

Challenges in Kazakh auxiliary selection

David Gyorfi 

University of Surrey

Proceedings of the 27th International Conference on
Head-Driven Phrase Structure Grammar

Online (Berlin/Seattle)


Stefan Müller, Anke Holler (Editors)

2020

Stanford, CA: CSLI Publications

pages 68–87

Keywords: HPSG, Online type construction, Kazakh, auxiliary verbs

Gyorfi, David. 2020. Challenges in Kazakh auxiliary selection. In Stefan Müller & Anke Holler (eds.), *Proceedings of the 27th International Conference on Head-Driven Phrase Structure Grammar, Online (Berlin/Seattle)*, 68–87. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2020.4. 

Abstract

This paper accounts for the four auxiliaries in Kazakh that express the imperfective aspect. The main factors – the auxiliary, the main verb, their inflections and the aspectual specifications reveal a complicated system, which can be captured with an appropriate monotonic, multiple inheritance type hierarchy using online-type construction with the implementation of Pāṇinian competition. This analysis sheds light to a very different auxiliary system that we find in Indo-European languages.

1 Factors in interaction – empirical observations

Kazakh¹ is a Turkic language from the Kypchak branch, which like its relatives, has a rich system of auxiliary verbs. There are over 25 auxiliary verbs in Kazakh, each with a number of characteristics in common. They are all morphologically and inflectionally identical with a corresponding lexical verb and they all can be used in finite clauses, acting as lexical verbs on their own. When they participate in an auxiliary verb construction (henceforth AVC, shown in example 1), they combine with a lexical verb. The lexical verb must be in one of two converbial forms (either CVB.A or CVB.B), and the auxiliary is inflected for person, number and tense, or can be of nonfinite categories, such as coordination, relativization or attributivization. The lexical semantics of the entire AVC is determined entirely by the lexical verb. The auxiliary verb contributes aspect or mood-like meanings to the AVC.

Since a number of auxiliaries will be mentioned, and most of them will be very similar in their semantic contribution, I will gloss them with their original lexical meaning for easy identification. When deemed necessary, I will also gloss the overall aspectual meaning of the AVC, following the auxiliary verb. In the example in (1), the auxiliary’s original lexical meaning is ‘lie’ or ‘lie down’, it is marked for the aorist tense and the overall aspectual meaning of the AVC is progressive.

- (1) Toyžan düken-ge bar-a žatır
Toyžan store-DAT go-CVB.A AUX(lie).AOR.3SG (PROG)
‘Toyžan is going to the store (right now).’

This paper focusses on the class of auxiliary verbs that express the imperfective aspect. These are *žat*, *otır*, *tur* and *žür* (their lexical meanings are ‘lie, sit,

¹ I am grateful for the helpful comments and guidance to Berthold Cysmann, Stefan Müller and the audience of the 2020 HPSG Conference. I have received further help from Oliver Bond and Greville G. Corbett at my home department, and of course, the examples in this paper could not have been valid without the devoted help of my informants, mainly Toyžan Turýanbayeva and Aytoša Abdigali.

stand and walk’). The other c. 21 auxiliary verbs are structurally similar, but express the perfective with the addition of modal meanings, and due to the limits of this paper, their analysis has to remain for a future study. Example (2) is just to exemplify the semantic contribution of one of the auxiliaries that will not be examined in this work.

- (2) (Muhamedowa 2015:119)
- | | | | |
|----------|----------|-----------|---------------------------------|
| kitap-tī | sömke-ge | sal-a | sal-dī-m |
| book-ACC | bag-DAT | put-CVB.A | AUX(put)-PST-1SG (INCIDENTALLY) |
- ‘I put the book incidentally into the bag.’

The aim is to provide a model of the combinations of auxiliary and lexical verbs, their inflections, and the aspectual specification and the distribution of the entire AVC. The following table illustrates the building blocks that are relevant factors in an AVC. The following sections will introduce the data that Section 1.6 aims to model. The table in (3) is a summary of the components that will be relevant.

(3)

<i>lexical verb</i>	<i>CVB</i>	<i>auxiliary verb</i>	<i>inflection of the auxiliary verb</i>	<i>resulting aspect</i>	<i>meaning</i>
ayt ‘say’	ip CVB.B	žat AUX(lie)	ir AOR	PROG	‘s/he is speaking’
qal ‘stay’	a CVB.A	tur AUX(stand)	uw-in-a NMLZ-3SG-DAT	IPFV	‘[so that s/he] stays [there]’

1.1 Aspect

Following Comrie (1976: 3), aspect specifies ‘different ways of viewing the internal temporal constituency of a situation’. In Kazakh, aspect is expressed in analytic constructions, and in AVCs as well. This paper focusses on three aspect values of the imperfective type, that is, ‘unbounded and internally homogeneous’ events (Langacker 2008: 147).

- *Progressive*: ‘a process ongoing at contextual occasion’ (Timberlake 2007: 304)
- *Habitual*: ‘some regular, repeated activity or event’ (Carlson 2012: 829)

- *Incremental*: similar to the progressive, but the process consists of defined units, some of which have been completed, and some have not yet, at a certain reference point.²

The aspectual specifications are determined based on descriptions (Muhamedowa 2015; Somfai Kara 2002), as well as adverb compatibility and contextual tests performed during elicitation sessions, some of which will be explicit in the examples.

1.2 The auxiliary lexeme

The four imperfective auxiliaries we focus on, *žat*, *otir*, *tur* and *žür* are similar in many respects, and sometimes they are interchangeable.

- (4) (Muhamedowa 2015: 132)
- | | | | |
|----|-----------------------------|------------|------------------------------------|
| a. | šegirtke | sekir-ip | žatir/tur/žür |
| | dragonfly | jump-CVB.B | AUX(lie/stand/walk).AOR.3SG (PROG) |
| | ‘The dragonfly is jumping.’ | | |
| b. | tamaq | že-p | žatir/otir |
| | food | eat-CVB.B | AUX(lie/sit).AOR.3 (PROG) |
| | ‘S/he is eating.’ | | |

The above examples are potential examples of overabundance (Thornton 2011, 2012), but admittedly, there could be factors I am not aware of at the moment. According to Muhamedowa (2015) and to my fieldwork, there is no elicitable difference in syntax, semantics and style. It must be made explicit that as long as there is no evidence for any kind of selectional factor, I will treat examples like (4) as overabundance with the narrower definition that the informants accept all versions without any comment, and they claim they would probably use all of them interchangeably.

1.3 The lexical verb lexeme

The lexical verb’s idiosyncratic peculiarities have a key role in determining what structural combinations are possible and what the overall semantics will be. It is clear that aktionsart and other lexeme-internal factors are in play, including semantic class memberships (e.g. motion verbs, internal change verbs etc.), but accounting for these peculiarities has to remain for future research. This paper takes into account a particularly clear phenomenon – a split in the lexicon, whereby the lexemes *come* and *go* group together against

² Incremental is not usually considered a type of aspect. This typology is data-driven, and the term is based on the discussion in Croft (2012, p. 41).

all other verbs. One of the four imperfective auxiliaries, *žat* ‘lie’, may only combine with the lexical verbs *come* and *go* when they are in the CVB.A form, while all other lexical verb lexemes must be marked for CVB.B. This distribution results in the progressive aspect, and any other combination will result in either unacceptable structures, or a different aspect.

As shown in example (5), the described combination results in the progressive aspect, while if the lexical verb is marked for CVB.B (6), the resulting aspect is the incremental. The informants, when asked in what sentence they could imagine the combination (*come/go*-CVB.B + AUX(*žat*)), said that it had to be a process that can be broken down to units, and some of the units have completed an action, while others are still in progress. Let us refer to this as the incremental aspect (Croft 2012; Dowty 1991).

- (5) qonaq-tar kel-e žatir
 guest-PL come-CVB.A AUX.(lie).AOR.3 (PROG)
 ‘The guests are coming (and none of them has arrived).’

- (6) qonaq-tar kel-ip žatir
 guest-PL come-CVB.B AUX.(lie).AOR.3 (INCR)
 ‘The guests are coming (and some of them have arrived).’

From the other three imperfective auxiliaries *žür* ‘walk’ freely combines with the lexical verbs *come* and *go*. The auxiliaries *otir* ‘sit’ and *tur* ‘stand’ can also combine with *come* and *go*, however, this combination seems to be slightly more restricted. One example is shown in (7).

- (7) Toyžan düken-ge bar-ïp tur-a-dï.
 Toyžan store-DAT go-CVB.B AUX(stand)-NPST-3SG (HAB)
 ‘Toyžan usually goes to the store.’

1.4 Inflection of the auxiliary verb

The inflection of the auxiliary verb has an impact on the AVC’s distribution, semantics and acceptability. In an AVC the auxiliary bears all the inflectional morphology that a finite verb would in the same distribution, while the lexical verb is marked for one of the converb forms. That is, an AVC’s distribution is in general equal to the distribution of main verbs. They both can appear in a range of finite and nonfinite forms: they can be relativized, attributivized, as well as they can appear in co- and subordinations, or in the conditional. As expected, the AVC, when attributivized, changes the aspectual specification. This is illustrated in , where the predicate is attributivized.

- (8) a. mektep-te oqi-yan bala
 school-LOC study-PTCP child
 ‘a child, who studied in school’
- b. mektep-te oq-ıp žat-qan bala
 school-LOC study-CVB.B AUX(lie)-PTCP (PROG) child
 ‘a child, who is studying in school [at the moment]’

Considering the four imperfective auxiliaries this paper focuses on, there is one striking deviation. While *žat* selects its lexical verb and converb according to the split in the lexicon (see Section 1.3 above), the other three auxiliaries select the converb with respect to the finiteness of the AVC. The auxiliaries *otir*, *tur* and *žür* can only combine with CVB.A in a nonfinite position³.

(9)

		Auxiliary verb	
		Finite usage	Nonfinite usage
Lexical verb	CVB.A	otir, tur, žür	✓
	CVB.B		✓
	CVB.A	žat	✓
	CVB.B		✓

The reason the generalization is worded using *finiteness* and not the type of converb, is that this applies to other nonfinite affixes as well. In such nonfinite positions, the auxiliary is frequently marked with CVB.B, since apart from marking the lexical verb in many AVCs, CVB.B also acts as a coordinator between VPs. This is shown in (10)a, while in (10)b another nonfinite form is shown, that is a nominalized, case marked complement of a VP.

- (10) a. ...batırlar-dı es-ke al-a otir-ıp ...
 ...heroes-ACC mind-DAT take-CVB.A AUX(sit)-CVB.B
 ‘Keeping the heroes in mind, [the soldiers marched forward].’

³ There are counterexamples, but for every thousand occurrences of LexV-CVB.A + AUX-CVB.B, there are only 2-3 occurrences of LexV-CVB.A + AUX-FINITE (Kilgariff et al. 2004). This argument is meant to be taken as a statistical fact.

- b. el-de qal-a tur-uw-ï-na
 country-LOC stay-CVB.A AUX(stand)-NMLZ-3SG-DAT
- ruqsat ber-di
 permission give-PST.3SG
- ‘S/he gave permission to him/her to stay in the country.’

1.5 Inflectional class and aspect

In this section I will be looking at two tenses, the nonpast and the aorist. I will argue that the four imperfective auxiliaries *žat*, *otïr*, *tur* and *žür* ‘lie, sit, stand, walk’, and their lexical verb counterparts constitute an inflectional class since they exhibit idiosyncratic morphological properties. This lines up with their idiosyncratic TAM specification, which is a strong argument for treating them as an inflectional class, and for assuming that the lexemes used in auxiliary and main verb constructions are one and the same.

Firstly, the four imperfective auxiliaries’ morphology is unique in the aorist, since the third singular form is equal to the lexeme’s stem, which is due to a process of haplology (*tur-ur* → *tur*, *otïr-ïr* → *otïr*, *žür-ür* → *žür*) whereby the aorist exponent’s similarity to the coda of the verb stem caused one of the identical syllables to drop (Johanson 2004). From the four imperfective auxiliaries *žat* has gone through a different path, as instead of having lost its coda, it is fossilized in the old aorist form (*žat* → *žatïr*; the modern Kazakh, productive aorist is *žatar*). Eventually, *žat*’s inflectional morphology is identical to the other three auxiliaries. One peculiarity of this special inflection is that the third singular of the haplogized aorist coincides with the verb’s stem (except *žat*), which in other verbs is reserved for the second singular imperative. The partial paradigms of the aorist forms in the haplogized and in the regular inflectional class are shown in (11).

(11)

	<i>hapl. aorist</i>	<i>reg. aorist</i>
<i>stem</i>	otïr ‘sit’	bar ‘go’
<i>1SG</i>	otïr-mïn	bar-ar-mïn
<i>2SG</i>	otïr-sïŋ	bar-ar-sïŋ
<i>3SG</i>	otïr	bar-ar

Now we turn to the semantic properties of the tense marked verbs. The nonpast tense, marked with the suffix $-A(y)^4$, can convey a number of present and future temporal specifications. The present progressive is usually not among these, as it is expressed in AVCs.

⁴ Capital letters indicate segments subject to consonant or vowel harmony.

The aorist, marked with the suffix $-(A)r$ (see the table above), has a number of meanings as well, such as general truth (as in “ice melts when the weather is warm”), unsure future (maybe something will happen) or near future. The above description holds for all lexical and auxiliary verbs, apart from the four imperfective auxiliaries – both in an AVC and as a main verb in their original meaning. Uniquely, when marked for the nonpast, they specify the habitual aspect (12), and when marked for the aorist, they express the present progressive (13).

- (12) a. Toyžan negizi aldıñyï qatar-da otır-a-dï
 Toyžan usually front row-LOC sit-NPST-3SG (HAB)
 ‘Toyžan usually sits in the front row.’
- b. *Toyžan negizi aldıñyï qatar-da otır
 Toyžan usually front row-LOC sit.AOR.3SG (*HAB)
 intended: ‘Toyžan usually sits in the front row.’
- (13) a. Toyžan qazir orındıq-ta otır
 Toyžan now chair-LOC sit.AOR.3SG (PROG)
 ‘Toyžan is sitting on a chair right now.’
- b. *Toyžan qazir orındıq-ta otır-a-dï
 Toyžan now chair-LOC sit-NPST-3SG (*PROG)
 intended: ‘Toyžan is sitting on a chair right now.’

This would allow for an analysis that is similar to Daniels and Corbett’s (2019), where a particular inflected form of a lexeme of a closed class (in that language the idiosyncrasy affects only one lexeme) shifts in semantics compared to all other lexemes inflected the same way. In Kazakh, four lexemes inflect for a particular tense, and the semantics changes in a systematic way compared to the rest of verbs. Compare the following examples where I assume that all verbs are in the aorist form (regardless of their haplogitized or regular inflectional class membership). In sentence a., the verbal lexeme is one of the four verbs from the haplogitized inflectional class, while in sentence b. and c. the verbal lexeme is in the regular inflectional class.

- (14) a. Toyžan qazir orındıq-ta otır
 Toyžan now chair-LOC sit.AOR.3SG (PROG)
 ‘Toyžan is sitting on a chair right now.’

- (Muhamedowa 2015: 199)
- b. erten žaᅇbir žaw-ar
tomorrow rain rain-AOR.3SG (FUT)
'It will rain tomorrow probably.'
- c. #qazir žaᅇbir žaw-ar
now rain rain-AOR.3SG (FUT/*PROG)
'It might be raining soon.' (intended: 'It is raining right now.')

This distinction is not only present in the lexical usage of these verbs, but also when they function as auxiliaries in AVCs. In the following examples, in sentences a. the auxiliary appears in the aorist form and conveys the present progressive TAM, while in sentences b. the auxiliary is inflected for the non-past tense, and it conveys the present habitual TAM.

- (15) a. qazir oyin oyna-p žür
now game play-CVB.B AUX(walk). AOR.3SG (PROG)
'The [child] is playing now.'
- b. *qazir oyin oyna-p žür-e-di
now game play-CVB.B AUX(walk)-NPST-3SG (*PROG)
intended: 'The [child] is playing now.'
- (16) a. *keyde ... de-p žatir
sometimes ... say-CVB.B AUX.(lie).AOR.3SG (*HAB)
intended: 'Sometimes [people] say ...'
- b. keyde ... de-p žat-a-di
sometimes ... say-CVB.B AUX(lie)- NPST-3SG (HAB)
'Sometimes [people] say ...'

1.6 Arguments that AVCs are periphrastic

AVCs might be treated as one-word (inflection, synthetic) or many-words (periphrasis, analytic) constructions. In this section I argue that Kazakh AVCs should be considered periphrases.

Firstly, let us look at semantic compositionality. As part of an AVC, the auxiliary's lexical meaning does not typically interact with the overall meaning of the AVC, although the speakers are aware of their being a separate semantic constituent. An informant told me at a session that her mother once criticized her because she used the lexical verb *eat* with the progressive auxiliary *žat* 'lie', since "it is impolite to eat while lying. You should use the auxiliary *otir* 'sit' instead." The two imperfective auxiliaries are indeed freely interchangeable in this case. However, it is clear that the auxiliary *žat* is the most productive one and it is completely acceptable to combine it with al-

most any lexical verb, even with ones that result in an “impolite” combination. *Žat* indeed combines with lexical verbs whose meaning is not quite compatible with a lying position, including *eating*, *running* or *standing*. In other words, the semantics of the individual components of an AVC do not add up to the entire semantics (Ackerman and Stump 2004; Spencer 2001; Vincent 2011)

Syntactically, Kazakh AVCs are rigid structures, which could be an argument for an inflectional analysis, however, as reported in (Muhamedowa 2015:129), one adverb can intervene. The bolded adverb in example (17)a generally follows the NP it modifies, but in the case of AVCs, it is positioned in between the lexical verb and the auxiliary, as in (17)b. The opposite of this phenomenon can be an argument for morphologization (Bonami & Samvelian, 2015: 354, also see Müller, 2010: 608–609), thus this is an argument for a syntactic treatment.

- (17) a. bir ret ***qana*** kör-di-m
 one time ***only*** see-PST-1SG
 ‘I saw it only once.’
- b. oyın oyna-p ***qana*** žür-mey
 game play-CVB.B ***only*** AUX(walk)-NEG.CVB
 ‘...[children] do not only play, [but also paint pictures and walk].’

Furthermore, as far as I can hear, the constituents of an AVC are pronounced as separate prosodic units and short pauses also occur when the speaker is hesitating. Regarding stress, Muhamedowa (2015: 124) notes that certain auxiliaries express different semantic distinctions when the lexical or the auxiliary verb is stressed.

Lastly, for phonology, auxiliaries can obey the ‘initial bilabial glide rule’. The phonemes /o/ and /ö/ are preceded by a bilabial glide in initial position, such that $_ \# /o/, /ö/ \rightarrow [{}^w o], [{}^w ö]$. In non-initial positions the glide does not appear. The imperfective auxiliary *otır* ‘sit’ many times obeys this rule, as my field recordings show (e.g. [${}^w oqip \ {}^w otır$] ‘S/he is reading’). If morphotactics treated AVCs as a single item, the glide would not be present (and the example above would be [${}^w oqip \ otır$]).

It is duly noted, however, that one auxiliary (*žat* ‘lie’) does fuse with the lexical verb (Muhamedowa 2015: 131), but this is only one auxiliary and the construction is considered to be of ‘low register’, and the fusion is never reflected in writing.

- (18) a. ol žumis iste-p žatır
 3SG work do-CVB.B AUX(lie).AOR.3SG (PROG)
 ‘S/he is working.’

- b. ol žumīs iste-vařır
 3SG work do-AUX(lie).AOR.3SG (PROG)
 ‘S/he is working.’

It must also be mentioned that other Turkic languages have gone further the grammaticalization path and ended up with suffixes that were once auxiliaries. The present progressive in Turkish, for example, is marked with a suffix that is a cognate to *žūr* ‘walk’, an imperfective auxiliary in Kazakh and many other Turkic languages. In Turkish, however, the converb has eroded into a linking vowel that is subject to vowel harmony (as all converbs are), however, the nucleus in the affix itself does not harmonize. The further grammaticalized Shor, a Siberian Turkic language, has auxiliary-origin suffixes whose nucleus is within the vowel harmony domain as well (Anderson 2004: 92–93).

It is likely that Kazakh auxiliaries could turn into suffixes in the future, but today, prosodic, syntactic and phonological evidence support that Kazakh auxiliaries are far from canonical affixes (Spencer and Luis 2013) and thus the balance turns into the analytic side, even if some traits point toward the synthetic direction.

Lastly, after having argued that Kazakh AVCs are analytic constructions, it should also be made clear that they are periphrases. Following Ackerman and Stump (2004, as well as Spencer and Popova, 2015; for a canonical analysis, see (Brown et al. 2012)), the following criteria apply:

1) The AVCs have a featurally intersective distribution, since they can freely combine with large part of the verbal lexicon, even considering the regular incompatibilities we know of.

2) The union of the morphosyntactic property sets of the elements building up an AVC is not equal to the morphosyntactic property sets of the entire construction. Consider the bleached semantics of the auxiliary, or the fact the CVB.B can appear in a progressive AVC – note that CVB.B, when not in an AVC, is a marker of coordination and it is affixed to verbs describing an event that is completed (example (18)a).

3) Lastly, the exponents of the morphosyntactic property set in an AVC are distributed. Consider any example where the converb, the auxiliary or the auxiliary’s inflection can change the aspectual specification of the AVC.

This section gave a summary of the data that will be analyzed in the next section. We have seen how the auxiliary and the lexical components of an AVC interact, how they can inflect, and what aspects emerge as a result. In the present section I argued for a periphrastic analysis.

2 Modelling in HPSG

2.1 Multiple inheritance with Paṇinian competition

The present analysis assumes the following statements, with respect to the above arguments:

- Kazakh AVCs are periphrases
- An auxiliary verb is the same lexeme as its main verb counterpart, and their differences are best described by derivational rules.
- Aspect is inherently carried by both the AVC, and the tense marker affixed to the AVC (if applicable, consider non-tensed examples)
- The aspect specification expressed by the tense markers should be considered a cross-cutting feature.
- There is overabundance (to the best of my knowledge)

The assumption that Kazakh AVCs are periphrases is implemented by following the feature geometry of periphrases in (Bonami, Borsley, and Tallerman 2016). In their approach, the auxiliary element's lexical identity is identical to the lexical verb's identity, and the auxiliary verb borrows its own phonological form from another lexical item. An alternative to this approach, that is not explicitly tailored for periphrases, would be where the lexical verb is specified in the COMPS list of the auxiliary (see e.g. Müller, 2010: 634).

The following analysis will use a monotonic, multiple inheritance type hierarchy of derivational rules combined with an online type construction approach (cf. Bonami & Crysmann, 2016; Crysmann & Bonami, 2016; or Malouf, 1998 but without the assumption of defaults).

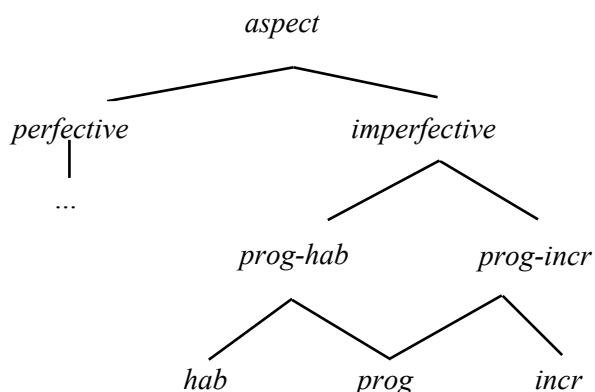
2.1.1 Assumed type hierarchy

I assume that both the auxiliary and the tense marker are specified for aspect. The auxiliary must be specified, since non-finite AVCs are specified for the same aspect values as finite AVCs, but do not bear tense markers; see example (8) above. The tense markers are individually specified for aspect (in the haplologized inflectional class), even when not in an AVC, but when affixed to a main verb, as shown in examples (12) and (13).

The tense marker can technically 'override' the AVC's aspect, as shown in Section 1.5. In order to keep this analysis free of defaults and any kind of overriding, the hierarchy recognizes that the nonpast tense marker can indicate both the *habitual* and the *progressive*, while the aorist tense marker can indicate the *progressive* and the *incremental*. This hierarchy is almost identical to the typology in (Comrie 1976), but this small additional complexity allows us to treat the lowest level of this hierarchy as part of a separate, orthog-

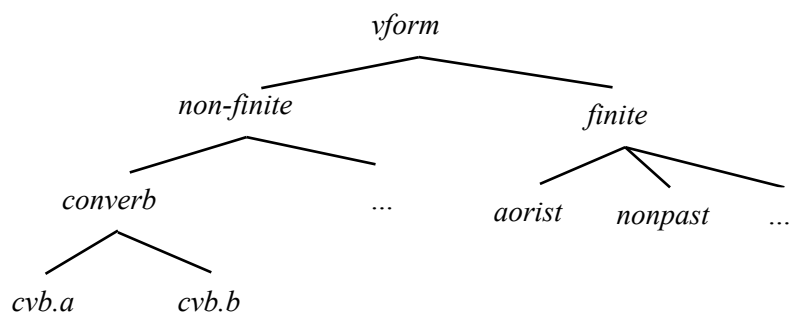
onal set of values that can only be accessed by one dimension, and not the other (more on the Online-type construction analysis in Section 2.1.3).

(19) Hierarchy of aspect values



In order to formalize generalizations on the structural characteristics, I will assume the hierarchy for VFORM values in (20). This step is important to explicitly model the data under question, but the same time it also represents a foresight for future research that will be looking at other auxiliaries that combine with lexical verbs of non-finite forms that are not converbs, or auxiliaries that are marked for tenses other than the aorist or the nonpast.

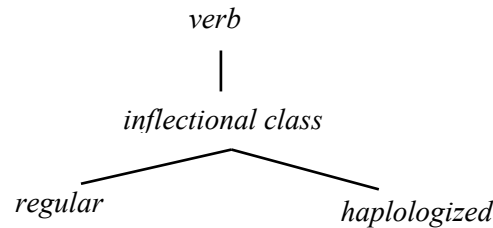
(20) Hierarchy of VFORM



And the last assumption formalizes the acknowledgement that there are at least two inflectional classes of Kazakh verbs, one of which (*haplologized*) characterizes the four imperfective auxiliaries. This formalization is similar to that of Koenig's *subregularities* (1999: 130–133). Koenig assumes a separate branch of the hierarchy that includes the inflectional properties of the exceptional classes. Since at the moment we only need to formalize the exis-

tence of a set of lexemes that constitute an inflectional class, I will not represent the peculiarities of the actual inflection.

(21) Hierarchy of inflectional classes



2.1.2 Analysis

This Section presents an analysis that uses a monotonic (does not require defaults), multiple inheritance type hierarchy. In addition, I follow recent papers (Bonami, Borsley, and Tallerman 2016; Crysmann and Bonami 2016) that argue for the implementation of Pāṇinian competition into HPSG.

At the top of the partial hierarchy in (22) is the type *derivational lexical rules*. The following type is *aux* which represents the common properties of all auxiliary verbs in Kazakh, and gives rise to the periphrastic construction. It is specified that the semantic content value of the auxiliary is identified with that of the lexical verb’s (in COMPS), but there is an added ASP(ect) feature as well. The INFL|LID feature, that is the lexical identity of the auxiliary, is underspecified, since it needs to allow the 25 auxiliaries to be identified with their main verb counterpart. The details of this selection need to be examined in future research. Lastly, since all AVCs require the main verb to take some sort of a converb form, this is specified in this description as well.

The subtype *aux-ipfv* is a restriction that is true for the four imperfective auxiliaries detailed in this paper. They are selected on the basis of their inflectional class membership (*hapl*), that contains only these four verbs. The ASP value is further specified to *ipfv*, which is a subtype of *aspect*. The sister of this type is just a reminder that there are perfective AVCs as well.

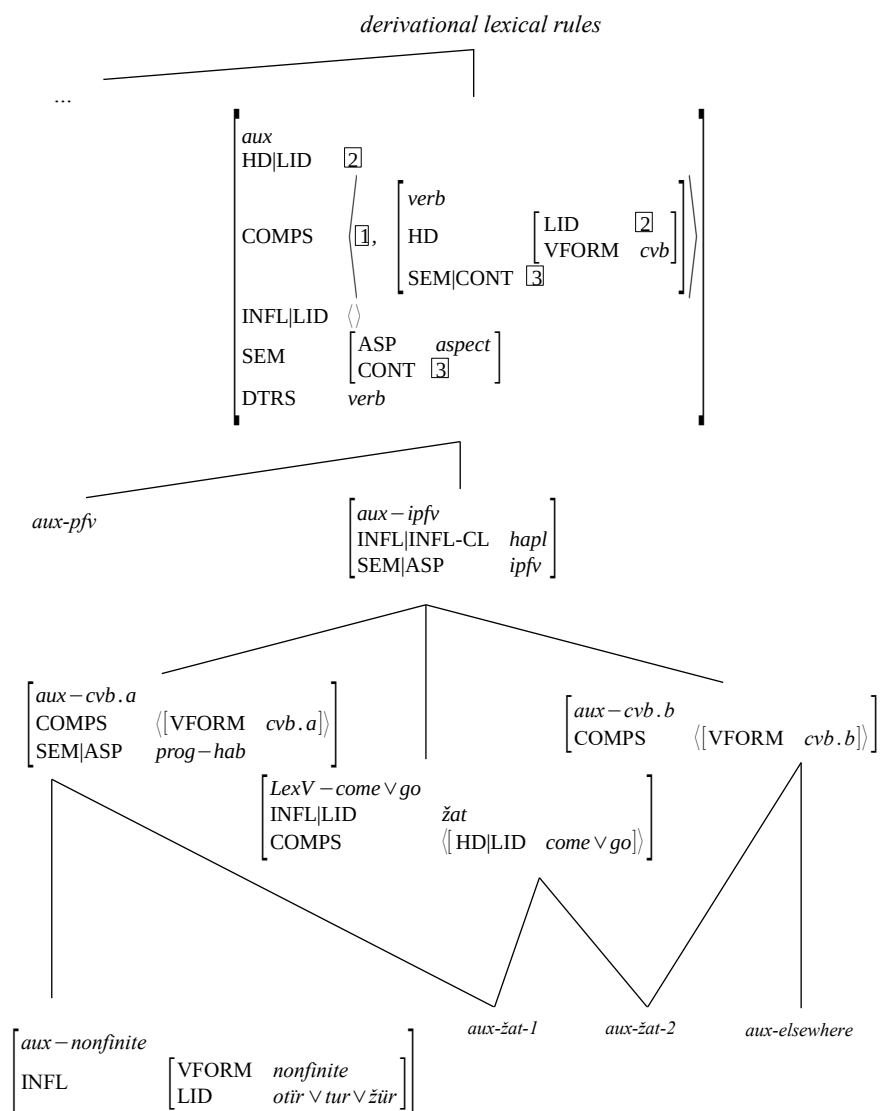
The following seven types describe how the converb forms, the auxiliary and the lexical verb identities are allowed to combine.

Aux-nonfinite describes the AVCs that could only appear in a nonfinite usage (recall Section 1.4). The types *aux-žat-1* and *aux-žat-2* account for the distribution of the auxiliary *žat* and the converb forms, as well as the lexical verbs *come* and *go*. *Aux-žat-1* licenses examples like the one in (1), while *aux-žat-2* licenses the example in (6). The type *aux-elsewhere* is a necessary point in order to implement Pāṇinian competition by licensing less specified constructions that are acceptable (Crysmann and Bonami 2016: 363–64). This type licenses all four auxiliaries to select CVB.B, unless one of the de-

scribed types is more specific. I believe the data I know of fits into this hierarchy.

Notice that the leaf nodes are specified for aspect. This is crucial to reflect the fact that AVCs on their own express aspect on the one hand, and on the other hand at this stage of my knowledge of this language, there are multiple constructions that express the same aspect (recall Section 1.2). At this point, all AVCs are specified for [ASP *imperfective*].

(22) Dimension 1: Auxiliary



2.1.3 The tense dimension - Online-type construction

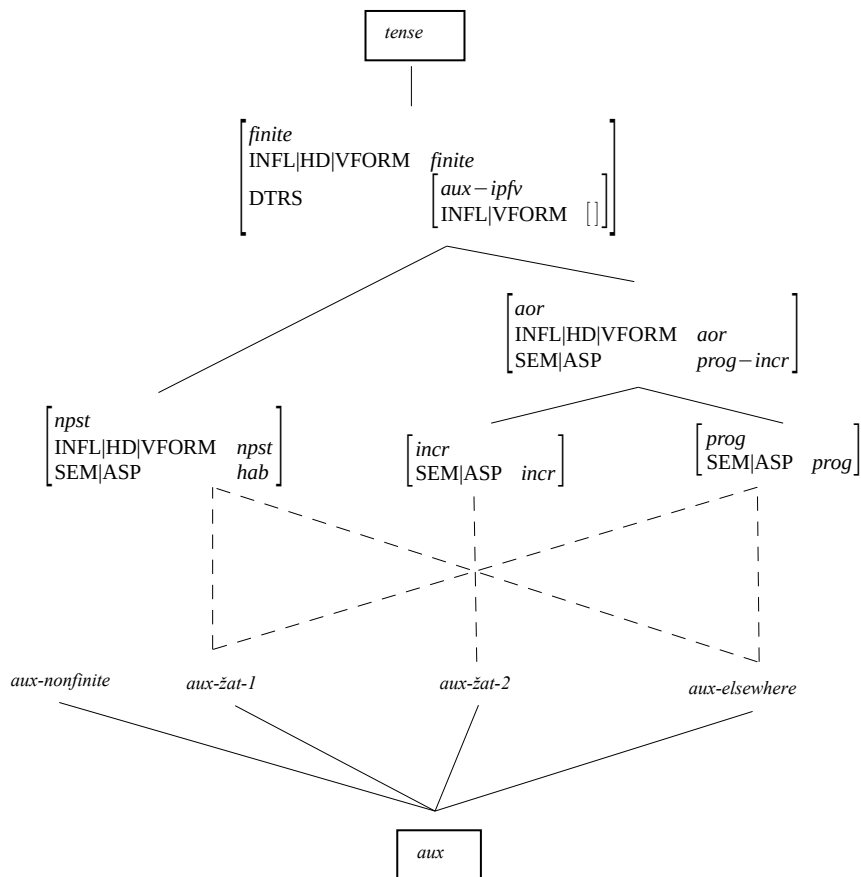
When the auxiliary is affixed by the aorist or the nonpast tense markers, the aspect value is ‘overridden’. In the present analysis, this phenomenon will be analyzed as instances of more specific aspect values. This alternation is formalized as an online type construction (Ackerman and Bonami 2017; Bonami and Crysmann 2016; Koenig 1999; Koenig and Jurafsky 1994). That is, leaf nodes of either dimension may be underspecified, and require to unify with the specifications of the other dimension’s leaf nodes.

The figure in (23) describes the tense dimension, with the addition of the of the leaf nodes of the auxiliary dimension in order to show the alternating properties (dashed lines). The mother node describes a rule that takes an item of type *aux-ipfv*, with the specification that its VFORM is not specified – this blocks the type *aux-nonfinite* to combine with it. While the *nonpast* tense is only associated with the *habitual* aspect, the *aorist* tense can both bear the *progressive* and the *incremental* aspects. This is duly represented by the intermediate node *aor*.

The leaf nodes in the tense dimension represent the relation between the two tenses, and the three most specific aspect values, shown in (19). The *progressive* and the *habitual* alternate with *aux-žat-1* and *aux-elsewhere*. *Aux-žat-2* can only combine with the *incremental* aspect. Lastly, *aux-nonfinite* cannot combine with any tense-marked type, since by definition, it can only take affixes marking nonfinite clauses.

This analysis might be unusual, considering that the two dimensions interacting involve the same feature, namely, aspect. However, it respects the requirements of online type construction to the extent that the alternating values are never part of both dimensions, they are thus, orthogonal. The tense dimension only specifies the *progressive*, the *incremental* and the *habitual*, while the auxiliary dimension can only access the less specific types of aspect.

(23) Dimension 2: Tense



3 Conclusions

This paper has proposed an analysis of the four imperfective auxiliary verbs in Kazakh. Section 1 presented the data and focused on the different behaviors of the four auxiliaries, a split in the lexicon that required the lexical verbs *come* and *go* to behave idiosyncratically, and showed how aspect emerges from these properties. The section concluded that AVCs should be treated as periphrases.

Section 2 proposed an analysis where aspect is carried both by the AVC and its tense marking separately, but they interact. This interaction is represented in an online-type construction approach, while the interacting nodes are part of monotonic, multiple inheritance type hierarchy. This description implements Pāṇinian competition, and does not rely on defaults or junk fea-

tures (Kathol 1994). It also reflects the overabundance that might be the best term for the optionality we see. Future research will aim to resolve this overabundance and to account for the rest of the AVCs in Kazakh.

4 References

- Ackerman, Farrell and Olivier Bonami. 2017. "Systemic Polyfunctionality and Morphology-Syntax Interdependencies." Pp. 233–68 in *Defaults in Morphological Theory*, edited by A. Hippisley and N. Gisborne. Oxford University Press.
- Ackerman, Farrell and Gregory Stump. 2004. "Paradigms and Periphrastic Expression: A Study in Realization-Based Lexicalism." Pp. 111–57 in *Projecting Morphology*, edited by L. Sadler and A. Spencer. Stanford, CA: CSLI Publication.
- Anderson, Gregory D. S. 2004. *Auxiliary Verb Constructions in Altai-Sayan Turkic*. Wiesbaden: Harrassowitz.
- Bonami, Olivier, Robert D. Borsley, and Maggie Tallerman. 2016. "On Pseudo-Non-Finite Clauses in Welsh." Pp. 104–124 in *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*. CSLI Publications.
- Bonami, Olivier and Berthold Crysmann. 2016. "Morphology in Constraint-Based Lexical Approaches to Grammar." Pp. 609–56 in *The Cambridge Handbook of Morphology*, edited by A. Hippisley and G. Stump. Cambridge University Press.
- Bonami, Olivier and Pollet Samvelian. 2015. "The Diversity of Inflectional Periphrasis in Persian." *Journal of Linguistics* 51(02):327–82.
- Brown, Dunstan, Marina Chumakina, Greville G. Corbett, Gergana Popova, and Andrew Spencer. 2012. "Defining 'Periphrasis': Key Notions." *Morphology* 22(2):233–75.
- Carlson, Greg. 2012. "Habitual and Generic Aspect." Pp. 828–51 in *The Oxford Handbook of Tense and Aspect*, edited by R. I. Binnick. New York: Oxford University Press.
- Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.
- Croft, William. 2012. *Verbs: Aspect and Causal Structure*. Oxford: Oxford University Press.
- Crysmann, Berthold and Olivier Bonami. 2016. "Variable Morphotactics in Information-Based Morphology." *Journal of Linguistics* 52:311–374.
- Daniels, Don and Greville G. Corbett. 2019. "Repartitioning." *Language* 95(4):711–50.
- Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection." *Language* 67:547–619.
- Johanson, Lars. 2004. "On Turkic Transformativizers and Nontransformativizers." *Turkic Languages* 8:180–90.

- Kathol, Andreas. 1994. "Passives without Lexical Rules." Pp. 237–272 in *Nerbonne et. al. (1994)*.
- Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž, and David Tugwell. 2004. "Itri-04-08 the Sketch Engine ([Http://Www.Sketchengine.Eu](http://www.sketchengine.eu))." in *Information Technology*.
- Koenig, Jean-Pierre. 1999. *Lexical Relations*. Stanford: CSLI Publications.
- Koenig, Jean-Pierre and Daniel Jurafsky. 1994. "Type Underspecification and Online Type Construction in the Lexicon." Pp. 270–285 in *West Coast Conference on Formal Linguistics*. Vol. 13. Stanford: CSLI Publications/SLA.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Malouf, Robert. 1998. "Categories, Prototypes, and Default Inheritance." Pp. 207–216 in *Proceedings of the Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorical Grammar*, edited by G. Bouma, G.-J. Kruijf, and R. Oehrle.
- Muhamedowa, Raihan. 2015. *Kazakh: A Comprehensive Grammar*. London, New York: Routledge.
- Müller, Stefan. 2010. "Persian Complex Predicates and the Limits of Inheritance-Based Analyses." *Journal of Linguistics* 46(3):601–55.
- Somfai Kara, Dávid. 2002. *Kazak (Languages of the World/Materials 417)*. München: Lincom Europa.
- Spencer, Andrew. 2001. "The Paradigm-based Model of Morphosyntax." *Transactions of the Philological Society* 99:279–313.
- Spencer, Andrew and Ana Luís. 2013. "The Canonical Clitic." Pp. 123–51 in *Canonical Morphology and Syntax*, edited by D. Brown, M. Chumakina, and G. G. Corbett. Oxford: Oxford University Press.
- Spencer, Andrew and Gergana Popova. 2015. "Periphrasis and Inflection." Pp. 197–230 in *The Oxford Handbook of Inflection*, edited by M. Baerman. Oxford: Oxford University Press.
- Thornton, Anna M. 2011. "Overabundance (Multiple Forms Realizing the Same Cell): A Noncanonical Phenomenon in Italian Verb Morphology." Pp. 358–81 in *Morphological autonomy: Perspectives from Romance inflectional morphology*, edited by M. Maiden and M.-O. Hinzen. Oxford: Oxford University Press.
- Thornton, Anna M. 2012. "Reduction and Maintenance of Overabundance. A Case Study on Italian Verb Paradigms." *Word Structure* 5:183–207.
- Timberlake, Alan. 2007. "Aspect, Tense, Mood." Pp. 280–333 in *Language typology and syntactic description. Vol. 3. Grammatical categories and the lexicon*, edited by T. Shopen. Cambridge: Cambridge University Press.
- Vincent, Nigel. 2011. "Non-finite Forms, Periphrases, and Autonomous Morphology in Latin and Romance." Pp. 417–35 in *Morphological Auto-*

nomy: Perspectives From Romance Inflectional Morphology, edited by M. Maiden, J. C. Smith, M. Goldbach, and M.-O. Hinzelin. Oxford: Oxford University Press.