Outline

- Introduction
- Typological Pragmaticon
- Conclusion
Introduction
Multilingual lexical resources

- (Multi)WordNet
- (Multi)FrameNet
- BabelNet
- MetaNet
- DatSemShift

+ Constructicons!
Discourse formulae (Rakhilina et al. 2021)

- multi-word, (mostly) idiomatic units
- used in a dialogue as a reaction to the speech act of an interlocutor
- express refusal, acceptance, negation, confirmation, disbelief etc.

- I’m making waffles. Want some? – I’m good.
- May I put on some music? – Be my guest.
- We’ve made it to the finals! – No way!
Russian Pragmaticon (Yaskevitch et al. 2021)

A database of Russian DF designed as a tool for language learners

Available at pragmaticon.ruscorpora.ru

Contains over 600 Russian DF from a list, compiled by manual mark-up of drama and fiction, followed by automatic extraction (Gerasimenko et al. 2020)

Can be used as a starting point for typological comparison
Gathering typological data

- Starting with Slavic languages: supposedly more similar structures
- Finding semantic (pragmatic) counterparts
  - searching parallel corpora
  - comparing contextual distributions with questionnaires (Bychkova 2020, on English data Koziuk, Badryzlova 2021)
- Finding formal counterparts by working through the list with a native speaker
Cross-linguistic variation

- **Target pragmatic meanings**
  - What meanings are consistently expressed with DF in all the languages?
  - What meanings are unusual for DF?
  - Which pragmatic meanings are contiguous? (polysemy in DF)

- **Source meanings**
  - Common structures with common meanings
  - “Unique” structures
  - Common structures with different meanings
Typological Pragmaticon
Database structure

- language,
- inner structure,
- glosses,
- lemmas,
- pragmatic function,
- additional semantics,
- speech acts,
- dialog structure,
- intonation,
- syntax,
- source construction,
- source construction syntax
- source construction intonation
# Pragmatic functions

- **agreement**
  - Slovene: *sliši se dobro* 'sounds good'

- **assessment**
  - Slovene: *brez veze* 'not relevant'

- **confirmation**
  - Serbian: *tako-tako* 'this is so'

- **indifference**
  - Slovene: *briga me* 'not bother me'

- **negation**
  - Russian: *da ćto tam* 'pointless to discuss'

- **prohibition**
  - Russian: *i ne dumaj* 'don't even think'

- **refusal**
  - Serbian: *Bog zna* 'lit.: God knows'

- **surprise**
  - Serbian: *Bog s tobom*  'holy cow!'  
    lit.: 'God with you'
Additional semantics

- More nuanced semantic characteristics, such as negative or positive assessment, disbelief, doubt or confidence

- Can be combined

\[\text{a to ty ne znaeš}\]
Speech acts and dialog structure

- bipartite
  - <polar question>
  - a gospod' ego znaet ptcl God him knows [REFUSAL]

- tripartite
  - <question>
  - <request to specify the reason>
  - da tak ptcl so [REFUSAL]
Inner structure

Provides a loose description of the literal meaning of a formula. This parameter is two-leveled: the main field corresponds to a more general classification, while the second field highlights additional distinctions.

ne можеть быть' — epistemic modality: impossibility

Pragmatics: NEGATION
Additional semantics: non-expert
Intonation: wh-question
Example queries

http://pragmaticon.ru
Conclusion
Possible applications

- Studies of pragmaticalization – a process of emergence of pragmatic markers which express the speaker's attitude towards something, including the interlocutor(s) and their speech acts (Diewald 2011)
- Classification of pragmatic meanings built on a typological basis
- Catalog of translational equivalents of the DF in different contexts
- Template for multilingual databases of routines (God bless you!) and interjections (Oh my God!)
Future plans

Extending the language sample

...so, we invite you to contribute to our database!
References

“[G]rammaticalization is [...] that subset of linguistic changes through which a lexical item in certain uses becomes a grammatical item, or through which a grammatical item becomes more grammatical” (Hopper & Traugott 2003: 2)


“World lexicon of grammaticalization” (Kuteva et al 2019) lists the inventory of lexical items that develop into grammatical markers.
Pragmaticalization (Aijmer 1997, Onodera 2011)

A process of emergence of pragmatic markers which express the speaker's attitude towards something, including the interlocutor(s) and their speech acts (cf. refusal).

This process remains understudied (see, however, Diewald 2011): it is yet to be discovered, which meanings should be considered pragmatic and what the sources for pragmatic units are.

The process should be studied on typological grounds, aiming at “world lexicon of pragmaticalization”.

Which kind of data could be taken as a basis?
Some observations
DF with verbs of speech

The inner structure tags highlight repetitive patterns:

- Russian čto ty + V-2p.PRES [verb of speech]
- Serbian šta + V-2p.PRES [verb of speech]
- Slovene kaj + V-2p.PRES [verb of speech]

DF appeared to be frozen (partially due to the most frequent ones that are), however bigger data shows variation.

Why so many alternatives that are close both semantically and formally?
Multiple formulae or one schematic construction?

A possible explanation: these are not multiple formulaic phrases but one construction with a slot. In Russian, there is a standard construction Čto Pron VP!

However:

- in DF we encounter only closely synonymous verbs. Constructions are normally more productive, cf. Čto Pron VP!

  Čto ty *begaeš'*! Čto vy tut *sidite!* Čto oni *šumjat!* Čto ty *delaesʼ*!

- this group of DF developed further pragmatic meanings:
  Čto Pron VP! — negative assessment / prohibition of the action
  Čto ty *govoriš!* — negative assessment / prohibition of the speech act + surprise, + negation

Initially, there was a construction with a slot, but only a finite number of its realizations developed into DF.
Another explanation

The pairs <DF structure> — <DF meaning> fit into larger, more basic schemas of pragmatic shifts: <source meaning> — <target meaning>.

Unlike grammatical markers that emerge from lexical items, DF emerge from sentences.

There are usually multiple means (compositional and idiomatic) to put the same proposition in a sentence:
cf. *Ne govori erundy VS. Ne meli čepuxi VS. Čto za erundu ty neseš' VS. Ladno erundu-to nesti*

Within one schema of the shift, the phrasing of the source meaning can vary, regardless of whether on a level of a lexical variable, or as an entirely different construction.
A global strategy:

<SOURCE> discontent with interlocutor’s speech act => <TARGET> negation

Accommodates multiple types of formally different sources, more or less mutually replaceable:

- verbal prohibition (RU: *perestanʹ, ne govori erundy*, SR: *ne lupaj gluposti, nemoj da lupaš*)
- direct assessment of the speech act (RU: *kakaja čepuxa*, SR: *ma to su gluposti*)
- rhetorical question about the speech act (RU: *čto ty govorišʹ;* SR: *ma šta pričaš / lupaš*)
- rhetorical question about the interlocutor (RU: *ty s uma sošel?* SR: *jesi li normalan?* SL: *a si nor?*), etc.

Ultimately, these are indirect speech acts: one message conveyed by means of another.
The more conflict the situation, the bigger variety of indirect means, cf. requests: V! Would you V? / Could you V? / Will you V? / I would like you to V., etc.)
Just like with indirect speech acts, multiple strategies can lead to the same target class of DF.

- enthusiastic agreement => refusal (EN: *like hell I will*, RU: *aga ščas, kak že, ešče čego* SR: *evo trčim, trčim pa se sapličem*)
- declaring the lack of need for an action => refusal (EN: *I'm good, No need*, RU: *ni k čemu*, SR: *nema potrebe, ne treba*)

The DF resulting from the two strategies differ in use: the first group are reactions to requests, and the second — reactions to offers.

The source **meaning** leaves a trace!
Pragmaticalization and variation

Does this variation mean that pragmaticalization differs from grammaticalization?

Despite this seemingly more significant flexibility, DF demonstrate the properties ascribed to grammaticalization:

- increase in frequency, context generalization
- syntactic reduction (parallel to phonetic erosion of grammaticalized lexical items), cf. RU čto ty [VP] (lit. ‘what are you …’), SR ma nemoj [VP] (lit. ‘don’t …’), SL seveda [VP] (lit. ‘it is known that …’)
- semantic bleaching, cf. nontransparent RU nado že ‘needed ptcl’, SR kamo sreče ‘where happiness-gen.sg’, SL v redu ‘in order’
- morphological decategorization of the parts, cf. RU čto / SR šta / SL kaj + V-2p, but not V-1p or V-3p)
Grammaticalization and variation

Additionally, variation is not characteristic of pragmaticalization only, it is found in grammaticalization as well:

- for a certain grammatical category, there is usually one purely grammatical marker
- however, besides it, there is a range of flexible constructions expressing the same grammatical meaning

cf. nominal plural in Russian:

stol-y ‘table-nom.pl’, dom-a ‘house-nom.pl’...

VS kuča ‘bunch’ / more ‘sea’ / buket ‘bouquet’ / gora ‘mountain’ N-gen.pl

see Russian Constructicon (Bast et al. 2021, available at constructicon.github.io)
DF are numerous (> 100 in every language of our sample).

It is natural to have many ways to say ‘yes’ or ‘no’.

Grammaticalization and pragmaticalization are similar phenomena, as they both go through the constructionalization process (see Bybee 2006, Noël 2007).

Expanding the language sample will help us to reveal new strategies of DF pragmaticalization and instances of lexical variation.

...so, we invite you to contribute to our database!