyses (i) derive evidential readings through semantics that do not explicitly encode evidential meanings in their denotations, and (ii) formally encode the situation in which the speaker acquired evidence for *p* in the semantics.[31] These two sets of analyses differ in that EAT theories of evidentiality solely pick out the **time** at which the speaker acquired evidence or learned that *p*; conversely, trace theories pick out a larger evidence acquisition **situation** (including e.g. a spatial component). Trace theories of evidentiality differ from EAT theories in that they turn on the notion of spatial displacement between the event described by the scope proposition and the time at which the speaker gained evidence for *p*. (Faller 2004's analysis explicitly refers to "spatiotemporal traces" of events, hence the term "trace analyses.") In contrast, EAT theories turn solely on temporal displacement between the time of the EAT/LE and the time of the event described by *p*.

In the following section, I evaluate the Tatar TAE data against the trace theories of Faller (2004) and Koev (2017). Matthewson (2011) largely assumes the formalism of Faller (2004)'s analysis to account for a sensory-non-visual evidential *lákw7a* in St'át'imcets. (St'át'imcets *lákw7a* does not mark tense/aspect, unlike the Tatar TAE suffixes.) Chung (2007) also assumes major components of Faller (2004)'s analysis. Like Lee (2011, 2013), Chung (2007) is primarily concerned with the semantics of the Korean verbal suffix *-te*. I discussed in §4.3.1.1 how the Korean *-te* data differs significantly from the Tatar TAE data; see also Lee (2011, 303-308) for a detailed critique of Chung (2007)'s analysis for Korean.

[30] These theories do not refer to situations in the terminology of Kratzer (2007).

[31] There are a number of primarily descriptive papers that use ideas in line with trace theories to account for their evidential data; these include Kalsang et al. (2013) for Tibetan (Tibeto-Burman) and Nikolaeva (1999) for Khanty (Uralic). However, since they do not give fully fleshed out semantic proposals for their data, I do not discuss them here.

4.3.3.1 Faller (2004): Cuzco Quechua

Faller (2004) provides the first formalized analysis of evidential data within the suite of EAT/trace theories of evidentiality. Faller's analysis concerns the Cuzco Quechua verbal suffix *-sqa*. Descriptively speaking, *-sqa* appears to contribute two things to an expression: (i) a past tense temporal component and (ii) an indirect evidential component. For a speaker to felicitously assert *-sqa*(p), they either must not have perceived the event described by p, or not remember the event described by p (as in the "blackout" example in (193)). (I follow Faller in separating the evidential contribution of *-sqa* from the scope proposition in the gloss of (192).) *-sqa* appears to contribute both temporal and evidential meanings; Faller (2004) therefore glosses it as a marker of 'non-experienced past' (NX.PST).

(192) CUZCO QUECHUA (Faller 2004, 46)

Para-sha-**sqa**.
rain-PROG-NX.PST
'It was raining.'

EVID: The speaker was told that it was raining/infers that it was raining.

(193) Macha-sqa-s imaymana-ta rimayu-**sqa**-ni.
drink-PP-REP whatever-ACC say-NX.PST-1
'Drunk, I said a lot of things.' (Faller 2004, 75, from Cusihuaman 1976)

Faller argues that *-sqa* is neither an epistemic modal (as in Izvorski 1997) nor an illocutionary operator (as in her previous work on other Cuzco Quechua evidentials; Faller 2002). Rather, she proposes that *-sqa* is a marker of spatiotemporal deixis that indicates that the event described by the expression occurred in the past and was located (at least partially) outside the speaker's perceptual field.

To formalize this account, Faller (2004) posits two spatiotemporal trace functions that refer to times and locations (of semantic type l). Roughly speaking, Faller's proposed denotation for *-sqa* asserts that the spatiotemporal trace of the described event is not contained within the spatiotemporal trace of the speaker's perception. (*-sqa* additionally asserts that the reference time i is in the past of the speaker's now, which accounts for its past tense contribution.)

I give Faller (2004, 70-75)'s semantics for the two spatiotemporal trace functions in (194). The event-trace (e-trace) function in (194a) applies to an event and maps the event onto its space-time coordinates, defined as a set of time-location pairs (<*i,l*>). The memory-perception-trace (mP-trace) function in (194b) applies to a person *sp* and maps that person onto their perceptual field (represented as time-location pairs) for each time in their runtime (i.e., their life); furthermore, the predicate M-PERCEIVE in (194b) is true for a given time-location pair if and only if *sp* currently remembers their perception of that time and location.[32]

(194) a. e-trace(e) = {<i,l> | i $\subseteq \tau$(e) & AT(e,i,l)}

b. mP-trace(sp) = {<i,l> | i $\subseteq \tau$(sp) & M-PERCEIVE(sp,i,l)}

Faller (2004) incorporates the spatiotemporal trace functions in (194) into the following denotation for *-sqa*. (*i* in (195) refers to the first temporal argument that *-sqa* combines with, which Faller assumes is the RT.) The denotation in (195) is compatible with there being no common time-location pairs between the e-trace of the event *e* and the memory-perception-trace of *sp*, thereby deriving the indirect evidential contribution of the morpheme.

(195) $[\![$*-sqa*$]\!]$ = λiλP$_{<v,t>}\lambda$e. P(e) & i < now & $\neg\forall$<i′,l>[i′ $\subseteq$ i & <i′,l> $\in$ e-trace(e) $\rightarrow$ <i′,l> $\in$ mP-trace(sp)]

Where *i* = RT

The Cuzco Quechua data fundamentally differs from Tatar in that the Tatar TAE suffixes make evidential distinctions in the future as well as the past. Faller (2004)'s data only involves a portmanteau evidential and past tense morpheme. Her analysis therefore is not intended to account for evidential distinctions in the future tense, as found in Tatar. If the described event will occur in the future and has not yet occurred at the speaking

[32]Faller defines M-PERCEIVE in this way to account for the availability of *-sqa* in "blackout" examples like (193). The speaker's spatiotemporal trace in this example clearly overlaps with the trace of the event of them being drunk; however, *-sqa* is still licensed when reporting (193) because the speaker does not remember their perceptions at that time.

time, then the speaker's spatiotemporal trace necessarily cannot overlap with it. Faller (2004) would therefore predict that the future oriented Tatar TAE suffixes would only be compatible with the speaker having "indirect evidence" for the scope proposition. Faller (2004)'s analysis therefore cannot account for the future-oriented Tatar TAE data; like Koev (2011)'s account, there would be no way to distinguish between the two different evidential readings associated with Tatar *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP.'

4.3.3.2 Koev (2017): Bulgarian

Koev (2017)'s analysis represents a combination of both EAT and trace analyses of evidentiality, and is formalized using an update semantics that is conceptually similar to Koev (2011).

Koev (2017) is concerned with accounting for roughly the same Bulgarian data as in Koev (2011); his empirical coverage differs from Smirnova (2013) in that he does not attempt to give an analysis for mirative uses of the Bulgarian evidential. (As noted previously in Chapter 2, none of the Tatar TAE suffixes express mirative readings on their own.) I again note that Tatar differs from Bulgarian with respect to the availability of delayed learning time contexts like (189). These delayed learning time examples are crucial to Koev (2017, 2011)'s analyses, since he uses them to argue that Bulgarian grammar is sensitive to the time at which the speaker learned the scope proposition. However, Tatar grammar does not demonstrate the same sensitivity.

Koev (2017) refers to Learning Events (LEs), rather than EATs. This enables him to refer to the spatial component of the LE as well as its temporal component. In (196), the evidential verbal suffix co-occurs with present tense marking. Expressions like (196) require that the speaker have indirect evidence for the scope proposition.[33] Under Koev (2017)'s analysis, (196) is infelicitous because the evidential suffix marks that the speaker's LE must

[33]Interestingly, Bulgarian present tense evidential expressions like (196) pattern oppositely to Korean expressions in which the evidential suffix *-te* co-occurs with present tense. Korean *-te* + present tense expressions require that the speaker have direct, not indirect, evidence for the scope proposition (Lee 2013, 2011; §4.3.1.1).

be **temporally and/or spatially disjoint** from the event described by the scope proposition. The evidentially marked expression in (196) is infelicitous because the speaker's LE overlaps with the described event both temporally and spatially.

(196) **Context**: You attended a party last night. Today, you tell your friend what happened there. (Koev 2017, 14)

#Na parti-to snošti Martin svir-e-**l** na kitara.
at party-DEF last.night Martin play-PRES-EVID on guitar
Intended meaning: 'At the party last night Martin played the guitar.'

Under Koev (2017)'s analysis, the interpretation of a Bulgarian evidential expression arises jointly through the contribution of the evidential suffix and tense marking, both of which manipulate relations between event times. He proposes that the Bulgarian evidential suffix always contributes the same temporal relations, namely: (i) the LE precedes or overlaps with the Speaking Event (SE), and (ii) the LE is spatiotemporally distant from the Described Event (DE).[34] This evidential suffix co-occurs with past, present, or future tense marking, which indicates the relation between the time of the DE and the LE. I summarize the components of Koev (2017)'s proposal in Table 4.4, where $\triangle$ indicates spatiotemporal distance.

	Tense component	**Evidential component**
Past tense sentences	DE $<$ LE	LE $\leq$ SE, LE$\triangle$DE
Present tense sentences	LE $=$ DE	LE $\leq$ SE, LE$\triangle$DE
Future tense sentences	LE $<$ DE	LE $\leq$ SE, LE$\triangle$DE

Table 4.4: Components of Koev (2017)'s trace analysis of evidentiality in Bulgarian (Koev 2017, 17).

Koev (2017, 27) shows that the evidential component of a Bulgarian expression projects above both negation and modality. Furthermore, when a Bulgarian evidential verb form is embedded under a verb of saying, the embedded evidential is obligatorily interpreted as speaker-oriented. (197) is therefore infelicitous since the speaker has direct evidence

[34]This proposal to treat the evidential morpheme as having a constant denotation is similar to Lee (2013, 2011) and Smirnova (2013, 2011), and differs from Koev (2011).

for the embedded proposition. (This differs from the embedded past oriented Tatar TAE suffixes, which I showed in Chapter 2 are compatible with matrix subject-oriented interpretations.)

(197) **Speaker:** Direct evidence

Matrix subject: Indirect evidence

Milena told Maria that Todor has red hair and Maria believes her. Maria says: "Todor imal červena kosa" [= EVID(Todor has red hair)]. I saw Todor's red hair with my own eyes (Koev 2017, 28, from Sauerland and Schenner 2007, 5).

*Maria kaza-$\varnothing_{PAST}$, če Todor ima-**l** červen-a kosa.
Maria say-PST that Todor have-EVID red-FEM hair

Koev takes this data to suggest that the evidential component of a Bulgarian expression is "invisible" to propositional operators; he treats it as informative, Not-At-Issue (NAI) content. He analyzes Bulgarian evidential expressions using an update semantics in which sentences express functions from input information states to output information states. In brief, this formalism permits Koev (2017) to keep the propositional content of an expression separate from the evidential content. Propositional content is only added to the context set (i.e., the shared assumptions of the speech participants) if it is accepted by the addressee. Conversely, the evidential content is added directly to the context set. It is therefore NAI and is not targeted by propositional operators.

This is accomplished formally as in (199), which provides a semantics for the evidential expression in (198) (modified slightly from Koev 2017, 33). Declarative expressions like (198) are preceded by a declarative operator DECL that is anaphoric to the context of utterance (represented by k) and introduces a discourse referent for the propositional content of the expression (represented by p). (I indicate the contribution of the evidential in orange, and the contribution of the past tense marker in blue.)

(198) DECL Ivan celuna-$\varnothing_{PAST}$-**l** Maria
DECL Ivan kiss-PST-EVID Maria
'Ivan kissed Maria, as I learned at a different time or location.'

(199) $\text{DECL}_k^p(\exists x \,\&\, x{=}\text{Ivan} \,\&\, \exists y \,\&\, y{=}\text{Maria} \,\&\, \exists e \,\&\, \text{kiss}_p(e,x,y) \,\&$
$\exists e_l \,\&\, \textbf{learn}_{cs(k)}(e_l, \textbf{sp(k), p}) \,\&\, \tau(e_l) \leq \textbf{time(k)} \,\&\, e \triangle e_l \,\&\, \tau(e) < \tau(e_l))$

The At-Issue content of (198) (namely, that Ivan kissed Maria) is introduced by the predicate *kiss*, which is marked with a subscript *p*. This is subject to an interpretation rule that updates the information state with the content that Ivan kissed Maria. Conversely, the speaker's learning event (indicated by the *learn* predicate in (199)) is marked with a subscript *cs(k)* (i.e., the context set of the context of utterance *k*), and is added directly to this context set.

Like Faller (2004) and the EAT theories discussed in §4.3.1, Koev (2017)'s proposal cannot account for the observed contrast in meaning between the two future oriented Tatar TAE suffixes.[35] Since the LE is always temporally distinct from the DE in a future tense utterance, there is no way to manipulate the relationship between these two events to encode a difference in meaning. Furthermore, I repeat that I have no morphosyntactic motivation to decompose evidentiality and tense/aspect in Tatar into separate syntactic heads, as proposed by Koev (2017) (schematized in Table 4.4). While Koev (2017)'s proposal is intuitively compelling, it cannot account for all of the observed Tatar data.

4.3.4 Recap

In this section, I showed that trace analyses of evidentiality also cannot account for all of the observed Tatar data described in Chapters 2 and 3. These theories encounter similar issues as the EAT analyses described in §4.3.1.

Overall, the EAT and trace theories of evidentiality reviewed in §4.3 encounter the following recurring issues with respect to the Tatar data, presented in order of significance:

1. **The inability to account for the contrast in "non-specific" versus "specific" evi-**

[35] Koev (2017, 9) only shows a reportative use of a Bulgarian evidential verb form in combination with the future tense auxiliary *šte* 'will.' This correlates with Smirnova (2013)'s description of Bulgarian future tense evidential expressions as only being compatible with reportative evidence. This reportative requirement contrasts with the future oriented Tatar TAE suffixes, as noted previously, which are compatible with reportative as well as inferential evidence for the scope proposition.

dential meaning between the future-oriented Tatar TAE suffixes *-(y)Er* and *-(y)AçAK* (as described in Chapter 2).

Both EAT and trace theories rely on the fact that past events are compatible with either overlap or non-overlap of the speaker's EAT/LE and the ET. These relations are represented graphically as in Figure 4.1 and Figure 4.2, and can be used to account for direct versus indirect evidential meaning, respectively.

However, when an expression describes a future event, there is only one possible temporal relation between the speaker's EAT/LE and the ET. I illustrate this relation in Figure 4.3. Since the temporal relation between the EAT and ET is fixed in future tense expressions, it cannot be manipulated along the lines of the EAT and trace theories to account for the differences in meaning between the two TAE suffixes.[36] That is, these theories do not answer Question 2 from §4.1.

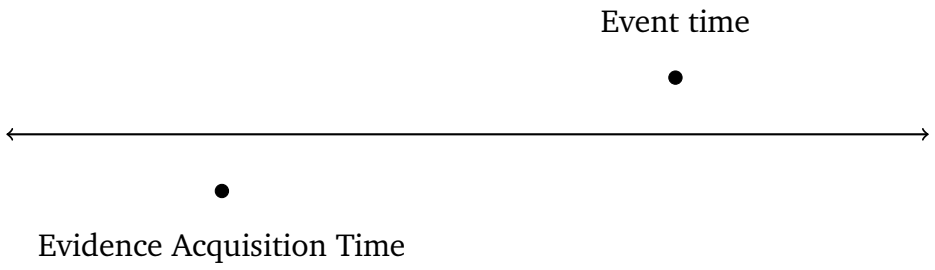


Figure 4.3: Only possible relation between the EAT and ET in expressions describing future events.

2. **The lack of morphosyntactic motivation to decompose tense/aspect and evidentiality into separate syntactic heads in Tatar.**

[36]We could propose that only one of the two future oriented Tatar TAE suffixes actually encodes evidential meaning; i.e., introduces an EAT. Perhaps only *-(y)AçAK* 'PROSP' introduces an EAT, and *-(y)Er* 'FUT' does not. However, it is still unclear how this could account for the observed contrast in meaning. Furthermore, this proposal could encounter similar issues as Lee (2013)'s account of future tense evidential expressions in Korean, which requires distinguishing between sensory and non-sensory evidence. This approach would require assuming that the grammar is sensitive to the type of evidence that the speaker has for *p*, and that only one kind of evidence is linked to the introduction of an EAT. The theory that I introduce in Chapter 5 does not encounter this issue.

Separating evidentiality and tense/aspect in the syntax is morphosyntactically well-motivated in e.g. Korean, which has a separate tense paradigm that co-occurs with the evidential suffix *-te* (Lee 2013, 2011). However, the Tatar suffixes are portmanteau tense/aspect and evidential morphemes. If we were to pursue an analysis in which (for example) only evidential meaning is contributed by the TAE suffixes, we would then have to posit a phonologically null tense paradigm to account for the temporal data in Chapter 3. This is simply not empirically motivated by the Tatar data.

Furthermore, while the theories discussed in this chapter all include a (spatio)temporal component in the semantics of their respective evidential morphemes, they do not explicitly answer the question of **why** tense and evidentiality might be morphologically linked. That is, they do not answer Question 1 from §4.1: What is the relationship between evidentiality and tense/aspect such that languages like Tatar have portmanteau tense/aspect and evidential morphemes?

3. **Evidence Acquisition Times/Learning Events are not independently motivated in the literature, and are not supported by the Tatar data.**

   To the best of my knowledge, EATs/LEs have only been proposed by the authors reviewed in this chapter to account for evidential data; i.e., they have not been independently motivated. This is not in itself a major issue, since new formal tools are proposed frequently in the literature.

   However, as I noted in this chapter, Tatar fails the primary diagnostic that Koev (2017, 2011) uses to motivate the presence of a Learning Event in the grammar. That is, Tatar *-GAn* is infelicitous in Koev's delayed learning time contexts, as shown in (190). This suggests that an LE analysis is not applicable to the Tatar data.

In an ideal world, we would like to have an analysis of the Tatar TAE data that can: (i) account for the contrast in meaning in both past- and future-oriented TAE suffixes using a similar theoretical mechanism, and (ii) explain **why** tense/aspect and evidentiality are morphologically linked. As I showed in this section, prior EAT and trace analyses of

evidentiality cannot do this. In §4.4, I conclude this chapter by foreshadowing the analysis that I propose, which addresses both of these desiderata.

4.4 Conclusion and foreshadowing

In this conclusion, I will briefly review the basic ideas behind my analysis, what data my analysis accounts for, and what is left for future research. Like the EAT and trace analyses of evidentiality discussed in §4.3, I propose an analysis of the Tatar TAE suffixes in which their evidential interpretations are not hardwired into their semantics (unlike e.g. Izvorski 1997, Faller 2002, Murray 2010, among others). My analysis is in the spirit of the EAT and trace analyses in that I, too, treat the evidential interpretation of the Tatar TAE suffixes as a byproduct of their temporal semantics.

As discussed in §4.2, I do not include a modal component in their semantics. This sets my analysis apart from those of Lee (2013, 2011) and Smirnova (2013, 2011). Additionally, unlike all of the EAT and trace analyses discussed in §4.3, I do not propose to include either an Evidence Acquisition Time or a Learning Event in the semantics of the Tatar TAE suffixes. This is motivated in part by the fact that Tatar fails Koev (2017, 2011)'s test for sensitivity to LEs in the grammar (as discussed in §4.3.1.3).

To account for the Tatar data discussed in Chapters 2 and 3, I assume an ontology of events in which all events are preceded by causally related **pre-states** and are followed by causally related **post-states**. (I define these formally in Chapter 5.) Some of the authors cited in this chapter have also informally observed a relationship between evidentiality and event pre-/post-states; these include descriptive accounts (Fleck 2007, Nikolaeva 1999) as well as formal analyses (Koev 2011, Matthewson 2011).[37] Such tripartite event ontologies

[37]While both Matthewson (2011) and Koev (2011) discuss event pre-/post-states in passing, neither of these authors explicitly incorporate them into their formal analyses. Matthewson (2011, 350) discusses event pre-/post-states when analyzing the sensory-non-visual St'át'imcets evidential *lákw7a*. Matthewson explicitly distinguishes between perception of the event itself and perception of the "results or precursors" of the event. She argues that *lákw7a* allows only non-visual perception of the event itself, but is compatible with the speaker having any kind of sensory evidence for its pre- or post-states.

Koev (2011, 129-130) also references event pre-states in his discussion of future tense evidential verb forms in Bulgarian. He suggests that when the event described by the scope proposition is non-past, the RT

have been previously independently motivated in the literature by e.g. Altshuler (2016, 2010), Moens and Steedman (1988), and Passonneau (1987), among others.

With this tripartite event ontology in hand, I argue that the Tatar TAE suffixes should be analyzed as markers of either tense or viewpoint aspect. I treat tenses as relating an event time to a reference time, typically the utterance time. I treat aspects as determining how the runtime of the event is "viewed" in time. I gloss and label each of the TAE suffixes my analysis addresses in Table 4.5, and provide full denotations for all of the suffixes in Chapter 5.

Suffix	**Temporal meaning**	**Evidential interpretation**
-DI	past tense	direct evidence
-GAn	resultative aspect	indirect evidence
-A	present tense	n/a
-(y)AçAK	prospective aspect	"specific" evidence
-(y)Er	future tense	"non-specific" evidence

Table 4.5: Temporal meanings of the Tatar TAE suffixes and their associated evidential interpretations, as motivated in Chapter 5.

I propose that the viewpoint aspects *-(y)AçAK* 'PROSP(ECTIVE)' and *-GAn* 'RESULT(ATIVE)' cause the described event to be "viewed" from within its pre- or post-state, respectively, and that the evidential meanings of the TAE suffixes arise as a side effect of of their underlying aspectual meanings. That is, when a speaker "views" an event from its causally related post-state, the kinds of propositions that are true at the speaker's utterance time are the same kinds of propositions that speakers cite as being indirect evidence for the event described by the scope proposition. Assuming that the speaker is cooperative (e.g. they are obeying Grice 1975's Maxims of Quantity and Quality), this leads to a reading that the speaker has (at best) indirect evidence for the described event. This accounts for both the past-oriented meaning of *-GAn* as well as its indirect evidential interpretation; i.e., it answers Question 1 from §4.1.

I use analogous reasoning to account for the "specific" evidential reading associated

that the evidential targets could be located within some "salient stages of the preparatory process" of the event described by p.

with *-(y)AçAK*. When a cooperative speaker "views" an event from its causally related pre-state, the propositions that are true at the speaker's utterance time describe state(s) and/or event(s) that the speaker believes will cause the event described by the scope proposition to occur. Informally speaking, speakers describe these states/events as "specific" evidence for the scope proposition; furthermore, the causal relationship between the pre-state and the event leads to a reading of increased speaker certainty regarding the described event. This accounts for both the future-oriented meaning of *-(y)AçAK* as well as its reading of "specific" evidence/increased speaker certainty. This analysis avoids the issues encountered by EAT/LE analyses (as described in §4.3.4), and answers both Question 1 and Question 2 from §4.1.

I derive the direct evidential reading of *-DI* and "non-specific" evidential reading of *-(y)Er* pragmatically, as implicatures, and show that they can be cancelled in certain contexts. To conclude Chapter 5, I review some of the tests for modality discussed in §4.2, and show that my analysis accounts for the observed Tatar data (e.g. the ability to assert TAE(p) if the speaker believes that p is true, and the inability to assert TAE(p) if the speaker believes that p is false). However, I note that my theory as it is currently presented requires some tweaking to account for the interaction of the TAE suffixes with sentential negation, as described in §4.2.2.3.

In Chapter 6, I use a scope analysis to account for the interpretations of the TAE suffixes in embedded CPs. I argue that (some of) the TAE suffixes can scope out of embedded clauses and be interpreted superior to the matrix attitude verb. When a TAE suffix scopes high, the embedded event is "viewed" from the speaker's utterance time; i.e., the evidential component of the embedded TAE suffix is interpreted as speaker-oriented. Conversely, when the TAE suffix remains in situ, the embedded event is "viewed" from the matrix subject's "now" and is therefore matrix subject-oriented.

I account for the lack of evidential reading in embedded verbal nominalizations by appealing to the fact that only *-GAn* and *-(y)AçAK* are grammatical in verbal nominalizations. Since their evidential meanings in matrix clauses and embedded CPs are partially derived through pragmatic competition with *-DI* and *-(y)Er*, I propose that such pragmatic compe-

tition does not occur in verbal nominalizations and therefore no evidential readings arise. (The exact mechanism by which pragmatic competition occurs in embedded CPs is left for future work.) This avoids the issues that would arise if evidentiality were encoded as part of the propositional content of the TAE suffixes, as noted in §4.2.2.1.

I conclude by noting that this theory is currently tailored to the Tatar TAE system. It is not applicable to languages in which tense/aspect and evidentiality are unrelated systems (e.g. Cheyenne, Cuzco Quechua, St'át'imcets, among many others). It is also not currently applicable to languages with interrelated tense/aspect and evidential systems, but which lack portmanteau TAE morphology (e.g. Bulgarian, Korean). At present, the analysis presented in Chapter 5 and 6 is geared towards accounting for Turkic TAE data. As I showed in Chapter 2, a large number of other Turkic languages have TAE systems that are very similar to Tatar. This aspectual analysis of evidentiality successfully explains **why** the Tatar TAE suffixes express both tense/aspect and evidentiality; i.e., why they are portmanteau morphemes.

CHAPTER 5

Deriving the temporal and evidential interpretations of the Tatar TAE suffixes in matrix clauses

5.1 Introduction

In this chapter, I give an analysis of the Tatar TAE morphemes in semantically unembedded environments, i.e., matrix clauses. This analysis is an attempt at joining, formalizing, and expanding upon two previously described intuitions from the evidential and tense/aspect literature. The first intuition is that languages can use "perfect"/"resultative"/"completive" aspectual morphemes to relate ongoing states to previous situations, perhaps through a causal relationship (e.g. Comrie 1976). The second intuition is that indirect evidentials—like these aspectual morphemes—indicate the speaker's observation of the results of some previous event (e.g. Nikolaeva 1999, Bashir 2006).

The organization of the chapter is as follows. In §5.2, I describe and motivate the ontology of events that I assume in this dissertation, and discuss some existing linguistic theories of event pre- and post-states, primarily Moens and Steedman (1988). In §5.3, I make the conceptual connection between linguistic reference to these pre- and post-states and their corresponding evidential readings. In §5.4, I spell out the semantics of the Tatar TAE morphemes, and give compositional derivations of their use in unembedded environments. I propose that these morphemes have underlyingly temporal semantics, and derive their evidential readings in part through pragmatic competition with one another. §5.5 concludes.

To increase the readability of this dissertation, I have written this chapter as a stand-

alone paper, separate from the subsequent chapter, in which I analyze the TAE suffixes in some semantically embedded environments. The theory that I present in this chapter can account for the interpretation and use of the TAE morphemes in matrix clauses. However, I note that some components of this theory will be revised in Chapter 6 to account for their use in embedded environments. More specifically, I will eventually posit a more complex semantic type for the TAE suffixes to enable an embedded time argument to be bound by the matrix clause. I will also revise the denotation of *-GAn* 'RESULT' to account for its interpretation in embedded verbal nominalizations, and update the denotations of *-(y)Er* and *-(y)AçAK* to include an indexical component. However, the general semantic and pragmatic arguments that I rely on in this chapter will be maintained across the analytical component of the dissertation.

5.2 Temporal ontology

I assume a temporal ontology in which events are followed by causally related post-states and are preceded by causally related pre-states. Similar temporal ontologies have been proposed and/or utilized by a number of authors. For instance, Comrie (1976, 56-65) describes the perfect aspect as relating some state to a previous situation, and the prospective aspect as relating some state to a subsequent situation. This descriptive work is followed by a series of computational linguistics publications in the late 1980s that posit event pre- and post-states as a means to account for e.g. the interpretation of the temporal progression of narratives and other multi-clausal phenomena (Moens and Steedman 1988, Webber 1988, Passonneau 1987). More recent theoretical work utilizing event pre- and post-states includes Bohnemeyer (2014), Altshuler (2016, 2010), and Brandt (2005), among others. As I noted in Chapter 4, some authors have also noted a connection between event pre- and post-states and evidentiality (Matthewson 2011, Koev 2011, Fleck 2007, Nikolaeva 1999).

The goal of this section is twofold. First, I briefly summarize Moens and Steedman (1988) to show that the inclusion of event pre- and post-states in the temporal ontology has been independently motivated by other linguists; as such, the the ontology I am assuming

does not need to be novelly motivated. Second, I spell out the details of the event ontology that I assume, which are largely in line with prior theories of event pre- and post-states.

5.2.1 Moens & Steedman (1988)

The overarching goal of Moens and Steedman (1988) is to provide a semantics of temporal categories and their use, particularly in multi-clausal expressions. The authors propose to ground their analysis in the causal relationships between events. They note that in narratives, the speaker's use of temporal categories like viewpoint aspects, adverbials, and *when*-clauses "will typically be colored by the fact that [the events that they are discussing] are involved in sequences that are planned, predicted, intended, or otherwise governed by agencies of one kind or another" (Moens and Steedman 1988, 16). The authors use the term **contingency** to refer to these dependencies between events. (I ultimately utilize their notion of contingency in my own analysis of the Tatar data.)

Moens and Steedman (1988) propose an temporal ontology that is based on this notion of contingency between events, i.e., relationships of causation and consequence. Their temporal ontology thereby differs from ontologies consisting of temporal primitives with no inherently causal relationship to each other, as in e.g. Reichenbachian theories of times. The authors motivate their ontology through data like (200). They note that if the *when*-clause in (200) is analyzed simply as locating the bridge building event relative to the time of the event described by the main proposition, *when* would appear to be multiply ambiguous (Moens and Steedman 1988, 15, following Ritchie 1979). (200a) seems to place the bridge building event after the main event, while (200c) places the bridge building event before the main event. (200b) suggests that the two events overlap.

(200) When they built the 39th Street bridge...

a. ... a local architect drew up the plans.

b. ... they used the best materials.

c. ... they solved most of their traffic problems.

To account for data like (200), the authors propose a temporal ontology in which tensed clauses refer to contingency-based event structures consisting of a preparatory process, a culmination, and a consequent state. Both the preparatory process and consequent state have temporal durations, as in Figure 5.1. They term this entire event structure a **nucleus**.

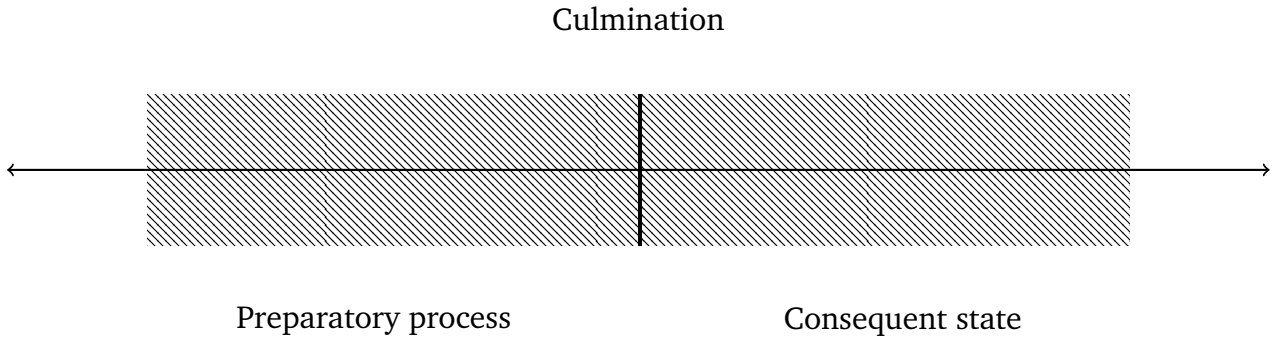


Figure 5.1: Graphic representation of the temporal ontology (nucleus) proposed by Moens and Steedman (1988, 18).

Moens and Steedman (1988) argue that *when*-clauses like (200) introduce a new temporal referent into the discourse. This temporal referent is an entire nucleus, as in Figure 5.1. The time of the main clause can then be located in any part of this *when*-clause nucleus. The specific location of the time of the main clause with respect to this nucleus is determined contextually through world knowledge or discourse specific knowledge.[1] For example, when interpreting (200a), the time of the architect drawing up the plans is located in the preparatory process of the bridge building event. When interpreting (200c), the time of solving their traffic problems is located in the consequent state of the bridge building event.

In addition to positing the event structure in Figure 5.1, Moens and Steedman (1988) argue that linguistic expressions like tenses, aspects, and *when*-clauses can coerce predi-

[1]Since the components of the event structure in Figure 5.1 are contingently related to each other, Moens and Steedman (1988) argue that this proposed ontology also accounts for the infelicity of data like (201). This infelicity arises because it is difficult to think of a part of either the preparatory process or consequent state of the event of the car breaking down that would involve the sun setting.
(201) # When my car broke down, the sun set.

cates into being interpreted as having preparatory processes or consequent states. They distinguish between the four classes of eventive verbal predicates in (202), which can in part be characterized by whether or not they are associated with a consequent state. Culminations (202a) and culminated processes (202b) necessarily have consequent states that can be referred to linguistically.

(202) **Classes of eventive verbal predicates defined by Moens and Steedman (1988):**

a. **Culmination:** Instantaneous change; causes a consequent state.
   Harry reached the top.

b. **Culminated process:** Event is extended in time; causes a consequent state.
   Harry climbed to the top.

c. **Point:** Indivisible event; consequent state is not relevant in the discourse.
   Harry hiccuped.

d. **Process:** Event is extended in time; no resulting consequent state.
   Harry climbed.

Points (202c) and processes (202d) do not necessarily have consequent states. However, they can be coerced into having consequent states when e.g. combined with the English present perfect, which Moens and Steedman (1988, 19) propose must combine with a culmination (i.e., a predicate class with a consequent state). An example of this coercion is shown in (203). When the present perfect combines with a point, the expression is infelicitous if it is uttered out of the blue, as in (203a). However, if the context is such that the result of the event is salient in the discourse, this expression is acceptable (203b). Moens and Steedman (1988) argue that this is because the ticking event in (203b) is coerced from a point into a culmination, and thereby has a consequent state.

(203) a. # The clock has ticked.

b. **Context:** You have a clock that only ticks once a day. You report for the day:
   ✓ The clock has ticked.

This notion of coercing events into having preparatory processes and consequent states is also relevant to my proposal. I similarly follow Moens and Steedman (1988) in arguing that the components of event structure (i.e., the event and its pre- and post-states) are linked by a contingent, causal relationship.

Moens and Steedman (1988) do not give a compositional semantic analysis of their proposal. However, given the ability of our current formalisms to manipulate events and states, it is possible to imagine how their tripartite ontology could be implemented compositionally. I adopt some conceptual components of their proposal—specifically their contingency relation—in my compositional analysis of the Tatar TAE data. I spell out some of the details of my formalism in the following section.

5.2.2 The ontology that I assume

Like Altshuler (2016, 2010), Moens and Steedman (1988), and Passonneau (1987), among others, I assume a tripartite event ontology in which events are preceded by **pre-states** and are followed by **post-states**. I depict this tripartite event ontology in Figure 5.2, borrowing the notion of contingency from Moens and Steedman (1988).

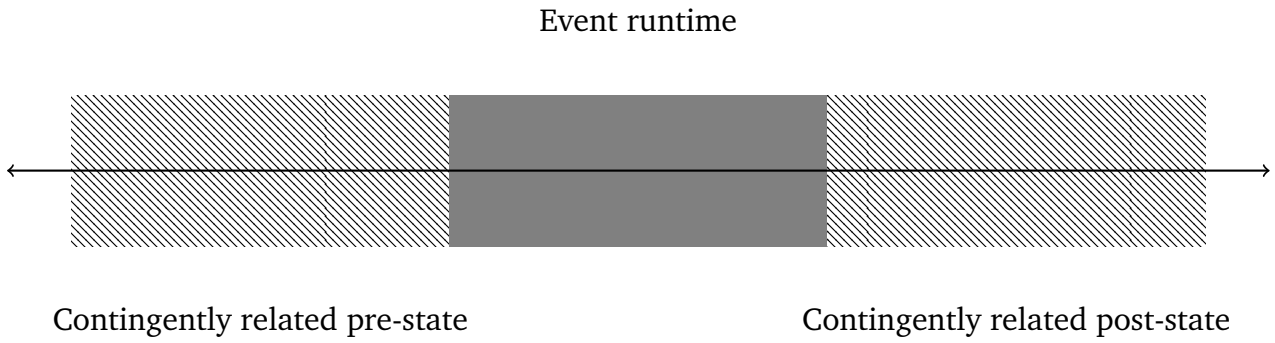


Figure 5.2: Basic graphic representation of the temporal ontology that I assume (to be revised).

My event ontology differs from Moens and Steedman (1988) primarily in that I treat the event pre- and post-states as distinct semantic objects from the event itself, rather than

grouping all three together into a nucleus. As I will show in §5.4, this enables me to derive the temporal interpretations of the TAE suffixes. I treat the event as a semantic object e of type v, and pick out the runtime (i.e., the temporal duration) of the event using Krifka's runtime function τ. The runtime of an event is therefore represented as $\tau(e)$. I treat the pre- and post-states as states s and refer to their runtimes using $\tau(s)$.

I assume an idealized model in which the boundaries between the pre- and post-state and event runtimes are crisp. I furthermore assume that the runtimes of the pre- and post-states immediately abut the runtime of the event itself, without any overlap: that is, the runtimes of the pre- and post-states and the event must "touch." (I give a formal definition of this adjacency requirement in §5.4 as part of the "suitability" requirement on event pre- and post-states.) Finally, the pre- and post-states necessarily have runtimes; i.e., they involve some temporal duration. There is no requirement on the length of the runtime of these pre- and post-states. For instance, if a Tatar speaker asserts an expression using *-GAn* 'RESULT' or *-(y)AçAK* 'PROSP,' it is not the case that their utterance time must be located close in time to the event that they are describing.[2] In (204), the speaker uses *-GAn* 'RESULT' to report an event that happened in 1939, long before they were born.[3] The pre- and post-states are therefore not **required** to be of a short temporal duration.

(204) Germaniya Polşa-nı bas-**qan**-∅.
Germany Poland-ACC invade-RESULT-3SG
'[I have indirect evidence that] Germany invaded Poland.'

Generally speaking, event pre-states must have a starting point.[4] This is required by my theory, which contrasts the use of *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' with respect to whether

[2]However, as I noted in Chapter 3, nonlinguistic facts about the world make it such that when a speaker describes an event that will happen in the near future of the utterance time, it is typically the case that they also have some specific evidence that the event will occur. Expressions describing near future events therefore correlate with the use of *-(y)AçAK* 'PROSP,' although *-(y)AçAK* is not strictly required by the grammar.

[3]Greed (2009, 18-19) similarly notes that *-GAn* is frequently used in biographical narratives and historical accounts; i.e., it is used to describe events that occurred in the far past of the utterance time.

[4]Exceptions to this are events that consistently reoccur based on properties of the world, as in (205). Since the speaker knows that the sun rises and sets every day (in Los Angeles), and as such it will set tomorrow, they always use *-(y)AçAK* 'PROSP' to report (205). (My primary Tatar consultant reports that the felicity of *-(y)Er* 'FUT' in (205) would change in locations in which the sun goes for extended periods of time without rising or setting, e.g. above the Arctic circle.)

or not the speaker locates themselves in a contingently related pre-state of the event. I again assume an idealized model in which there is a crisp boundary at the beginning of the event pre-state. Given all times i prior to the event runtime, i is located either in the runtime of the event pre-state or not; the pre-state boundary is not fuzzy.

Conversely, my theory requires that event post-states extend indefinitely following the event. This results from two assumptions: first, that cooperative speakers must have **some** perceptual evidence for the events described by their assertions; and second, that this perceptual evidence can take the form of perceiving either some portion of the runtime of the event itself, or its contingently related post-state. (I expand on this point in §5.4.2.4.)

In summary: event pre-states are bounded, closed intervals that have both minimum and maximum values. Conversely, event post-states are bounded, half-open intervals. More specifically, they are left-closed and right-open: they have a minimum value, but no defined maximum value. I define these intervals in (206), where a, b and i are instants (i.e., points in time), and $a < b$. I provide an updated, final graphic representation of the temporal ontology that I assume in Figure 5.3.

(206) **Conditions on the intervals picked out by event pre- and post-states**

a. Event pre-state: Bounded, closed interval

$[a,b] = \{i\colon a \leq i \leq b\}$

b. Event post-state: Bounded, left-closed, right-open interval

$[a,b) = \{i\colon a \leq i < b\}$

Aside from these properties, for the purpose of this dissertation, I am not concerned with giving a formally rigorous definition of what constitutes an event pre- or post-state. I assume the relation of contingency informally defined by Moens and Steedman (1988) to describe the relationship between the event and its pre- and post-states. Intuitively

(205) qoyaş irtägä {bat-**açaq**-∅ / #bat-**ar**-∅}.
sun tomorrow set-PROSP-3SG / set-FUT-3SG
'[I have specific/#non-specific evidence that] The sun will set tomorrow.'

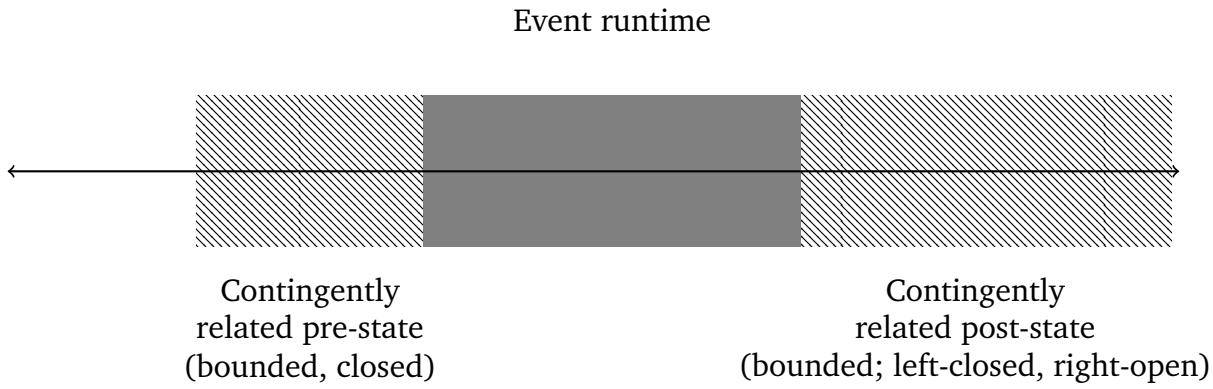


Figure 5.3: Graphic representation of the final temporal ontology that I assume.

speaking, we can think of the event and its pre- and post-states standing in a relation of causation or, in Moens and Steedman (1988)'s terms, "enablement." We can do this by considering an example sentence like *John gave the cat catnip*. There are some events or states preceding the event of John giving the cat catnip that will cause the event to occur or enable it to occur. For example, John could decide to give the cat catnip, or could walk over to the cabinet where the catnip is kept to take it out to give to the cat. Similarly, there are some events or states following the event of John giving the cat catnip that are caused by it. These could include the event of the cat running around the house wildly, or the state of there being catnip on the floor.

Moens and Steedman (1988) assert that this contingent relationship is not strictly causal. They note that "if Event A stands in a contingent relation to Event B, then an occurrence of A will not automatically lead to an occurrence of B: John laying the foundations of the house is a prerequisite for or enables him to build the walls and roof, but does not cause it in the more traditional sense of the word and does not automatically or inevitably lead to him building the walls" (Moens and Steedman 1988, 26). Nonetheless, the event of John laying the foundation of the house could be classified as being in a contingent relation to the event of John building the walls and roof. I follow them in assuming that this contingency relation is primarily, but not strictly, causal.

I discuss the link between pre- and post-states and evidentiality in the following section,

focusing particularly on the the kinds of propositions ("evidence") that lead speakers to believe that they are in a pre- or post-state of an event.

5.3 The link between speaker perception of pre- and post-states and evidentiality

I ground the link between speaker perception of pre- and post-states and the evidential readings of the Tatar TAE morphemes in Grice (1975)'s Conversational Maxims. More specifically, I assume that when a Tatar speaker cooperatively utters an expression, they obey Grice's Maxims of Quantity ("make your contribution (only) as informative as required") and Quality ("be truthful"). For a speaker to obey the Maxim of Quality, the speaker must believe that what they are expressing is true. This belief in the truth of their expressions must come from some source, whether through direct sensory perception or internal reasoning based on their sensory perception. The speaker's source of their belief (that is, their "way of knowing") is typically considered to be the domain of evidentiality, as discussed in Chapter 2. In this section, I propose that all assertions (and therefore, beliefs) are fundamentally based at least in part on speakers' perceptions of the world.[5]

A speaker can come to believe a proposition in a number of different ways. For example, a speaker could see for themselves that Kirk's house is red, and then state (207). Alternately, they could see a photograph of Kirk's house, or hear from Kirk's neighbor that his house is red. All of these are possible ways of knowing the proposition in (207); some languages (e.g. Tariana (Aikhenvald 2004, 2-3)) appear to grammatically distinguish between these different ways of knowing through their evidential systems.

(207) Kirk's house is red.

For my analysis of the Tatar data, I do not need to distinguish between different ways

[5]I differ from some other authors in this respect; several theories of evidentiality rely on a distinction between propositions that speakers come to believe through internal reasoning/inference versus through sensory perception, e.g. Lee (2013, 2011) and Matthewson (2011).

of knowing. All that matters for my proposal is that when a speaker asserts a proposition, I assume that their belief in their assertion is based on their perception of some event or state in the world. In a context in which the speaker is told by Kirk's neighbor that his house is red, they perceive the event of the neighbor reporting that his house is red, and they believe the proposition that the neighbor reports. In a context in which the speaker sees that Kirk's clothes are covered in red paint, they could infer through internal reasoning both the propositions that Kirk painted his house and that the house is (now) red. In this case, the speaker perceives the state of Kirk's clothes having red paint on them, and their belief in the proposition that Kirk's house is red results from internal reasoning based on their perception of this state. Fundamentally, I propose that all beliefs stem from some perception of events and/or states in the world.[6]

I consider a slightly more subtle example in (208). The expression in (208) refers to a future event. Nonetheless, I propose that when a speaker utters (208), they are basing their assertion on their belief that some proposition(s) are true at the moment of speaking, as obtained through sensory perception of some events or states in the world. Events or states that a speaker could perceive to lead them to assert (208) could include that Kirk bought plane tickets to London, that Kirk is applying for a visa to go to the UK, or that Kirk told the speaker that they are going to go to London. The only kind of perception that is unavailable with respect to (208), as opposed to (207), is witnessing the event of Kirk going to London; this is because the event is located in the future of the utterance time.

(208) Kirk is going to go to London.

I discuss these examples to illustrate that speakers always base their beliefs in the content of their assertions (at least in part) on their sensory perception of the world. The

[6]The idea that speakers base all of their beliefs off of their sensory perception requires some elaboration for beliefs in propositions involving abstract concepts. For example, we can imagine that a speaker believes the proposition *Love is beautiful*. The speaker could base their belief in this proposition off of e.g. romantic movies that they have watched, love stories that they have read, etc. In fact, it is not even necessary that the speaker have been in love before to believe that love is beautiful. In cases like this, I assume that speakers' beliefs stem from sensory perception (of e.g. romantic narratives) **as well as** their own reasoning regarding subjective concepts like beauty.

specifics of this perception are irrelevant to my analysis of the Tatar TAE data. What is important to my proposal is the notion that speakers base their beliefs—and therefore, their assertions—on their perception of events or states in the world around them. In evidential terminology, I propose that all "evidence" is fundamentally obtained through sensory perception.

To summarize, I assume:

1. Speakers are cooperative and obey the Gricean conversational Maxims.
2. Given the Gricean Maxim of Quality: speakers believe the content of their assertions; i.e., they assert (what they believe) is truthful.
3. Speakers base their beliefs off of their sensory perception of the world around them. This can involve:
   (a) Perception of the event or state that their assertion describes.
   (b) Perception of another individual's report regarding the event or state that their assertion describes.
   (c) Perception of an event or state that causes them to infer the truth of another proposition through internal reasoning.

In the following sections, I discuss the application of these assumptions to expressions containing the Tatar aspectual suffixes *-GAn* 'RESULT' and *-(y)AçAK* 'PROSP.' These suffixes are underlyingly viewpoint aspects. However, given this link between the events/states that speakers perceive and their assertions, I show that we can understand (in part) how the evidential readings of these suffixes are obtained.

5.3.1 Unformalized application to *-GAn* 'RESULT'

I proposed in Chapter 3 that the Tatar TAE suffix *-GAn* 'RESULT' is a marker of resultative aspect: it picks out a reference time in the contingently related post-state of the event from

which the event is "viewed." In semantically unembedded environments, this reference time is valued by the utterance time. (I give a full compositional semantics for *-GAn* in §5.4.)

I give an example of a *-GAn*(*p*) expression in (209). When a speaker says (209)—given our understanding of *-GAn* as a marker of resultative aspect—they assert that their utterance time is located within the contingently related post-state of the event of Güzäl making cookies.

(209) Güzäl peçeniye yas-**kan**-∅.
Güzäl cookie make-RESULT-3SG
'[I have indirect evidence that] Güzäl made cookies.'

To cooperatively assert (209), the speaker must believe that this proposition is true. As I described in the preceding section, the speaker comes to believe that this proposition is true by perceiving events and states in the world around them (i.e., by obtaining "evidence" for the scope proposition "Güzäl make cookies" through their sensory perception). There are a range of events and states that the speaker could perceive that would lead them to believe that they are in a contingently related post-state of Güzäl making cookies. I list propositions describing some of these possible events and states in (210).[7]

(210) Events or states that could occur in the post-state of "Güzäl make cookies":

a. There are cookies in Güzäl's kitchen.

b. Güzäl's kitchen smells like cookies.

c. There is a dirty baking sheet and mixing bowl in Güzäl's sink.

d. Güzäl has cookie batter on her clothes.

e. Läylä (Güzäl's roommate, who hates baking) is eating a freshly baked cookie.

Interestingly, the propositions in (210) are precisely the kinds of propositions that descriptions of evidential systems (e.g. Willett 1988, Aikhenvald 2004) would describe as

[7]I note that for a speaker to believe that they are in a contingently related post-state of Güzäl making cookies, they cannot perceive the event of Güzäl making cookies itself. Observing this event would cause them to believe that they are located within the runtime of the event of Güzäl making cookies, rather than its post-state.

"indirect evidence" for an evidential assertion with the scope proposition "Güzäl made cookies." If *-GAn* were strictly a marker of indirect evidentiality, some subset of the propositions in (210) (or similar propositions) would have to be true in order for the expression in (209) to be uttered felicitously. Given this connection between the felicity of (209) and the truth of the propositions in (210), it becomes apparent why *-GAn* has previously been described as a marker of indirect evidentiality by e.g. Greed (2014, 2009).

Obtaining the use of *-GAn* as a suffix that is **strictly** compatible with indirect evidence (and not direct sensory perception) requires one further step. This is based on the fact that when a speaker reports a past event, they could in principle have directly perceived the event itself, rather than solely its post-state. (Past events differ from future events in this respect.) That is, when the speaker of (209) reports that Güzäl made cookies, they could theoretically have directly perceived the event of Güzäl making cookies. I assume (by the Gricean Maxim of Quality) that speakers make the most informative assertions possible; I further assume (following e.g. Faller 2012, Davis et al. 2007) that speakers utilize an "evidential hierarchy" in which expressing direct perception of an event is considered to be "stronger" than expressing indirect perception. The speaker's choice to use *-GAn* 'RESULT' in (209) necessarily invokes a contingently related event post-state. By opting to invoke this event post-state, the speaker of (209) therefore signals that their "strongest" evidence for Güzäl making cookies is based on perception of the post-state of the event, as opposed to perception of the event itself.[8] In evidential terms, this conveys that the speaker has (only) indirect evidence for the described event, and not direct evidence. (I discuss the pragmatics of *-GAn* 'RESULT' and its interaction with *-DI* 'PST' further in §5.4.2.3.)

[8]Assuming something like Faller (2012)'s evidential hierarchies does not necessarily commit me to treating the Tatar TAE morphemes as underlyingly evidential. I could assume a conceptually identical hierarchy in which expressions that assert that a speaker observed an event's post-state are considered "weaker" than other expressions; this avoids using any evidential terms, but is based on the same conceptual grounds as Faller (2012).

5.3.2 Unformalized application to *-(y)AçAK* 'PROSP'

I proposed in Chapter 3 that *-(y)AçAK* 'PROSP' is a marker of prospective aspect. I treat prospective aspect as a "mirror image" of resultative aspect; it introduces a reference time within a contingently related event pre-state from which the event is viewed. In semantically unembedded environments, this reference time is valued by the utterance time. (I give a full compositional semantics for *-(y)AçAK* 'PROSP' in §5.4.1.3.)

I give an example of a *-(y)AçAK*(*p*) expression in (211). When a speaker says (211), they assert that their utterance time is located within the contingently related pre-state of the event of Güzäl making cookies.

(211) Güzäl peçeniye yas-**açaq**-∅.
Güzäl cookie make-PROSP-3SG
'[I have specific evidence that] Güzäl will make cookies.'

To cooperatively assert (211), the speaker must believe that this proposition is true; i.e., they must believe that their utterance time is located in a contingently related pre-state of the event described by "Güzäl make cookies." They come to believe that this proposition is true by perceiving events and states in the world around them (in evidential terminology, "evidence"). I give examples of propositions describing possible events or states that the speaker could perceive to lead them to believe that (211) is true in (212). These events and states stand in a contingent relationship (in the terminology of Moens and Steedman 1988) to the event of Güzäl making cookies.

(212) Events or states that could occur in the pre-state of "Güzäl make cookies":

a. Güzäl tells you, "I will make cookies."

b. Güzäl makes a shopping list of ingredients that go into cookies.

c. Güzäl's boss asks her to make cookies to bring into the office.

d. All of the ingredients that go into cookies are laid out on Güzäl's kitchen counter.

As I discussed in Chapter 2, Tatar speakers report that saying (211) conveys that the speaker is "confident" or "certain" that the event described by the proposition will occur. My consultants typically describe *-(y)AçAK* as communicating that the speaker has "specific" evidence that the scope proposition will occur, and list propositions like in (212) as examples of "specific" evidence for the scope proposition in (211). However, "specific" evidence is not a cross-linguistically documented evidential category in the typologies of Aikhenvald (2004) and Willett (1988). This fact is not necessarily fatal to a purely evidential analysis of *-(y)AçAK*, since it could be the case that "specific" evidentials are simply very rare, and are unaccounted for in the previous typologies. However, this does provide a strike against analyzing *-(y)AçAK* as an evidential. Furthermore, typologists note that evidentiality is much less commonly marked in the future, as opposed to the past (e.g. Aikhenvald 2004).

The need to posit a new category of "specific" evidence is done away with under a purely temporal analysis of *-(y)AçAK* 'PROSP.' I propose that the reading of increased speaker certainty with respect to the scope proposition in (211) stems directly from the meaning of *-(y)AçAK* as a prospective aspect. *-(y)AçAK* is used to locate the utterance time in a contingently related pre-state of the described event; for a speaker to utter (211) felicitously, they therefore must believe that some events or states at their utterance time will eventually cause or enable the event described by the scope proposition to occur. Informally speaking, these events or states can be described as "specific" evidence that the scope proposition will occur. Since *-(y)AçAK* asserts that the scope proposition is contingently related to these ongoing events or states, *-(y)AçAK* expressions are associated with a reading of increased speaker certainty about its occurrence.

In sum, under this treatment of *-(y)AçAK* as a marker of prospective aspect, we do not need to revise our understanding of the possible evidential categories to include a category of "specific" evidentials.

5.3.3 Recap

If we conceptualize speakers' beliefs in their expressions as being fundamentally grounded in their perception of events and states in the world around them, we can explain the link between event pre- and post-states and evidentiality. In this section, I tie up some remaining loose ends regarding the implementation of this proposal with respect to the Tatar TAE suffixes.

First, I note that that speakers may differ with respect to how they interpret the events or states that they perceive: even if their sensory input is (roughly) the same, their internal reasoning regarding this input may differ. This can lead Tatar speakers to choose different TAE suffixes to report the same event. I illustrate this through the Tatar example in (213).

(213) **Context:** Aygöl and Mansur have a colleague, Alsu, who tells them that she is going to go on holiday to Paris. Aygöl believes Alsu and expects that she will go; i.e., Aygöl locates herself in a pre-state of the event of Alsu going to Paris.

However, Mansur thinks that Alsu frequently boasts about her big plans without actually going through with them. He isn't convinced that Alsu is actually planning the trip, or will go; i.e., Mansur does not locate himself in a pre-state of the event of Alsu going to Paris.

Both Aygöl and Mansur felicitously report their conversation with Alsu as follows:

a. **Aygöl:**

Alsu Parij-ğa bar-**açaq**-∅.
Alsu Paris-DAT go-PROSP-3SG
'[I have specific evidence that] Alsu will go to Paris.'

b. **Mansur:**

Alsu Parij-ğa bar-**ır**-∅.
Alsu Paris-DAT go-FUT-3SG
'[I have non-specific evidence that] Alsu will go to Paris.'

The same basic set of propositions are available to the two speakers in (213) at their utterance time, primarily the proposition that Alsu says that she is going to go to Paris.

Aygöl's assessment of this proposition leads her to believe that she is located in the contingently related pre-state of the described event, as in (213a); she therefore reports the event of Alsu going to Paris using *-(y)AçAK* 'PROSP.' However, Mansur's assessment of this same proposition does not lead him to believe that he is located in a contingently related pre-state of this event. Since he does not fully expect Alsu to carry out the plans that she reports, he does not believe that he is in a contingently related pre-state of the event. He therefore reports the event using *-(y)Er* 'FUT,' as in (213b). This example shows that one speaker can count a proposition as "specific" evidence, whereas another speaker doesn't.

Second, I follow Moens and Steedman (1988) (in part in the spirit of Nedjalkov 1988) in assuming that all events can have associated pre- or post-states, even if the event is not typically associated with either.[9] I give an example of such an event in (214). In a neutral context, an event of knocking does not typically have any contingently related post-state. However, the context in (214) is such that if a knocking event occurs, the post-state of the knocking event will include the event of the dog barking.

(214) **Dog barking context**: Your friend has a dog that barks very loudly whenever anyone knocks on their door. You and your husband Ali go over to their house for dinner. Ali gets out of the car ahead of you and you remind him not to knock on the door. However, a few moments later, you hear the dog barking. You infer from this that Ali knocked on the door anyway. You say:

Ali işek-ne şaqı-**ğan**-∅.
Ali door-ACC knock-RESULT-3SG
'[I have indirect evidence that] Ali knocked on the door.'

Data like (214) is also addressed by Nedjalkov (1988), who gives a descriptive typology of resultative expressions cross-linguistically. Nedjalkov (1988, 499) distinguishes between propositions that he terms **trivially** and **non-trivially true** in the post-states of events.

[9] Moens and Steedman (1988) only discuss consequent states (post-states); I extend their proposal to pre-states. They are in part concerned with organizing predicates into lexical classes depending on the presence/absence of a consequent state in their lexical semantics, as described in §5.2.1. They frame the issue as one of coercion. That is, they argue that predicates without lexically specified consequent states can nonetheless be coerced into readings in which they have an associated consequent state by the context (as in example (203b)).

Generally speaking, trivially true propositions are linked to the lexical semantics of the verb, whereas non-trivially true propositions are not. (That is, the use of the term "trivial" here refers to trivial truth in the linguistic, lexical semantic sense, as opposed to trivial truth in the logical sense.) In evidential terms, both types of propositions can serve as "evidence" for the scope proposition.

In the example in (215), the proposition "The kitchen is dirty" is trivially true because if a kitchen-dirtying event occurs, the post-state of this event will necessarily involve the kitchen being dirty. (Barring any subsequent events that cause the kitchen to become clean again.) Conversely, the proposition "Howard's girlfriend is mad at him" is non-trivially true. This is because nothing about the event of dirtying the kitchen inherently causes Howard's girlfriend to get mad. Nonetheless, a Tatar speaker could reason based on the state of Howard's girlfriend being mad at him that they are in a post-state of Howard dirtying the kitchen (given their additional knowledge that Howard's girlfriend hates it when the kitchen is dirty). Similarly, nothing about the event of someone knocking on the door in (214) inherently causes the dog to bark. However, a Tatar speaker in the context in (214) could reason based on the event of the dog barking that they are in a post-state of someone knocking on the door.

(215) Howard dirtied the kitchen.

a. **Trivially true:** The kitchen is dirty.

b. **Non-trivially true:** Howard's girlfriend is mad at him.

Framed another way: Nedjalkov (1988)'s trivially true propositions are true independently of the context of utterance, whereas non-trivially true propositions are not. If the context is such that Howard's girlfriend is a slob and doesn't care about the state of the kitchen, then we would predict that (215b) would be false in the post-state of (215). However, (215a) would be true regardless of the context of utterance.

Since all eventive predicates can be interpreted as having contingently related pre- or post-states, we correctly predict that all Tatar eventive predicates can combine with *-GAn* 'RESULT' and *-(y)AçAK* 'PROSP.' This additionally accounts for the variety of "evidence"

that speakers can cite for a given proposition: event pre- and post-states can contain both trivially and non-trivially true propositions, all of which can be interpreted as "evidence" for the scope proposition.[10] In the following section, I spell out more formal details of the semantics and pragmatics of my proposal.

5.4 Semantics and pragmatics of the Tatar TAE morphemes

I repeat the set of Tatar TAE suffixes from Chapters 2 and 3 in Table 5.1, including their temporal and evidential interpretations in semantically unembedded environments. Descriptively speaking, these are portmanteau morphemes; that is, they encode multiple components of meaning in a single morpheme. My theory therefore has three goals: first, to account for their evidential interpretations; second, to account for their temporal interpretations; and third, to provide an analysis in which both of these components of meaning are represented in a single semantic denotation, i.e., in a single syntactic head.

	Morpheme	**Interpretation in semantically unembedded environments**
Tenses	*-DI* 'PST'	≈ past, direct evidence
	-A 'PRES'	≈ present; evidentially neutral
	-(y)Er 'FUT'	≈ future, "non-specific" evidence
Aspects	*-GAn* 'RESULT'	≈ past, indirect evidence
	-(y)AçAK 'PROSP'	≈ future, "specific" evidence

Table 5.1: Interpretation of the Tatar TAE suffixes in semantically unembedded environments.

I begin my discussion of the semantics and pragmatics of the Tatar TAE morphemes by spelling out my theoretical assumptions. First, as discussed in Chapter 4, I do not include a modal component in my semantics. I therefore do not include world arguments in any of my following denotations. This is intended only to make my proposal easier to read; worlds could be included without any negative effects.

[10] Nikolaeva (1999) also refers to Nedjalkov (1988)'s notion of trivially and non-trivially true propositions in her discussion of evidentiality in Khanty.

Second, I formally represent Moens and Steedman (1988)'s contingency relation between the event and its pre- and post-states using the notation $\ll$.[11] Given e.g. an event pre-state *s* and an event *e*, the contingency relation between *s* and *e* is represented as $s \ll e$. I define $\ll$ as requiring three major components of meaning: **contingency**, **temporal precedence**, and **adjacency**. All three of these components must be met in order for a state to be a **suitable** pre- or post-state for a given event.

Contingency is defined informally as in Moens and Steedman (1988). Furthermore, I assume that if an event or state *x* is contingent on an event or state *y*, then *y* must temporally precede *x*. I further assume an idealized model in which events and states are strictly ordered with respect to each other, without overlap. This is encoded in part by the adjacency requirement. This requires that contingently related events and states must "touch"; i.e., there cannot be any intervening times between the runtimes of events and their pre- or post-states, and the runtime of an event cannot overlap with the runtimes of either its pre- or post-state.

I formalize this adjacency requirement using the functions *min* and *max* as well as Krifka's runtime function, τ. The *min* and *max* functions combine with an interval (runtime) picked out by τ, and return the earliest and latest values of the interval, respectively (if such values exist).[12] For my purposes, I use *min* and *max* to pick out the beginning and end times of the runtimes of events and states, if applicable. For example, $min(\tau(e))$ returns the minimum (beginning) time of the runtime of *e*, whereas $max(\tau(s))$ returns the maximum (end) time of the runtime of *s*.

I define the adjacency requirement on event pre-states by stating that the maximum time of the runtime of the pre-state *s* must be equal to the minimum time of the runtime of the event *e*, as in (216). I define adjacency for event post-states by stating that the

[11] This notation is borrowed from Bohnemeyer (2014)'s analysis of viewpoint aspect versus relative tense. However, the properties that I ascribe to $\ll$ differ slightly from Bohnemeyer's account.

[12] I noted in §5.2.2 that events and event pre-states are both closed on both upper and lower bounds; hence, it is possible to apply both *min* and *max* to them, since they have defined minimum and maximum values. However, since I assume that event post-states are only closed on their lower bound, only *min* can be applied to them. Since they have no defined maximum value, *max* cannot apply to them.

minimum time of the runtime of the post-state *s* must be equal to the maximum time of the runtime of the event *e*, as in (217). (This is a necessary condition on pre- and post-states, not necessary and sufficient, since contingency and temporal precedence must also be met.)

(216) **Adjacency requirement on pre-states and events:**

s is a suitable pre-state for *e* only if $\max(\tau(s)) = \min(\tau(e))$

(217) **Adjacency requirement on post-states and events:**

s is a suitable post-state for *e* only if $\min(\tau(s)) = \max(\tau(e))$

I assume uncontroversially that the TAE morphemes combine high in the tree, above the bare proposition (i.e., the proposition absent any tense or aspect). I analyze the Tatar tense and aspect suffixes as being of the same semantic type. In this chapter, I treat the TAE suffixes as combining with properties of events and outputing properties of times, i.e., I treat them as being of type $<<v,t>,<i,t>>$.[13] This differs from prior theories of tenses as quantifiers over times, e.g. Prior (1967). (As noted previously, I will eventually propose a more complex type for the Tatar TAE suffixes in the following chapter. This more complex type is motivated by data in which the TAE suffixes are semantically embedded. For the purpose of the unembedded data addressed in this chapter, this more complex type is unnecessary and all of the data can be accounted for by treating them as being of type $<<v,t>,<i,t>>$.)

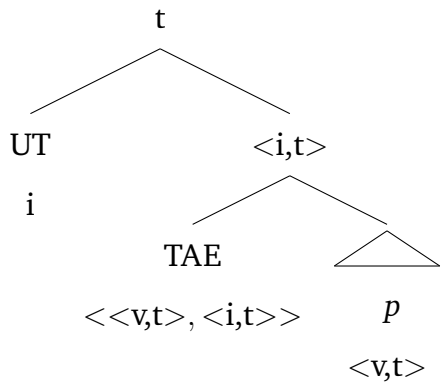
As I showed in Chapters 2 and 3, Tatar expressions are generally not overtly marked for both tense and aspect; it is a morphological property of Tatar that there is only one "slot" for a tense/aspect suffix on the verb.[14] Expressions including only *-DI* 'PST' or only

[13]The properties (and denotations) of the tense suffixes differ from the free past tense morpheme *ide*, which can only combine with stative predicates and is not the main focus of this dissertation. I propose that *ide* relates sets of times, i.e., it is an existential quantifier over times of type $<<i,t>,<<i,t>,t>>$. I discuss this in §5.4.2.

[14]Exceptions to this are when finite verbs are embedded under the free past tense morpheme *ide*, as discussed in Chapter 3. Assuming a left-branching syntax for Tatar, the linear order of *ide* following the verb suggests that it merges higher than the TAE suffixes. This is consistent with it being a tense morpheme, since tenses merge higher than aspects.

-(y)Er 'FUT' are therefore unmarked for aspect. Conversely, expressions including only *-GAn* 'RESULT' or only *-(y)AçAK* 'PROSP' are unmarked for tense. I assume that in both tense-only and aspect-only expressions in matrix clauses, the open time argument is saturated by the utterance time, which I represent extensionally in the tree. I give a basic schema for a matrix Tatar clause in (218).

(218) **Basic schema for a matrix Tatar clause marked (only) for tense/aspect**



Given these basic assumptions, I will now define the semantics that I propose for the Tatar TAE suffixes. I argue that their evidential interpretations arise in part through pragmatic competition, and in part through the link between pre- and post-states and evidentiality as described in the preceding sections.

5.4.1 Present- and future-oriented TAE morphemes

I begin by addressing the semantics of the present- and future-oriented Tatar TAE suffixes, since their interpretations are somewhat simpler to derive than the past-oriented suffixes. As noted in the preceding section, I do not include a modal component in the semantics of these expressions. However, I am sympathetic to modal analyses of the future (e.g. Klecha 2014), and modality could be added to these denotations without any negative effects. My choice to not include a modal component in these expressions is not intended to endorse a non-modal view of the future; it is solely for simplicity.

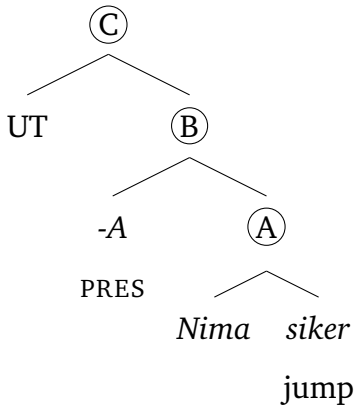
5.4.1.1 Semantics of *-A* 'PRES'

As I noted in Chapter 2, the verbal suffix *-A* 'PRES' is not associated with any evidential reading. I propose that *-A* is simply a marker of present tense, as it has been described previously by Poppe (1961) and Tatevosov (2007) (for Mishar Tatar). I provide a denotation for *-A* in (219). According to this denotation, *-A* combines with a property of events *p* and a time *i* (in matrix clauses, this is valued by the utterance time). *-A* asserts that there exists an event *e*; *i* is located within (or is equal to) the runtime of *e* and *p* is a property of *e*.

(219) $[\![\text{-}A]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[i \subseteq \tau(e)\ \&\ p(e)]$

I give a compositional semantics for the expression in (220) in (221). The semantics for the entire expression are as in (221c).

(220) Nima siker-**e**-∅.
Nima jump-PRES-3SG
'Nima jumps/is jumping.'



(221) a. $[\![\text{Ⓐ}]\!] = \lambda e.\ \text{jump}(\text{Nima})(e)$

b. $[\![\text{Ⓑ}]\!] = \lambda i.\ \exists e[i \subseteq \tau(e)\ \&\ \text{jump}(\text{Nima})(e)]$

c. $[\![\text{Ⓒ}]\!] = 1$ iff $\exists e[\text{UT} \subseteq \tau(e)\ \&\ \text{jump}(\text{Nima})(e)]$

The denotation in (221c) asserts that there exists an event *e* of Nima jumping, and that the utterance time is located within (or is equal to) the runtime of this jumping event.

Since *-A* does not have any evidential reading associated with it, its semantics are not the focus of this dissertation. I therefore do not derive the habitual reading that is available for *-A* expressions (e.g. 'Nima jumps [habitually]'), or its highly productive use as a futurate (as described in Chapter 3). The denotation in (219) simply accounts for the progressive reading of expressions like (220).[15]

5.4.1.2 Semantics of *-(y)Er* 'FUT'

I analyze the TAE suffix *-(y)Er* 'FUT' as a marker of future tense. In Chapter 2, I noted (descriptively speaking) that *-(y)Er*(*p*) expressions tend to have a reading of lower speaker certainty that the described event will occur than *-(y)AçAK*(*p*) 'PROSP' expressions. However, I do not encode any modal semantics in the denotation of this suffix. I propose simply that *-(y)Er* is a "default" suffix for reporting future events, akin to the English simple future tense, and that it pragmatically competes with *-(y)AçAK* 'PROSP' to result in a reading that is compatible with lowered speaker certainty.

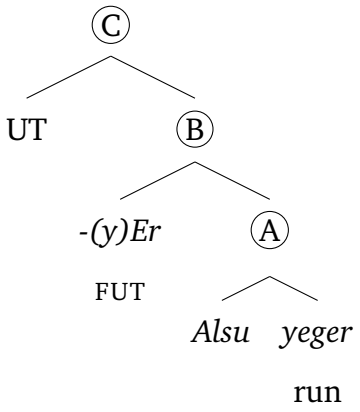
I give a denotation for *-(y)Er* in (222). *-(y)Er* combines with a property of events *p* and a time *i*. It asserts that there exists an event *e*; the time *i* (in matrix clauses, valued by the utterance time) precedes the runtime of *e*, and *p* is a property of *e*.

(222) $[\![\text{-(y)Er}]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[i < \tau(e)\ \&\ p(e)]$

I give a compositional semantics for the expression in (223) in (224). The semantics for the entire expression are as in (224c).

(223) Alsu yeger-**er**-∅.
Alsu run-FUT-3SG
'[I have non-specific evidence that] Alsu will run.'

[15] In Chapter 6, I will use the denotation in (219) to motivate pragmatic competition between *-A* 'PRES' and *-GAn* 'RESULT' in some embedded environments. As noted previously, I will update the semantics of *-GAn* in the following chapter.



(224) a. $[\![\text{Ⓐ}]\!] = \lambda e.\ \text{run}(\text{Alsu})(e)$

b. $[\![\text{Ⓑ}]\!] = \lambda i.\ \exists e[i < \tau(e)\ \&\ \text{run}(\text{Alsu})(e)]$

c. $[\![\text{Ⓒ}]\!] = 1$ iff $\exists e[\text{UT} < \tau(e)\ \&\ \text{run}(\text{Alsu})(e)]$

(224c) asserts that there exists an event e of Alsu running, and that the utterance time precedes the runtime of e. This obtains the future time reading of *-(y)Er*(p) expressions. Unlike *-(y)AçAK*(p) expressions, *-(y)Er*(p) expressions do not invoke a contingently related event pre-state. Since the speaker does not assert that events or states at their utterance time will eventually lead the described event to occur, *-(y)Er*(p) expressions are interpreted as making a "weaker" claim about the future event than *-(y)AçAK*(p) expressions (in a non-technical sense). I ultimately derive the reading of lower speaker certainty associated with *-(y)Er* through explicit pragmatic competition with *-(y)AçAK*; I describe this pragmatic competition in §5.4.1.4.

5.4.1.3 Semantics of *-(y)AçAK* 'PROSP'

I analyze the TAE suffix *-(y)AçAK* as a marker of prospective aspect. Informally speaking, it introduces a contingently related event pre-state (as discussed in §5.3.2), and asserts that the described event is "viewed" from a time within the runtime of this pre-state. As I noted in Chapter 2, *-(y)AçAK*(p) expressions tend to have a reading of high speaker certainty that the described event will occur.

I give a denotation for *-(y)AçAK* in (273e). It combines with a property of events p and

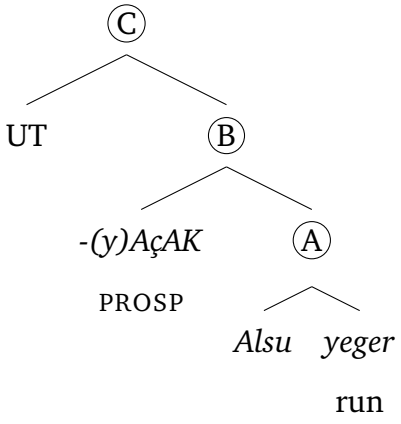
a time *i*, and asserts that there exists an event *e* that is contingent on a state *s*, which is a suitable pre-state for *e*. (I defined what makes a state a "suitable" event pre- or post-state in §5.4; since the contingency relation $\ll$ encodes both contingency and temporal precedence, the suitability requirement ensures that adjacency is met. From now on, I will explicitly state the suitability requirement on *s* only when initially defining the denotations of morphemes that involve pre- or post-states. I will omit it elsewhere for readability, as in e.g. (227).) *-(y)AçAK* additionally asserts that the time *i* is located within the runtime of the pre-state *s*, and that *p* is a property of *e*. In matrix clauses, *i* is valued by the utterance time.

(225) $[\![$*-(y)AçAK*$]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\exists s[s \text{ is a suitable pre-state for } e \ \&\ s \ll e \ \&$
$i \subset \tau(s) \ \&\ p(e)]]$

As I noted previously, the contingency relation $\ll$ conveys both contingency and temporal precedence. If *e* is contingent on *s*, then *s* temporally precedes *e*. This assumption is crucial for both obtaining the future time reading of *-(y)AçAK*(*p*) expressions, and for motivating pragmatic competition between *-(y)AçAK* and *-(y)Er*.

I give a compositional semantics for the expression in (226) in (227). The semantics for the entire expression are as in (227c).

(226) Alsu yeger-**äçäk**-∅.
Alsu run-PROSP-3SG
'[I have specific evidence that] Alsu will run.'



(227) a. $[\![\text{Ⓐ}]\!] = \lambda e.\ \text{run(Alsu)(e)}$

b. $[\![\text{Ⓑ}]\!] = \lambda i.\ \exists e[\exists s[s \ll e \ \&\ i \subset \tau(s)\ \&\ \text{run(Alsu)(e)}]]$

c. $[\![\text{Ⓒ}]\!] = 1$ iff $\exists e[\exists s[s \ll e \ \&\ \text{UT} \subset \tau(s)\ \&\ \text{run(Alsu)(e)}]]$

(227c) asserts that there exists an event *e* and a state *s*; *e* is contingently related to *s* (i.e., *s* is a pre-state of *e*). The utterance time is located in the runtime of this pre-state *s*, and *e* is an event of Alsu running. Events and states that could lead the speaker to believe that they are in a pre-state of the event of Alsu running could be described by propositions like "Alsu says that she is going to run" or "Alsu put on her running clothes and shoes." These stand in a contingent relationship to the event of Alsu running. Observation of these events and states lead to increased speaker certainty that the event of Alsu running will occur.

I note that readings akin to those described for Tatar *-(y)AçAK(p)* expressions (i.e., evidence of a "plan" for the described event, specific evidence that the described event will occur, increased speaker certainty regarding the described event, and so on) are documented for a number of prospective aspectual markers cross-linguistically (Korn and Nevskaya 2017, Nevskaya 2005). Such readings are associated with many of the cognate Turkic TAE suffixes described in Chapter 3. I provide additional examples of prospective markers in Northeastern Neo-Aramaic (Coghill 2010), Syrian Arabic (Jarad 2014), and Khakas (Nevskaya 2005) in (228)-(230) (in the Khakas example in (230), the prospective co-occurs with past tense). Similar prospective data—though lacking clearly defined contexts—is also described for Cherokee (Schreiner and Stone 2014), Karuk (Carpenter 2014), and Tundra Yukaghir (Maslova 2003).

(228) NORTHEASTERN NEO-ARAMAIC (SEMITIC) (modified from Coghill 2010, 11)

Context: The speaker sees gathering clouds.

zi-lə ʾāθə məṭrɒ.
PROSP-L.3MS come.3MS rain
'It's going to rain.'

(Lit. 'Rain is going to come.')

(229) SYRIAN ARABIC (SEMITIC) (modified from Jarad 2014, 106-107)

Context: Plans have been made for a taxi to take the speaker to the airport.

t-taksi **rah** tāxid-na ʕa-l-matār.
the-taxi PROSP take-us to-the-airport
'Taxi's going to take us to the airport.'

(230) KHAKAS (TURKIC) (modified from Nevskaya 2005, 113-114)

Context: The speaker saw the first signs of approaching frosts (e.g. there were no clouds in the sky and a cold wind was blowing from the north).

Soox-tar pol-**arɣa** tur-d-ɪ.
frost-PL be-INFIN/PROSP stand:AUX-PST-3SG
'The frosts were about to come.'

In all of the examples in (228)-(230), the speaker requires some sort of specific evidence that the described future event will occur. This typological data gives cross-linguistic support for a category of prospective aspect, i.e., a grammatical means of reflecting the speaker's belief in some causal relationship between ongoing events/states and a described future event.

5.4.1.4 Deriving evidential readings of *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' through pragmatic competition

I repeat the denotations of *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' in (231)-(232).

(231) $[\![\text{-}(y)Er]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[i < \tau(e)\ \&\ p(e)]$

(232) $[\![\text{-}(y)AçAK]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\exists s[s \ll e\ \&\ i \subset \tau(s)\ \&\ p(e)]]$

Only *-(y)AçAK*(*p*) assertions assert the existence of a causally related event pre-state; *-(y)Er*(*p*) assertions do not. As I noted previously in §5.3.2, when a speaker locates themselves in the pre-state of an event, they believe that some events and states that are true at their utterance time will eventually cause the described event to occur. Since they have access to this "specific" (i.e., contingently related) evidence for the described event, they can be reasonably certain that it will occur. I assume that by using *-(y)AçAK* to assert the

existence of an event pre-state—and given the option of asserting *-(y)Er*(*p*), thereby **not** invoking such an event pre-state— *-(y)AçAK*(*p*) expressions therefore convey high speaker certainty regarding the scope proposition.

I note that this reading of high speaker certainty regarding the scope proposition is always present for *-(y)AçAK*(*p*) expressions in matrix clauses. For instance, it is infelicitous to assert *-(y)AçAK*(*p*) in a context in which the speaker does not believe that they have specific evidence for *p*, as in (233).

(233) **Speculation context**: You have no training in meteorology. While talking with your friends, you look up at the sky and see some clouds. You speculate that it will rain next week.

kiläse atna yañgır {#yaw-**açaq**-∅ / yaw-**ar**-∅}.
next week rain rain-PROSP-3SG / rain-FUT-3SG
'[I have #specific/non-specific evidence that] It will rain next week.'

I will now turn to the issue of accounting for the reading of lower speaker certainty associated with *-(y)Er*(*p*) expressions. According to the denotations in (231)-(232), *-(y)AçAK*(*p*) assertions entail *-(y)Er*(*p*) assertions. This falls out from the assumption that the contingency relation $\ll$ also indicates temporal precedence. Given this assumption, the denotation for *-(y)AçAK* 'PROSP' in (232) expresses that the reference time is located in a state that temporally precedes the event runtime. This entails that the reference time precedes the event runtime; i.e., the temporal relation between reference time and event runtime that is encoded in the denotation for *-(y)Er* 'FUT' in (231).

Given this entailment relation between *-(y)AçAK* and *-(y)Er*, we predict that *-(y)Er*(*p*) expressions can undergo pragmatic strengthening through scalar implicature. Scalar implicatures are a variety of quantity implicature based on the assumption (following Grice's Maxim of Quantity) that cooperative speakers will make the most informative truthful assertion possible, given their beliefs. According this assumption, when speakers make less informative assertions than they theoretically could have, we can infer that the speaker believes that the stronger alternative(s) to their assertion are false. Scalar implicature has

been discussed at length in the pragmatic literature (Hirschberg 1985; Grice 1989, 1975; Horn 1972, among many others).

For the purpose of this dissertation, I assume Horn (1972)'s recipe for calculating scalar implicature. According to Horn (1972), lexical items can belong to "scales" of relative strength, which are determined by their entailment relations. A classic example of a Horn scale is one containing the quantifiers *some* and *all*, where *all* entails *some* (i.e., *some* is "weak").[16] The standard recipe for scalar implicature is one in which weak expressions are strengthened by being conjoined with the negation of the expression containing their stronger scalar alternative. I give a basic schema for this in (234), and show an example of this in (235).[17] (I use $\rightsquigarrow$ to indicate an implicature.)

(234) **Pragmatically strengthened meaning of *some* expressions, by scalar implicature**

$\text{some}_{strengthened} \rightsquigarrow$ some & ¬all

(235) Kirk ate $\text{some}_{strengthened}$ of the satsumas.

$\rightsquigarrow$ Kirk ate some of the satsumas & ¬Kirk ate all of the satsumas

Following the schema in (234), "weak" *-(y)Er*(*p*) expressions are strengthened through conjunction with the negation of the stronger *-(y)AçAK*(*p*) expression. Pragmatically strengthened *-(y)Er*(*p*) expressions therefore convey the meaning in (236).

(236) **Pragmatically strengthened meaning of *-(y)Er(p)* assertions, by scalar implicature**

$\text{-(y)Er(p)}_{strengthened} \rightsquigarrow$ -(y)Er(p) & ¬ -(y)AçAK(p)

With this pragmatic tool in hand, we can now account for the reading of lowered speaker certainty associated with Tatar *-(y)Er*(*p*) expressions. According to (236), when

[16] For the purpose of this discussion, I assume a simplified version of this scale that **only** contains *some* and *all*.

[17] I omit the component of speaker belief in (234)-(235) for simplicity.

a speaker asserts *-(y)Er*(*p*), they assert that they believe that the described event will occur in the future of the utterance time, and implicate that they do not believe that the utterance time is located in a contingently related pre-state of the event. That is, (236) implicates that the speaker does not believe that any events or states at the utterance time will cause the described event to eventually occur. As a result, there is a reading of lower speaker certainty that the described event will occur. As I showed in Chapter 2, this level of speaker certainty is not as weak as in an assertion of epistemic possibility (i.e., a *might(p)* expression); indeed, *-(y)Er(p)* assertions are the "default" strategy by which Tatar speakers talk about future events. However, they do not make as strong of a claim about future events as *-(y)AçAK(p)* assertions do.

Since this strengthened meaning of *-(y)Er*(*p*) is derived through scalar implicature, we predict that it is optional (Grice 1989). That is, we predict that the strengthened reading of *-(y)Er*(*p*) should be absent in some contexts. This prediction is borne out; I show in (237) that the implicature is cancellable. In (237b), the speaker cancels the strengthened reading of *-(y)Er* (as in (236)) by explicitly asserting *-(y)AçAK*(*p*).

(237) **Casino context:**

a. You and Güzäl are going to a casino in Las Vegas tomorrow. You know that all the games are run by chance, and that you have no way of predicting whether or not you will win. You say:

Güzäl irtägä aqça ciñ-**er**-∅.
Güzäl tomorrow money win-FUT-3SG
'[I have non-specific evidence that] Güzäl will win some money tomorrow.'

b. However, Timur knows how to rig the games, and will do so that she wins. He responds by saying:

Güzäl irtägä aqça ciñ-**er**-∅... Güzäl älbättä aqça ciñ-**eçek**-∅!
Güzäl tomorrow money win-FUT-3SG Güzäl of.course money win-PROSP-3SG
'[I have non-specific evidence that] Güzäl will win some money tomorrow... in fact, [I have specific evidence that] Güzäl will win some money!'
(Alternate translation: 'Güzäl might win some money tomorrow... in fact, Güzäl will definitely win some money!')

In (237a), the speaker uses *-(y)Er* 'FUT' to assert that Güzäl will win money. Their choice to use *-(y)Er* expresses that they have lower certainty that this event of winning money will occur than had they used *-(y)AçAK*; that is, they do not locate themselves in a contingently related pre-state of the event of Güzäl winning money. However, Timur (i.e., the speaker of (237b)) believes that he **is** located in a pre-state of the event of Güzäl winning money, since he plans to rig the games. As a result, Timur is able to felicitously assert a *-(y)Er*(*p*) expression and then cancel its pragmatically strengthened reading by overtly asserting that *-(y)AçAK*(*p*). I take (237) to suggest that a pragmatic analysis is appropriate for the Tatar data.[18]

[18]While the strengthened meaning of *-(y)Er* is cancellable, it is not contextually defeasible. That is, a speaker cannot assert *-(y)Er*(*p*) if the context is such that they have "specific" evidence for *p* (i.e., if they have evidence that would license the use of *-(y)AçAK*). I show this in (238).

(238) **Moscow trip context**: Your friend Mansur bought plane tickets to Moscow (i.e., you have specific evidence that he is going to go to Moscow). While discussing his upcoming travel plans, you say:
Mansur Mäskäü-gä {#bar-**ır**-∅ / bar-**açaq**-∅}.
Mansur Moscow-DAT go-FUT-3SG / go-PROSP-3SG
'[I have #non-specific/specific evidence that] Mansur will go to Moscow.'

I suggest that the lack of defeasibility of the implicature associated with *-(y)Er*(*p*) expressions is linked to the use of *-(y)Er*(*p*) to convey lower speaker certainty regarding the scope proposition. If a speaker asserts *-(y)Er*(*p*) in a context in which they have specific evidence for the scope proposition, they are making a claim that is "too weak" for the evidence at hand; i.e., they are violating the Gricean Maxim of Quantity. Thus, while the pragmatically strengthened reading of a *-(y)Er*(*p*) expression can be explicitly cancelled in discourse, it is not contextually defeasible.

5.4.2 Past-oriented TAE morphemes

I give the semantics of the past-oriented Tatar TAE morphemes in the following sections. As I noted in Chapter 3, the two past tense morphemes *-DI* 'PST' and *ide* 'PST' contrast with respect to their ability to combine with eventive versus stative predicates. *-DI* can only combine with eventive predicates, whereas *ide* is primarily restricted to combining with stative (i.e., nominal and adjectival) predicates. I account for this distribution by proposing that *-DI* must combine with a property of events, whereas *ide* must combine with a property of times.

I analyze the TAE suffix *-GAn* 'RESULT' as a marker of resultative aspect. Conceptually speaking, I treat resultative aspect as a "mirror image" of the prospective aspect. (As noted previously, I will eventually revise the denotation of *-GAn* in Chapter 6 to account for its interpretation in embedded verbal nominalizations.)

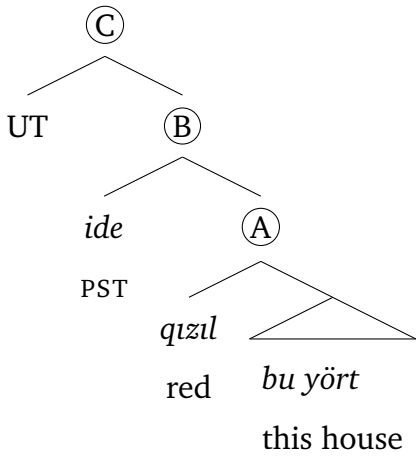
5.4.2.1 Semantics of *ide* 'PST'

This dissertation is primarily concerned with the semantics and pragmatics of the Tatar TAE suffixes; as such, the free morpheme *ide* is of secondary concern. I treat *ide* as a marker of simple past tense. I give a denotation for *ide* in (239) in which I analyze it as an existential quantifier over times, of type $<<i,t>,<<i,t>,t>>$. It combines with a property of times p and a time i'; it asserts that there exists a time i that precedes the time i', and that p is true at this earlier time i. In matrix clauses, i' is valued by the utterance time, which I represent extensionally in the tree.

(239) $[\![ide]\!] = \lambda p_{<i,t>} \lambda i'.\ \exists i[i < i' \ \&\ p(i) = 1]$

I give a compositional semantics for the *ide* expression in (240) in (241). The semantics for the entire expression are as in (241d). I propose that stative properties like *red* combine with a time argument that *ide* can bind, as in (241a).

(240) Bu yört qızıl **ide**-∅.
this house red PST-3SG
'This house was red.'



(241) a. $[\![qızıl]\!] = \lambda x \lambda i.\ \text{red}(x)(i)$

b. $[\![\text{Ⓐ}]\!] = \lambda i.\ \text{red}(\text{this house})(i)$

c. $[\![\text{Ⓑ}]\!] = \lambda i'.\ \exists i[i < i' \ \&\ \text{red}(\text{this house})(i) = 1]$

d. $[\![\text{Ⓒ}]\!] = 1$ iff $\exists i[i < \text{UT} \ \&\ \text{red}(\text{this house})(i) = 1]$

Informally speaking, (241d) asserts that there exists a time *i* that precedes the utterance time, and that the house is red at this past time *i*. I assume that the cessation implicature associated with *ide* (described in Chapter 3) arises pragmatically. (While I don't spell out the details of how this cessation implicature is derived, I note that such temporal implicatures have been previously described and analyzed by Altshuler and Schwarzschild 2013 and Thomas 2012.)

5.4.2.2 Semantics of *-DI* 'PST'

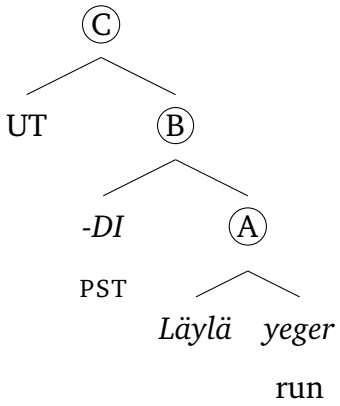
The distribution of *-DI* 'PST' contrasts with *ide* 'PST' in that it can only combine with eventive predicates, i.e., properties of events of type <*v,t*>. I give a denotation for *-DI* in (242); I treat it as a simple past tense, analogous to the future tense *-(y)Er* 'FUT.' *-DI* combines with a property of events *p* and a time *i*. It asserts that there exists an event *e*; the time *i* (in

matrix clauses, valued by the utterance time) follows the runtime of *e*, and *p* is a property of *e*.

(242) $[\![\text{-}DI]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\tau(e) < i\ \&\ p(e)]$

I give a compositional semantics for the *-DI* expression in (243) in (244). The semantics for the entire expression are as in (244c).

(243) Läylä yeger-**de**-∅.
Läylä run-PST-3SG
'[I have direct evidence that] Läylä ran.'



(244) a. $[\![\text{A}]\!] = \lambda e.\ \text{run(Läylä)(e)}$

b. $[\![\text{B}]\!] = \lambda i.\ \exists e[\tau(e) < i\ \&\ \text{run(Läylä)(e)}]$

c. $[\![\text{C}]\!] = 1$ iff $\exists e[\tau(e) < \text{UT}\ \&\ \text{run(Läylä)(e)}]$

(244c) asserts that there exists an event *e* of Läylä running, and that the runtime of *e* precedes the utterance time. This successfully derives the past tense reading of (243) by locating the the event of Läylä running prior to the utterance time.

I note that the semantics that I propose for both *-DI* and *ide* do not encode any component of direct evidential meaning. The evidential interpretation of both morphemes is derived via pragmatic competition with other morphemes. In §5.4.2.4, I derive the pragmatic competition between *-DI* 'PST' and its alternative *-GAn* 'RESULT.' (The direct evidential interpretation of *ide* likely arises through competition with the free indirect evidential

particle *ikän*; in this dissertation, I focus primarily on deriving the evidential interpretations of *-DI* and *-GAn*, and leave the interaction of *ide* and *ikän* for future work.)

5.4.2.3 Semantics of *-GAn* 'RESULT'

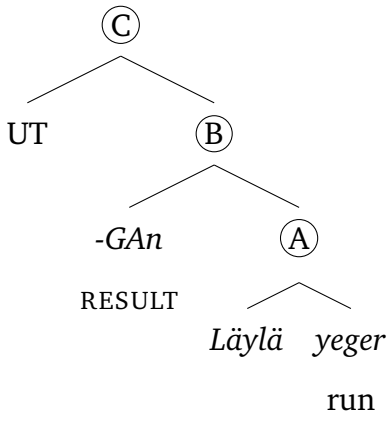
I analyze the Tatar TAE suffix *-GAn* as a marker of resultative aspect; I effectively treat it as a "mirror image" of the prospective aspect *-(y)AçAK* 'PROSP.' Informally speaking, *-GAn* introduces a contingently related event post-state (as discussed in §5.3.1), and asserts that the time from which the described event is "viewed" is located within the runtime of this post-state.

I give a denotation for *-GAn* in (245). It combines with a property of events *p* and a time *s*, and asserts that there exists an event *e* and state *s*, where *s* is a suitable post-state for *e*. *s* is contingent on *e*, as indicated by the notation $\ll$. Since the contingency relation also encodes temporal precedence, this also encodes that *e* temporally precedes *s*. *-GAn* asserts that *i* (in matrix clauses, the utterance time) is located within the runtime of *s*, and that *p* is a property of *e*.

(245) $[\![\text{-GAn}]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\exists s[s \text{ is a suitable post-state for } e \ \&\ e \ll s \ \&$
$i \subset \tau(s)\ \&\ p(e)]]$

I give the compositional semantics of the *-GAn*(*p*) expression in (246) in (247); the semantics for the entire expression are as in (247c).

(246) Läylä yeger-**gän**-∅.
Läylä run-RESULT-3SG
'[I have indirect evidence that] Läylä ran.'



(247) a. $[\![\text{A}]\!] = \lambda e.\ \text{run}(\text{Läylä})(e)$

b. $[\![\text{B}]\!] = \lambda i.\ \exists e[\exists s[e \ll s \ \&\ i \subset \tau(s)\ \&\ \text{run}(\text{Läylä})(e)]]$

c. $[\![\text{C}]\!] = 1$ iff $\exists e[\exists s[e \ll s \ \&\ \text{UT} \subset \tau(s)\ \&\ \text{run}(\text{Läylä})(e)]]$

(247c) asserts that there exists an event *e* and a state *s*; *s* is contingently related to *e* (i.e., *s* is a post-state of *e*). The utterance time is located within the runtime of the post-state *s*, and *e* is an event of Läylä running. This denotation successfully derives the past-oriented meaning of *-GAn* in matrix clauses; given the temporal precedence relation encoded by $\ll$, the runtime of *e* necessarily precedes the runtime of *s*. Since the utterance time is located within the runtime of the post-state *s*, the utterance time therefore follows *e*.

Events and states that could lead the speaker to believe that they are in a post-state of the event of Alsu running could be described by propositions like "Alsu is sweaty" or "Alsu came home wearing running shoes." These stand in a contingent relationship to the event of Alsu running. In evidential terminology, these propositions could be described as "indirect evidence" for the scope proposition.

Such indirect evidential readings are commonly associated with past-oriented morphemes that are glossed (synchronically or diachronically) as "resultatives," "perfects," "completive aspects," and so on (Lindstedt 2000, Bybee and Dahl 1989, Nedjalkov 1988, among many others). Izvorski (1997) explicitly states that languages can exhibit "perfect of evidentiality," in which so-called perfect expressions also express indirect evidentiality.

In addition to many of the cognate Turkic suffixes noted in Chapter 3, indirect evidential readings have also been described for past-oriented expressions in Cherokee (Pulte 1985), Dogon (Plungian 1988), Balkan languages (Friedman 1986), Baltic languages (Lindstedt 2000), and Scandinavian languages (Dahl 2000). In (248)-(249), I provide examples of portmanteau morphemes that pattern similarly to Tatar *-GAn* in conveying both past time reference and indirect evidential meaning.

(248) KALAALLISUT (ESKIMO-ALEUT) (modified from Fortescue 2003, 293)

Context: The speaker goes outside and sees a pool of water.

siallir-**sima**-vuq.
rain-PERF-3SG.INDIC
'[I have indirect evidence that] It rained.'

(249) KHANTY (URALIC) (modified from Nikolaeva 1999, 10)

Context: The speaker sees the cut down tree.

Ma werl-əm-em-na aś-em tăm jŭx ewət-**m**-al.
I wake.up-PART-1SG-LOC father-1SG this tree cut.down-EVID.PST-3SG
'When I woke up, [I have indirect evidence that] my father already cut down that tree.'

Like Tatar *-GAn*(*p*) expressions, the examples in (248)-(249) are such that the speaker must have (what is described in the evidential literature as) "indirect evidence" for the scope proposition. For example, Nikolaeva (1999) notes that (249) is inappropriate if the speaker witnessed their father cutting down the tree. The cross-linguistic prevalence of such portmanteau indirect evidential and past-oriented morphemes supports the project of encoding both temporal and evidential meaning in a single denotation, as in (245).

5.4.2.4 Deriving evidential readings of *-DI* 'PST' and *-GAn* 'RESULT' through pragmatic competition

I repeat the denotations of *-DI* 'PST' and *-GAn* 'RESULT' in (250)-(251).

(250) $[\![\text{-}DI]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\tau(e) < i\ \&\ p(e)]$

(251) $[\![\text{-}GAn]\!] = \lambda p_{<v,t>} \lambda i.\ \exists e[\exists s[e \ll s \ \& \ i \subset \tau(s) \ \& \ p(e)]]$

Of the two past-oriented TAE suffixes, only *-GAn* asserts the existence of a causally related event post-state. I assume that cooperative speakers obey Grice (1975)'s conversational Maxims, specifically the Maxims of Quality ("be truthful") and Quantity ("make your contribution (only) as informative as required"). Asserting *-GAn*(*p*) therefore conveys that the speaker's belief in the event described by *p* is based on their perception of the contingently related event post-state. Since a *-GAn*(*p*) expression describes an event that the speaker could in principle have directly perceived—and which the speaker could have described using *-DI*, which does not invoke an event post-state— *-GAn* 'RESULT' conveys that the speaker has **only** perceived the event's post-state, rather than the event itself. That is, in evidential terms, the use of *-GAn* conveys that the speaker has indirect evidence for the scope proposition.

In matrix clauses, the indirect evidential reading associated with *-GAn* is always present. For instance, it is impossible for a speaker to explicitly cancel its indirect evidential interpretation by asserting a *-GAn*(*p*) expression, and then following it with a *-DI*(*p*) expression. I show this in (252).

(252) **Context**: You found a train ticket to Moscow in Mansur's desk, and you infer from this that he went to Moscow. Then, you remember that you actually went to the train station with him and saw him leave.

Mansur Mäskäü-gä bar-**ğan**-∅... #Mansur älbettä Mäskäü-gä bar-**dı**-∅.
Mansur Moscow-DAT go-RESULT-3SG Mansur of.course Moscow-DAT go-PST-3SG
'[I have indirect evidence that] Mansur went to Moscow... #in fact, [I have direct evidence that] Mansur went to Moscow.'

The indirect evidential reading associated with *-GAn* is also not contextually defeasible. It is infelicitous for a speaker to assert *-GAn*(*p*) in a context in which they directly perceived the event described by *p*, as shown in (253).[19]

[19] *-GAn* is unavailable in direct perception contexts like (253) regardless of the speaker's belief in the scope proposition. That is, even in a context in which the speaker "cannot believe their eyes," they still must use *-DI* to report the observed event. I gave examples of such contexts in Chapter 2, in which I showed that Tatar speakers use *-DI* even when reporting outlandish hallucinations.

(253) **Fishing trip context**: You go fishing with your family. You see your son, Nima, catch a fish. Later, you tell your mother that Nima caught a fish. You say:

Nima balıq {#tot-**qan**-∅ / tot-**tı**-∅}.
Nima fish catch-RESULT-3SG / catch-PST-3SG
'[I have #indirect/direct evidence that] Nima caught a fish.'

I now turn to deriving the direct evidential reading of *-DI*, which I noted previously is slightly more complex. I ground this evidential reading in two major conceptual points. First, I reiterate my previous observation that the past oriented Tatar TAE suffixes fundamentally differ from the future oriented TAE suffixes; since the described event of a past-oriented expression occurred in the past of the utterance time, it is therefore possible that the speaker **could** have directly witnessed the event. This is impossible for expressions describing events in the future of the utterance time. In evidential terminology, speakers can have both direct and indirect evidence for past events, but can only have indirect evidence for future events.

Second, as discussed previously in §5.3, I assume that **all** speaker beliefs fundamentally stem from sensory perception.[20] With respect to events in the past of the utterance time, this could either take the form of perception of the described event itself, or perception of the event's post-state. In evidential terminology: for a speaker to believe that *p*, I assume that they must have either direct or indirect evidence (or both) for the event that *p* describes.

Given these assumptions, we can now derive the direct evidential reading of *-DI*(*p*) expressions. Like *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' in §5.4.1.4, the denotations of *-DI* and *-GAn* in (250)-(251) are such that *-GAn* entails *-DI*: if the utterance time is located in the runtime of a contingently related event post-state, then the utterance time necessarily follows the event runtime. Since *-GAn* entails *-DI*, *-DI*(*p*) expressions can undergo pragmatic strengthening through scalar implicature. Assuming a cooperative speaker, a "weak" *-DI*(*p*)

[20]My theory differs in this respect from e.g. Lee (2013, 2011) and Matthewson (2011). These authors differentiate between propositions that the speaker believes due to sensory perception and propositions that the speaker believes due to internal reasoning/inference.

assertion is strengthened by being conjoined with the negation of the stronger *-GAn*(p) expression. Pragmatically strengthened *-DI*(p) expressions therefore convey the meaning in (254).

(254) **Pragmatically strengthened meaning of *-DI*(p) assertions, by scalar implicature**

$$\textit{-DI}(p)_{strengthened} \rightsquigarrow \textit{-DI}(p) \;\&\; \neg\, \textit{-GAn}(p)$$

When a speaker asserts *-DI*(p), by the strengthening mechanism in (254), they assert that the event occurred in the past of the utterance time, and implicate that their utterance time is not located in a contingently related event post-state. As noted above, *-GAn*(p) expressions convey that the speaker has indirect evidence for the described event. In evidential terms, the pragmatically strengthened meaning of *-DI*(p) in (254) implicates that the speaker does not have indirect evidence for the described event. Since I assume that all assertions are fundamentally based on either perception of the described event ("direct evidence") or perception of its post-state ("indirect evidence"), and (254) denies that the speaker has perceived the event post-state, it therefore implicates that the speaker has perceived the event itself. In evidential terms, $\textit{-DI}(p)_{strengthened}$ implicates that the speaker has direct evidence for the scope proposition.

Since this direct evidential reading is derived through scalar implicature, we again predict that it should be optional; that is, we predict that it should be absent in some contexts (Grice 1989). This prediction is borne out; the direct evidential reading of *-DI* is defeasible in certain contexts.[21] I show examples of such contexts in (256)-(257) (repeated from

[21] *-DI*(p) expressions pattern oppositely from *-(y)Er*(p) expressions in that their pragmatically strengthened reading is defeasible, but not cancellable. That is, a speaker cannot assert a *-DI*(p) expression and then explicitly cancel the implicature in (254) by asserting *-GAn*(p). I show this in (255).

(255) **Context**: You go to the train station with Mansur and see him leave to go to Moscow. Then, you find a receipt for Mansur's train ticket to Moscow. You say:
Mansur Mäskäü-gä bar-**dı**-∅... #Mansur älbettä Mäskäü-gä bar-**ğan**-∅.
Mansur Moscow-DAT go-PST-3SG Mansur of.course Moscow-DAT go-RESULT-3SG
'[I have direct evidence that] Mansur went to Moscow... #in fact, [I have indirect evidence that] Mansur went to Moscow.'

At present, I do not have a full account of the infelicity of examples like (255) (though I note that similar data has been described in other languages, e.g. Matthewson et al. 2007, 30-31 for St'át'imcets). However,

Chapter 2). In these contexts, the speaker is highly confident (based on their reasoning with respect to the events and states that they perceive) that the event described by the scope proposition occurred. Despite the fact that the speaker did not witness the event described by the scope proposition, they nonetheless assert the proposition using *-DI*. (*-GAn* is also felicitous in the contexts in (256)-(257).)

(256) **Highly reliable source context:** Läylä's friend Wäğıyz works for a news channel as a sports broadcaster. Wäğıyz told her that Michael Phelps won a race in the Rio Olympics. Läylä did not watch the race herself, but she considers Wäğıyz as a good source on events in sports. She can tell another person:

Michael Phelps ciñ**-de**-∅.
Michael Phelps win-PST-3SG
'[I have direct evidence that] Michael Phelps won.'

(257) **Widely known facts context:** You are a history professor lecturing your class on WWII.

Germaniya Polşa-nı bas**-tı**-∅.
Germany Poland-ACC invade-PST-3SG
'[I have direct evidence that] Germany invaded Poland.'

I take data like (256)-(257) to suggest that evidentiality is not strictly encoded as part of truth-conditional meaning in Tatar, and involves some pragmatic component.

---

this inability to follow a *-DI*(*p*) assertion with a *-GAn*(*p*) assertion could fall out from something like Davis et al. (2007)'s or Faller (2012)'s concept of evidential strength. When a speaker asserts a *-DI*(*p*) expression, they convey that they have direct evidence for the scope proposition. Direct evidence for a proposition is considered "stronger" than indirect evidence for the same proposition; since cooperative speakers should make the strongest claims possible, it is infelicitous to follow a "strong" expression of direct evidence for *p* with a "weak" expression of indirect evidence for *p*.

This contrasts with *-(y)Er*(*p*) and *-(y)AçAK*(*p*) assertions, as described in §5.4.1.4. As I showed in this section, *-(y)Er*(*p*) assertions are interpreted as making "weaker" claims than *-(y)AçAK*(*p*) assertions with respect to the speaker's belief that the scope proposition will occur. It is therefore possible for a speaker to felicitously follow a "weak" *-(y)Er*(*p*) assertion with a "strong" *-(y)AçAK*(*p*) assertion. However, the contrast between *-DI*(*p*) and *-GAn*(*p*) expressions is not only one of entailment, but also—in the opposite direction—one of evidential strength.

5.4.3 Evaluating my theory against some proposed tests for modality in the Tatar TAE suffixes

In Chapter 4, I laid out some tests that have been previously used to diagnose modality in evidential expressions, and showed how the Tatar TAE expressions pattern relative to these tests. Now that I have spelled out the details of my theory, I can evaluate the ability of my theory to account for the data described in this section. (I address the embedding data in Chapter 6; I omit an account of the modal subordination data, since my discussion in Chapter 4 showed how it can be accounted for due to a mismatch between the speaker's evidence and choice of TAE suffix.)

5.4.3.1 Ability to assert TAE(*p*) if the speaker believes that *p* is true

According to the denotations of the TAE suffixes that I proposed in this chapter, when a speaker asserts a proposition using a TAE suffix, they also assert the proposition itself. This accounts for the ability of speakers to assert TAE(*p*) expressions when they believe that *p* is true (contrary to Matthewson et al. 2007, 20's proposed test for modality in evidential expressions).

I show an example of *-DI* being used to report a proposition that the speaker believes is true in (258), and an example of *-GAn* being used to report a proposition that the speaker believes is true in (259).

(258) **Fishing trip context**: You go fishing with your family. You see your son, Nima, catch a fish. Later, you tell your mother that Nima caught a fish. You say:

Nima balıq tot-**tı**-∅.
Nima fish catch-PST-3SG
'[I have direct evidence that] Nima caught a fish.'

(259) Germaniya Polşa-nı bas-**qan**-∅.
Germany Poland-ACC invade-RESULT-3SG
'[I have indirect evidence that] Germany invaded Poland.'

5.4.3.2 Inability to assert TAE(*p*) if the speaker believes that *p* is false

Another consequence of my analysis of TAE(*p*) expressions as asserting *p* is that speakers should be unable to cooperatively assert a TAE(*p*) expression if they believe that *p* is false. I showed in Chapter 2 that in reportative contexts, Tatar speakers can use *-GAn* to express a proposition even if they are not certain that the proposition is true (due to e.g. not trusting the source of the report). However, it is infelicitous for a speaker to assert *-GAn*(*p*) if they are **certain** that the proposition is false.[22]

I show an example of this infelicity in (261). In this context, the speaker is certain that the scope proposition "Aliens abducted Alsu" is false. It is therefore infelicitous for the speaker to assert (261) without the inclusion of the clause-final particle *imeş*. (Tatevosov 2007 (writing on Mishar Tatar) describes *imeş* as a marker of uncertainty and indirect evidentiality.)

[22] AnderBois (2014) shows that in a range of unrelated languages, it is possible for speakers to follow REPORTATIVE.EVID(*p*) assertions with either ¬DIRECT.EVID(*p*) or plain ¬*p* continuations. My primary Tatar consultant reports that analogous Tatar expressions are infelicitous without including the particle *imeş* in the initial assertion. However, in reportative contexts, speakers can follow *-GAn*(*p*) expressions with assertions that they do not believe that *p*. I show an example of this in (260).

(260) Unreliable source context: Wäğıyz told you that Läylä went to Kazan, but you think that Wäğıyz is a compulsive liar and you don't really trust him. Later, you say:
Läylä Qazan-ğa kit-**kän**-∅, läkin min aña ışan-m-ıy-m.
Läylä Kazan-DAT leave-RESULT-3SG but 1SG.NOM 3SG.DAT believe-NEG-PRES-1SG
'[I have indirect evidence that] Läylä went to Kazan, but I don't believe it.'

AnderBois (2014) argues that reportative evidentials facilitate pragmatic perspective shift from the speaker to the source of the reported proposition, and speculates that the availability of expressions like (260) could be universal to all languages with reportative evidentials. While I do not provide an analysis of Tatar data like (260) in this dissertation, I note that it is typologically unexceptional and could potentially be accounted for by applying AnderBois (2014)'s theory to Tatar.

(261) **Alien abduction context** (adapted from Smirnova 2011, 277): You visited your friend Alsu in a psychiatric clinic. Alsu was hospitalized because of severe hallucinations and thinks that she was abducted by aliens. You do not believe that this actually occurred, since you do not believe that aliens exist. When a friend asks you what is going on with Alsu, you say:

başqa planeta-da yäşäüçe-lär Alsu-ne url-ap al-ıp kit-**kän**-när
other planet-LOC resident-PL Alsu-ACC steal-IP take-IP leave-RESULT-3PL
#(**imeş**).
IMEŞ
'[I have indirect evidence that] Aliens abducted Alsu [and I don't think it happened].'

The infelicity of (261) (without *imeş*) is expected if TAE(p) expressions assert p.

5.4.3.3 Inability to scope under negation

I showed in Chapter 2 that the evidential contribution of the Tatar TAE suffixes scopes above sentential negation, as illustrated in (262). Similar data is described for evidential expressions in a range of other languages (Murray 2010 for Cheyenne; Matthewson et al. 2007 for St'át'imcets; Faller 2002 for Cuzco Quechua; Izvorski 1997 for Bulgarian, among others). Accounting for data like (262) is therefore a common issue for theories of evidentiality.

(262) Mansur Mäskäü-gä bar-**ma-ğan**-∅.
Mansur Moscow-DAT go-NEG-RESULT-3SG
a. = '[I have indirect evidence that] Mansur didn't go to Moscow.'
b. ≠ '[It is not the case that I have indirect evidence that] Mansur went to Moscow.'

Negation data like (262) is not straightforwardly predicted by my current theory of Tatar evidentiality; however, my theory can be adjusted to account for such data.

The issue encountered by the theory I propose in §5.4 is akin to the issues observed by Partee (1973) on the interaction of tense and negation in English. Partee (1973, 602-603)

noted that when an English speaker says the sentence in (263), they are communicating that there is a particular time in the past at which they did not turn off the stove.

(263) I didn't turn off the stove.

However, this reading is not expected if tenses are analyzed as existential quantifiers (as by e.g. Prior 1967). If tenses are treated as existential quantifiers, the predicted interpretation of (263) would instead be one of the following two readings in (264). Neither of the possible scope relations between negation and the existential quantifier in (264) accurately reflect the meaning of (263).[23]

(264) a. $\exists > \neg$: There exists a time in the past at which I did not turn off the stove

b. $\neg > \exists$: There does not exist a time in the past at which I turned off the stove

In §5.4, I utilized existential quantifiers over events and states in my denotations of the Tatar TAE suffixes. The analysis I have currently posited therefore encounters the same issues that Partee (1973) described for English data like (263), and does not predict the ability of the evidential meaning to scope above negation, as in (262).

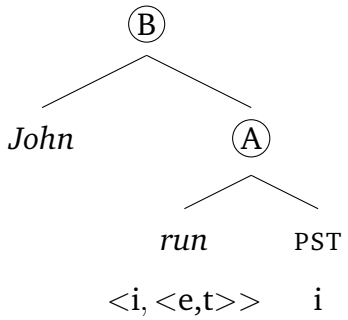
In Partee (1973) (and in later work on tense, e.g. Kratzer 1998, Heim 1994), tenses are treated as **pronouns** rather than existential quantifiers. Generally speaking, tenses are treated as being pronouns of type *i*. They are valued by an assignment function *g* and are accompanied by a presupposition that locates them in time. In the basic pronominal denotation for past tense in (265), the presupposition for the time i' locates i' in the past of a time *i* that is valued by a parameter on the interpretation function that sets *i* equal to the utterance time.

(265) $[\![\text{PST}]\!]^{g,i} = \text{i}'\text{: i}' < \text{i}$

[23]This problem is avoided if we treat quantifiers as having restricted domains (von Fintel 1994); however, this cannot save the evidential interpretation of (262).

I give a simple derivation for a past tense expression in (267). The denotation for the entire expression is as in (267b).

(266) John ran.

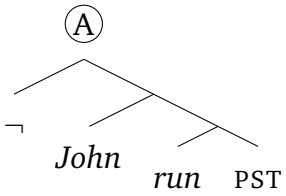


(267) a. $[\![\text{A}]\!]^{g,i}$ = λx. run(x) at i′: i′ < i

b. $[\![\text{B}]\!]^{g,i}$ = 1 iff run(John) at i′: i′ < i

Given a pronominal, presuppositional analysis of tense, we can now explain the data in (263). When an expression like *I turned off the stove* is negated, the presupposition of the tense projects above negation. That is, an expression like *John didn't run* in (268) refers to a specific time *i*′ in the past of the utterance time at which John did not run. I give a denotation for (268) in (269).

(268) John didn't run.



(269) $[\![\text{A}]\!]^{g,i}$ = 1 iff if it is not the case that run(John) at i′: i′ < i

This presuppositional mechanism can also be utilized to account for the interpretation of the Tatar TAE suffixes in negated clauses.[24] In brief: we can treat the Tatar TAE suffixes as pronouns of type *i*, akin to the denotation for the English past tense in (265). The temporal relations that I have encoded as assertions in the denotations in §5.4 would instead

[24]This is similar to what Izvorski (1997) proposes for Bulgarian; Izvorski accounts for the interaction of the Bulgarian indirect evidential with negation by encoding its indirect evidential meaning as a presupposition.

be encoded as presuppositions. As presuppositions, they would project above negation. Events and states would be treated as free variables that would be valued by parameters on the assignment function.

Under this pronominal analysis, the TAE suffixes would all make the same assertion. This analysis would therefore require assuming the availability of pragmatic competition between presuppositions. This is not proposed by canonical theories of pragmatic competition (e.g. Grice 1975, Horn 1972), which assume that pragmatic competition occurs between assertions. While this proposal is not fully fleshed out, it provides a conceptual sketch of how the negation data in (262) can be accounted for under an aspectual analysis of Tatar evidentiality. As I noted previously, the question of how to account for the ability of evidentials to scope above negation is a common problem for theories of evidentiality.

5.5 Conclusion

In this chapter, I gave a semantic and pragmatic analysis of the Tatar TAE system, focusing primarily on the set of four core portmanteau verbal suffixes. My analysis joins and formalizes two ideas discussed previously in the evidential and tense/aspect literature. The first idea (discussed previously by e.g. Nikolaeva 1999) is that indirect evidentials indicate the speaker's observation of the results of some preceding situation. The second idea (discussed previously by e.g. Comrie 1976) is that so-called "perfect"/"resultative"/"completive" aspects relate ongoing states to previous situations, and that prospective aspects relate ongoing states to future situations. I gave an analysis of these previously unformalized intuitions. I argued that the Tatar TAE suffixes all have underlyingly temporal semantics, and that their evidential readings are largely derived pragmatically. This analysis accounts for both major evidential and temporal components of the morphemes' meaning; i.e., it answers Question 1 from Chapter 4.

Unlike other analyses of evidentiality (e.g. Murray 2010, Faller 2002, Izvorski 1997), I do not hardwire the evidential interpretations of the morphemes into any component of their meanings. Unlike the Evidence Acquisition Time analyses of evidentiality discussed

in Chapter 4 (e.g. Koev 2017, Smirnova 2013, Lee 2013, among others), I do not posit any novel component in the semantics, but rather utilize independently motivated event pre- and post-states (following e.g. Moens and Steedman 1988). Finally, unlike prior analyses of portmanteau temporal and evidential morphemes in which the two components of meaning are represented separately in the syntax (e.g. Şener 2011 for Turkish), this analysis reflects the portmanteau nature of the Tatar TAE morphemes by accounting for both components of meaning with a single denotation.

This chapter solely addressed the interpretations of the TAE suffixes in semantically unembedded environments, i.e., in matrix clauses. In the following chapter, I work towards a formal analysis of the interpretation and use of the TAE suffixes in some embedded environments, as described previously in Chapter 2 and 3.

CHAPTER 6

Deriving the temporal and evidential interpretations of the Tatar TAE suffixes in clausal embeddings

6.1 Introduction

This chapter accounts for the descriptive generalizations laid out in Chapters 2 and 3 regarding the interpretation and distribution of the Tatar TAE suffixes when embedded under the attitude predicate *äytergä* 'to say.' I summarize this data in Table 6.1. My analysis focuses chiefly on deriving the evidential interpretations of the suffixes when they are embedded, although I also address some of the temporal interpretations of the relevant expressions.

TAE suffix	**Meaning**	**Grammaticality in embedded CPs**	**Grammaticality in embedded verbal nominalizations**
-DI	past tense	✓ (can shift)	*
-GAn	resultative aspect	✓ (can shift)	✓ (no evid. reading)
-A	present tense	✓	*
-(y)AçAK	prospective aspect	✓ (must shift)	✓ (no evid. reading)
-(y)Er	future tense	✓ (must shift)	*

Table 6.1: Distribution and interpretation of the Tatar TAE suffixes in clausal embeddings, where "shift" refers to evidential shift.

As I showed in these preceding chapters, there are a number of puzzles to be accounted for with respect to the interpretation and use of the TAE suffixes in these clausal embeddings. I summarize the primary puzzles here:

1. **Puzzles regarding the use of TAE suffixes in embedded CPs:**

(a) Why the evidential component of past oriented TAE suffixes can optionally undergo evidential shift: i.e., why they generally seem to be able to be interpreted relative either to the speaker or to the matrix subject.

(b) Why the evidential component of future oriented TAE suffixes **must** undergo evidential shift; i.e., why they can only be interpreted relative to the matrix subject.

2. **Puzzles regarding the use of TAE suffixes in embedded verbal nominalizations:**

(a) Why only a subset of the TAE suffixes are grammatical (only the aspectual TAE suffixes *-GAn* 'RESULT' and *-(y)AçAK* 'PROSP' can occur).

(b) Why the evidential component of these TAE suffixes is absent.

I address the full CP embedding data in §6.2, and address the embedded verbal nominalization data in §6.3. I sketch a compositional derivation of the relevant expressions, and show how my initial analysis of the TAE suffixes in Chapter 5 will have to be revised to account for some quirks of the embedding data. An important takeaway from this chapter is that in environments in which only some of the TAE suffixes are grammatical (i.e., the embedded verbal nominalizations in §6.3), pragmatic competition is unavailable and evidential readings do not arise. Finally, I note that in this chapter, I focus entirely on the distribution and use of the five verbal suffixes. I do not discuss embedded uses of the free past tense morpheme *ide*, which I leave for future work.

6.2 Tatar TAE suffixes in embedded CPs

I begin by giving a short conceptual summary of the mechanics that I propose drive the speaker-oriented and matrix subject-oriented evidential interpretations of the TAE suffixes in embedded CPs. This summary has two purposes. First, it will provide a conceptual background to the formalized theory presented later in this section. Second, it will in part motivate my proposal to revise the denotations of the Tatar TAE suffixes from their original formulation in Chapter 5, and introduce some of the updated formalisms that I will use.

In a nutshell, I propose that the speaker-oriented and matrix subject-oriented interpretations of the TAE suffixes in embedded CPs reflect **wide scope** versus **narrow scope** readings of the suffixes with respect to the matrix attitude verb, respectively. I draw the motivation for this wide scope/narrow scope distinction from prior accounts of de dicto and de re readings of intensional lexical items. According to scope analyses of de dicto/de re readings, the location of a lexical item in the syntax determines which worlds it is interpreted relative to (inspired by Russell 1905, among many others). I propose an analogous syntactic analysis of the Tatar TAE suffixes in which their location in the syntax determines which **times** they are interpreted relative to.

De dicto and de re readings of intensional lexical items are well documented in the semantic literature; I give an example of de dicto and de re readings of *some student* embedded under the attitude verb *believe* in (270). I give the de dicto/narrow scope reading of *some student* in (270a), and the de re/wide scope reading of *some student* in (270b). (De dicto and de re readings are also observed for a range of other lexical items, including proper nouns, quantifiers, modals, and tenses (Keshet and Schwarz 2014).)

(270) Maren believes that some student is a spy.

a. **De dicto/narrow scope reading:**
   Maren believes that there is a person who is a student and a spy; a student may or may not exist in the actual world.

b. **De re/wide scope reading:**
   There is some student in the actual world who Maren believes is a spy; she may or may not know that the individual is a student.

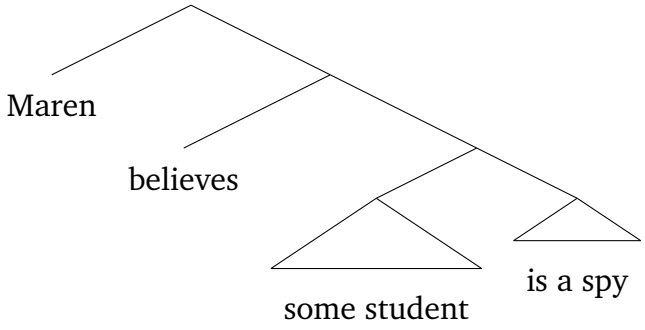
The de dicto/narrow scope reading of *some student* in (270a) is one in which both *some student* and the predicate *be a spy* are interpreted relative to Maren's belief worlds. That is, there is an individual in Maren's context of evaluation to whom she assigns the properties of both being a student and being a spy. However, neither property need be true with respect to the individual in the actual world.

Conversely, the de re/wide scope reading of *some student* in (270b) is one in which

there is an individual who is a student in the actual world; Maren believes of that individual that they are a spy. Under a scope analysis of (270b), *some student* is interpreted higher than *believe* and is therefore interpreted relative to the actual world, as opposed to Maren's belief worlds. This reading therefore entails that there exists a student in the actual world. However, on a de re reading of *some student* in (270), it need not be the case that Maren believes that the individual is a student.

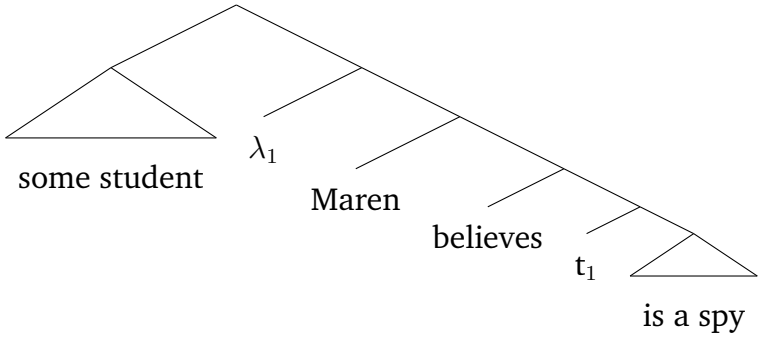
I sketch a scope analysis of the de dicto/narrow scope and de re/wide scope readings of *some student* in (270) in (271)-(272). (I omit the complementizer for simplicity.) In the de dicto/narrow scope reading in (271), *some student* is interpreted within the scope of *believe* and is therefore interpreted relative to Maren's belief worlds. In the de re/wide scope reading in (272), *some student* moves to a higher position in the clause and is therefore interpreted outside the scope of the attitude verb, relative to the actual world. (I follow Heim and Kratzer 1998 in assuming that adjunction of a raised lexical item co-occurs with lambda-abstraction over its trace.)

(271) **De dicto/narrow scope reading of *some student*:**
Maren believes that there is a person who is a student and a spy.



(272) **De re/wide scope reading of *some student*:**

There is some student who Maren believes is a spy.



I utilize a scope analysis like in (271)-(272) to account for the availability of evidential shift when the TAE suffixes are embedded under attitude verbs like *äytergä* 'to say.' As I showed in Chapter 5, the Tatar TAE suffixes are underlyingly temporal morphemes; their evidential interpretations are derived in part from their temporal semantics. Generally speaking, when the TAE suffixes remain embedded under an attitude verb, they are interpreted relative to the matrix subject's "now," and as such their evidential reading is also matrix subject-oriented. However, if they scope above the attitude verb, they are interpreted relative to the utterance time. As a result, their evidential reading is speaker-oriented. Henceforth, I will primarily refer to my analysis of the TAE suffixes as referring to a wide scope/narrow scope distinction, as opposed to a de dicto/de re distinction.

6.2.1 Updated denotations

My wide scope/narrow scope analysis of the Tatar TAE suffixes in embedded CPs hinges on their ability to move out of the embedded clause. This movement-based analysis therefore requires revising their denotations as initially posited for matrix clauses in Chapter 5.

In Chapter 5, I treated the TAE suffixes as being functions from properties of events to properties of times, i.e., of type $<<v,t>,<i,t>>$. However, under a theory of movement in which moved lexical items leave behind a trace, it is unclear how to move objects of this type. This problem stems from the difficulty of determining the type of the remaining trace. Traces are traditionally assumed to be of simple types; for instance, when we move

quantified noun phrases of type $<<e,t>,t>$, we stipulate that the trace that is left behind is of type e. However, it is less clear what kind of trace would remain after moving something of type $<<v,t>, <i,t>>$.

To resolve this issue, I modify the types of the Tatar TAE suffixes to be of type $<<v,<i,t>>, <i,t>>$. I propose that when they undergo movement, they leave behind a trace of type v; i.e., their trace is of the same type as the first argument of their first argument. I provide updated denotations for the five TAE suffixes in (273). (I will further update the denotation of *-GAn* in (273b) in §6.3.2 to account for its interpretation in embedded verbal nominalizations.)

(273) **Updated denotations for the TAE suffixes:**

a. $[\![\text{-}DI]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[\tau(e) < i \ \&\ p(e)(i)]$

b. $[\![\text{-}GAn]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[\exists s[s \text{ is a suitable post-state for } e \ \&\ e \ll s \ \&\ i \subset \tau(s) \ \&\ p(e)(i)]]$

c. $[\![\text{-}A]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[i \subseteq \tau(e) \ \&\ p(e)(i)]$

d. $[\![\text{-}(y)Er]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[i < \tau(e) \ \&\ p(e)(i)]$

e. $[\![\text{-}(y)AçAK]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[\exists s[s \text{ is a suitable pre-state for } e \ \&\ s \ll e \ \&\ i \subset \tau(s) \ \&\ p(e)(i)]]$

Since the TAE suffixes are now of a more complex semantic type, I am required to also treat verbs as being of a more complex type than I previously posited in Chapter 5. Henceforth, I assume that verbs combine with both event and time arguments; i.e., I treat intransitive verbs as being of type $<e,<v,<i,t>>>$. However, this crucially comes with the stipulation that the time argument of verbs is present **solely for type purposes**. That is, these time arguments do not supply the time of the event described by the verb.

I indicate this formally by distinguishing between verbs in the object language and predicates in the metalanguage. I use small capitals to indicate metalanguage predicates. For example, the metalanguage predicate WIN corresponds to the object language predicate

ciñergä 'to win'; the former is written in small capitals. I give the denotation of *ciñergä* in (274a), and provide its truth conditions in (274b). Crucially, the truth conditions for *ciñergä* in (274b) do not refer to the time of the winning event, despite the fact that it combines with a time argument. I include a time argument in part for the compositional reasons described above. However, this time argument additionally provides a slot that can be bound by higher operators, as I will show later in this chapter.

(274) a. $[\![\textit{ciñergä}]\!] = \lambda x \lambda e \lambda i.\ \textsc{win}(x)(e)(i)$

b. $\textsc{win}(x)(e)(i) = 1$ iff e is a winning event by x

I treat *äytergä* as being of type $<<i,t>, <e,<v,<i,t>>>>>$. It combines with a property of times p, an individual x, an event e', and a time i'; it asserts that the set of times compatible with what x says is a subset of the set of times i of which p is predicated.[1] i is the time at which the matrix subject x locates themselves; i.e., it is the matrix subject's "now." When *äytergä* combines with an embedded clause (of type $<i,t>$), the matrix subject's "now" binds the time argument of the embedded clause. As I will show later in this chapter, the value of the time argument of the embedded clause determines whether its evidential contribution is speaker oriented or matrix subject oriented.

Like other verbal predicates in Tatar, as described above, I distinguish between the object language predicate *äytergä* 'to say' and the metalanguage predicate SAY, which I again indicate using small capitals. I give a denotation for *äytergä* in (275a), and reiterate that, like all other Tatar predicates, the time argument of *äytergä* in (275a) does not contribute any temporal meaning. The metalanguage predicate SAY applied to x, e', and i' outputs a set of times, as I show in (275b).

[1] See Sharvit (2018), among others, for the semantics of *say*. In these proposals, *say* is typically treated as ranging over sets of world-time pairs rather than just times. That is, *say* asserts that the set of world-time pairs compatible with what x says is a subset of the set of world-time pairs $<w,i>$ of which p is predicated. As discussed previously, I do not include worlds in my denotations for simplicity; however, I am amenable to fully intensional analyses, and worlds could be included in my denotations with no negative effects. The denotation that I give for *äytergä* in (275a) could therefore be read as shorthand for ranging over sets of world-time pairs.

(275) a. $[\![\text{äytergä}]\!] = \lambda p_{<i,t>} \lambda x \lambda e' \lambda i'.$ SAY(x)(e′)(i′) $\subseteq$ {i: p(i)}

b. SAY(x)(e′)(i′) = {i: i is compatible with what x says in the saying event e′}

Finally, I assume that the Tatar complementizer *dip* is semantically vacuous.[2] I therefore omit it from the following derivations.

6.2.2 Narrow scope readings of future oriented TAE suffixes in embedded CPs

I begin by giving a rough compositional derivation of the future oriented TAE suffixes. I showed in Chapter 2 that when the future oriented TAE suffixes *-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP' occur in CPs embedded under attitude predicates, their evidential contribution is obligatorily interpreted as reflecting the matrix subject's evidence for the embedded proposition. That is, embedded future oriented TAE suffixes **obligatorily** undergo evidential shift. Furthermore, as shown in Chapter 3, such expressions are incompatible with future-in-the-past readings of the embedded event.

I repeat an example of *-(y)Er* in an embedded CP in (276). I indicate in the context the kind of evidence that the speaker and matrix subject have for the embedded proposition, and underline the individual whose evidence the embedded TAE suffix reflects. In this example, *-(y)Er* must be interpreted as reflecting the matrix subject's lack of specific evidence for the embedded proposition. *-(y)AçAK* is infelicitous in the context in (276).

(276) **Swimming fan context (Narrow scope reading of *-(y)Er* 'FUT')**

Speaker: Specific evidence

Matrix subject: Non-specific evidence

You are watching an Olympic swimming event with your friend Läylä. You are a huge fan of competitive swimming, and know the statistics about all of the swimmers. You believe that Michael Phelps has trained the best out of everyone competing at the Olympics this year: his diet, practice regimen, etc. are all impeccable.

[2] However, see e.g. Özyıldız (2016) for discussion as to why the Turkish complementizer *diye* (and by analogy, perhaps Tatar *dip*) is not semantically vacuous.

However, Läylä has no strong opinion about Michael Phelps's ability to win. Before the race starts, Läylä speculates that he will win. You tell someone else:

Läylä [Michael Phelps ciñ-**er**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-FUT-3SG COMP] say-PST-3SG
'Läylä said that [and has non-specific evidence that] Michael Phelps will win.'

I repeat a minimal pair example with embedded *-(y)AçAK* in (277). In this example, *-(y)AçAK* must be interpreted as reflecting that the matrix subject has specific evidence for the embedded proposition, as indicated by the underline. *-(y)Er* is infelicitous in the context in (277).

(277) **Swimming fan context (Narrow scope reading of *-(y)AçAK* 'PROSP')**

Speaker: Non-specific evidence

Matrix subject: Specific evidence

Läylä is watching an Olympic swimming event. She is a huge fan of competitive swimming, and knows the statistics about all of the swimmers. She believes that Michael Phelps has trained the best out of everyone competing at the Olympics this year: his diet, practice regimen, etc. are all impeccable. She tells you he will definitely win. However, you have no strong opinion about his ability to win. Later, you tell someone else:

Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG
'Läylä said that [and has specific evidence that] Michael Phelps will win.'

Both (276) and (277) are incompatible with future-in-the-past interpretations. That is, it is impossible to use either (276) or (277) to express that the event of Michael Phelps winning happened in the past of the utterance time but in the future of Läylä's saying time; i.e., they cannot mean something like the English expression *Läylä said that Michael Phelps would win*. Given this data, I therefore have two primary objectives for my theory: first, to account for the obligatory evidential shift of the embedded TAE suffixes; and second, to account for the lack of availability of future-in-the-past readings.

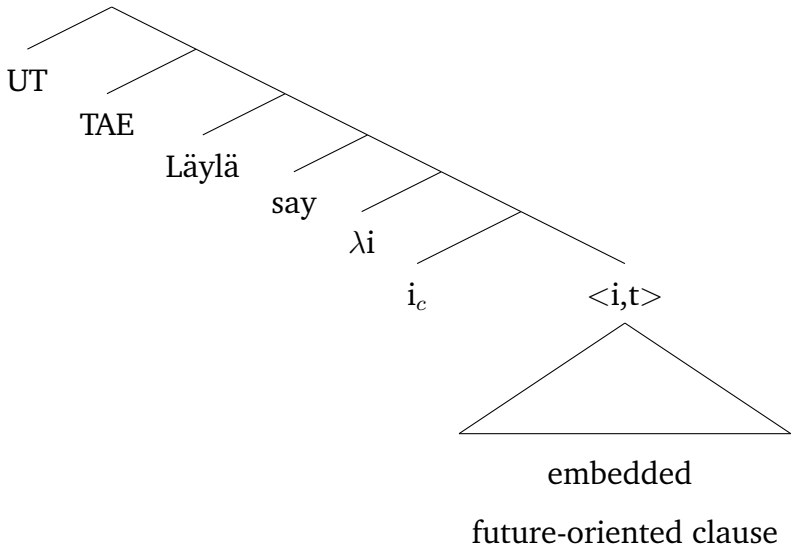
I account for these empirical facts using two primary mechanisms. First, to account for the lack of future-in-the-past, I propose to treat the time argument of the embedded CP as **indexical**, in the spirit of Kaplan (1989). Second, I utilize the wide scope/narrow scope (i.e., de dicto/de re) distinction discussed in §6.2 to account for the evidential interpretation of the embedded CP, relying on a proposal for the de re interpretation of embedded tenses from Abusch (1997, 1991).

Indexical expressions are interpreted relative to properties of the context of utterance. Following Kaplan (1989), I assume that utterance contexts can include information such as the time of the utterance, the speaker of the utterance, the location of the utterance, and so on. Indexicals are referential expressions that refer to these components of the context. Canonical English indexicals include lexical items like *I* and *you*; *I* picks out the speaker of the utterance, while *you* picks out the addressee. I propose to treat the time argument of embedded future oriented TAE suffixes as an indexical.[3]

I implement an indexical treatment of the Tatar future oriented TAE suffixes by introducing an indexical time argument i_c, which values the open time argument of the embedded future-oriented clause. As an indexical, i_c is valued by the context of the utterance and (typically) refers to the utterance time (as I will discuss momentarily, I posit a de re interpretation of i_c in which it picks out a larger time interval that includes the utterance time). I include lambda-abstraction over times immediately above this indexical time argument for type purposes. I give a simplified schema for a Tatar expression with an embedded future oriented TAE suffix in (278).

[3]In this sense, the Tatar future oriented TAE suffixes are similar to the English future and present tenses, which are also typically treated as indexicals. The English future tense marker *will* is frequently treated as consisting of a modal component WOLL in combination with an indexical present tense (Abusch 1988). The inclusion of the indexical present tense also accounts for the indexical interpretation of the English future tense. Since I do not posit any underlying relationship between Tatar present tense and future tense/prospective aspect, my proposal that the Tatar future oriented TAE suffixes are indexical is a stipulation.

(278)



Since i_c is unbound within the scope of the matrix attitude verb *äytergä* 'to say,' I propose that it is interpreted de re; i.e., it is interpreted above the matrix attitude verb. This is in the spirit of e.g. Abusch (1997, 1991)'s proposal for the double access reading of embedded present-under-past tenses in English. Abusch (1991) proposes that the embedded present tense in present-under-past sentences like (279) is interpreted de re.

(279) John believed that Mary **is** pregnant.

A de re interpretation of (the time argument of) *is* in (279) is roughly one in which it picks out a time interval that covers Mary's pregnancy, according to John (Mary need not be pregnant in the actual world). This interval minimally includes the utterance time (as perceived by John) and John's believing time as he perceives it to be (i.e., his "now"). This accounts for the double access reading of (279), in which the time of Mary being pregnant overlaps with the utterance time.

Similarly, a de re interpretation of i_c in (278) is one in which it minimally picks out a time interval including the utterance time as Läylä perceives it and Läylä's saying time as she perceives it (i.e., her "now"). Under a de re interpretation of i_c, the speaker assigns a temporal description to i_c. In (276)-(277), the speaker describes the time i_c as a time in the future of which Michael Phelps will win. This accounts for the lack of future-in-the-past, since the de re interval picked out by i_c includes the utterance time. The event of Michael

Phelps winning therefore cannot be located in the past of the utterance time.

I account for the obligatory matrix subject oriented evidential interpretation of the future oriented TAE suffixes by proposing that the suffixes must remain in situ; that is, they cannot undergo movement to a position superior to the matrix attitude verb. In Chapter 5, I argued that the evidential readings of the Tatar TAE suffixes are byproducts of how they cause the event described by the scope proposition to be "viewed" in time. Since the future oriented TAE suffixes remain in situ, only the matrix subject "views" the embedded event from the time picked out by i_c. Since the embedded event is always "viewed" by a time that is relative to the matrix subject (i.e., the matrix subject's "now"), the evidential contribution of these suffixes is therefore always interpreted as being matrix subject oriented. (In Chapter 5, I argued that the evidential readings of the TAE suffixes are derived in part pragmatically; I return to the issue of how to fully derive these readings in embedded environments later in this section.)

Ultimately, this requirement that the future oriented TAE suffixes take narrow scope/remain in situ is a stipulation; there is nothing inherent to the semantics of the future oriented TAE suffixes that should restrict their scope taking ability in embedded CPs. This narrow scope requirement could be due to some morphological restriction on the future oriented suffixes. It could also be viewed as a reflecting a "default" syntactic setting for the embedded TAE suffixes. That is, undergoing movement could be more syntactically "costly" than remaining in situ; remaining low in the clause would therefore be the syntactic default. This would make taking narrow scope with respect to the attitude verb—and as a side effect, reflecting the matrix subject's evidence/evidential shift—the unmarked interpretation of embedded TAE suffixes.

With this background in mind, I will now sketch compositional derivations for the Tatar expressions in (276) and (277), omitting a complete compositional analysis of the de re component.[4] I derive the matrix subject oriented readings of embedded *-(y)Er* in §6.2.2.1

[4]I note that this proposal does not rule out a reading in which the time of the embedded event is interpreted as overlapping with the utterance time, i.e., a reading like *Läylä said that Michael Phelps is winning*. This is because if the embedded time is interpreted de re, Läylä must have **some** "acquaintance" relation to it such that she can have a belief about it; that is, the embedded time can be her past or her present, just not

and *-(y)AçAK* in §6.2.2.2. I discuss their pragmatic competition in §6.2.2.3.

6.2.2.1 Matrix subject oriented (narrow scope) reading of embedded *-(y)Er* 'FUT'

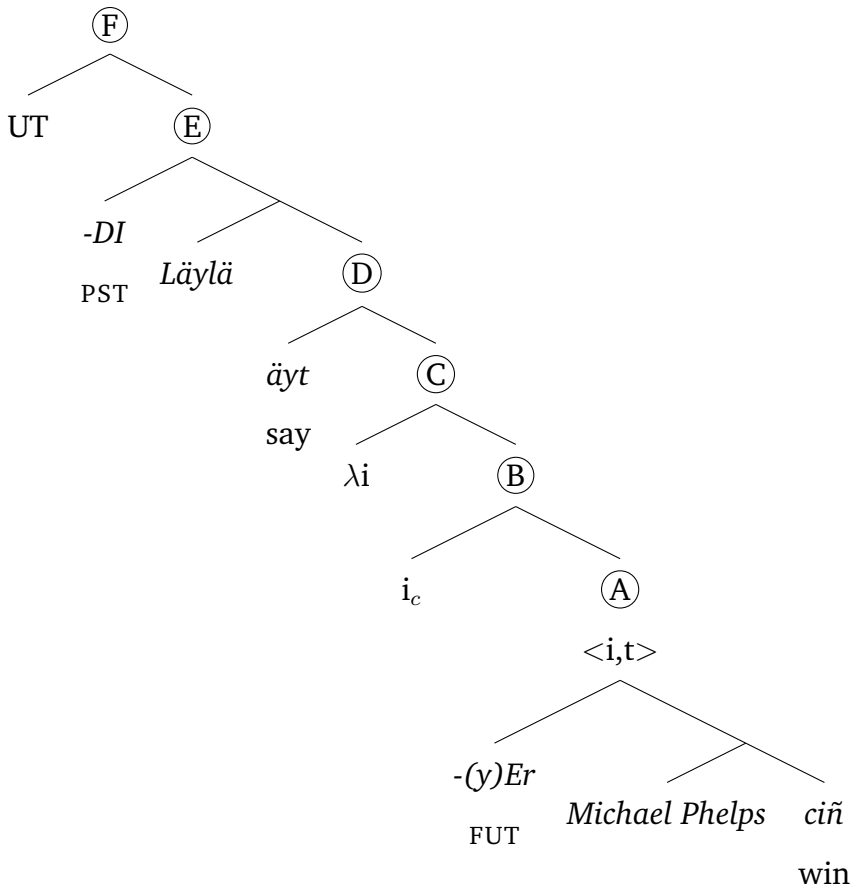
I give a tree structure for the expression in (276) in (280). In this tree, the embedded future tense suffix *-(y)Er* remains in situ, taking narrow scope with respect to the matrix attitude verb. *-(y)Er* is interpreted relative to the indexical time argument i_c, which is interpreted de re and therefore picks out an interval including the utterance time (as Läylä perceives it) and the time of Läylä's saying (as she perceives it; i.e., her "now"). Following lambda-abstraction over times, the type of the embedded clause is $<i,t>$.

(280) Läylä [Michael Phelps ciñ**-er**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-FUT-3SG COMP] say-PST-3SG
'Läylä said that [and has non-specific evidence that] Michael Phelps will win.'

---

her future. However, such a reading is not attested for expressions like (276) or (277).

Abusch (1997) proposes an Upper Limit Constraint (ULC) on the interpretation of tenses, which is formulated in these terms regarding de re interpretation. The ULC can be used to rule out e.g. future readings of embedded present tense clauses in English present-under-past expressions. It could perhaps be the case that Tatar has an even more restrictive constraint that rules out present oriented readings of embedded de re tenses.



I give a semantics for (280) in (281). The denotation of the clause embedded under *äytergä* 'to say' is in (281c). The denotation of the entire expression is as in (281f); I give a denotation enriched by the de re interpretation of i_c in (281g). Recall from §6.2.1 that I now use more complex semantics for the TAE suffixes and the verbs; the updated denotation of *-(y)Er* is given in (273d). Furthermore, both metalanguage predicates SAY and WIN in (281) combine with time arguments, but I reiterate that these arguments do not contribute any temporal meaning; the temporal locations of the saying and winning events are determined by the semantics of the TAE suffixes.

(281) a. $[\![Ⓐ]\!]^c = \lambda i.\ \exists e[i < \tau(e)\ \&\ \textsc{win}(MP)(e)(i)]$

b. $[\![Ⓑ]\!]^c = 1$ iff $\exists e[i_c < \tau(e)\ \&\ \textsc{win}(MP)(e)(i_c)]$

c. $[\![Ⓒ]\!]^c = \lambda i.\ \exists e[i_c < \tau(e)\ \&\ \textsc{win}(MP)(e)(i_c)]$

d. $[\![Ⓓ]\!]^c = \lambda x \lambda e' \lambda i'.\ \textsc{say}(x)(e')(i') \subseteq \{i: \exists e[i_c < \tau(e)\ \&\ \textsc{win}(MP)(e)(i_c)]\}$

e. $[\![\textcircled{E}]\!]^c = \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i\text{: } \exists e[i_c < \tau(e) \ \&\ \textsc{win}(\text{MP})(e)(i_c)]\}]$

f. $[\![\textcircled{F}]\!]^c = 1$ iff $\exists e'[\tau(e') < \text{UT} \ \&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i\text{: } \exists e[i_c < \tau(e) \ \&\ \textsc{win}(\text{MP})(e)(i_c)]\}]$

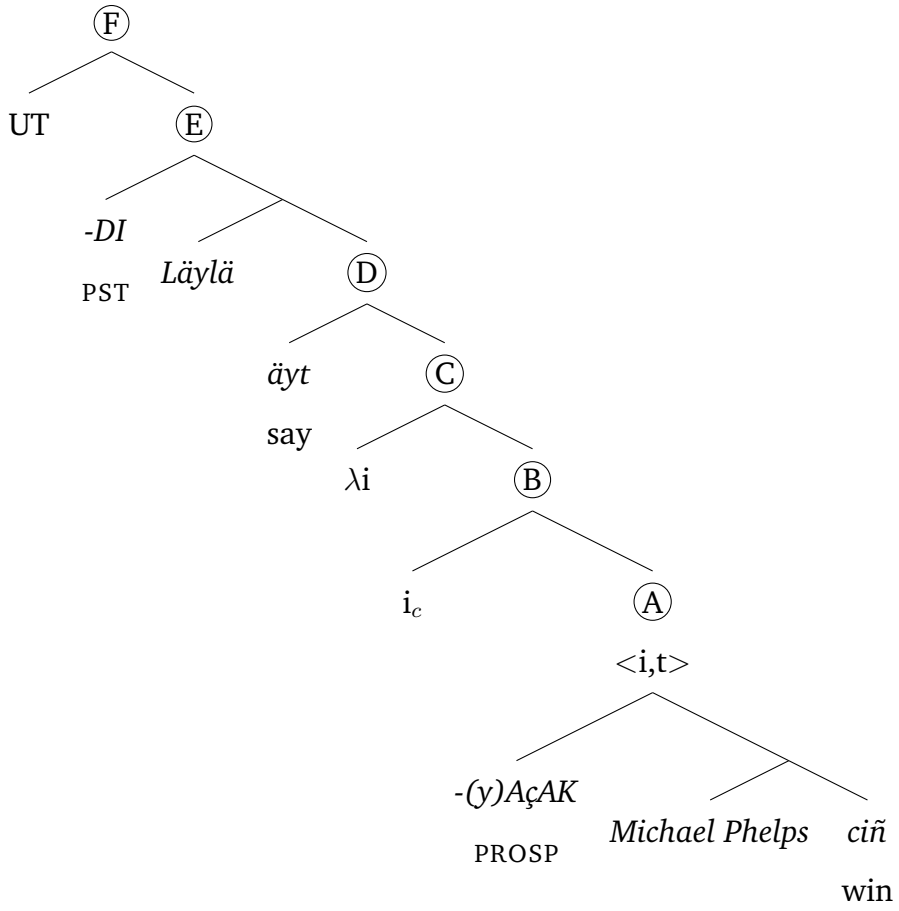
g. $[\![\textcircled{F'}]\!]^c$ (enriched with a de re mechanism) = 1 iff there is a (time) description d such that $\exists e'[d(\tau(e')) = i_c \ \&\ \tau(e') < \text{UT} \ \&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i\text{: } \exists e[d(i) < \tau(e) \ \&\ \textsc{win}(\text{MP})(e)(d(i))]\}]$

By "time description" here, I mean a function from times to times (also called a time concept). The enriched denotation in (281g) is true iff there is a description d such that there exists an event e'; the description applied to the runtime of e' is equal to i_c (where d is a description of a time interval like "this week" or "this month"). The runtime of e' precedes the utterance time, and e' is an event of Läylä saying. The set of times compatible with what Läylä says in e' is a subset of the set of times i, where i is Läylä's "now"; the description applied to Läylä's "now", *d(i)*, is such that there exists an event e and the runtime of e follows *d(i)* and e is an event of Michael Phelps winning. Läylä's saying event is therefore located in the past of the utterance time, and Läylä describes to herself the time interval i_c (including the utterance time, as she perceives it) as a time in the future of which Michael Phelps will win. This de re interpretation of i_c rules out future-in-the-past.

6.2.2.2 Matrix subject oriented (narrow scope) reading of embedded *-(y)AçAK* 'PROSP'

I give a tree structure for the expression in (277) in (282). In this expression, *-(y)AçAK* remains low in the embedded CP. The indexical time argument i_c in the embedded clause is interpreted de re.

(282) Läylä [Michael Phelps ciñ-**eçek**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PROSP-3SG COMP] say-PST-3SG
'Läylä said that [and has specific evidence that] Michael Phelps will win.'



I give a semantics for (282) in (283). The denotation of the embedded clause containing *-(y)AçAK* is given in (283c), and the denotation of the entire expression is as in (283f). I give an enriched de re interpretation of (282) in (283g). I again note that both metalanguage predicates SAY and WIN in (283) combine with time arguments; however, these arguments do not contribute any temporal meaning.

(283) a. $[\![\text{Ⓐ}]\!]^c = \lambda i.\ \exists e[\exists s[s \ll e \ \&\ i < \tau(s)\ \&\ \text{WIN}(\text{MP})(e)(i)]]$

b. $[\![\text{Ⓑ}]\!]^c = 1$ iff $\exists e[\exists s[s \ll e \ \&\ i_c < \tau(s)\ \&\ \text{WIN}(\text{MP})(e)(i_c)]]$

c. $[\![\text{Ⓒ}]\!]^c = \lambda i.\ \exists e[\exists s[s \ll e \ \&\ i_c < \tau(s)\ \&\ \text{WIN}(\text{MP})(e)(i_c)]]$

d. $[\![\text{Ⓓ}]\!]^c = \lambda x \lambda e' \lambda i'.\ \text{SAY}(x)(e')(i') \subseteq \{i: \exists e[\exists s[s \ll e \ \&\ i_c \subset \tau(s)\ \&\ \text{WIN}(\text{MP})(e)(i_c)]]\}$

e. $[\![\text{Ⓔ}]\!]^c = \lambda i'.\ \exists e'[\tau(e') < i'\ \&\ \text{SAY}(\text{Läylä})(e')(i') \subseteq \{i: \exists e[\exists s[s \ll e$

& $i_c \subset \tau(s)$ & WIN(MP)(e)(i_c)]]}]

f. $[\![\textcircled{F}]\!]^c$ = 1 iff $\exists e'[\tau(e') < UT$ & SAY(Läylä)(e')(UT) $\subseteq$ {i: $\exists e[\exists s[s \ll e$
& $i_c \subset \tau(s)$ & WIN(MP)(e)(i_c)]]}]

g. $[\![\textcircled{F'}]\!]^c$ (enriched with a de re mechanism) = 1 iff there is a (time) description d such that $\exists e'[d(\tau(e')) = i_c$ & $\tau(e') < UT$ & SAY(Läylä)(e')(UT) $\subseteq$ {i: $\exists e[\exists s[s \ll e$ & $d(i) \subset \tau(s)$ & WIN(MP)(e)(d(i))]]}]

The enriched denotation in (283g) is true iff there is a description d such that there exists an event e'; the description applied to the runtime of e' is equal to i_c (where d is a description of a time interval like "this week" or "this month"). The runtime of e' precedes the speaker's utterance time, and e' is an event of Läylä saying. The set of times compatible with what Läylä says in e' is a subset of the set of times i, where i is Läylä's "now"; the description applied to Läylä's "now", *d(i)*, is such that there exists an event e and a state s; e is contingently related to s and *d(i)* (the time that Läylä describes as her "now") is located in the runtime of s and e is an event of Michael Phelps winning. That is, this denotation locates Läylä's "now" in the pre-state of the event of Michael Phelps winning; this is the time from which the embedded event is "viewed." (I refer the reader to Chapter 5 for a definition and discussion of event pre-states.)

As in the derivation in (280), the de re interpretation of i_c is such that i_c picks out an interval including the utterance time (as Läylä perceives it) and Läylä's saying time as she perceives it (i.e., her "now"). Because the interval picked out by i_c is located in the runtime of the pre-state of Michael Phelps winning, and the suitability conditions of event pre-states are such that they temporally precede the runtime of the event, (283f) is incompatible with future-in-the-past. That is, Läylä describes i_c as a time in the future of which Michael Phelps will win. Because this interval includes the time that Läylä perceives to be the utterance time, this successfully rules out future-in-the-past.

6.2.2.3 Deriving evidential readings of matrix subject oriented (narrow scope) *-(y)Er* and *-(y)AçAK* through pragmatic competition

In Chapter 5, I analyzed the so-called "specific" and "non-specific" evidential readings of *-(y)AçAK* 'PROSP' and *-(y)Er* 'FUT,' respectively, as arising (in part) pragmatically. I began by assuming that cooperative speakers assert what they believe is true, based on their sensory perception of the world (i.e., they obey Grice 1975's Maxim of Quality). When a Tatar speaker asserts an expression using *-(y)AçAK*, they therefore must have some sensory evidence that they are located in a contingently related pre-state of the event that their utterance describes. I roughly assume Moens and Steedman (1988)'s definition of "contingency," which requires that the pre-state stand in a causal relationship to the event; i.e., some event or state within the pre-state will cause the described event to occur.

The sensory evidence that the speaker uses to locate themselves in an event pre-state takes the form of what we can descriptively label "specific" evidence. That is, to cooperatively assert *-(y)AçAK*(*p*), the speaker must perceive some event or state that they believe will cause the event described by *p* to occur. This in turn leads to a reading of increased speaker certainty that the event will occur, since they believe that something will cause it to occur.

I proposed in Chapter 5 that *-(y)AçAK* entails *-(y)Er*. As a result, when a speaker asserts a *-(y)Er*(*p*) expression, the resulting implicature is that *-(y)AçAK*(*p*) is false. This implies that the speaker does not locate themselves in a causally related pre-state of the event. Assuming the Gricean Maxims of Quality and Quantity, this implies that the speaker does not perceive an event or state that they believe will cause the described event to occur. However, *-(y)Er*(*p*) expressions are compatible with all other sensory evidence that the described event will occur, i.e., "non-specific" evidence. Since *-(y)Er* competes with *-(y)AçAK*, this leads to a reading of lowered speaker certainty that the event will occur.

I propose that the evidential interpretations of *-(y)AçAK* and *-(y)Er* arise through the same pragmatic calculations when they are embedded under attitude predicates, since the same entailment relationship holds. The only contrast between their interpretation in ma-

trix environments and embedded environments is that when the TAE suffixes remain in situ under an attitude verb, the time from which the embedded event is "viewed" is from the perspective of the matrix subject, rather than the speaker; i.e., it is (a time that includes) the matrix subject's "now." This results in an interpretation in which the evidential contribution of the embedded TAE suffix is relative to the matrix subject, rather than the speaker.

I note that this proposal requires positing pragmatic competition in semantically embedded environments. However, the availability of pragmatic competition in embedded environments is not straightforwardly predicted by Gricean theories of pragmatics, which operate at the utterance level. To illustrate this, we can compare the monoclausal expression in (284) to the embedding data in (285) (from Chierchia 2004, 44). Both expressions contain the existential quantifier *some*, which is typically assumed to compete pragmatically with the universal quantifiers *all/every* (Horn 1972, among many others). The expression in (284) can receive a strengthened interpretation (via scalar implicature) that some, but not all, students are waiting for John. I show this strengthened reading in (284a).

(284) Some students are waiting for John.

a. $\rightsquigarrow$ Some students are waiting for John &
¬Every student is waiting for John

The strengthened interpretation that is reported for the multiclausal expression in (285) is given in (285a); i.e., this is a way in which it is possible for speakers to interpret (285) (Chierchia 2004, 44-45; Chemla and Spector 2011).

(285) John believes that some students are waiting for him.

a. $\rightsquigarrow$ John believes that some students are waiting for him &
John believes that ¬every student is waiting for him

However, (285a) is not predicted by the traditional Gricean account of how implicatures are calculated, which assumes that pragmatic competition occurs at the utterance

level. Gricean theories predict that the strengthened meaning of (285) should instead be (286a), which makes a much weaker statement than (285a). (285a) states that John believes that not every student is waiting for him, whereas (286a) simply states that it is not the case that John believes that every student is waiting for him (in fact, John may have no belief at all regarding students).

(286) John believes that some students are waiting for him.

a. $\rightsquigarrow$ John believes that some students are waiting for him &
¬John believes that every student is waiting for him

Despite this issue, a number of authors have observed that implicatures seem to be available in a range of semantically embedded environments (including under attitude verbs, under quantifiers, and inside the coordinands of disjunctive expressions) (Chemla and Spector 2011, Chierchia 2004, Sauerland 2004, among many others).

The question of how best to theoretically account for embedded implicatures is far from settled. However, speaker judgments regarding data like (285a) suggests that embedded implicatures are in fact available. I take this to mean that while my proposal for embedded implicature in Tatar may be theoretically tricky to implement, it is not out of the ordinary, empirically speaking. Since the availability of embedded implicatures are well-documented in e.g. English, it is not unreasonable to expect that they should also be available in Tatar.

For the purposes of this dissertation, I am not concerned with comparing and evaluating different theories of embedded implicature. I note simply that my analysis of the TAE suffixes requires adopting **some** mechanism for pragmatic competition in semantically embedded environments. At present, I leave the precise formal implementation of this proposal for future work. As I show in §6.3, utilizing this pragmatic mechanism is further motivated by embedded environments in which some of the TAE suffixes are ungrammatical; since pragmatic competition is unavailable in these environments, no evidential interpretations arise.

6.2.3 Narrow scope readings of past oriented TAE suffixes in embedded CPs

I showed in Chapter 2 that the past oriented TAE suffixes (*-DI* 'PST' and *-GAn* 'RESULT') differ from the future oriented TAE suffixes (*-(y)Er* 'FUT' and *-(y)AçAK* 'PROSP') with respect to their evidential interpretation in embedded CPs. When either of the past oriented TAE suffixes are embedded under an attitude predicate, it appears that their evidential contribution **optionally** shifts: they can be interpreted as reflecting the evidence of either the matrix subject or the speaker. That is, an expression with *-DI* or *-GAn* embedded in a full CP is ambiguous with respect to its evidential interpretation. In this section, I focus solely on deriving the matrix subject oriented (i.e., narrow scope) evidential interpretations of embedded *-DI* and *-GAn*. These proceed similarly to the derivations for narrow scope *-(y)Er* and *-(y)AçAK* in §6.2.2.

I repeat an example of embedded *-DI* in (287). In this context, the matrix subject is interpreted as having direct evidence for the embedded proposition.

(287) **Olympics attendee context (Narrow scope reading of *-DI* 'PST')**

Speaker: Indirect evidence

Matrix subject: Direct evidence

Läylä went to the Olympics and saw Michael Phelps compete. She calls you and tells you he won; you did not see the race yourself. You then tell someone else:

Läylä [Michael Phelps ciñ-**de**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and has direct evidence that] Michael Phelps won.'

I repeat a minimal pair example with embedded *-GAn* in (288). In this context, the matrix subject is interpreted as having indirect evidence for the embedded proposition. However, recall from Chapter 2 that speakers frequently judge examples like (288) as pragmatically odd. This pragmatic oddness stems from the semantics of the matrix verb; my consultants remark that it is odd that the speaker is reporting someone else's account of an event that they (the speaker) personally witnessed. When prompted with examples like (288), my primary Tatar consultant typically remarks that "If you witnessed [the event

described by the embedded proposition], you wouldn't be saying the sentence to begin with," i.e., if the speaker has direct evidence for the embedded proposition, then they should relay their own observation of it rather than someone else's report of it.

With this major caveat in mind, I note, however, that my consultants have at times accepted **and** produced examples like (288). Since judgments regarding this data vary, and analogous examples may be felicitous with other matrix verbs and contexts I have been unable to construct, I will nonetheless lay out an analysis for (288).

(288) **Neighboring sports fan context (Narrow scope reading of *-GAn* 'RESULT')**

Speaker: Direct evidence

Matrix subject: Indirect evidence

You are attending the Rio Olympics; you see Michael Phelps compete and win. Your friend Aygöl did not attend the Olympics; however, her next door neighbor is a huge fan of Michael Phelps. She knows that the Olympics are going on right now, and she hears her neighbor cheering through the wall. She calls you and says that she inferred that Michael Phelps won. You then tell someone else:

?Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and has indirect evidence that] Michael Phelps won.'

I analyze the narrow scope/matrix subject oriented readings of *-DI* in (287) and *-GAn* in (288), like the future oriented TAE suffixes in §6.2.2, as a result of their location in the tree. Their matrix subject oriented readings again arise due to taking narrow scope with respect to the attitude verb *äytergä* 'to say.' The open time argument of the embedded clauses in these expressions is valued by the time that Läylä identifies as her "now," rather than the utterance time. As a result, this is the time from which the embedded event is "viewed."

Unlike the embedded future oriented TAE suffixes, the embedded past oriented TAE suffixes are not interpreted indexically. As I noted previously in Chapter 3, past-under-past embeddings in Tatar are only compatible with backshifted interpretations in which the

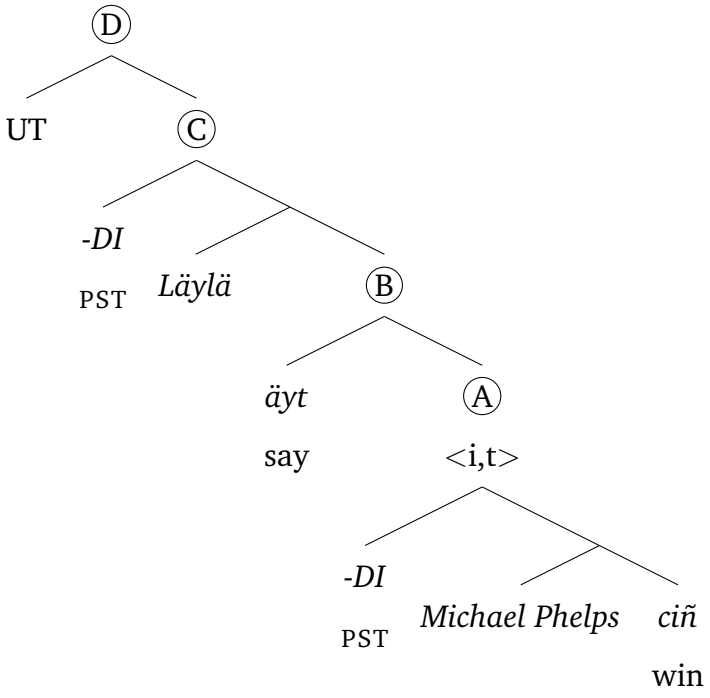
time of the embedded event is in the past of both the matrix event and the utterance time. As a result, I do not introduce any indexical time argument into the embedded clause.

I compositionally derive the matrix subject oriented readings of embedded *-DI* in §6.2.3.1 and *-GAn* in §6.2.3.2, and then discuss their pragmatic competition in §6.2.3.3.

6.2.3.1 Matrix subject oriented (narrow scope) reading of embedded *-DI* 'PST'

I provide a tree structure for the narrow scope/matrix subject oriented reading of *-DI* in (289). In this tree, the embedded *-DI* remains in situ, within the embedded clause. The open time argument of the embedded clause (of type <*i,t*>) is therefore valued by the time that Läylä identifies as her "now."

(289) Läylä [Michael Phelps ciñ**-de**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and has direct evidence that] Michael Phelps won.'



I give a compositional semantics for (287) in (290). The denotation of the embedded clause is in (290a), and the denotation of the entire expression is in (290d). Again, the metalanguage predicates SAY and WIN in (290) combine with time arguments; how-

ever, these time arguments are present solely for type purposes and do not contribute any temporal meaning.

(290) a. $[\![\text{Ⓐ}]\!] = \lambda i.\ \exists e[\tau(e) < i\ \&\ \textsc{win}(\text{MP})(e)(i)]$

b. $[\![\text{Ⓑ}]\!] = \lambda x \lambda e' \lambda i'.\ \textsc{say}(x)(e')(i') \subseteq \{i: \exists e[\tau(e) < i\ \&\ \textsc{win}(\text{MP})(e)(i)]\}$

c. $[\![\text{Ⓒ}]\!] = \lambda i'.\ \exists e'[\tau(e') < i'\ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \exists e[\tau(e) < i\ \&\ \textsc{win}(\text{MP})(e)(i)]\}]$

d. $[\![\text{Ⓓ}]\!] = 1$ iff $\exists e'[\tau(e') < \text{UT}\ \&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i: \exists e[\tau(e) < i\ \&\ \textsc{win}(\text{MP})(e)(i)]\}]$

The denotation in (290d) asserts that there exists an event e'; the runtime of e' precedes the utterance time, and e' is an event of Läylä saying. The set of times compatible with what Läylä says in e' is a subset of the set of times i, where i is Läylä's "now" and there exists an event e; e is an event of Michael Phelps winning and the runtime of e precedes i. Läylä's saying event is therefore located in the past of the utterance time, and the event of Michael Phelps winning is located in the past of the time that Läylä identifies as her "now."

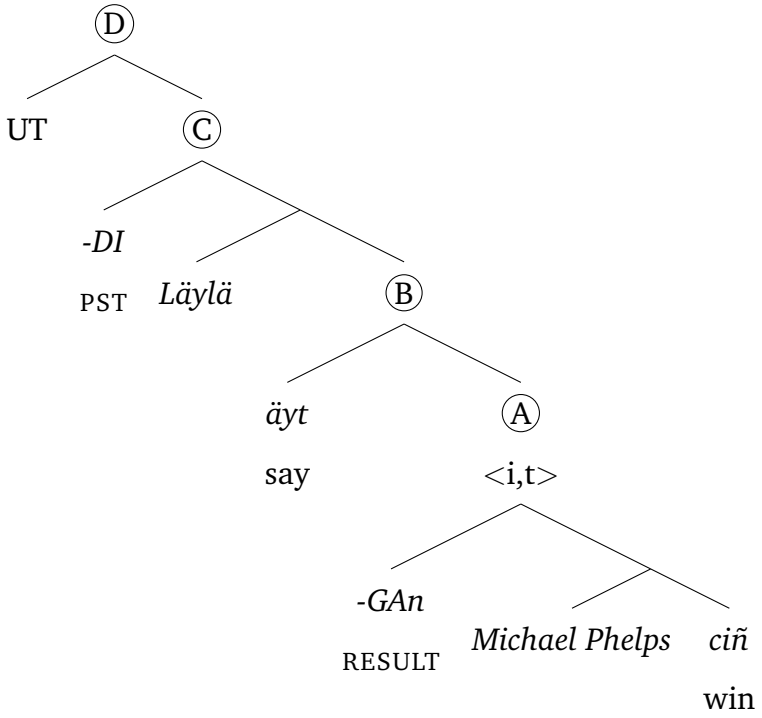
I noted in Chapter 3 that expressions like (287) are incompatible with future-in-the-past readings of the embedded event. That is, (287) cannot mean something like the English expression *Läylä said that Michael Phelps would win*. This is accounted for by the fact that the embedded *-DI* always locates the event runtime in the past of the time that Läylä takes to be her "now."

6.2.3.2 Matrix subject oriented (narrow scope) reading of embedded *-GAn* 'RESULT'

I provide a tree structure for (288) in (291). In this expression, *-GAn* remains in situ within the embedded clause and therefore takes narrow scope with respect to *äyt* 'say.' Again, the narrow scope of *-GAn* relative to this matrix attitude verb is such that that time from which the embedded event is "viewed" (i.e., the time valuing the open time argument

of the embedded clause) is Läylä's "now."

(291) ?Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and has indirect evidence that] Michael Phelps won.'



I give a compositional semantics for (291) in (292). The denotation for the embedded clause containing *-GAn* is in (292a), and the denotation for the entire expression is in (292d).

(292) a. $[\![\text{Ⓐ}]\!] = \lambda i.\ \exists e[\exists s[e \ll s \ \&\ i \subset \tau(s)\ \&\ \textsc{win}(MP)(e)(i)]]$

b. $[\![\text{Ⓑ}]\!] = \lambda x \lambda e' \lambda i'.\ \textsc{say}(x)(e')(i') \subseteq \{i: \exists e[\exists s[e \ll s \ \&\ i \subset \tau(s)\ \&\ \textsc{win}(MP)(e)(i)]]\}$

c. $[\![\text{Ⓒ}]\!] = \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(Läylä)(e')(i') \subseteq \{i: \exists e[\exists s[e \ll s \ \&\ i \subset \tau(s)\ \&\ \textsc{win}(MP)(e)(i)]]\}]$

d. $[\![\text{Ⓓ}]\!] = 1$ iff $\exists e'[\tau(e') < UT \ \&\ \textsc{say}(Läylä)(e')(UT) \subseteq \{i: \exists e[\exists s[e \ll s \ \&\ i \subset \tau(s)\ \&\ \textsc{win}(MP)(e)(i)]]\}]$

The denotation in (292d) asserts that there exists an event e'; the runtime of e' precedes the utterance time, and e' is an event of Läylä saying. The set of times compatible with what Läylä says in e' is a subset of the set of times i, where i is Läylä's "now" and there exists an event e and a state s; s is contingent on e and Läylä's "now" is located in the runtime of s and e is an event of Michael Phelps winning. I take s to be a contingently related post-state of e; see Chapter 5 for definition and discussion of event post-states. According to this semantics, Läylä's saying event is located in the past of the utterance time, and Läylä's "now" is located in the runtime of the post-state of the event of Michael Phelps winning; that is, Läylä's "now" is the time from which the embedded event is "viewed."

I noted in Chapter 3 that expressions like (288) are also incompatible with future-in-the-past readings. This falls out similarly to the account for embedded *-DI*. Given the semantics of *-GAn*, the time that Läylä identifies as her "now" is located in the embedded event's post-state. As a result, the runtime of the embedded event is always in the past of Läylä's "now." (In Chapter 5, I defined event post-states as necessarily temporally following the associated event.)

6.2.3.3 Deriving evidential readings of matrix subject oriented (narrow scope) *-DI* and *-GAn* through pragmatic competition

In Chapter 5, I derived the direct and indirect evidential readings of *-DI* 'PST' and *-GAn* 'RESULT,' respectively, through pragmatics. This accounted for the evidential readings of these suffixes in matrix (i.e., semantically unembedded) clauses. In brief, I proposed that cooperative speakers assert what they believe is true, based on their sensory perception of the world (i.e., they obey Grice 1975's Maxim of Quality). Furthermore, I assume that speakers make the "strongest" claims possible regarding past events, given their beliefs and the choice of asserting *-DI* versus *-GAn* (Grice 1975's Maxim of Quantity). Asserting *-GAn* conveys that the "strongest" evidence that the speaker has for the past event is (only) observation of its contingently related post-state (i.e., "indirect evidence").

I additionally argued that speakers must have some sensory evidence for all of the

propositions that they assert; in evidential terms, they must have either direct or indirect evidence (or both) for all of their assertions. Given this assumption, and the fact that *-GAn* entails *-DI*, *-DI*(*p*) assertions implicate that the speaker directly perceived the past event itself (i.e., they have "direct evidence" for the scope proposition). (I refer the reader to Chapter 5 for a more thorough discussion of how the evidential readings are pragmatically derived.)

I propose that the evidential readings of *-DI* and *-GAn* are derived identically in both embedded and unembedded environments. The only contrast when *-DI* and *-GAn* take narrow scope with respect to an attitude verb is that the time from which the embedded event is "viewed" is from the perspective of the matrix subject, rather than the speaker. Since the open time argument of the embedded TAE suffix is valued by the matrix subject's "now," the evidential reading of the TAE suffix is interpreted relative to the matrix subject, rather than the speaker. In this sense, the evidential readings of embedded *-DI* and *-GAn* proceed identically to the derivation of embedded *-(y)Er* and *-(y)AçAK*, as discussed in §6.2.2.3, modulo the de re interpretation of the embedded indexical future oriented TAE suffixes.

As I noted previously, implementing my pragmatic analysis when the embedded TAE suffixes remain in situ crucially requires assuming some mechanism of pragmatic competition in embedded environments. Pragmatic competition in embedded environments is not straightforwardly predicted by Gricean theories of pragmatics. However, a number of authors have shown that embedded implicatures are in fact available (e.g. Chemla and Spector 2011). Furthermore, I note that in embedded environments in which *-DI* is ungrammatical (i.e., verbal nominalizations, discussed in §6.3), evidential readings of *-GAn* do not arise. I take this data to support my theory in which the pragmatic interaction of the two TAE suffixes results in their evidential meanings, including in semantically embedded environments.

For the purpose of this dissertation, I do not evaluate the applicability of different theoretical proposals for embedded implicature to the Tatar data. I leave the formal implementation of this proposal for future work.

6.2.4 Wide scope readings of past oriented TAE suffixes in embedded CPs

In this section, I derive the speaker oriented (i.e., wide scope) evidential interpretations of *-DI* and *-GAn* in embedded CPs. I propose that in these expressions, the embedded TAE suffixes undergo movement to a position above the matrix attitude verb. (As described in §6.2.1, this movement-based analysis in part required updating the initially proposed denotations of the Tatar TAE suffixes to better motivate the semantic type of their trace.)

I repeat an example of speaker oriented embedded *-GAn* in (293). In this context, the speaker is interpreted as having indirect evidence for the embedded proposition.

(293) **Olympics attendee context (Wide scope reading of *-GAn*)**

Speaker: Indirect evidence

Matrix subject: Direct evidence

Läylä went to the Olympics and saw Michael Phelps compete. She calls you and tells you he won; you did not see the race yourself. You then tell someone else:

Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] say-PST-3SG
'Läylä said that [and I have indirect evidence that] Michael Phelps won.'

(Speaker's comment: "It's possible to say (293) even if Läylä saw it happen, because you are the speaker reporting about Läylä.")

I repeat a minimal pair example of speaker oriented embedded *-DI* in (294). In this context, the matrix subject is interpreted as having direct evidence for the embedded proposition. However, recall again from Chapter 2 that my consultants frequently judge expressions like (294) to be somewhat pragmatically odd due to the semantics of *äytergä* 'to say.' This is again because it is somewhat odd for an individual to report a third party's report of an event that they (the speaker) personally witnessed. However, since expressions like (294) are sometimes accepted and produced, I will nonetheless provide a compositional derivation.

(294) **Neighboring sports fan context (Wide scope reading of *-DI*)**

Speaker: Direct evidence

Matrix subject: Indirect evidence

You are attending the Rio Olympics; you see Michael Phelps compete and win. Your friend Läylä did not attend the Olympics; however, her next door neighbor is a huge fan of Michael Phelps. She knows that the Olympics are going on right now, and she hears her neighbor cheering through the wall. She calls you and says that she inferred that Michael Phelps won. You then tell someone else:

?Läylä [Michael Phelps ciñ-**de**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [I have direct evidence that] Michael Phelps won.'

I analyze the wide scope/speaker oriented readings of embedded *-GAn* in (293) and *-DI* in (294) as a result of their location within the tree. When the suffixes undergo movement out of the embedded clause, they adjoin at a position that is superior to the matrix attitude verb. As a result, their time argument is valued by the utterance time, rather than (as in the matrix subject oriented expressions in §6.2.3) the matrix subject's "now." This results in a speaker oriented interpretation of the embedded past oriented TAE suffixes. Their evidential reading is then derived through the same pragmatic mechanism as I proposed in Chapter 5 for the TAE suffixes in matrix clauses; in matrix clauses, the open argument of the Tatar TAE suffixes is also valued by the utterance time.

As I noted previously, past-under-past CP embeddings consistently receive backshifted interpretations in which the time of the embedded event is in the past of both the matrix event and the utterance time. Like all other embedded CPs, they are incompatible with future-in-the-past readings. I propose to account for this using the same rule of de re interpretation of free variables discussed in §6.2.2 for the interpretation of the embedded future oriented TAE suffixes.

I extend Abusch (1997, 1991)'s proposal that free variables in the scope of attitude verbs are interpreted de re to include event variables. When a TAE suffix undergoes movement out of an embedded CP, its event-denoting trace (of type *v*) is free within the scope of

the matrix attitude verb. As a result, it is interpreted de re. This de re interpretation is such that the speaker describes the event as a past event of Michael Phelps winning. Furthermore, I assume a mechanism similar to Abusch (1997)'s Upper Limit Constraint (see also footnote 4) that requires that the matrix subject stand in some "acquaintance" relation to the event such that they can have some belief about it. As such, the event can be located in the matrix subject's past or present, but not their future. This rules out future-in-the-past readings of (293) and (294). (As I noted previously in §6.2.2, this de re approach does not rule out a reading in which the embedded event is interpreted as overlapping with the matrix subject's "now"; i.e., a reading like *Läylä said that Michael Phelps is winning*. This reading perhaps could be ruled out in future work through an even more restrictive implementation of Abusch 1997's Upper Limit Constraint.)

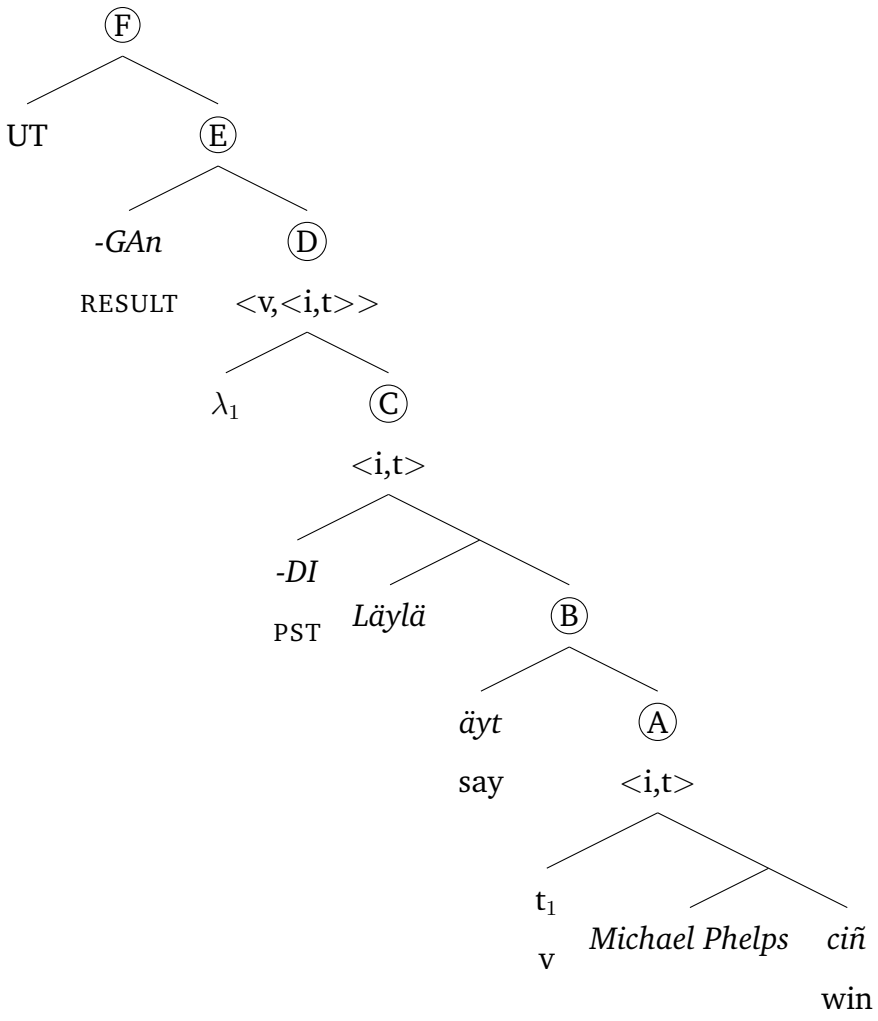
I give a rough compositional derivation of the speaker oriented readings of embedded *-GAn* in §6.2.4.1 and *-DI* in §6.2.4.2, omitting a fully compositional analysis of the de re component. I discuss their pragmatic competition in §6.2.4.3.

6.2.4.1 Speaker oriented (wide scope) reading of embedded *-GAn* 'RESULT'

I provide a tree structure for (293) in (295). In this expression, *-GAn* undergoes movement out of the embedded clause to a position superior to the matrix attitude verb. As a result, the embedded event of Michael Phelps winning is "viewed" from the utterance time.

In §6.2.1, I stipulated that when the TAE suffixes undergo movement, they leave behind traces of type *v*. This assumption enables the matrix verb *äyt* 'to say' to consistently combine with an embedded clause of type <*i,t*>, regardless of whether or not the embedded TAE suffix undergoes movement. Furthermore, since this event-denoting trace is free within the scope of *äyt* 'to say,' it is interpreted de re. This accounts for the lack of future-in-the-past.

(295) Läylä [Michael Phelps ciñ-**gän**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-RESULT-3SG COMP] say-PST-3SG
'Läylä said that [and I have indirect evidence that] Michael Phelps won.'



I give a semantics for (295) in (296). The denotation for the embedded clause is in (296a); this includes the trace, and as such is interpreted relative to the assignment function *g*. The denotation for the entire expression is in (296f). I provide a denotation for (295) that is enriched by a de re mechanism in (296g), akin to my treatment of the indexical time argument of the embedded future oriented TAE suffixes in §6.2.2. Again, the time arguments of the metalanguage predicates do not contribute any temporal meaning.

(296) a. $[\![\text{Ⓐ}]\!]^g = \lambda i.\ \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)$

b. $[\![\text{Ⓑ}]\!]^g = \lambda x \lambda e' \lambda i'.\ \textsc{say}(x)(e')(i') \subseteq \{i: \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)\}$

c. $[\![\text{Ⓒ}]\!]^g = \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)\}]$

d. $[\![\text{Ⓓ}]\!]^g = \lambda e \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \textsc{win}(\text{MP})(e)(i)\}]$

e. $[\![\textcircled{E}]\!]^g = \lambda i'.\ \exists e[\exists s[e \ll s \ \&\ i' \subset \tau(s)\ \&\ \exists e'[\tau(e') < i'\ \&$
$\textsc{say}(\text{Läylä})(e')(i') \subseteq \{i\text{: }\textsc{win}(\text{MP})(e)(i)\}]]]$

f. $[\![\textcircled{F}]\!]^g = 1$ iff $\exists e[\exists s[e \ll s \ \&\ \text{UT} \subset \tau(s)\ \&\ \exists e'[\tau(e') < \text{UT}\ \&$
$\textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i\text{: }\textsc{win}(\text{MP})(e)(i)\}]]]$

g. $[\![\textcircled{F'}]\!]^g$ (enriched with a de re mechanism) = 1 iff there is an (event) description d such that $\exists e[\exists s[e \ll s \ \&\ \text{UT} \subset \tau(s)\ \&\ \exists e'[\tau(e') < \text{UT}\ \&\ d(\tau(e')) = e$
$\&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i\text{: }\textsc{win}(\text{MP})(d(i))(i)\}]]]$

By "event description" here, I refer to a function from times to events (also called an event concept). The enriched denotation in (296g) is true iff there is a description d such that there exists an event *e* and a state *s*, where *s* is contingent on *e* and the speaker's utterance time is in the runtime of *s*. There exists an event e' and the runtime of e' precedes the utterance time; the description applied to the runtime of e' is equal to *e* and e' is an event of Läylä saying. The set of times compatible with what Läylä says in e' is a subset of the set of times *i*, where *i* is Läylä's "now"; the description applied to Läylä's "now", *d(i)*, yields an event of Michael Phelps winning (as does the description applied to the runtime of the event of Läylä saying in the matrix clause).

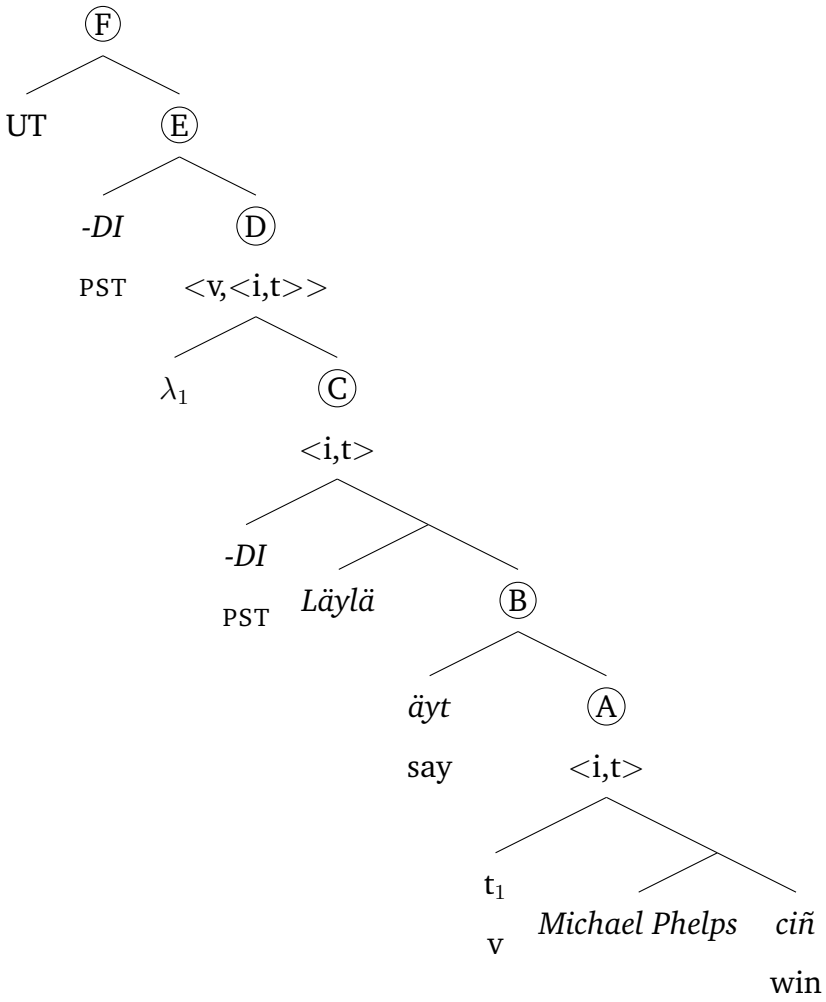
In this denotation, the utterance time is located in the runtime of the contingently related post-state of Michael Phelps winning. Since the event is "viewed" from the utterance time, this results in a speaker-oriented evidential interpretation of the embedded event.

According to the suitability conditions on event post-states discussed in Chapter 5, (296f) locates both the event of Läylä saying and the event of Michael Phelps winning in the past of the utterance time. Since the embedded event-denoting trace is interpreted de re, I again assume something akin to Abusch (1997)'s ULC in which Läylä must be "acquainted" with the event of Michael Phelps winning in some way; i.e., the event must be in either her past or present. (296f) is therefore compatible with a reading in which both Läylä and the speaker interpret the embedded event as being in their past; i.e., a backshifted reading of the embedded clause.

6.2.4.2 Speaker oriented (wide scope) reading of embedded *-DI* 'PST'

I provide a tree structure for (294) in (297). In this expression, embedded *-DI* undergoes movement to a position superior to the matrix attitude verb. As a result, the time from which the embedded event is "viewed" is the utterance time. *-DI* leaves behind a trace of type *v*, which is free in the scope of the matrix attitude verb and as such is interpreted de re. This rules out a future-in-the-past reading of (297).

(297) ?Läylä [Michael Phelps ciñ-**de**-∅ dip] äyt-te-∅.
Läylä [Michael Phelps win-PST-3SG COMP] say-PST-3SG
'Läylä said that [and I have direct evidence that] Michael Phelps won.'



I give a semantics for (297) in (298). The denotation for the embedded clause is in (298a), including the event-denoting trace. The denotation for the entire expression is

in (298f). I provide a denotation for (297) that is enriched by a de re interpretation in (298g).

(298) a. $[\![\text{Ⓐ}]\!]^g = \lambda i.\ \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)$

b. $[\![\text{Ⓑ}]\!]^g = \lambda x \lambda e' \lambda i'.\ \textsc{say}(x)(e')(i') \subseteq \{i: \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)\}$

c. $[\![\text{Ⓒ}]\!]^g = \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \textsc{win}(\text{MP})([\![t_1]\!]^g)(i)\}]$

d. $[\![\text{Ⓓ}]\!]^g = \lambda e \lambda i'.\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \textsc{win}(\text{MP})(e)(i)\}]$

e. $[\![\text{Ⓔ}]\!]^g = \lambda i'.\ \exists e[\tau(e) < i' \ \&\ \exists e'[\tau(e') < i' \ \&\ \textsc{say}(\text{Läylä})(e')(i') \subseteq \{i: \textsc{win}(\text{MP})(e)(i)\}]]$

f. $[\![\text{Ⓕ}]\!]^g = 1$ iff $\exists e[\tau(e) < \text{UT} \ \&\ \exists e'[\tau(e') < \text{UT} \ \&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i: \textsc{win}(\text{MP})(e)(i)\}]]$

g. $[\![\text{Ⓕ}']\!]^g$ (enriched with a de re mechanism) = 1 iff there is an (event) description d such that $\exists e[\tau(e) < \text{UT} \ \&\ \exists e'[\tau(e') < \text{UT} \ \&\ d(\tau(e')) = e \ \&\ \textsc{say}(\text{Läylä})(e')(\text{UT}) \subseteq \{i: \textsc{win}(\text{MP})(d(i))(i)\}]]$

The enriched denotation in (298f) is true iff there is an event description *d* such that there exists an event *e*; the runtime of *e* precedes the utterance time. There also exists an event *e′*; the runtime of *e′* also precedes the utterance time. The description applied to the runtime of *e′* is equal to *e*, and *e′* is an event of Läylä saying. The set of times compatible with what Läylä says in *e′* is a subset of the set of times *i*, where *i* is Läylä's "now." The description applied to Läylä's "now", *d(i)*, is an event of Michael Phelps winning (as is the description applied to the runtime of the event of Läylä's saying in the matrix clause).

In this denotation, the time argument of the moved *-DI* is valued by the utterance time; i.e, the embedded event of Michael Phelps winning is "viewed" from the utterance time. This results in a speaker-oriented evidential interpretation of the embedded event. Since the embedded event-denoting trace is interpreted de re, (298f) is compatible with a reading in which both Läylä and the speaker interpret the embedded event as being in their past; i.e., a backshifted reading of the embedded clause.

6.2.4.3 Deriving evidential readings of speaker oriented (wide scope) *-DI* and *-GAn* through pragmatic competition

As discussed previously in §6.2.3.3, I proposed in Chapter 5 that the evidential readings of *-DI* and *-GAn* are derived pragmatically. (I refer the reader to Chapter 5 for more in-depth discussion of how these pragmatic readings are derived.) In the derivations in (295) and (297), the embedded TAE suffixes scope out of the embedded clause. As such, I do not need to utilize any mechanism of pragmatic competition in embedded environments, as discussed previously with respect to the embedding data in §6.2.2.3 and §6.2.3.3.

In sum, I propose that the pragmatic competition of wide scope *-DI* and *-GAn* proceeds identically as in matrix clauses. Since the time argument of the embedded TAE suffixes is valued by the utterance time, their evidential interpretation is evaluated with respect to the speaker, not the matrix subject. The resulting meaning is that instances of wide scope *-DI* convey that the speaker has direct evidence for the embedded event, whereas instances of wide scope *-GAn* convey that the speaker has indirect evidence for the embedded event.

6.2.5 Recap

I proposed a movement-based account of the evidential interpretations of the Tatar TAE suffixes in CPs embedded under attitude verbs. In brief, I proposed that the evidential interpretation of a TAE suffix in an embedded CP is determined by its location in the clause relative to the matrix attitude verb.

Generally speaking, if a TAE suffix remains in situ in the embedded clause, its open time argument will be valued by (an interval that includes) the matrix subject's "now," as introduced by the semantics of the attitude verb. This results in a reading of matrix subject oriented evidentiality with respect to the event described by the embedded clause, i.e., evidential shift. Conversely, if a TAE suffix undergoes movement out of the embedded clause to a position that is superior to the matrix attitude verb, its open time argument is valued by the utterance time. (This is the same configuration as for the interpretation of TAE suffixes in matrix clauses, as I discussed in Chapter 5.) In such expressions, the

embedded event is "viewed" from the utterance time. This results in a reading of speaker oriented evidentiality with respect to the event described by the embedded clause.

I showed in Chapter 2 that future oriented TAE suffixes in embedded CPs are obligatorily interpreted as reflecting the matrix subject's evidence for the embedded event; i.e., they necessarily undergo evidential shift. As a result, I stipulated in §6.2.2 that the future oriented TAE suffixes must take narrow scope with respect to the attitude verb, perhaps for syntactic reasons. Conversely, I showed in Chapter 2 that past oriented TAE suffixes in embedded CPs can reflect either the speaker's or matrix subject's evidence for the embedded event. I therefore proposed in §6.2.3 and §6.2.4 that in embedded CPs, the past oriented TAE suffixes can take either narrow or wide scope. This accounts for their optional evidential shift.

In sum, I proposed that the evidential readings of the TAE suffixes in embedded CPs arise through the same general mechanisms of pragmatic competition as in matrix clauses. I provide additional support for this pragmatic proposal in §6.3. In the following section, I sketch an analysis of the TAE suffixes in embedded verbal nominalizations. I show that in verbal nominalizations, only a subset of the TAE suffixes are grammatical. Since only a subset of the suffixes are grammatical, pragmatic competition cannot occur. As a result, no evidential readings arise. This lends additional support to a pragmatic account of evidentiality in the Tatar TAE suffixes, as opposed to an account in which their evidential meaning is hardwired into their semantics.

6.3 Tatar TAE suffixes in embedded verbal nominalizations

I showed in Chapter 2 that the Tatar embedded verbal nominalizations differ from the embedded CPs in two main respects. First, only a subset of the Tatar TAE suffixes (only *-GAn* and *-(y)AçAK*) are grammatical in these expressions. Second, the embedded verbal nominalizations are evidentially "neutral"; they do not contribute any evidential meaning. I summarize these empirical facts in Table 6.2.

TAE suffix	Grammaticality in verbal nominalizations
-DI 'PST'	*
-GAn 'RESULT'	✓ (no evidential reading)
-(y)AçAK 'PROSP'	✓ (no evidential reading)
-(y)Er 'FUT'	*

Table 6.2: Grammaticality and evidential interpretation of the Tatar TAE suffixes in embedded verbal nominalizations (repeated from Chapter 2).

I take the ungrammaticality of *-DI* and *-(y)Er* in verbal nominalizations to support treating *-DI/-(y)Er* and *-GAn/-(y)AçAK* as two separate natural classes of morphemes. This falls out naturally from my proposal in Chapter 5 to treat *-DI* and *-(y)Er* as tenses and *-GAn* and *-(y)AçAK* as aspects, and cross-cuts their very different evidential contributions in matrix clauses.

Kornfilt and Whitman (2011) and Borsley and Kornfilt (2000) (writing on Turkish) give an account under which verbal nominalizations can occur at different "heights" in the syntax. The location of the nominal functional head in the syntax determines what material is included within the nominalization, and what is excluded from it. Without spelling out the precise syntax of Tatar embedded nominalized clauses, I propose that Tatar verbal nominalizations truncate at such a point that aspectual morphology (i.e., *-GAn* and *-(y)AçAK*) is included within the nominalization, and tense morphology (i.e., *-DI* and *-(y)Er*) is excluded. This accords with the fact that tense is syntactically higher than aspect. Furthermore, no person/number agreement is marked on the verb within the nominalization; this accords with the absence of a TP projection.[5] (See Şahan 2002 for an in-depth discussion

[5]Verbs in embedded nominalized clauses host possessive marking reflecting the person and number of the subject, as in (299a). However, this possessive marking is not part of the verbal agreement paradigm, which I described in Chapter 2. I show in (299b) that substituting verbal agreement for possessive marking is ungrammatical.

(299) a. Timur [minem yeger-gen-**em**-ne] äyt-te-∅.
Timur [1SG.GEN run-RESULT-1SG.POSS-ACC] say-PST-3SG
'Timur said that I ran.'

b. * Timur [minem yeger-gen-**min**-ne] äyt-te-∅.
Timur [1SG.GEN run-RESULT-1SG-ACC] say-PST-3SG
Intended: 'Timur said that I ran.'

of Tatar verbal nominalizations.)

In Chapter 5, I proposed that the evidential interpretations of the TAE suffixes turn on the availability of pragmatic competition between the sets of past- and future-oriented tense and aspect suffixes. In the absence of pragmatic competition with the tense suffixes, we predict that the aspectual suffixes should not undergo any pragmatic strengthening. As a result, they should be evidentially "neutral." This is precisely what we find in embedded verbal nominalizations. I take this data to provide support for a pragmatic account of evidentiality in the Tatar TAE suffixes. If the evidential meaning of the TAE suffixes were hardwired into their semantics, we would not expect that it could be absent in some environments.

I give rough analyses of *-(y)AçAK* and *-GAn* in embedded verbal nominalizations in §6.3.1 and §6.3.2, respectively. In Chapter 2, I noted that there are significant differences in interpretation between embedded CPs and embedded verbal nominalizations in Tatar. For instance, embedded verbal nominalizations generally convey that the speaker believes that the embedded proposition is true. Conversely, embedded CPs do not convey any speaker belief in the truth of the embedded proposition. (I refer the reader to Chapter 2 for a more thorough discussion of these semantic differences.)

For the purpose of this dissertation, I focus solely on deriving the evidential and temporal interpretations of the TAE suffixes in verbal nominalizations, and do not compositionally derive the semantic contribution of the clause type itself.[6] I treat the genitive case marking on the subject and the possessive and accusative case marking on the nominalization as being semantically vacuous, just as I treated the complementizer *dip* as being semantically vacuous in §6.2. This is a simplification, and is not intended to suggest that there are no further semantic distinctions between embedded CPs and nominalizations.

[6]See Özyıldız (2016) for an analysis of embedding strategies in Turkish. Özyıldız links the analogous semantic distinction in Turkish to the choice of clausal complementizer, i.e., whether it is a phonologically (and semantically) null nominalizing morpheme or the overt complementizer *diye* (the Turkish equivalent of Tatar *dip*).

6.3.1 *-(y)AçAK* in embedded verbal nominalizations

I repeat an example of *-(y)AçAK* in an embedded verbal nominalization in (300). I provide contexts to illustrate that *-(y)AçAK* is compatible with the matrix subject having either specific or non-specific evidence for the embedded proposition. I noted previously in Chapter 2 that the use of a verbal nominalization with *-(y)AçAK* is somewhat marginal if the speaker is aware that the matrix subject does not have specific evidence for the embedded event, as in (300b). This is due to the fact that embedded verbal nominalizations convey that the speaker believes that the embedded proposition is true. However, if the speaker does not believe that the matrix subject has strong grounds for their assertion (e.g. they lack specific evidence), they are less likely to report the assertion using an embedded nominalization. I believe that this is due to the semantics of the nominalization strategy itself, rather than any evidential contribution by *-(y)AçAK*.

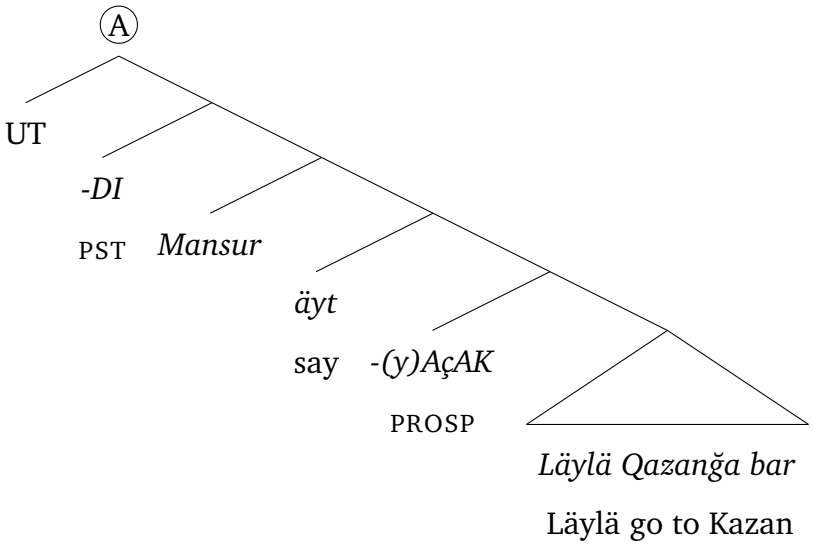
(300) a. **Specific evidence context:** Mansur is a close friend of Läylä's. He knows that Läylä has bought tickets to go to Kazan next week, and has taken time off work specifically to go on vacation there. He tells you about her plans. You express Mansur's report of what Läylä will do.

b. ?**Non-specific evidence context:** Mansur is a colleague of Läylä's. He overhears another colleague saying that Läylä will go to Kazan next week, and tells you this. He does not know about any particular plans of hers, and he has no reason to think this colleague is an authority on Läylä's actions. You express Mansur's report of what Läylä will do.

Mansur [Läylä-nen Qazan-ğa bar-**açağ**-ı-n] äyt-te-∅.
Mansur [Läylä-GEN Kazan-DAT go-PROSP-3SG.POSS-ACC] say-PST-3SG
'Mansur said Läylä will go to Kazan.'

I sketch a tree structure for (300) in (301), omitting the genitive case marking on *Läylä* and the possessive and accusative case marking on the embedded nominalization.

(301)



The denotation for the top node in (301) is roughly as in (302). Again, the time arguments of the metalanguage predicates are present solely for type purposes, and do not contribute any temporal meaning.

(302) $[\![\text{Ⓐ}]\!] = 1$ iff $\exists e'[\tau(e') < \text{UT} \;\&\; \textsc{say}(\text{Mansur})(e')(\text{UT}) \subseteq \{i\text{: } \exists e[\exists s[s \ll e \;\&\; i \subset \tau(s) \;\&\; \textsc{go-to-Kazan}(\text{Läylä})(e)(i)]]\}]$

The denotation in (302) asserts that there exists an event e'; the runtime of e' precedes the utterance time, and e' is an event of Mansur saying. The set of times compatible with what Mansur says in e' is a subset of the set of times i, where i is Mansur's "now" and there exists an event e and a state s; e is contingent on s and Mansur's "now" is located in the runtime of s and e is an event of Läylä going to Kazan. I take s to be a contingently related pre-state of e.

The temporal contribution of *-(y)AçAK* locates Mansur's "now" in a pre-state of the event of Läylä going to Kazan. I proposed in Chapter 5 that event pre-states temporally precede the event itself. (302) is therefore compatible with a temporal interpretation in which the event of Läylä going to Kazan is in the future of the utterance time as well as the time of Mansur's saying event (i.e., *Mansur said (yesterday) that Läylä will go to Kazan (two weeks from now)*). I showed in Chapter 3 that such a temporal interpretation is available.

The denotation in (302) does not include any indexical time in the embedded nominalization; this differs from the analysis I posited for future-oriented TAE suffixes in embedded CPs in §6.2.2. This is because, as I showed in Chapter 3, expressions like (300) are compatible with future-in-the-past interpretations. That is, (300) is compatible with a reading in which the event of Läylä going to Kazan is in the future of Mansur's saying event but in the past of the utterance time. This is akin to the English expression *Mansur said (two weeks ago) that Läylä would go to Kazan (yesterday)*. In embedded nominalizations, *-(y)AçAK* is therefore able to locate the embedded event time in the past of the utterance time. I take this to suggest that it is not interpreted indexically in verbal nominalizations, unlike in embedded CPs.

I will now address the lack of any evidential reading associated with the embedded nominalization in (300). If a speaker wishes to express that they believe that Mansur's report that Läylä will go to Kazan is true, they will express it using an embedded nominalization, as in (300). Since Tatar verbal nominalizations exclude tense marking, they can only express the future event using *-(y)AçAK*, not *-(y)Er*. Tatar speakers only have one choice of future oriented TAE suffix in nominalizations, unlike in matrix clauses and in embedded CPs. As a result, no pragmatic competition occurs and no evidential reading arises.

6.3.2 *-GAn* in embedded verbal nominalizations

I repeat an example of *-GAn* in an embedded verbal nominalization in (303). I provide contexts to demonstrate that (303) is compatible with the matrix subject having either direct or indirect evidence for the embedded proposition; that is, it is evidentially neutral.

(303) a. **Matrix subject has direct evidence for embedded proposition:**
Mansur tells you that he accompanied Läylä to the train station yesterday and watched her get on a train to go to Kazan. You express Mansur's report of what Läylä did yesterday.

b. **Matrix subject has indirect evidence for embedded proposition:**
Mansur finds a receipt for a train ticket to Kazan in Läylä's desk. He tells you that he infers that Läylä went to Kazan yesterday. You express Mansur's report of what Läylä did.

Mansur [Läylä-nen Qazan-ğa kit-ep bar-**ğan**-ı-n] äyt-te-∅.
Mansur [Läylä-GEN Kazan-DAT leave-IP go-RESULT-3SG.POSS-ACC] say-PST-3SG
'Mansur said Läylä left for Kazan.'

I noted in Chapter 3 that in embedded verbal nominalizations, *-GAn* is compatible with two different temporal interpretations. The first is a backshifted reading in which *-GAn* is interpreted as in embedded CPs; that is, it places the time of the embedded event in the past of the time of the matrix event. An equivalent English expression would be *Mansur said (yesterday) that Läylä left for Kazan (two weeks ago)*.

The second available temporal interpretation is one in which *-GAn* expresses overlap of the runtime of the embedded event and the utterance time. An equivalent English expression would be something like *Mansur said that Läylä is leaving for Kazan (right now)*. I repeat (304) from Chapter 3 to demonstrate that such a reading of *-GAn* is possible in a nominalization.

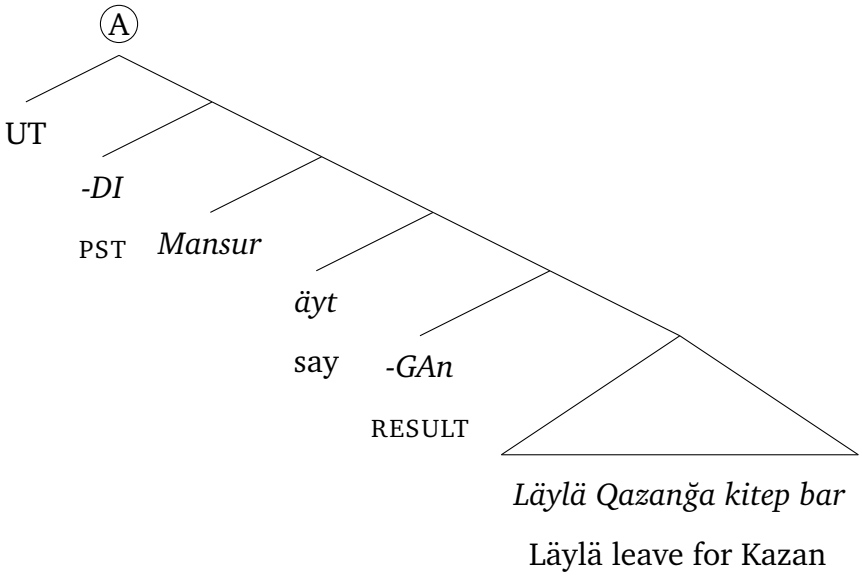
(304) Timur [Alsu-nın (**xäzer**) çäkçäk-ne aşı-**ğan**-ı-n] äyt-te-∅.
Timur [Alsu-GEN right.now çäkçäk-ACC eat-RESULT-3SG.POSS-ACC] say-PST-3SG
'Timur said that Alsu is eating the çäkçäk (right now).'

This present temporal interpretation of *-GAn* requires modifying the denotation that I posited in (273b). I modify the denotation of *-GAn* such that it locates its time argument i in the combined runtimes of the event and its post-state, represented as $\tau(e+s)$ in (305). That is, it locates i at some point in time after the beginning of the event.

(305) $[\![\text{-}GAn]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[\exists s[e \ll s \ \& \ i \subset \tau(e{+}s) \ \& \ p(e)(i)]]$

With this updated denotation in mind, I now give a tree structure for (303) in (306). I again omit the contribution of the genitive case marking on *Läylä* and the possessive and accusative case marking on the embedded nominalization.

(306)



The denotation of the top node of (306) is roughly as in (305).

(307) $[\![\text{Ⓐ}]\!] = 1$ iff $\exists e'[\tau(e') < \text{UT} \ \& \ \textsc{say}(\text{Mansur})(e')(\text{UT}) \subseteq \{i\text{: } \exists e[\exists s[e \ll s \ \&$
$i \subset \tau(e{+}s) \ \& \ \textsc{leave-for-Kazan}(\text{Läylä})(e)(i)]]\}]$

(307) asserts that there exists an event e'; the runtime of e' precedes the utterance time, and e' is an event of Mansur saying. The set of times compatible with what Mansur says in e' is a subset of the set of times i, where i is Mansur's "now" and there exists an event e and a state s. The post-state s is contingent on e and Mansur's "now" is located in the combined runtimes of s and e; e is an event of Läylä leaving for Kazan.

The lack of evidential reading associated with (303) arises through the same mechanism as in §6.3.1. If the speaker wishes to convey that they believe that Mansur's report that Läylä left/is leaving for Kazan is true, they will express it using an embedded nominalization, as in (303). Since tense marking is ungrammatical in verbal nominalizations, the

speaker only has one choice of TAE suffix that they can use. That is, they can only express this event of leaving using *-GAn*, not *-DI* (or *-A*). As a result, no pragmatic competition occurs between the TAE suffixes and no evidential reading arises.

I will now address the possible temporal interpretations of (303). The denotation in (307) is compatible with a backshifted reading of the time of the embedded event; this reading arises if Mansur's "now" is located in the runtime of the post-state of Läylä leaving for Kazan. Since post-states necessarily temporally follow their associated event, this places the event of Läylä leaving for Kazan in Mansur's past. I showed in Chapter 3 that this reading is attested. (307) is also compatible with a present interpretation of embedded *-GAn*, as in the example in (304). This reading arises if Mansur's "now" is located within the runtime of the event of Läylä leaving for Kazan.

The updated denotation for *-GAn* in (305) therefore successfully accounts for its temporal interpretation in embedded verbal nominalizations. However, this denotation is at odds with how *-GAn* is interpreted in matrix clauses and in embedded CPs. I showed in Chapter 3 that in these environments, only a past oriented interpretation of *-GAn* is available.

I tentatively propose that this past oriented interpretation of *-GAn* arises from pragmatic competition with the present tense suffix *-A* 'PRES.' I gave an updated denotation for *-A* in (273c); I repeat it here in (308).

(308) $[\![\text{-}A]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[i \subseteq \tau(e)\ \&\ p(e)(i)]$

As a tense morpheme, *-A* is ungrammatical in verbal nominalizations, i.e., the environments in which the "non-future" reading of *-GAn* occurs. In the environments in which *-A* is grammatical, *-GAn* must have a past oriented temporal meaning. This suggests to me that the temporal contributions of these suffixes may interact in some way, i.e. they may be pragmatically related.

In matrix clauses and embedded CPs, *-GAn* 'RESULT' and *-A* 'PRES' compete for the same syntactic "slot" on the verb. Both of these morphemes can locate their time argument *i* within the runtime of the described event. However, *-GAn* 'RESULT' is also compatible with *i* being located in the post-state of the event; i.e., *-GAn* entails *-A*. The availability of *-A* in

matrix clauses and embedded CPs therefore pragmatically "blocks" the present use of *-GAn* in these expressions, as shown graphically in Figure 6.1.

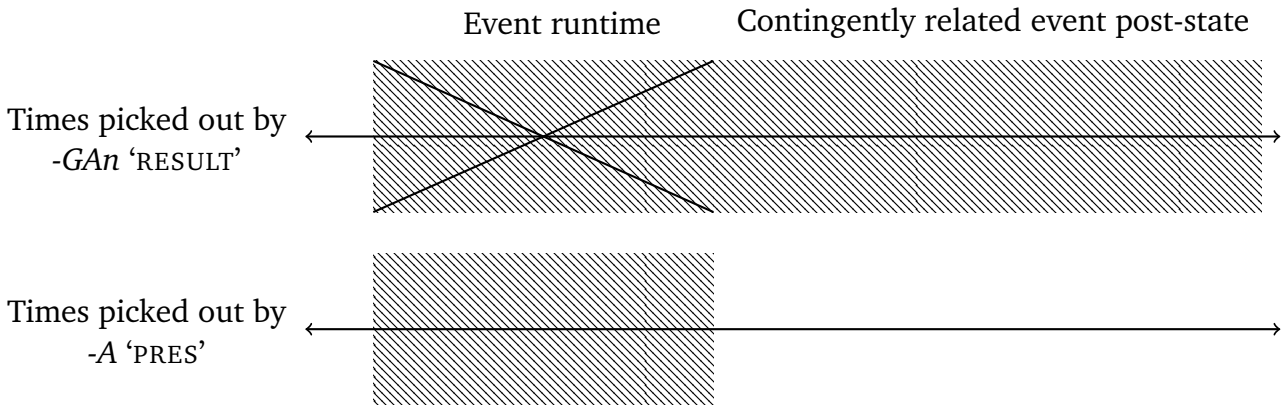


Figure 6.1: Graphic representation of the result of pragmatic competition between *-GAn* 'RESULT' and *-A* 'PRES.'

Given this competition between *-GAn* and *-A*, the pragmatically strengthened meaning of *-GAn*(*p*) expressions is as in (309). The pragmatically strengthened denotation of *-GAn* in (310) is the same as the (strictly past oriented) denotation that I posited in (273b), i.e., it locates its time argument *i* within the runtime of the post-state *s*.

(309) $\textit{-GAn}(p)_{strengthened} \rightsquigarrow \textit{-GAn}(p)\ \&\ \neg\ \textit{-A}(p)$

(310) $[\![\textit{-GAn}_{strengthened}]\!] = \lambda p_{<v,<i,t>>} \lambda i.\ \exists e[\exists s[e \ll s\ \&\ i \subset \tau(s)\ \&\ p(e)(i)]]$

I acknowledge that this pragmatic account is highly stipulative. At present, I have not collected any data that suggests that the past oriented interpretation of *-GAn* can be cancelled (either contextually or explicitly) in either matrix clauses or embedded CPs. However, I am reluctant to posit multiple denotations for *-GAn*, depending on whether it occurs in a verbal nominalization or a matrix clause/embedded CP. As I noted above, this "non-future" use of *-GAn* correlates with the absence of *-A*. When speakers are unable to use *-A* to express the inclusion of the utterance time within the runtime of the embedded event, they must use some other strategy to convey this temporal relationship; they therefore turn to *-GAn*.

Adopting the updated denotation of *-GAn* in (305) requires making a further assumption regarding my previous proposals. In order to preserve the pragmatic account of evidentiality that I laid out in Chapter 5, I must assume that *-DI* competes with the already strengthened, past oriented denotation of *-GAn* in (310). Without this assumption, there is no entailment relationship between *-DI* and *-GAn*.

6.3.3 Recap

I sketched an analysis of the Tatar TAE suffixes in embedded verbal nominalizations. I focused on accounting for the evidential and temporal interpretations of the TAE suffixes in these constructions; for the purpose of this dissertation, I did not address the semantic contribution of the nominalization itself.

The distribution and interpretation of the Tatar TAE suffixes in embedded verbal nominalizations supports the proposal I laid out in Chapter 5 in two major ways. First, as I showed in Chapter 2, the TAE suffixes in embedded verbal nominalizations do not contribute any evidential meaning. This would be unexpected if evidentiality is hardwired into their semantics. However, this is compatible with my pragmatic proposal from Chapter 5. Since the tense suffixes (i.e., *-DI* and *-(y)Er*) are ungrammatical in verbal nominalizations, they cannot pragmatically compete with the aspectual suffixes (i.e., *-GAn* and *-(y)AçAK*) in these environments. As a result, no evidential readings arise.

Second, the distribution of the suffixes in nominalizations supports treating the Tatar TAE suffixes as two separate natural classes. This falls out naturally from my proposal in Chapter 5 to treat *-DI* and *-(y)Er* as tenses, and *-GAn* and *-(y)AçAK* as aspects. I showed in this section that only the aspectual suffixes are grammatical in verbal nominalizations. This follows from syntactic grounds, given the typological observation that aspects are lower than tenses and the assumption that Tatar (like other SOV languages) has a left-branching syntax.

6.4 Conclusion

In this chapter, I addressed the distribution and use of the Tatar TAE suffixes in some semantically embedded environments. I showed in §6.2 that the pragmatic analysis I proposed for the TAE suffixes in Chapter 5 can be maintained in embedded CPs, assuming the availability of some mechanism for pragmatic competition in embedded environments. I derived the speaker-oriented versus matrix-subject oriented readings of TAE suffixes in embedded CPs through their location in the syntax, i.e., whether they take wide scope or narrow scope with respect to the matrix attitude verb. In §6.3, I showed that the lack of evidentiality associated with the TAE suffixes in verbal nominalizations can be accounted for by a lack of pragmatic competition. This lack of pragmatic competition arises due to the fact that verbal nominalizations in Tatar exclude tenses.

In the following brief chapter, I summarize what my proposal does and does not accomplish. I discuss some implications of my proposal for cross-linguistic variation in tense/aspect and evidential systems, and lay out some future directions for this project.

CHAPTER 7

Conclusion

7.1 What this dissertation does and does not do

The theory that I proposed in this dissertation is based solely on Tatar data, and at present is only intended to account for Tatar's TAE system. (See §7.2 for possible extensions of this work to other languages.) As a result, there are limitations to my proposal; that is, there is data it can and cannot account for. In lieu of a complete recap of the theory I have proposed in this dissertation, I summarize the puzzles that my proposal addresses in §7.1.1. In §7.1.2, I spell out the kinds of data that fall outside the scope of my analysis.

7.1.1 Data this dissertation accounts for

This dissertation is geared towards accounting for portmanteau tense/aspect and evidential morphemes of the kind found in Tatar, among many other Turkic languages. In Chapter 5, I proposed that the semantics of the Tatar TAE suffixes are underlyingly temporal; i.e., they locate events in time. I did not encode any evidential meaning in their denotations; rather, I proposed that the evidential meanings of these suffixes arise pragmatically as a byproduct of their temporal meanings.

This accounted for the following puzzles raised by the Tatar data:

- **Why tense/aspect and evidentiality are fused in Tatar.**

  My analysis in Chapter 5 accounts for the portmanteau nature of the Tatar TAE suffixes by treating them as asserting temporal meaning and pragmatically implicating evidential meaning.

- **Why it is that the future-oriented Tatar TAE suffixes have apparently evidential readings.**

  I showed in Chapter 2 that the future-oriented TAE suffixes *-(y)AçAK* and *-(y)Er* appear to make a non-standard evidential distinction that can be descriptively characterized as marking "specific" versus "non-specific" evidence, respectively. I showed in Chapter 4 that EAT and trace theories of evidentiality cannot account for these readings. Furthermore, Aikhenvald (2004) notes that evidentiality is relatively uncommon in future oriented expressions.

  In Chapter 5, I account for this typologically unusual data by proposing that the future-oriented TAE suffixes can invoke causally related event pre-states. If the speaker makes an assertion including the prospective aspectual suffix *-(y)AçAK*, they thereby locate themselves in a pre-state of the described event. I showed that the kind of events and states that speakers perceive in event pre-states could be descriptively labeled as being "specific" evidence for the scope proposition.

- **Why the Tatar TAE suffixes have the evidential meanings that they do.**

  As I noted in the Introduction to this dissertation, it is a striking cross-linguistic fact that evidential meanings are as restricted as they are. My proposal in Chapter 5 explains why the Tatar TAE suffixes have direct/indirect and "specific"/"non-specific" evidential readings. I argue that the kinds of events and states that speakers perceive in event pre-states and post-states can be descriptively thought of as "specific" and indirect evidence for the scope proposition, respectively.

- **Why the evidential contribution of the TAE suffixes can sometimes be cancelled (either contextually or explicitly).**

  In Chapter 5, I showed that the "non-specific" evidential reading associated with *-(y)Er* can be explicitly cancelled in discourse. I also showed that the direct evidential reading of *-DI* is contextually defeasible. This is predicted if we treat the evidential readings of these suffixes as arising pragmatically.

- **Why the past- versus future-oriented TAE suffixes differ with respect to their evidential interpretations in embedded CPs.**

  I showed in Chapter 2 that in embedded CPs, the evidential component of past oriented TAE suffixes can optionally undergo evidential shift, whereas the evidential component of future oriented TAE suffixes must undergo evidential shift.

  In Chapter 6, I accounted for the shifted versus non-shifted evidential interpretations of the embedded TAE suffixes by proposing that they can take either wide scope or narrow scope with respect to the matrix attitude verb. Their relative scope determines whether they are interpreted relative to the speaker or to the matrix subject. I argued that past oriented TAE suffixes can take either wide or narrow scope, whereas the future oriented TAE suffixes must take narrow scope.

- **Why only a subset of the Tatar TAE suffixes are grammatical in verbal nominalizations.**

  In Chapter 5, I proposed to treat the TAE suffixes as consisting of two natural classes: *-DI* and *-(y)Er* are tenses and *-GAn* and *-(y)AçAK* are aspects. Only the aspectual suffixes are grammatical in verbal nominalizations. This accords with our knowledge of the relative heights of tenses and aspects cross-linguistically.

- **Why the TAE suffixes in verbal nominalizations are evidentially neutral.**

  I proposed in Chapter 5 that the evidential readings of the Tatar TAE suffixes arise through pragmatic competition. In verbal nominalizations, only the aspectual suffixes are grammatical. In the absence of pragmatic competitors, the aspectual suffixes are felicitous with any evidence type; that is, they are evidentially neutral.

7.1.2 Data this dissertation does not account for

This dissertation is not intended to account for all evidential systems. My proposal is geared solely towards languages with portmanteau tense/aspect and evidential morphemes (see §7.2 for discussion of extensions of this proposal to other languages).

My theory is not applicable to languages with separate tense/aspect and evidential systems. Such systems are described for e.g. Cheyenne (Murray 2010), Cuzco Quechua (Faller 2002), and St'át'imcets (Matthewson 2011), among many other languages. The temporal component of my analysis in Chapter 5 is simply not relevant to these languages.

My proposal also does not account for languages that make fine-grained evidential distinctions beyond simply "direct" and "indirect" evidence. The use of event pre-states and post-states in my analysis cannot distinguish between fine-grained evidential categories such as visual versus aural evidence for the scope proposition. Typological work suggests that such languages do exist, e.g. Foe, Mamainde, Tariana, Tuyuca, Wintu, among others (Aikhenvald 2004, 42-66). I give some relevant Foe data in (311).

(311) FOE (KUTUBUAN) (Aikhenvald 2004, 62)

a. aiya bare wa-**boba'ae**.
air plane come-VIS.EVID
'An airplane is coming [I can see it].'

b. aiya bare wa-**bida'ae**.
air plane come-NONVIS.EVID
'An airplane is coming [I can only hear it].'

c. Kabe Irabo wa-**ada'ae**.
Mr. Irabo come-DEDUCTIVE.EVID
'Mr. Irabo is coming [I can hear him speaking and can recognize his voice].'

d. Kabe Maduane minage wa-**bubege**.
Mr. Maduane still come-PREVIOUS.EVIDENCE.EVID
'Mr. Maduane is still coming [both left together, but the speaker came faster than Maduane, and so he knows he's still on the way].'

The verbal suffixes in (311) appear to make four different evidential distinctions. The suffix in (311a) indicates that the speaker has visual evidence for the scope proposition, while the suffix in (311b) indicates that speaker has nonvisual evidence for the scope proposition. (311c) indicates that the speaker is inferring the truth of the scope proposition through indirect evidence, whereas (311d) indicates that the speaker previously had, but does not currently have, evidence for the scope proposition.

Interestingly, it appears to be the case that tense/aspect and evidentiality can be linked even in languages with fine-grained evidential systems like Foe. In the Tariana data in (312), the verbal suffixes—like the Tatar TAE suffixes—convey both temporal and evidential meanings.

(312) TARIANA (ARAWAK) (Aikhenvald 2004, 2)

a. Juse iɾida di-manika-**ka**.
José football 3SG.NF-play-RECENT.PST.VIS
'José has played football [we saw it].'

b. Juse iɾida di-manika-**nihka**.
José football 3SG.NF-play-RECENT.PST.INF
'José has played football [we infer it from visual evidence].'

c. Juse iɾida di-manika-**sika**.
José football 3SG.NF-play-RECENT.PST.ASSUME
'José has played football [we assume this on the basis of what we already know].'

In these examples, all of the Tariana suffixes mark recent past tense in addition to some evidential meaning. The proposal that I laid out in Chapter 5 cannot currently account for these distinctions. As my analysis is currently formulated, my use of event pre- and post-states can only make a binary distinction between the kinds of evidence that the speaker has access to. However, further fieldwork on languages like Tariana could reveal e.g. an epistemic modal component in the semantics of some of their TAE suffixes, introducing an additional dimension of variation with respect to their interpretation.

In sum, my proposal is limited in its scope. However, this isn't necessarily a shortcoming. Aikhenvald (2018, 2004) notes that evidentials are a highly morphosyntactically heterogenous class; they can occur as free particles, clitics, or affixes. Furthermore, evidential meanings can co-occur with a variety of other grammatical categories, including tense and aspect systems (Izvorski 1997), complementizers (Noonan 1985), noun class marking (Gluckman and Bowler 2016), and pronominal systems (Bashir 2006), among others. I take this morphosyntactic diversity to suggest that the theoretical mechanisms

through which evidential meanings arise may be similarly diverse.[1] I therefore see no issue in introducing a novel account of evidentiality that targets some portion of the range of evidential data.

7.2 Cross-linguistic observations and predictions

The theory proposed in this dissertation is based on Tatar data. However, a sound linguistic theory should in principle be applicable to a number of languages. While I have tailored my proposal to account specifically for the quirks of Tatar, the conceptual basis of my proposal should be able to be extended to other languages with similar TAE systems. My analysis also makes some predictions regarding expected cross-linguistic variation, which I will now address.

7.2.1 Languages with tense/aspect pairs

In Chapter 5, I proposed that the evidential readings of the Tatar TAE suffixes arise through pragmatic competition. This predicts that if a language has a "pair" of tense and aspect morphemes (e.g. both prospective aspect and future tense morphology), these expressions can be associated with evidential readings like those I described for Tatar in Chapter 2.[2]

The cross-linguistic applicability of this proposal is supported most strongly by the existence of a number of similar TAE suffixes with similar morphosyntactic distributions and evidential interpretations in other Turkic languages. As I noted in Chapters 2 and 3, other Turkic languages that have tense/aspect pairs with similar evidential readings include Azerbaijani (Öztopçu 2000), Bashkir (Poppe 1964), Gagauz (Pokrovskaya 1964), Karachay (Seegmiller 1996), Karaim (Mysaev 1964), Kumyk (Johanson and Csató 1998),

[1] If we assume—as I do—that evidentiality can arise through different theoretical mechanisms, the question is then how these theories converge on such similar meanings. As Korotkova (2016) observes, evidentials share a striking number of semantic similarities across languages.

[2] Here I refer strictly to viewpoint aspects that invoke event pre- and post-states; i.e., prospective and resultative aspects.

Kyrgyz (Abduldaev and Zakharova 1987), Salar (Dwyer 2000), Tofa (Rassadin 1978), Turkish (Kornfilt 1997, Slobin and Aksu 1982), Tuvan (Harrison 2000, Krueger 1977), and Uzbek (Straughn 2011), among others. Extension of this theory to other Turkic languages is therefore a natural next step. Doing so would first involve checking for properties such as the cancellability of the evidential readings associated with these morphemes, and their distribution and interpretation in embedded environments (e.g. do they pattern similarly to the Tatar TAE suffixes in verbal nominalizations?).

The proposal that I laid out for semantically embedded TAE suffixes in Chapter 6 used the relative scope of the TAE suffixes and the matrix attitude verb to account for their available evidential reading(s). I proposed that the future oriented Tatar TAE suffixes must remain in situ under matrix attitude verbs; as a result, their evidential contribution necessarily reflects the matrix subject's evidence for the embedded proposition. However, once could imagine a system in which **all** TAE suffixes could (or must) scope out of embedded clauses. One could also imagine a system in which all embedded TAE suffixes must remain in situ. Extension of this proposal to other Turkic languages could involve checking for cross-linguistic variation in this respect.

If a language has a pair of tense/aspect expressions but no evidential reading associated with them, it could be the case that the two expressions are too morphosyntactically distinct to be pragmatic competitors. This could follow from a theory like Katzir (2007) in which possible scalar alternatives are computed based on morphosyntactic similarity. I remain agnostic as to the denotation of English present perfects; however, as a thought experiment, we could posit that English present perfects, like Tatar *-GAn(p)* expressions, invoke a causally related event post-state. The morphosyntactic differences between English present perfects (*Leroy **has eaten** the satsuma*) and past tense expressions (*Leroy **ate** the satsuma*) are such that we would not expect them to pragmatically compete. As a result, we would not predict English present perfects to be associated with indirect evidentiality.

7.2.2 Languages without tense/aspect pairs

In Chapter 5, I proposed that the evidential reading of the Tatar TAE suffixes arises in part from the semantics of the aspectual suffixes. More specifically, I argued that if a speaker has the **option** to invoke an event pre- or post-state (by using a prospective or resultative aspect), their choice to do so leads to readings of "specific" evidentiality/indirect evidentiality, respectively.[3] This predicts an absence of portmanteau TAE morphology if a language lacks either resultative or prospective aspects. (Nothing in the system prevents a language from having only a resultative aspect or only a prospective aspect; however, if a language lacks e.g. a resultative aspect, we would expect that the language would also lack any past-oriented portmanteau TAE expression.)

I argued in Chapter 6 that if pragmatic competition between a tense/aspect pair is unavailable, no evidential readings should arise. I showed that this is observed in verbal nominalizations in Tatar. The tense suffixes are ungrammatical in verbal nominalizations; as a result, no pragmatic competition occurs between the tense and aspectual suffixes and no evidential readings arise. This proposal therefore predicts that if a language **only** has resultative and prospective aspects, those expressions should not be associated with any evidential readings.

This pragmatic proposal also predicts that a language will lack portmanteau TAE morphemes—even if it has resultative/prospective aspects—if it lacks any appropriate pragmatic competitors. For example, if a language has both resultative aspect and past tense, but their realizations are morphosyntactically very distinct, we would not expect them to compete.

In short, this proposal states that if a language has TAE morphology, the language must: (i) have resultative/prospective aspects; (ii) have alternate means for locating events in time (e.g. past/future tenses); and (iii) that those alternate strategies be morphosyntactically similar enough to the resultative/prospective aspects such that they can pragmatically

[3]This proposal could also offer a way to diagnose resultative and prospective aspects cross-linguistically. If a past-oriented morpheme also conveys indirect evidentiality, we could label it a resultative aspect. Conversely, if a future-oriented morpheme also conveys "specific" evidentiality, we could label it a prospective aspect. In both cases, we would expect that the language would have other tense/aspect morphemes that the aspectual morphemes compete with.

compete.

7.3 Future directions

This dissertation provides another step towards untangling the relationship between tense/aspect and evidentiality. However, there is a **lot** of work still to be done. I provide a wishlist of future points to address.

- Providing a fully formalized pragmatic account of how asserting prospective/resultative aspect conveys "specific"/"indirect" evidentiality, respectively. In Chapter 5, I proposed that choosing to locate oneself in an event pre- or post-state—given the option not to—results in these evidential readings. However, as I noted, the challenge of motivating a fully pragmatic account for this data stems from the fact that these evidential readings (unlike the direct/"non-specific" evidential readings) are not cancellable.

- Making the connection between event pre-states and speaker certainty more explicit. More specifically, explaining why it is that if a speaker locates themselves in the pre-state of an event, they are more certain regarding the future occurrence of that event. This is something I assumed intuitively in Chapter 5, but I did not provide any formal account of this connection.

- Possible route: Incorporating a formalized causal component into the relationship between events and their pre- and post-states. This would be akin to what Hara et al. (2018) (following Davis and Hara 2014) propose for the semantics of the Japanese indirect evidential *youda*.

- Another possible route: Linking the interpretation of the TAE suffixes (in matrix clauses) to de se attitudes invoked by aspects, as proposed by Bittner (2005) to account for prospective aspectual data in Kalaallisut. (This is in line with Korotkova 2016's proposal to treat evidentials as de se attitude reports.)

- Extending the analysis to account for the interpretation of the TAE suffixes in questions and in combination with clausemate negation.

- Motivating why it is that the future- and past-oriented TAE suffixes differ with respect to their scope-taking abilities in embedded CPs. Do some languages pattern oppositely (i.e., are their past-oriented TAE morphemes required to stay in situ)?

- Providing better descriptive coverage of the interpretation and use of the TAE suffixes in embedded environments. As I noted in Chapter 2, collecting embedding data in Tatar can be quite challenging. Determining the felicity of a multiclausal expression in Tatar involves calculating the contribution of the matrix verb, the contribution of the embedded clause type, and the contribution of the embedded TAE suffix. I would like to expand the descriptive coverage of clausal embeddings to include other matrix verbs, such that evidential shift is more natural with respect to the embedded past-oriented TAE suffixes.

- Describing the ability of the Tatar TAE suffixes to co-occur with modals (e.g. *bälki* 'might,' *tiyeş* 'must') and the free evidential particles (e.g. *ikän*, *imeş*).

- Providing a more detailed typology of TAE morphology in Turkic languages. A preliminary typology of temporal suffixes in $\sim$25 Turkic languages suggests that evidential readings are available only if the language has more than one past- or future-oriented suffix. This observation is in line with my proposal that evidential readings of temporal morphemes are driven by pragmatic competition.

I look forward to the future work of improving our understanding of evidentiality and how it relates to tense and aspect.

Bibliography

Abduldaev, E. and Zakharova, O. (1987). *Grammatika Kigirzskogo Literaturnogo Yazyka*. Frunze.

Abusch, D. (1988). Sequence of tense, intensionality and scope. In Borer, H., editor, *Proceedings of the Seventh West Coast Conference on Formal Linguistics*.

Abusch, D. (1991). The present under past as de re interpretation. In Bates, D., editor, *Proceedings of the 10th West Coast Conference on Formal Linguistics*, pages 1–12. CSLI Publications.

Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and Philosophy*, 20(1):1–50.

Aikhenvald, A. (2004). *Evidentiality*. Oxford University Press.

Aikhenvald, A., editor (2018). *The Oxford Handbook of Evidentiality*. Oxford Handbooks in Linguistics. Oxford University Press.

Aikhenvald, A. and Dixon, R., editors (2003). *Studies in Evidentiality*. John Benjamins Publishing Company.

Altshuler, D. (2010). Meaning of 'now' and other temporal location adverbs. In Aloni, M., Bastiaanse, H., de Jager, T., and Schulz, K., editors, *Amsterdam Colloquium 2009, LNAI 6042*, pages 183–192. Springer Verlag.

Altshuler, D. (2016). *Events, States, and Times: An essay on narrative discourse in English*. De Gruyter Open Ltd., Warsaw/Berlin.

Altshuler, D. and Schwarzschild, R. (2013). Moment of change, cessation implicatures and simultaneous readings. In Chemla, E., Homer, V., and Winterstein, G., editors, *Proceedings of Sinn und Bedeutung 17*, pages 45–62.

AnderBois, S. (2014). On the exceptional status of reportative evidentials. In Snider, T., D'Antonio, S., and Weigand, M., editors, *Proceedings of Semantics and Linguistic Theory 24*, pages 234–254.

AnderBois, S., Brasoveanu, A., and Henderson, R. (2010). Crossing the appositive/at-issue meaning boundary. In Li, N. and Lutz, D., editors, *Proceedings of Semantics and Linguistic Theory 20*, pages 328–346.

Arregui, A. and Kusumoto, K. (1998). Tense in temporal adjunct clauses. In Strolovitch, D. and Lawson, A., editors, *Proceedings of Semantics and Linguistic Theory 8*, pages 1–18.

Bashir, E. (2006). Evidentiality in South Asian languages. In Butt, M. and King, T., editors, *Proceedings of the LFG06 Conference, Workshop on South Asian Languages*. CSLI Publications.

Berta, Á. (1998a). Tatar and Bashkir. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 283–300. Routledge.

Berta, Á. (1998b). West Kipchak Languages. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 301–317. Routledge.

Bittner, M. (2005). Future discourse in a tenseless language. *Journal of Semantics*, 22:339–387.

Boeschoten, H. (1998). Uzbek. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 357–378. Routledge.

Bohnemeyer, J. (2014). Aspect vs. relative tense: the case reopened. *Natural Language and Linguistic Theory*, 32:917–954.

Borsley, R. and Kornfilt, J. (2000). Mixed extended projections. *Syntax and Semantics*, 32:101–131.

Bowler, M. and Özkan, S. (2017). A perfect mess: towards a typology of the "present perfect". Handout, UCLA.

Brandt, P. (2005). On the common format of eventives and *too*-comparatives. In Maier, E., Bary, C., and Huitink, J., editors, *Proceedings of Sinn und Bedeutung 9*, pages 44–56.

Bybee, J. and Dahl, Östen. (1989). The creation of tense and aspect systems in the languages of the world. *Studies in Language*, 13(1):51–103.

Carpenter, K. (2014). Aspectual future in Karuk. Manuscript, University of California, Berkeley.

Chemla, E. and Spector, B. (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics*, 28:359–400.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Belletti, A., editor, *Structures and Beyond: The Cartography of Syntactic Structures*, volume 3, pages 39–103. Oxford University Press.

Chomsky, N. (1970). Deep structure, surface structure, and semantic interpretation. In Jakobson, R. and Kawamoto, S., editors, *Studies in Oriental and General Linguistics*, pages 52–91. TEC Corporation for Language and Educational Research.

Chung, K. (2007). Spatial deictic tense and evidentials in Korean. *Natural Language Semantics*, 15(3):187–219.

Coghill, E. (2010). The grammaticalization of prospective aspect in a group of Neo-Aramaic dialects. *Diachronica*, 27(3):359–410.

Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.

Comrie, B. (1997). Tatar (Volga Tatar, Kazan Tatar) phonology. In Kaye, A., editor, *Phonologies of Asia and Africa*, pages 899–911. Eisenbrauns.

Comrie, B. and Polinsky, M. (2007). Evidentials in Tsez. In Guentchéva, Z. and Landaburu, J., editors, *L'Enonciation Mediatisee II: Le traitement epistemologique de l'information: illustrations amerindiennes et caucasiennes*, pages 335–350. Éditions Peeters.

Condoravdi, C. (2002). Temporal interpretation of modals: modals for the present and for the past. In Beaver, D., Kaufmann, S., Clark, B., and Casillas, L., editors, *Stanford Papers on Semantics*. CSLI Publications.

Copley, B. (2008). The plan's the thing: deconstructing futurate meanings. *Linguistic Inquiry*, 39(2):261–274.

Copley, B. (2009). *The Semantics of the Future*. Outstanding Dissertations in Linguistics. Routledge, New York.

Şahan, F. (2002). *Nominal Clauses in Kazan Tatar*. PhD thesis, University of Wisconsin-Madison.

Csató, E. and Karakoç, B. (1998). Noghay. In Johanson, L. and Csató, E., editors, *The Turkic Languages*. Routledge.

Şener, N. (2011). *Semantics and Pragmatics of Evidentials in Turkish*. PhD thesis, University of Connecticut.

Cusihuaman, A. (1976). *Gramática Quechua: Cuzco-Collao*. Ministerio de Educación/Instituto de Estudios Peruanos, Lima.

Dahl, Östen. (1985). *Tense and Aspect Systems*. Oxford University Press.

Dahl, Östen., editor (2000). *Tense and Aspect in the Languages of Europe*. Empirical Approaches to Language Typology. Mouton de Gruyter.

Davis, C. and Hara, Y. (2014). Evidentiality as a causal relation: A case study from Japanese *youda*. In Piñón, C., editor, *Empirical Issues in Syntax and Semantics*, volume 10, pages 179–196.

Davis, C., Potts, C., and Speas, M. (2007). The pragmatic values of evidential sentences. In Friedman, T. and Gibson, M., editors, *Proceedings of Semantics and Linguistic Theory 17*, pages 71–88.

Davliyeva, A. (2011). An Investigation of Kazan Tatar Morphology. Master's thesis, San Diego State University.

de Haan, F. (1999). Evidentiality and epistemic modality: setting boundaries. *Southwest Journal of Linguistics*, 18:83–101.

Dwyer, A. (2000). Direct and indirect experience in Salar. In Johanson, L. and Utas, B., editors, *Evidentials: Turkic, Iranian, and Neighbouring Languages*. Mouton de Gruyter.

Erdal, M. (2004). *A Grammar of Old Turkic*. Brill, Leiden.

Faller, M. (2002). *Semantics and Pragmatics of Evidentials in Quzco Quechua*. PhD thesis, Stanford University.

Faller, M. (2004). The deictic core of 'non-experienced past' in Cuzco Quechua. *Journal of Semantics*, 21:45–85.

Faller, M. (2011). A possible worlds semantics for Cuzco Quechua evidentials. In Li, N. and Lutz, D., editors, *Proceedings of Semantics and Linguistic Theory 20*, pages 660–683.

Faller, M. (2012). Evidential scalar implicatures. *Linguistics and Philosophy*, 35:285–312.

Fleck, D. (2007). Evidentiality and double tense in Matses. *Language*, 83(3):589–614.

Fortescue, M. (2003). Evidentiality in West Greenlandic: A case of scattered coding. In Aikhenvald, A. and Dixon, R., editors, *Studies in Evidentiality*, pages 291–306. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Friedman, V. (1986). Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In Nichols, J. and Chafe, W., editors, *Evidentiality: The Linguistic Coding of Epistemology*, pages 186–187. ABLEX.

Gluckman, J. and Bowler, M. (2016). Expletive agreement, evidentiality, and modality in Logooli. In *Proceedings of Semantics and Linguistic Theory 26*, pages 1063–1082.

Göksel, A. and Kerslake, C. (2005). *Turkish: A Comprehensive Grammar*. Routledge.

Greed, T. (2009). Evidentiality in Tatar. Master's thesis, University of Helsinki.

Greed, T. (2014). The expression of knowledge in Tatar. In Aikhenvald, A. and Dixon, R., editors, *The Grammar of Knowledge: A Cross-Linguistic Typology*, Explorations in Linguistic Typology, chapter 3, pages 69–88. Oxford University Press.

Grice, P. (1975). Logic and conversation. In Cole, P. and Morgan, J., editors, *Syntax and Semantics*, volume 3: Speech Acts, pages 41–58. Elsevier.

Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press.

Groenendijk, J. and Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100.

Hammarström, H., Forkel, R., and Haspelmath, M. (2018). Glottolog 3.3. Max Planck Institute for the Science of Human History.

Hara, Y., Orita, N., and Sakai, H. (2018). Evidentials in Causal Premise Semantics: Theoretical and experimental investigation. In *JSAI International Symposium on Artificial Intelligence*, pages 282–298. Springer.

Harrison, K. D. (2000). *Topics in the Phonology and Morphology of Tuvan*. PhD thesis, Yale University.

Heim, I. (1994). Comments on Abusch's theory of tense. Manuscript, Massachusetts Institute of Technology.

Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell Publishers.

Heine, B. (1994). On the genesis of aspect in African languages: The proximative. In *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special Session on Historical Issues in African Linguistics*, pages 35–46.

Hirschberg, J. (1985). *A theory of scalar implicature*. PhD thesis, University of Pennsylvania.

Horn, L. (1972). *On the Semantic Properties of Logical Operators in English*. PhD thesis, University of California, Los Angeles.

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

Izvorski, R. (1997). The present perfect as an epistemic modal. In *Proceedings of Semantics and Linguistic Theory 7*.

Jarad, N. (2014). The grammaticalization of the motion verb 'rah' as a prospective aspect marker in Syrian Arabic. *Al-Arabiyya: Journal of the American Association of Teachers of Arabic*, 47:101–118.

Johanson, L. (1998). The structure of Turkic. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 30–66. Routledge.

Johanson, L. (2000). Turkic indirectives. In Johanson, L. and Utas, B., editors, *Evidentials: Turkic, Iranian, and Neighbouring Languages*. Mouton de Gruyter, Berlin.

Johanson, L. (2003). Evidentiality in Turkic. In Aikhenvald, A. and Dixon, R., editors, *Studies in Evidentiality*, pages 273–291. John Benjamins, Amsterdam.

Johanson, L. and Csató, Éva., editors (1998). *The Turkic Languages*. Routledge.

Kalsang, Garfield, J., and Speas, M. (2013). Direct evidentials, case, tense and aspect in Tibetan: evidence for a general theory of the semantics of evidential. *Natural Language and Linguistic Theory*, 31:517–561.

Kaplan, D. (1989). Demonstratives. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes from Kaplan*, pages 481–563. Oxford University Press.

Karakoç, B. (2017). Prospectives, proximatives and avertives in Noghay. In Korn, A. and Nevskaya, I., editors, *Prospective and Proximative in Turkic, Iranian and Beyond*, pages 57–76. Reichert Verlag, Wiesbaden.

Karttunen, L. (1973). Presuppositions of compound sentences. *Linguistic Inquiry*, 4(2):169–193.

Katz, G. (2003). A modal account of the English present perfect puzzle. In Young, R. and Zhou, Y., editors, *Proceedings of Semantics and Linguistic Theory 13*, pages 145–161.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy*, 30:669–690.

Keshet, E. and Schwarz, F. (2014). De re / de dicto. Manuscript.

Key, G. and Schreiner, S. (2014). The prospective marker in Turkish: A unified treatment. Handout, Workshop on Altaic Formal Linguistics (WAFL) 10. May 2-4, 2014. MIT.

Kirchner, M. (1998). Kirghiz. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 344–356. Routledge.

Klecha, P. (2014). Diagnosing modality in predictive expressions. *Journal of Semantics*, 31:443–455.

Klein, W. (1992). The present perfect puzzle. *Language*, 68:525–552.

Klein, W. (1994). *Time in Language*. Routledge, London.

Klose, C. (2014). A temporal evidential in Aymara. In *Proceedings of ConSOLE XXII*, pages 114–131.

Koev, T. (2011). Evidentiality and temporal distance learning. In *Proceedings of Semantics and Linguistic Theory 21*, pages 115–134.

Koev, T. (2017). Evidentiality, learning events and spatiotemporal distance: the view from Bulgarian. *Journal of Semantics*, 34:1–41.

Korn, A. and Nevskaya, I., editors (2017). *Prospective and Proximative in Turkic, Iranian and Beyond*. Reichert Verlag, Wiesbaden.

Kornfilt, J. (1997). *Turkish*. Routledge.

Kornfilt, J. and Whitman, J. (2011). Afterword: Nominalizations in syntactic theory. *Lingua*, 121:1297–1313.

Korotkova, N. (2016). *Heterogeneity and uniformity in the evidential domain*. PhD thesis, University of California, Los Angeles.

Kratzer, A. (1998). More structural analogies between pronouns and tenses. In *Proceedings of Semantics and Linguistic Theory 8*. MIT.

Kratzer, A. (2007). Situations in natural language semantics. In Zalta, E., editor, *The Stanford Encyclopedia of Philosophy (Winter 2017 Edition)*.

Krueger, J. (1977). *Tuvan Manual*, volume 126 of *Uralic and Altaic Series*. Indiana University.

Kubota, Y., Lee, J., Smirnova, A., and Tonhauser, J. (2009). On the cross-linguistic interpretation of embedded tenses. In Riester, A. and Solstad, T., editors, *Proceedings of Sinn und Bedeutung 13*, pages 307–320.

Lee, J. (2011). The Korean evidential *-te*: a modal analysis. In Bonami, O. and Hofherr, P., editors, *Empirical Issues in Syntax and Semantics*, volume 8, pages 287–311. Colloque de Syntaxe et Sémantique à Paris.

Lee, J. (2013). Temporal constraints on the meaning of evidentiality. *Natural Language Semantics*, 21:1–41.

Lindstedt, J. (2000). The perfect—aspectual, temporal and evidential. In Dahl, Östen., editor, *Tense and Aspect in the Languages of Europe*, Empirical Approaches to Language Typology, pages 365–383. Mouton de Gruyter.

Littell, P., Matthewson, L., and Peterson, T. (2010). On the semantics of conjectural questions. In Peterson, T. and Sauerland, U., editors, *Evidence from Evidentials*, volume 28 of *University of British Columbia Working Papers in Linguistics*, pages 89–104.

Maslova, E. (2003). *Tundra Yukaghir*, volume 372 of *Languages of the World*. LINCOM Europa, Munich.

Matthewson, L. (2011). On apparently non-modal evidentials. In Bonami, O. and Hofherr, P., editors, *Empirical Issues in Syntax and Semantics*, volume 8, pages 333–357. Colloque de Syntaxe et Sémantique à Paris.

Matthewson, L. (2012a). Evidence about evidentials: where fieldwork meets theory. In Stolterfoht, B. and Featherston, S., editors, *Empirical Approaches to Linguistic Theory: Studies in Meaning and Structure*, pages 85–114. Mouton de Gruyter.

Matthewson, L. (2012b). On the (non-)future orientation of modals. In Guevara, A. A., Chernilovskaya, A., and Nouwen, R., editors, *Proceedings of Sinn und Bedeutung 16*, pages 431–446.

Matthewson, L., Davis, H., and Rullmann, H. (2007). Evidentials as epistemic modals: Evidence from St'át'imcets. *Linguistic Variation Yearbook*, 7:201–254.

McCawley, J. (1981). Notes on the English perfect. *Australian Journal of Linguistics*, 1:81–90.

McCoard, R. (1978). *The English Perfect: Tense-choice and Pragmatic Inferences*. Elsevier, Amsterdam/New York.

McCready, E. and Ogata, N. (2007). Evidentiality, modality and probability. *Linguistics and Philosophy*, 30(2):147–206.

Menz, A. (2000). Indirectivity in Gagauz. In Johanson, L. and Utas, B., editors, *Evidentials: Turkic, Iranian, and Neighbouring Languages*, volume 24 of *Empirical Approaches to Language Typology*, pages 103–114. Mouton de Gruyter.

Meriçli, B. (2016). Modeling Indirect Evidence. Master's thesis, University of California, Santa Cruz.

Moens, M. and Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28.

Murray, S. (2010). *Evidentiality and Structure of Speech Acts*. PhD thesis, Rutgers, the State University of New Jersey.

Murray, S. (2017). *The Semantics of Evidentials*, volume 9 of *Oxford Studies in Semantics and Pragmatics*. Oxford University Press.

Mysaev, K. (1964). *Grammatika Karaimskogo Yazyka*. Akademiya Nauk SSSR, Institut Yazykoznaniya.

Nedjalkov, V., editor (1988). *Typology of Resultative Constructions*, volume 12 of *Typological Studies in Language*. John Benjamins Publishing Company.

Nedjalkov, V. and Jaxontov, S. (1988). The typology of resultative constructions. In Nedjalkov, V., editor, *Typology of Resultative Constructions*, volume 12 of *Typological Studies in Language*, pages 3–62. John Benjamins Publishing Company.

Nevskaya, I. (2005). The typology of the prospective in Turkic languages. *Sprachtypologie und Universalienforschung*, 58(1):111–123.

Nevskaya, I. (2018). Introduction. In Nevskaya, I. and Korn, A., editors, *Prospective and Proximative in Turkic, Iranian and Beyond*, pages 13–27. Reichert Verlag, Wiesbaden.

Nikolaeva, I. (1999). The semantics of Northern Ostyak evidentials. *Journal de la Société Finno-Ougrienne*, 88:131–159.

Noonan, M. (1985). Complementation. In Shopen, T., editor, *Language Typology and Syntactic Description, Volume II: Complex Constructions*, pages 42–140. Cambridge University Press, Cambridge.

Öztopçu, K. (2000). *Elementary Azerbaijani*. Türk Dilleri Araşırmaları Dizisi.

Özyıldız, D. (2016). Factivity defeated. University of Massachusetts, Amherst, handout.

Partee, B. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy*, 70(18):601–609.

Passonneau, R. (1987). Situations and intervals. In *Proceedings of the 25th Annual Meeting on Association for Computational Linguistics*, pages 16–24. Association for Computational Linguistics.

Plungian, V. (1988). Resultative and apparent evidential in Dogon. In Nedjalkov, V., editor, *Typology of Resultative Constructions*, volume 12 of *Typological Studies in Language*, pages 481–493. John Benjamins Publishing Company.

Pokrovskaya, L. (1964). *Grammatika Gagauzskogo Yazyka: Fonetika i Morfologiya*. Akademiya Nauk SSSR, Institut Yazykoznaniya.

Poppe, N. (1961). *Tatar Manual*. Research and Studies in Uralic and Altaic Languages. Indiana University Publications.

Poppe, N. (1964). *Bashkir Manual*. Indiana University Publications.

Prior, A. (1967). *Past, Present, and Future*. Oxford University Press.

Pulte, W. (1985). The experienced and nonexperienced past in Cherokee. *International Journal of American Linguistics*, 51:543–544.

Rassadin, V. (1978). *Morfologia Tofalarskogo Yazyka v Sravnitel'nom Osveshchenii*. Akademiya Nauk SSSR, Sibirskoe Otdelenie.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. Macmillan & Co., New York.

Rett, J. and Murray, S. (2013). A semantic account of mirative evidentials. In *Proceedings of Semantics and Linguistic Theory 23*, pages 453–472.

Ritchie, G. (1979). Temporal clauses in English. *Theoretical Linguistics*, 6:87–115.

Rivero, M. L. (2014). Spanish inferential and mirative futures and conditionals: An evidential gradable modal proposal. *Lingua*, (151):197–215.

Rivero, M. L. and Sheppard, M. M. (2015). The Slovenian future auxiliary *biti* as a tenseless gradable evidential modal: Inferential and Concessive readings.

Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12:683–721.

Russell, B. (1905). On denoting. *Mind*, 14(56):479–493.

San Roque, L., Floyd, S., and Norcliffe, E. (2015). Evidentiality and interrogativity. *Lingua*.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27:367–391.

Sauerland, U. and Schenner, M. (2007). Embedded evidentials in Bulgarian. In Puig-Waldmüller, E., editor, *Proceedings of Sinn und Bedeutung 11*, pages 525–539.

Schönig, C. (1998). Turkmen. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 261–272. Routledge.

Schönig, C. (1999). The internal division of modern Turkic and its historical implications. *Acta Orientalia Academiae Scientiarum Hungaricae*, 52(1):63–95.

Schreiner, S. and Stone, M. (2014). Aspect outside the stem: Prospective morphology in Cherokee. Handout, SSILA Winter Meeting.

Sebba, M. (2006). Ideology and alphabets in the former USSR. *Language Problems & Language Planning*, 30(2):99–125.

Seegmiller, S. (1996). *Karachay*. Lincom Europa.

Sharvit, Y. (2018). Sequence of tense: syntax, semantics, pragmatics. In Patel-Grosz, P., Grosz, P., and Zobel, S., editors, *Pronouns in Embedded Contexts at the Syntax-Semantics Interface*, volume 99 of *Studies in Linguistics and Philosophy*, pages 215–247. Springer.

Simeonova, V. and Zareikar, G. (2015). The syntax of evidentials in Azeri, Bulgarian, and Persian.

Simons, G. and Fennig, C., editors (2018). *Ethnologue: Languages of the World, Twenty-first edition*. SIL International, Dallas.

Slobin, D. and Aksu, A. (1982). Tense, aspect and modality in the use of the Turkish evidential. In Hopper, P., editor, *Tense-Aspect: Between Semantics and Pragmatics*. John Benjamins Publishing Company.

Smirnova, A. (2011). The meaning of the Bulgarian evidential and why it cannot express inferences about the future. In *Proceedings of Semantics and Linguistic Theory 21*, pages 275–294.

Smirnova, A. (2013). Evidentiality in Bulgarian: temporality, epistemic modality, and information source. *Journal of Semantics*, 30:479–532.

Stachowski, M. and Menz, A. (1998). Yakut. In Johanson, L. and Csató, Éva., editors, *The Turkic Languages*, pages 417–433. Routledge.

Straughn, C. (2011). *Evidentiality in Uzbek and Kazakh*. PhD thesis, University of Chicago.

Strawson, P. (1950). On referring. *Mind*, 59(235):320–344.

Tatevosov, S. (2007). Evidentiality and mirativity in the Mishar dialect of Tatar. In Guentchéva, Z. and Landaburu, J., editors, *L'Enonciation Mediatisee II: Le traitement epistemologique de l'information: illustrations amerindiennes et caucasiennes*, pages 407–433. Éditions Peeters.

Tatevosov, S. (2017). Budet, budet i budet: o semantike vremeni v misharskom dialekte tatarskogo yazyka. *Uralo-altajskie issledovaniya*, pages 77–92.

Thomas, G. (2012). *Temporal Implicatures*. PhD thesis, Massachusetts Institute of Technology.

Vander Klok, J. (2014). Questionnaire on modality for cross-linguistic use. http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/cross-linguistic-use.php.

Vendler, Z. (1957). Verbs and times. *Philosophical Review*, 66(2):143–160.

von Fintel, K. (1994). *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts, Amherst.

von Fintel, K. and Gillies, A. (2010). *Must*... stay... strong! *Natural Language Semantics*, 18(4):351–383.

Webber, B. (1988). Tense as discourse anaphor. *Computational Linguistics*, 14(2):61–73.

Wertheim, S. (2002). Language "purity" and the de-russification of Tatar. Berkeley Program in Soviet and Post-Soviet Studies, Working Paper Series.

Willett, T. (1988). A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language*, 12(1):51–97.

Winans, L. (2016). *Inferences of will*. PhD thesis, University of California, Los Angeles.

Wolvengrey, A. (2006). Prospective aspect in the Western dialects of Cree. *International Journal of American Linguistics*, 72(3):397–407.

Woodbury, A. (1986). Interactions of tense and evidentiality: a study of Sherpa and English. In Chafe, W. and Nichols, J., editors, *The Linguistic Coding of Epistemology*, pages 188–202. ABLEX.

Yalcin, S. (2007). Epistemic modals. *Mind*, 116(464):983–1026.