Construction $za + infinitive$ – evidence from the Croatian corpora

Matea Filko, matea.filko@ffzg.hr
Krešimir Šojat, ksojat@ffzg.hr
Marko Tadić, marko.tadic@ffzg.hr
Outline of the talk

- prepositions in Croatian
  - preposition za
- infinitive in Croatian
- construction za + infinitive
  - evidence from the corpora
  - za + infinitive and syntactic structures
    - the replacement (im)possibilities
    - syntactic functions
- concluding remarks
Prepositions in Croatian

- function words which denote relations among entities
- prepositions complement case system (Raguž 1997)

- always followed by nouns or pronouns in specific cases > prepositional phrases
  - one preposition = one, two or three cases
    - only prepositions *u* ‘in’ and *za* ‘for’ can be followed by three cases
      - *u* + genitive, accusative, locative
      - *za* + genitive, accusative, instrumental

- according to Croatian grammars:
  - prepositions are never followed by a noun or a pronoun in nominative or vocative case
  - prepositions are never followed by verbal or adverbial or prepositional phrases
    - corpora analysis shows that the preposition *za* is a counterexample
Preposition za

- „genuine preposition“ (Barić et al. 2003), diachronically non-motivated word
- very diverse meanings
  - **temporal** (+G, + A, + I)
    - + G: za rata ‘during war’, za moje mladosti ‘during my youth’
    - + A: za Božić ‘at Christmas’; za pet minuta ‘in five minutes’
    - + I: dan za danom ‘day after day’
  - **locational**
    - location; usually at the back (+ A, + I): sakriti se za ormar ‘to hide behind the closet’, sjediti za stolom ‘to sit at the table’
    - directionality (+ A, + I): ići za Zagreb ‘to go to Zagreb’, dođite za mnom ‘follow me, lit. come after me’
Preposition za

- **purpose, attribute (+ A)**
  - meso za juhu ‘meat for the soup’; stroj za pranje ‘washing machine’
- **intentions, efforts (+ A)**
  - boriti se / žrtvovati se za slobodu ‘to fight / sacrifice for freedom’
- **indirect object (+ A)**
  - rekao sam mu za tebe ‘I told him about you’
- **price, measure, exchange (+ A)**
  - sve za 10 kuna ‘everything for 10 kunas’, duži za metar ‘one meter longer’, daj mi lutku za autić ‘give me the doll for the car’
- **juncture point (+ A)**
  - drži se za mene ‘hold on to me’
- **cause (+ A, + l)**
  - platit ćeš mi za to ‘you’ll pay me that’, poludjela je za njim ‘she went crazy after him’
- **aspirations (+ l)**
  - žudnja za životom ‘lust for life’
Preposition za

- can co-occur with prepositional phrases (Raguž 1997)
  - *Dao mu je to za u kuću.* ‘He gave him that for the inside of the house.’
- can co-occur with non-declensional parts of speech – verbs and adverbs (Pranjković 1987):
  - *za ovdje* ‘for here’
    - *kava za van* ‘coffee to go, lit. coffee for out’
  - *za danas* ‘for today’
  - *za deset* ‘for ten’
  - *za plakati* ‘to cry for, lit. for to cry’

- evidence from Croatian corpora: only the preposition *za* co-occurs with verbs in infinitive
- moreover, the infinitive is the only verbal form it co-occurs with – WHY?
Infinitive in Croatian

- indefinite verbal form (Babić et al. 2007)
- in the borderline area between nouns and verbs
  - infinitive denotes verbal actions without marking person, tense or mood
    - this is why it is close to nouns
  - infinitive marks aspect and voice – inherent verbal categories
    - this is why it is considered to be a verb
  - Proto-Indo-European (PIE): **no infinitives** (Matasović 2006):
    - verbal roots + different suffixes > verbal nouns
    - infinitive in IE languages originates from the generalization of one type of verbal nouns in PIE
    - infinitive in Slavic languages < nouns derived with suffix *-t-; dative singular
- both imperfective and perfective forms can co-occur with the preposition
  - *nešto za jesti* – *nešto za pojesti* ‘something to eat’
Construction *za* + *infinitive*

- Croatian prescriptive linguists claim that this construction is **ungrammatical** (cf. Rišner 2007, Pranjković 1987, Vukojević 2009)
  - Pranjković (1987) lists numerous linguists who claim that this construction is syntactically incorrect, starting from the beginning of 20th century
- the construction is not dealt with in Croatian grammars
- however, it has been used since the 12th century
  - first attested in 1198
  - older Croatian grammars recognize it as a special verbal form
    - e.g. Relković (1767) claims this construction is gerund
  - it is still in use, both in literary texts and oral speech
- linguistic experts claim that this construction should not be used
  - pure infinitive or gerund should be used instead
    - is it always possible?
Construction *za + infinitive*

- research questions:
  - Is there evidence of the contemporary usage of this construction in the Croatian corpora?
  - In which syntactic structures is this construction used?
  - Can the construction *za + infinitive* be replaced with some other syntactic elements without the change in meaning?
    - as prescribed by 'language advisors'
  - What are the syntactic functions of these structures?
  - Should this construction be considered as ungrammatical?
Evidence from the corpora - frequency

- two biggest corpora of the Croatian language:
  - HNK – Croatian National Corpus (Tadić 2009)
    - balanced corpus mainly consisting of edited and proofread texts
    - newspapers, literature, textbooks,...
  - hrWaC (Ljubešić & Erjavec 2011)
    - Croatian web corpus
    - texts crawled from the .hr domain

- frequency of the construction:
  - HNK – absolute frequency: 3,337; relative frequency: 15.4 per million
  - hrWaC – absolute frequency: 147,211; relative frequency: 105.3 per million
    - Croatian speakers tend to use za + infinitive more frequently in informal discourse
    - proofreading interventions in texts from HNK – cause of low frequency?
Evidence from the corpora – syntactic structures

- the data from the corpora indicates that this construction is still used in Croatian
- therefore: this construction should be more thoroughly described
- the analysis is based on:
  - 1,500 sentences containing the construction za + infinitive
  - two large Croatian corpora
- the analysis shows that za + infinitive can be used:
  - as a nominal complement
  - as a verbal complement
  - as a part of an idiom (as nominal or verbal complements)
Evidence from the corpora – syntactic structures

Nominal complement I

- **noun + za + infinitive**
  - teški *uvjeti* za *igrati* ‘difficult conditions to play’
  - zalihe za *preživjeti* godinu dana ‘supplies for a whole year’
  - *uvjeti* za *dobiti* lokacijsku dozvolu ‘conditions required to get a location permit’
  - tri *lekcije* za *naučiti* ‘three lessons to learn’

- **replacement possibilities:**
  - construction za + N<sub>ACC</sub>:
    - teški *uvjeti* za *igrati* > teški *uvjeti* za *igr* (igra = ‘play’); *teški uvjeti igrati*
    - *uvjeti* za *dobiti* lokacijsku dozvolu > *uvjeti* za *dobivanje* lokacijske dozvole > gerund, IMPERFECTIVE verb
  - attributive clause:
    - tri *lekcije* za *naučiti* > tri *lekcije* *koje se moraju naučiti*; *tri lekcije naučiti*
Evidence from the corpora – syntactic structures

Nominal complement II

- **pronoun + za + infinitive**
  - *imati što za prodati/reći* ‘to have something to sell/say’
  - *nemati što za izgubiti* ‘to have nothing to lose’
  - *nešto za raditi/jesti/pojesti* ‘something to do/eat’
  - *nemati ništa za prijaviti* ‘nothing to declare’

- **replacement possibilities:**
  - bare infinitive when it complements relative/indefinite pronouns
    - *imati što za reći* > *imati što reći*
    - it cannot be replaced by the construction za + \(N_{ACC}\)
  - construction za + \(N_{ACC}\) when the pronoun is followed by an adjective
    - *imati što toplo za popiti* ‘to have something warm to drink’ > *imati što toplo za piće* (piće = ‘drink’),
      *imati što toplo popiti*
Evidence from the corpora – syntactic structures
Nominal complement III

- **adjective + za + infinitive**
  - *idealna za utažiti žed* ‘ideal to quench your thirst’
  - *stvorena za osvježiti se* ‘made to refresh’

- **replacement possibilities:**
  - construction za + N_{ACC}
    - *idealna za utažiti žed > idealna za utaživanje žed* 
      - the gerund formed from the imperfective verb has to be used; *utaženje*
  - adverbial clause
    - *stvorena za osvježiti se > stvorena kako biste se osvježili* 
      - purpose, intention
Evidence from the corpora – syntactic structures
Verbal complement I

- za + *infinitive* as a *copula* complement
  - *bilo je za očekivati* ‘it could be expected, lit. it was to expect’
  - *nije za zamjeriti* 'not to blame smb.'
  - *vijesti su nam za povratiti* 'the news were disgusting'

- replacement possibilities:
  - different constructions with modal verbs
    - *bilo je za očekivati* > *moglo se očekivati* ‘it could be expected’
    - *nije za zamjeriti* > *ne može se zamjeriti* ‘it cannot be blamed’
  - bare infinitives or verbal nouns / gerunds cannot be used as a nominal part of the predicate
Evidence from the corpora – syntactic structures
Verbal complement II

- **za + infinitive** as a part of a complex predicates
  - *ostaje nam za pitati se* ‘we were left wondering, lit. it remains to us to wonder’
  - *realno je za očekivati* ‘it is plausible to expect’
  - *smiješno je za čuti* ‘it is funny to hear’
  - *bilo je ugodno za gledati* ‘it was pleasant to watch, lit. it was pleasant for watching’

- **replacement possibilities:**
  - bare infinitive:
    - *ostaje nam za pitati se* > *ostaje nam pitati se*, *ostaje nam za pitanje*
    - *realno je za očekivati* > *realno je očekivati*, *realno je za očekivanje*
  - construction za + N_{acc}:
    - *bilo je ugodno za gledati* > *bilo je ugodno za gledanje* (but also: *bilo je ugodno gledati*)
      - in this case the construction determines the meaning of the nominal part of the predicate
  - in some cases it cannot be replaced at all, e.g. *smiješno za čuti* ‘funny to hear’
Evidence from the corpora – syntactic structures

Idiomatic constructions

- **za + infinitive** as a part of idiomatic constructions
  - *to je za krepat od smijeha* ‘this is to laugh one's head off’
  - *to je Bogu za plakati / ovo je za plakati* ‘it’s a crying shame’
  - *ovo je za poludjeti* ‘this drives me crazy’
  - *vonja za popizditi* ‘it fucking smells’
  - *ekipa/nedjelja/dijete za poželjeti* ‘team/Sunday/child to wish for’
  - *ružno za vidjeti* ‘ugly to see’

- replacement possibilities:
  - *za + infinitive* cannot be replaced with other syntactic structures in these examples
Syntactic functions

- construction za + \textit{infinitive} can have different syntactic functions:
  - **Subject**
    - \textit{ostaje nam za pitati se} ‘it remains to wonder’
    - some authors (e.g. Smailagić) consider this construction as a predicate
      - \textit{to be / to remain} + za + \textit{infinitive} expresses modallity = a complex predicate
  - **Object**
    - \textit{koliko mentalne energije treba za napraviti bilo što} ‘how much mental energy is needed to do anything’
  - **Predicate**
    - \textit{to je za krepat od smijeha} ‘this is to laugh one's head off’
  - **Adverbial**
    - \textit{ali danas za preživjeti u umjetnosti, treba raditi na promidžbi} ‘nowadays, you have to promote yourself if you want to survive in art’
    - purpose, intention
  - **Attribute**
    - \textit{teški uvjeti za igrati} ‘difficult conditions to play’
Syntactic functions

- according to Croatian grammars, infinitives can function as subjects and predicates
  - they exclude other three possibilities: attribute, object, adverbial
  - however, our corpus driven research has shown that this is also possible

■ preposition za ”nominalizes” infinitives
  - by adding a case
  - by adding additional meaning components
  - za in NP phrases – all three above mentioned syntactic functions are possible (attribute, adverbial, indirect object)
Conclusion

- construction \textit{za + infinitive} is frequently found in contemporary Croatian corpora
  - regardless of the "instructions" of linguistic experts to avoid it
- it is very frequent in spoken language
- it occurs in specific syntactic environments and it can have all major syntactic functions
  - unlike the „bare infinitive”
- in some cases it can be replaced by other syntactic structures
  - the replacement possibilities are not straightforward
  - when it can be replaced, the meaning of constructions is usually altered
- in numerous cases it cannot be substituted by any other variants
  - especially in idiomatic structures
- this construction should not be regarded as ungrammatical:
  - the combination of the preposition and the infinitive expresses meanings that can hardly be expressed by other syntactic elements
  - it has been in use for more than 800 years and it survived
Thank you!