

## Challenges in Kazakh Auxiliary selection

Dávid Győrfi (PhD researcher)

Surrey Morphology Group, University of Surrey

The literature has been very interested in auxiliary verb constructions' (AVC) morphosyntax (Bonami and Webelhuth 2012; Bonami and Samvelian 2015), the selection of the auxiliaries BE and HAVE in European languages (McFadden 2007; Kailuweit and Rosemeyer 2015; Ackema and Sorace 2017), and the general description of AVCs including challenges for formal frameworks, such as the NICE(R) properties in English (see a great summary in Sag et al., 2020). The data I present in this paper extends the observations of above-mentioned papers with a single system where a number of selection types are observed, as detailed below.

The construction in question: Kazakh (Kipchak, Turkic) tense and aspect expressions are realized in a proliferation of AVCs. The c. 25 auxiliary verbs combine with lexical verbs in one of two converbial forms (CVB\_A or CVB\_B). This means that there are c. 50 distinct, potential AVCs, most of which are actually attested. The exact rules and distribution of which lexical verb and AVC may combine, and what precisely determines the semantics of a given construction, is still an open question (Somfai Kara 2002; Muhamedowa 2015; Johanson 2004). Every auxiliary is used both as a lexical verb, and as a semantically bleached auxiliary in AVCs, as shown in (1).

- (1) a. temeki-ni                      **tasta-dï-m**  
cigarette-ACC                      **throw**-PST-1SG  
'I threw away the cigarette [I quit smoking].'
- b. kitap-tï              oqï-p              **tasta-dï-m**  
book-ACC              read-CVB\_B      **AUX(PFV 'throw')**-PST-1SG  
'I have read the entire book.'

This paper focusses on the particular behavior of three auxiliary verbs, all of which represent the imperfective aspect. The aim is to formalize the following facts:

1. Three imperfective auxiliaries can convey the progressive aspect.
2. Auxiliaries govern the main verb's form in the following ways:
  - a. some govern only one converbial form
  - b. some govern both converbial forms, resulting in the same semantics
  - c. some govern both converbial forms, resulting in different semantics
  - d. the three imperfectives show a particular split in the lexicon
3. All auxiliary verbs have a lexical meaning and can be used as main verbs.

Firstly, in order to account for the synonymy of three progressive auxiliaries and for cases of optional converb selection, I review the notion of 'defeasible features' (Sag et al, 2003) and starting from their formalization, argue that the type of optionality attested in Kazakh should be formalized with the addition of the logic operator *exclusive or* (or *xor*)  $\vee$ , which allows us to restrict the valence values to

be either option 1 or option 2 from a set of more than 2 choices. This analysis does not require any new machinery or assumption, since the xor operator is understood as a logical term, such as the concatenation operator.

Some auxiliaries show a particular case of converb selection, since the lexicon appears to be split into a compartment (lexemes *go* and *come*), that takes one converb, while the other compartment (all verbs apart from *go* and *come*) take the other converb, when combining with auxiliaries that result in the progressive meaning. I propose a valence specification that separately restricts the appropriate main verb's LID values to cooccur with the appropriate converb, and this is achieved with the xor operator listing the acceptable combinations of main verb and converb form, allowing both to be specified simultaneously. The compartments are defined with a list of LID values and in doing so I utilize the negation operator, following Crysmann (2006) in HPSG and Dalrymple et al. (2019) in LFG. The constraints in (2), thus, correctly predicts a progressive auxiliary, while allows any main verb to obey the distributional rules.

$$(2) \left[ \begin{array}{l} \text{PHON} \\ \text{SYN} \\ \text{SEM | RESTR | RELN} \end{array} \right] \left[ \begin{array}{l} \text{žatır} \\ \text{HD} \\ \text{VAL} \end{array} \right] \left[ \begin{array}{l} \left[ \begin{array}{l} \text{verb} \\ \text{LID} \quad \text{aux} - \text{prog} \\ \text{VFORM} \quad \text{fin} \\ \text{AUX} \quad + \end{array} \right] \\ \left\langle \left[ \begin{array}{l} \text{HD} \\ \text{VFORM} \quad \text{cvb}(a) \\ \text{LID} \quad \{go, come\} \end{array} \right] \vee \left[ \begin{array}{l} \text{verb} \\ \text{VFORM} \quad \text{cvb}(b) \\ \text{LID} \quad \{\neg go, come\} \end{array} \right] \right\rangle \end{array} \right] \left[ \begin{array}{l} \text{prog} \end{array} \right]$$

Three progressive auxiliaries appear to be equal in semantics, style, frequency and all functional factors apart from the selected auxiliary. I propose to treat these AVCs as optional choices. The optional distribution can be formalized by listing the auxiliaries' LID values and specify the valence restrictions the same way as we did when analyzing the split converb selection above – but now we have to diverge from a lexicalist view and assume that there is a construction that treats the overabundant auxiliaries and their valence specifications.

$$(3) \left[ \begin{array}{l} \text{MTR} \\ \text{HD - DTR} \\ \text{DTRS} \end{array} \right] \left[ \begin{array}{l} \text{SYN} \\ \text{SEM | RESTR | RELN} \end{array} \right] \left[ \begin{array}{l} \left[ \begin{array}{l} \text{HD} \quad [2] \\ \text{VAL} \quad \langle [1] \rangle \end{array} \right] \\ \text{prog} \oplus L \\ \left[ \begin{array}{l} \text{SYN} \\ \text{SEM | RESTR | RELN} \end{array} \right] \left[ \begin{array}{l} \text{HD} \quad [2] \\ \text{VAL} \quad \langle [1], [3] \rangle \end{array} \right] \left[ \begin{array}{l} \text{verb} \\ \text{AUX} \quad + \\ \text{LID} \quad \langle \text{žatır} \vee \text{otır} \vee \text{tur} \rangle \end{array} \right] \\ \left\langle \left[ \begin{array}{l} \text{HD} \quad [3] \\ \text{VAL} \quad \langle [1] \rangle \end{array} \right] \left[ \begin{array}{l} \text{verb} \\ \text{AUX} \quad - \\ \text{VFORM} \quad \text{cvb}(a) \\ \text{LID} \quad \{go, come\} \end{array} \right] \vee \left[ \begin{array}{l} \text{verb} \\ \text{AUX} \quad - \\ \text{VFORM} \quad \text{cvb}(b) \\ \text{LID} \quad \{\neg go, come\} \end{array} \right] \right\rangle \end{array} \right]$$

This treatment allows the progressive auxiliaries to be described with highly underspecified lexical entries, and it obviates the need for a complicated type hierarchy that assumes raising verbs and several types of auxiliaries. The SEM values of the lexical entries of the auxiliaries will specify for a defeasible value (their original, lexical meanings) and thus allow the construction presented above to override the lexical meaning and specify for aspect instead. A strictly lexicalist analysis would require 6 lexical entries for the 3 auxiliaries (one for the lexical use and one for the auxiliary use), which is less economical than the present proposal. Being a highly lexicalist approach, the HPSG+PFM analyses (e.g. Bonami and Webelhuth 2012) of periphrases could give similar results, but they would still require a larger number of lexical entries, and it would require the maintenance of the assumption that a HPSG+PFM can work side by side, while the analysis proposed represents a rather conservative view.

## References

- Ackema, Peter, and Antonella Sorace. 2017. "Auxiliary Selection." In *The Wiley Blackwell Companion to Syntax, Second Edition*, edited by Martin Everaert and Henk van Riemsdijk, 1–32. John Wiley & Sons, Inc.
- Bonami, Olivier, and Pollet Samvelian. 2015. "The Diversity of Inflectional Periphrasis in Persian." *Journal of Linguistics* 51 (02): 327–82.
- Bonami, Olivier, and Gert Webelhuth. 2012. "The Phrase-Structural Diversity of Periphrasis: A Lexicalist Account." In *Periphrasis: The Role of Syntax and Morphology in Paradigms*, edited by Marina Chumakina and Greville G. Corbett, 141–68. Oxford University Press/British Academy.
- Crysmann, Berthold. 2006. "Floating Affixes in Polish." In *The Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna*, edited by Stefan Müller, 123–39. Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John J. Lowe, and Louise Mycock. 2019. *The Oxford Reference Guide to Lexical Functional Grammar*. Oxford: Oxford University Press.
- Johanson, Lars. 2004. "On Turkic Transformativizers and Nontransformativizers." *Turkic Languages* 8: 180–90.
- Kailuweit, Rolf, and Malte Rosemeyer, eds. 2015. *Auxiliary Selection Revisited: Gradience and Gradualness (Vol. 44)*. Walter de Gruyter.
- McFadden, Thomas. 2007. "Auxiliary Selection." *Language and Linguistics Compass* 1 (6): 674–708.
- Muhamedowa, Raihan. 2015. *Kazakh: A Comprehensive Grammar*. London, New York: Routledge.
- Sag, Ivan A., Rui P. Chaves, Anne Abeillé, Bruno Estigarribia, Dan Flickinger, Paul Kay, Laura A. Michaelis, et al. 2020. "Lessons from the English Auxiliary System." *Journal of Linguistics* 56 (1): 87–155.
- Somfai Kara, Dávid. 2002. *Kazak (Languages of the World/Materials 417.)*. München: Lincom Europa.