

Laura Becker

Aspectuality in Hungarian, German, and Slavic. A Parallel Corpus Study

Abstract The present paper compares verbal prefixation in Hungarian and German with the expression of aspect in Russian and Czech based on parallel movie subtitles. In order to account for interactions between lexical (actional) properties and aspect, four classes of verbs are considered: relative-statives, activities, gradual-terminatives, and total terminatives. The other factors examined with respect to their relation to aspect marking are: presence of a prefix, presence of a suffix, tense, mood, negation, transitivity (presence of an accusative argument). Results show that Hungarian patterns with Slavic for relative-statives and total-terminatives, while it is similar to German for activities and gradual-terminatives. This hybrid behavior of Hungarian is confirmed by the importance of the factors: in both Slavic languages, the presence of the prefix has the greatest influence on the aspect choice, followed by actionality, tense, and mood. In Hungarian and German, however, actionality is the most relevant factor; therefore, despite many similarities between Hungarian and Slavic, aspect cannot be viewed as grammatical in Hungarian.

Keywords Aspect, Slavic, Hungarian, prefix, parallel corpus

1 Introduction

Aspectuality in Slavic is a well-known and widely discussed topic, as it has been argued to be expressed grammatically and, at the same time, involve derivation (e. g. Dahl 1985; Lehmann 1999). The latter is dominantly expressed by verbal prefixes¹, which are found in a very similar form and function in Hungarian as

1 It is not uncontroversial to assume that prefixation is part of aspect formation proper in Slavic. Isačenko (1960), for instance, argued that real aspectual pairs are only those formed by suffixation. In this paper, I follow the more liberal tradition assuming that at least some prefixes are able to function as proper perfectivizers in combination with certain verbs.

well. For Hungarian, different analyses of the verbal prefixes have been proposed. Some authors argued that they are perfectivizers (e. g. Soltész 1959; Piñón 1995; Kiefer 2006; É. Kiss 2006), while others attributed only a telicizing function to prefixes, delimiting the situation due to the lexical content of the prefix in interaction with the verbal semantics (e. g. Dahl 1985; Eördögh 1986; Csató 1994). Therefore, it is still not clear to what extent aspectuality is grammaticalized in Hungarian. The aim of the present study is to address this issue empirically and to determine to what extent the presence of verbal prefixes and the expression of aspectuality are correlated in Hungarian.

To do so, verbal prefixation in Hungarian will be compared with that of German, as well as with the expression of aspectuality in Russian and Czech. The latter two languages will serve as “aspect” base line and ensure that potential inner-Slavic variation between East- and West-Slavic is accounted for (Dickey 2000; Wiemer 2008). German will be considered for its formally similar system of verbal prefixation, which is not involved in the marking of aspect.

The corpus used consists of parallel movie subtitles from the four languages. By using parallel texts, semantics and pragmatics can be controlled for, which makes aspectual marking directly comparable across languages. Also, the similarity of form-function mapping of aspect in the different languages can be measured, so that Hungarian verbal prefixation and the expression of aspect can be situated between German (no aspect marking) and Czech/Russian (Slavic aspect).

2 Aspect (in Slavic)

2.1 General remarks on aspectuality

There is a general consensus that aspectuality, especially with respect to Slavic, is primarily a matter of ‘boundaries’, meaning that we deal with temporal boundaries of situations (Sasse 2001). Aspectuality can be coded grammatically. In that case, we speak of aspect, which must represent a grammatical category (Dahl 1985:23; Lehmann 1999:218). In order to constitute a grammatical category, the following (idealized) criteria should hold: (i) aspectual values must be abstract and not concrete, (ii) aspect must affect the entire verbal system, and, with respect to the Slavic aspect, (iii) it must feature a binary opposition of imperfective and perfective values. Since the present paper addresses the Slavic aspect type, the following paragraphs will focus on the properties of the latter type only.

The perfective value marks situations as bound in time, its core functions cover the expression of sequences of situations and single events. The imperfective value, on the other hand, presents situations as unbound in time, and is typically used to denote parallel and repeated situations.

The most frequent formal pattern² to derive perfectives and imperfectives in Slavic begins with a simple verb that has been reinterpreted as imperfective. A perfective counterpart can be derived by prefixation from such a verb as *pisat'* 'write', e. g. *na-pisat'* 'write (pfv)'. Since, in some cases, the prefix might add lexical semantics to the verb meaning, we also find new lexemes derived by prefixation (e. g. *pere-pisat'* 'write anew' from *pisat'* 'write'). To form an imperfective counterpart of the perfective, lexically-modified verb, Slavic features suffixation to form "secondary imperfectives" such as *pere-pis-yvat'* 'write anew', which constitute lexical counterparts of the imperfective form.

The Slavic aspect system is highly intertwined with tense. What is formally a present tense perfective has been reinterpreted as future (with a few exceptions). Therefore, the perfective aspect is incompatible with the present tense meaning. Imperfectives, on the other hand, have developed an analytic future tense.

2.2 Aspect and actionality

The notion of aspect, denoting a grammatical phenomenon, is usually employed in opposition to aktionsart as lexical phenomenon.³ For the present purposes, we will distinguish between aspect, i. e. externally set boundaries of a situation independent of inherent semantics of the verb, and actionality (cf. Tatevosov 2002), the latter marking telicity⁴, the inherent boundaries of a situation dependent on the semantics of the verb. These two levels have to be distinguished from each other, since they can combine in the ways displayed in Table 1.

Table 1: Combination of aspect (terminativity) and actionality (telicity).

		telicity	
		telic	atelic
terminativity	perfective	<i>po-stroit'</i> 'build (up)'	<i>po-rabotat'</i> 'work for some time'
	imperfective	<i>na-xodit'</i> 'find'	<i>igrat'</i> 'play'

There are different proposals to integrate actional properties into the selection of aspectual values for Slavic (e. g. Breu 2000; Tatevosov 2002; Lehmann 2009).

2 Note that aspectual pairs can also be marked by other mechanisms: suffix opposition (*stučat'* "knock (ipf) vs. *stuknut'* "knock (pfv)") and suppletion (*brat'* "take (ipf)" vs. *vzjat'* "take (pfv)").

3 Another prominent approach that distinguishes between lexical (situation) and grammatical (viewpoint) aspect is found in Smith (1997).

4 For this use of the term telicity, also see Arkadiev (2015).

The present study adapts the classification of interactions between aspect and actionality from Breu (1994, 2000). The author distinguishes the following actional classes: (i) total-statives, (ii) relative-statives, (iii) activities, (iv) total-terminatives, (v) gradual-terminatives, (vi) inceptive-statives, and (vii) inchoatives. Section 4.2 will discuss the classes in detail.

3 Verbal derivation in Hungarian and German

3.1 Verbal prefixation and suffixation in Hungarian

Hungarian verbal prefixes have mostly been studied for their syntactic properties, as they are separable from the rest of the verb under certain syntactic, formal, or pragmatic conditions (e. g. É. Kiss 2006; Ladányi 2015). Therefore, they have also been referred to as verbal particles. For the sake of comparison, they are labeled as prefixes in the present paper.

Similar to verbal prefixes in many languages, most Hungarian prefixes originate from spatial expressions (Ladányi 2015). The most frequent ones are: *be* ‘into’, *ki* ‘out’, *fel* ‘up’, *le* ‘down’, *el* ‘away’, *meg* ‘completely’⁵.

A prominent function of prefixes, especially of *meg*, is to mark applicatives (1) and upgrade oblique arguments to direct objects (2):

- (1) a. ajándékoz egy könyvet
 give.as.present a book-ACC
 ‘give a book as present’
 (Hungarian)
- b. **meg**-ajándékoz egy barát-ot
 PFX:APPL-give.as.present a friend-ACC
 ‘make a present to a friend’
 (Hungarian)
- (2) a. beszél a helyzet-ről⁶
 talk the situation-DELAT
 ‘talk about the situation’
 (Hungarian)

5 This prefix originates from an expression for ‘behind’, but has lost its lexical semantics almost completely in the current language.

6 The glossing of examples follows the Leipzig Glossing Rules (<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>). Other less common abbreviations used are: DELAT = delative, PFX = prefix, SUPERESS = superessive.

- b. **meg**-beszéli a helyzet-et
PFX:APPL-talk.DEF the situation-ACC
 ‘address the situation’
 (Hungarian)

The other main function is telicization. By adding a goal/delimitation to the verbal meaning, the prefixes telicize the denoted situation, as is shown in the example below:

- (3) a. épít egy város-t
 build a city-ACC
 ‘build a city’
 (Hungarian)
- b. **fel**-épít egy város-t
PFX:UP-build a city-ACC
 ‘build up a city’
 (Hungarian)

What does not seem to be clear until now is whether those telicizing prefixes also perfectivize the situation, i. e. whether delimitation only operates on a lexical or on a more systematic, grammatical level. To illustrate this, two examples from the corpus are given in (4) and (5).

- (4) Persze nem történt volna **meg**, [...]?⁷
 of.course NEG happen.PST.3SG IRREAL **PFX**
 ‘Of course, none of it would have happened’
 (Hungarian, Frozen)
- (5) Hogyan találta **meg**?
 how find.PST.3SG.DEF **PFX**
 ‘How did you find it?’
 (Hungarian, Inception)

In the examples above, the verbs for ‘happen’, and ‘find’ are telic, which means that the function of the prefix cannot be to telicize the situation. Rather, they seem to point out and highlight the telic semantics of the verb. Whether this occurs on a more abstract and systematic level, which would be required to label it aspect, cannot be discussed based on these few examples alone, but is an empirical question and will be addressed in section 5.

7 This example, as well as the following ones except (6) and (7), are taken from the subtitle corpus. The English translations given are the original subtitle lines.

In addition to prefixes, Hungarian also features a few derivational suffixes on the verb that are somewhat involved in the expression of actionality, e. g. deriving frequentatives. However, those suffixes are lexically restricted and occur idiosyncratically. Therefore, they are not considered in the present study.

3.2 Verbal prefixation in German

Verbal particles in German (labeled prefixes in this paper) have also been addressed in previous research with respect to their syntactic status and semantic functions (e. g. Stiebels 1996, Lüdeling 2001); however, they are usually not associated with aspect. In combination with many verbs, prefixes in German add spatial orientation to the verbal semantics, as is shown in (6) below.

- (6) **hinein**-legen
PFX:INTO-put
 ‘put into’
 (German)

Also, applicatives and upgrading of oblique arguments into direct object positions are marked by prefixes on the verb:

- (7) a. ein Buch schenken
 a.ACC book.ACC offer
 ‘offer a book’
 (German)
- b. einen Freund **be**-schenken
 a.ACC friend.ACC **PFX:APPL**-offer
 ‘give a friend a present’
 (German)

Prefixes can also be used to telicize situations, especially if the verbal semantics include an endpoint or limit of the situation that can but does not have to be reached in a given instance. In these cases, the prefix points to that endpoint and hence delimits the situation expressed. Examples (8) and (9) from the corpus below illustrate this:

- (8) Wärm dich **auf**.
 warm.IMP yourself.ACC **PFX:UP**
 ‘Get warmed up.’
 (German, Black Swan)

- (9) Das Herz ist nicht leicht zu **ver**-ändern.
 the.ACC heart.ACC is NEG easily to **PFX**-change
 ‘The heart is not so easily changed.’
 (German, Frozen)

The two simple verbs *wärmen* ‘warm’ and *ändern* ‘change’ refer to situation with no endpoint inherently implied. When combining with a prefix, the latter points to that endpoint so that the situation necessarily is presented as telic.

4 Methodology

4.1 Corpus and annotation

The corpus used for the present study includes subtitles from the movies *Avatar*, *Black Swan*, *Frozen*, *Noah*, and *Inception* (Levshina 2016). From these subtitles, the first 1000 sentences with different verbal lexemes which fulfilled certain requirements (see below) were extracted. Finally, 578 verbs in Russian, Czech, Hungarian, and German were manually annotated for the four languages, so that, in total, 2312 data points could be considered. The choice of tokens was not restricted to certain lexemes to avoid potential bias by particular lexemes. Also, no restriction on the verb classes was made to determine the frequency distribution of those classes is in natural usage (which proved to be fairly equally distributed). Crucial for the choice of tokens, however, was that the meanings of the verbs in the four languages were sufficiently similar.⁸

The predicates selected were annotated for the lexeme, actionality, aspect, presence of a prefix, presence of a suffix (only for Russian and Czech), negation, tense, mood, presence of an accusative object.

We will now address the annotation and values of the factors considered in more detail. As for the prefix, only the presence or absence was noted, independently of whether it derives a new lexeme and/or is no longer separable from the rest of the verb on the synchronic level (e. g. Russian *ubit* ‘kill’, Czech *najít* ‘find’, Hungarian *befejez* ‘end’, and German *erzählen* ‘tell’).

It has also been noted in which cases it is the presence of the prefix that perfectivizes; for a verb like *sozdat* ‘create (pfv)’ the prefix (*soz-*) was counted

8 Although the texts are parallel in the four languages and are used to accompany the same movie scenes, the languages use other constructions, predicates, and sentence types in some contexts. Only those contexts with verbs of shared lexical semantics and the same participants were considered in this study. Note that the verbs in each language, even if sharing lexical meaning, do not necessarily belong to the same actional class.

in. However, it was not counted as perfectivizing prefix, since, synchronically, it is not the prefix itself that perfectivizes the simple verb *dat* ‘give’ without deriving a new lexeme.

I differentiated between the presence of a prefix and a perfectivizing prefix to control for potential similarities in the prefixational systems between German / Hungarian and Slavic due to lexical factors other than aspectuality. In all the four languages addressed, prefixation functions to derive (synchronically and diachronically) lexemes that are lexically more complex. This distinction was made to ensure that the distribution of prefixes in the four languages is not due to lexical effects other than aspectuality.

While that distinction is crucial with respect to the analysis of single verb forms, the two parameters did not influence with respect to the tests applied in this study. Therefore, the following sections will only list the parameter “presence of prefix”.

Due to its lack in German and Hungarian⁹, the presence of an imperfectivizing suffix has only been considered for Russian and Czech.

For coding purposes, I distinguish between four tenses: present, past, future, and infinitive¹⁰, the latter referring to dependent infinitives. For German, I additionally distinguished between preterit and perfect, however, it did not show any effect and will not be considered in the remainder of this paper. For mood, indicative, imperative, and irrealis have been annotated.

Transitivity is tied to telicity, since direct objects often delimit the situation. Therefore, the presence (yes) or absence (no) of an accusative object was annotated to consider the transitivity of the verb (the same notation was applied to the presence/absence of a suffix, prefix, and the negation). Table 2 on page 191 summarizes the most important factors with their values.

The following section will elaborate on how the values for the factors aspect and actionality have been annotated in the four languages.

4.2 Annotation of aspect and actionality

Since aspect in Russian and Czech is systematically marked, its value could simply be determined by the form of the verb. In Hungarian and German, on the

9 As was mentioned in section 3.1, Hungarian has several verbal suffixes that change the actionality of the verb, e. g. derive frequentatives. As this is no systematic process (different suffixes, different compatibilities with verb roots, a high number of lexicalized forms), it has not been considered here.

10 The infinitive was grouped with other tense values for practical rather than linguistic reasons.

Table 2: The factors relevant to aspect marking.

factor	value
actionality	relative-stative (relstative) activity gradual-terminative (gradual) total-terminative (total)
aspect	perfective (pfv) imperfective (ipfv)
presence of a prefix	y n
presence of a suffix	y n
negation	y n
tense	present past future infinitive
mood	indicative imperative irrealis
presence of an accusative object	y n

other hand, aspectuality could not be expected to be marked in a systematic way. Therefore, the context of the situation was taken into account to determine whether a given token refers to a situation as temporarily bound (pfv) or unbound (ipfv).¹¹ Examples for this classification of predicates in German and Hungarian from the corpus are given in (10) and (11).

(10) Wir graben hier. (ipfv)
we dig.PRS.1PL here
'We mine here.'
(German, Noah)

(11) Senki sem fogja meg-látni. (pfv)
no.one NEG will.3SG.DEF PFX-see
'No one will see it.'
(Hungarian, Black Swan)

11 Inevitably, this choice is also influenced by e. g. the actionality of the verb, tense and mood marking, and might vary across annotators. This issue cannot be taken up here, but should be addressed in a future study, e. g. in form of inter-rater agreement, in order to ensure the validity of such subtle semantic judgments.

We will now turn to a more detailed discussion of the four values of actionality adopted from Breu (1994, 2000). Since total-statives (e. g. *weigh*, *be called*) do not have a perfective counterpart in Slavic, only relative-statives have been considered in this study. Relative-stative predicates are defined by the following semantic properties: (i) the situation can but does not have to be inalienably bound to its participants; (ii) a temporal delimitation is possible, but not implied; and (iii) no supply of energy is required to maintain the situation.

Example (12) from the corpus illustrates a relative-stative verb.

- (12) Either way, you'll shine. (Black Swan)
- a. Tak ili inače, no ty budeš' blistat'.
like.this or like.that but you will.2SG shine.IPFV
(Russian)
- b. Ať to dopadne jakkoli, budeš zářit.
whether it turn.out.PFV.3SG so will.2SG shine.IPFV
(Czech)
- c. Így vagy úgy, de ragyogni fogsz.
like.this or like.that but shine. will.2SG
(Hungarian)
- d. Auf die ein oder andere Weise, du
on the one or other way you
wirst auf der Bühne strahlen.
will.2SG on the stage shine
(German)

The next value of actionality, activity, corresponds to activities in the Vendlerian sense. It comprises situations that (i) are non-culminating, homogeneous.; (ii) with a possible but not implied temporal delimitation, and (iii) require a constant supply of energy to maintain the situation, as in (13) below:

- (13) I was dancing the White Swan. (Black Swan)
- a. Ja tancevala partiju beloĵ lebedi.
I dance.IPFV.PST part.ACC white.GEN swan.GEN
(Russian)
- b. Tancovala jsem roli bílé labutě.
dance.PTCP was.1SG role.ACC white.GEN swan.GEN
(Czech)

- c. Én táncoltam a Fehér Hattyút.
I dance.PST.1SG the white swan.ACC
(Hungarian)
- d. Ich tanzte den weißen Schwan.
I dance.PST.1SG the.ACC white.ACC swan.ACC
(German)

We will now turn to total-terminatives. They are similar to what is traditionally understood as achievement verbs, although some differences exist. The semantic criteria for total-terminatives are: (i) the situation is culminating; (ii) a temporal delimitation is inherently given by the lexical semantics; (iii) the situation is not necessarily punctual. An example for the verb ‘kill’ is provided in (14).

(14) Are you here to kill me? (Inception)

- a. Ty prišël ubit’ menja?
you come.PFV.PST kill.PFV me.ACC
(Russian)
- b. Jste zde, abyste mě zabil?
be.PRS.2SG here COMP.2SG me.ACC kill.PST.PTCP
(Czech)
- c. Idejött, hogy meg-öljön?
here.come.PST.3SG COMP PFX-kill.COND.3SG
(Hungarian)
- d. Sind Sie hier, um mich um-zu-bringen?
are you here for me.ACC PFX-to-kill
(German)

The last class of verbs considered are gradual-terminatives. Unlike the previous ones addressed, gradual-terminatives represent a complex class, i. e. consists of two phases. The first one is activity-like, but can lead to a point of culmination, the second phase, which is similar to total-terminatives. Aspectual marking can be used to point to either of the two phases: the imperfective highlights the activity (atelic) part, while the perfective aspect focuses on the culmination (the telic part). An example from the corpus is given below.

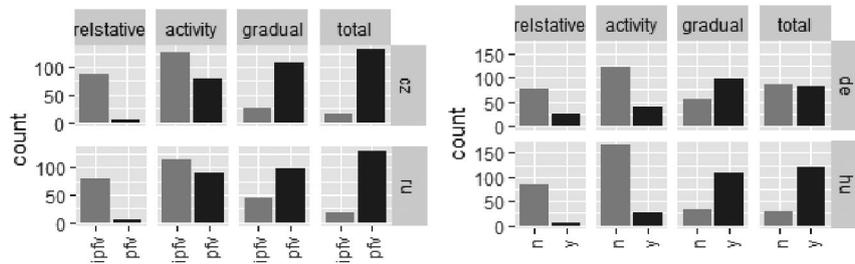


Figure 1: Aspectual values / presence of the prefix across actional classes.

In both Slavic languages, the perfective and imperfective forms occurred as expected: relative-stative verbs are almost exclusively imperfectives (the few perfective forms found were imperatives), also holding for activities as a weak tendency. Gradual-terminatives occurred more often as perfectives. As for total-terminatives, only a few instances of imperfectives are attested, almost all occurrences are perfectives. This reflects the compatibility of actionality and aspectual values: the two atelic classes (relative-stative, activity) are inherently more compatible with the imperfective value, hence, it is more frequent. The telic classes (gradual-terminative and total-terminative), on the other hand, are more compatible with the perfective value, the one attested in most instances.

As for Hungarian, the distribution of the prefix seems to follow the distribution of the aspectual forms in Slavic: almost no prefixes for relative-statives, and a strong trend for prefixed forms with gradual-terminatives and total-terminatives. Activity verbs seem to be less compatible with prefixes than imperfectives in Slavic, which also holds for German. Example (16) shows that the activity verb ‘help’ with a future meaning is perfective in Slavic, but lacks a prefix in Hungarian and German.

- (16) Will He help us? (Noah)
- a. On nam pomožet?
he us.DAT help.PFV.FUT.3SG
(Russian)
 - b. Pomůže nám?
help.PFV.3SG us.DAT
(Czech)
 - c. Segít rajtunk?
help.PRES.3SG us.SUPERESS
(Hungarian)

- d. Wird Er uns helfen?
will.3SG he us.DAT help
(German)

Also in German, gradual-terminatives tend to be more frequent with prefixes than without. The following example shows how the situation is expressed by a perfective form in Slavic, and features a verbal prefix in Hungarian and German:

- (17) You slipped on ice. (Frozen)
- a. Vy poskol'znulis' na l'du.
you.POL slip.PFV.PST.REFL on ice.ACC
(Russian)
- b. Uklouzl jste na ledu.
slip.PFV.PTCP AUX.2SG on ice.ACC
(Czech)
- c. Csak **meg**-csúszott!
only PFX-slip.PST.3SG
(Hungarian)
- d. Du bist **aus**-gerutscht.
you AUX.2SG PFX-slip.PTCP
(German)

As for total-terminatives, German contrasts with Hungarian and Slavic; both forms with and without prefixes occur with no preference. This suggests that there is no aspectual function involved in the prefixation for this class of verbs in German. In (18) below, Hungarian patterns with Slavic perfectives which have a prefix, while German has a simple verb.¹³

- (18) How did you find it? (Inception)
- a. Kak vy eë našli?
how you her.ACC find.PFV.PST
(Russian)
- b. Jak jste to našel vy?
how AUX.2SG it.ACC find.PFV.PTCP you
(Czech)

13 The perfect marker *-ge* in German is not considered as a prefix that can be linked to aspect marking for the purposes of the present paper.

- b. Oheň vše ničí.
 fire all destroy.**IPFV.PRS.3SG**
 (Czech)
- c. A tűz mindent **fel**-emészt.
 the fire all **PFX:UP**-process
 (Hungarian)
- d. Feuer **ver**-zehrt alles.
 fire **PFX-consume** all
 (German)

Transitivity, annotated here as the presence (y) and absence (n) of an accusative object in a given instance, is expected to have an effect on prefixation in German and less so in Hungarian, whereas it should not play a role in aspect marking in Russian and Czech. Figure 3 shows the distribution of (im)perfective verbs and verbs with(out) prefixes across the presence of an accusative object.

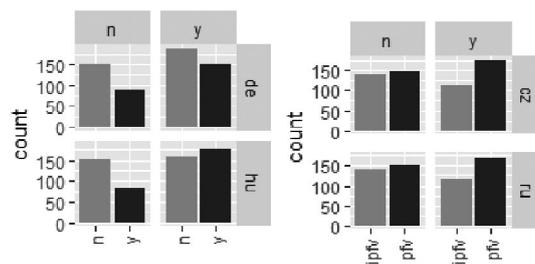


Figure 3: Perfectivity / presence of a prefix with transitivity.

As for verbs with an accusative object, no strong preference can be observed in Hungarian and German. Intransitive verbs, however, occurred with verbs without prefixes more frequently, which suggests that transitivity has the effect of making prefixes be more available to verbs.

Similarly, we do not find an effect for Slavic intransitives. Transitive verbs, on the other hand, show a very weak trend towards perfective forms.

5.2 Factor importance for the marking of aspectuality

After looking at the raw frequency distributions of aspectual forms in Slavic and the verbal prefix in Hungarian and German, we will now address the factors annotated and their importance with respect to the expression of aspect in the

Table 3: Performance of the random forest model for Russian, Czech, Hungarian, and German

Russian			Czech		
	Reference			Reference	
Prediction	ipfv	pfv	Prediction	ipfv	pfv
ipfv	213	12	ipfv	211	22
pfv	43	309	pfv	41	302
Accuracy: 0.9047			Accuracy: 0.9006		
No Information Rate: 0.5563			No Information Rate: 0.5625		
Hungarian			German		
	Reference			Reference	
Prediction	ipfv	pfv	Prediction	ipfv	pfv
ipfv	237	20	ipfv	218	37
pfv	24	291	pfv	23	299
Accuracy: 0.9231			Accuracy: 0.896		
No Information Rate: 0.5437			No Information Rate: 0.5823		

four languages. I used a random forest model to measure the importance of the factors.

Random forests (e. g. Baayen & Tagliamonte 2012; Baayen et al. 2008) can help to determine the strength of factors, i. e. to what extent they are correlated with the dependent variable (aspect). Random forests are based on a large number of conditional inference trees of random sub-samples of the data. Trees split the data according to the factor that makes the purest groups with the smallest p-value with respect to the dependent variable. Random forests (a large number of trees) have some advantages that are crucial for this study. They allow to control for factors that influence each other, as, e. g. tense and actionality, and to observe smaller effects, which would be hidden by more influential factors otherwise. As was noted in section 4.1, the factors considered for the models are: actionality, presence of a prefix, presence of a suffix, negation, tense, mood, and the presence of an accusative object.

Before turning to the results, the accuracy of the model will be addressed, i. e. the question of how well the model is able to capture the data. This is important, since it provides information on how reliable the results of the model are. To determine its accuracy, we let the model predict the values of the dependent variable (perfective, imperfective) based on the factors annotated. These predictions are then compared to the attested forms, providing information as to how well the model performs. Table 3 shows this by way of a confusion matrix for each forest modelling aspect marking in Russian, Czech, Hungarian, and German. The confusion matrix shows the number of tokens the model predicts as pfv/ipfv, while the reference marks the number of attested tokens. Taking Russian as

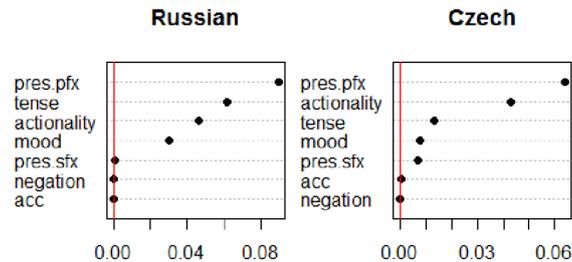


Figure 4: Conditional variable importance for aspect in Russian and Czech.

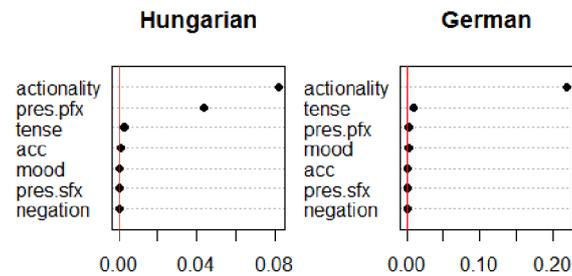


Figure 5: Conditional variable importance for aspect in Hungarian and German.

an example, the model correctly predicts 213 tokens as imperfectives, while 12 perfectives were predicted to be imperfectives. As for perfectives, the model correctly identified 309 tokens, but predicted 43 imperfective forms to be perfective. Given that the model is able to predict the majority of tokens correctly and the accuracy of 0.9047 being clearly above the no information rate¹⁴, we can assume that the random forest for Russian with the factors considered is able to capture aspect marking. The same holds for the other languages, with an accuracy of approx. 0.9. This means that we can model the marking of aspectuality with the same factors in the four languages.

In Figures 4 and 5, we see the conditional variable importance of the factors examined. The conditional variable importance (e. g. Strobl et al. 2008, Baayen & Tagliamonte 2012) indicates how strongly a given factor is correlated with aspect. It is determined by randomly permuting the values of a single factor so that it is no longer linked to aspect. Then, the model's performance is tested: the greater the effect, i. e. the loss of accuracy, the higher the factor's importance.

¹⁴ The No Information Rate is the accuracy the model would have with the levels of the factors randomly manipulated.

The numbers in Figure 4 should not be understood in an absolute way, but are to be interpreted relative to each other. The red line marks significance.¹⁵ The factors that fall to the left of it can be excluded from having a significant effect on the marking of aspect; the factors to the right show a significant correlation with the expression of aspect.

In both Russian and Czech, the presence of the prefix, followed by actionality, tense, and mood are relevant factors to predict whether an instance of a verb is likely to be perfective or imperfective. In Czech, in addition, the presence of a verbal suffix is significant. A more detailed discussion of the role of suffixes in Czech compared to Russian would surpass the scope of the present paper; however, it should be noted that previous work has argued for suffixation to be more productive in East-Slavic than West-Slavic (e. g. Wiemer & Seržant Forthc.; Arkadiev 2015). The present results rather suggest the opposite; as this surpasses the scope of the present paper, this issue will not be discussed in more detail here.

The presence of an accusative object and negation do not have an influence on the marking of aspect in either Russian or Czech. Hence, aspectual marking is highly correlated to the presence of a prefix, to lexical properties of the verb (actionality), and to other verbal categories.

In Hungarian and German, on the other hand, the most influential factor clearly is actionality. Hungarian shows a hybrid-like behaviour. On the one hand, it shares the high significance for actionality and significance of a much lower degree for tense with German. On the other hand, Hungarian patterns with Slavic for the high significance of the presence of the prefix to aspect marking. Thus, verbal prefixation in Hungarian is systematically involved in aspectual marking.

These findings support the initial hypothesis that aspect is systematically expressed in Hungarian to a certain extent, while it is not in German, so that Hungarian can be positioned between German (no aspect marking) and Slavic (aspect as a grammatical category). However, Hungarian also patterns with German in contrast to Russian and Czech, since the main factor correlated to aspectual functions is lexically determined. Although actionality plays a significant role in Slavic as well, it is less relevant than in Hungarian or German. Moreover, Slavic showed a significant influence for mood and tense, which means that aspect interacts with other verbal categories. For both German and Hungarian, the system is less complex as it depends on fewer factors and is more directly correlated to actionality, the lexical properties of the verb.

15 Following Strobl et al. (2008) to determine which factors are significant, their values were compared to the absolute value of the lowest negative value, the latter being indicated by the red line.

5.3 Similarity between the four languages with respect to aspect marking

Since the previous section showed that Hungarian patterns with German but also with Slavic with respect to different properties, this section will address the similarity between the four languages with respect to aspect marking in more detail. The similarity is determined also based on the factors considered for random forests, repeated here: aspect, negation, tense, mood, acc, presence of the prefix, presence of the suffix. Taking these factors in the four languages, we can measure the difference between them by clustering the languages according to their value distributions of the factors.

The cluster in Figure 6 confirms the results discussed in the previous sections. Russian and Czech pattern together, however, cutting the cluster at a higher point, Hungarian also patterns with Slavic, being situated between Slavic and German.

However, if we look at the four actional classes separately, we find that the languages cluster in two different ways. For activity and gradual-terminative verbs (see figure 7), we find a cluster of Slavic on the one hand, and German and Hungarian on the other. Relative-stative and total-terminative verbs in figure 8, however, show that Hungarian clearly patterns with Slavic instead of German.

We will now consider the clustering for each class in more detail. In section 5.1, it was observed for activity verbs that in Hungarian and German, many verbs do not combine with a prefix, so that there is no formal opposition available, which sets them apart from Slavic, featuring both imperfective and perfective forms (cf. (16)). This can explain why German and Hungarian pattern together for activity verbs.

Gradual-terminatives cluster in the same way, although they are compatible with both perfective and imperfective values and would be expected to be the first group of verbs showing aspect marking in an emerging aspect system, so that Hungarian would have been expected to pattern with Slavic for this group. A possible explanation for the clustering with German could be that, although in both languages prefixes are available, their distribution differs from perfective and imperfective forms in Slavic. Example (19) in section 5.1 showed that Slavic used the imperfective due to the present tense, whereas in both Hungarian and German the prefix was present. We also find cases in which Slavic uses a perfective form, with no prefix being present in Hungarian and German, possibly because of the direct object delimiting the situation, as in (20) below.

All verbs

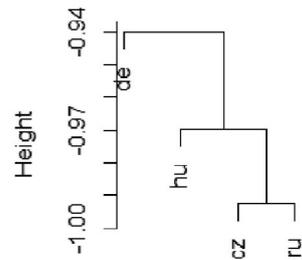
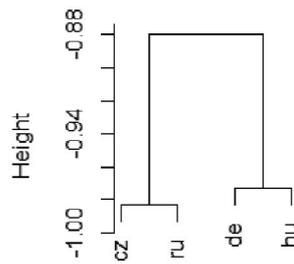


Figure 6: Similarity of the four languages for all verbs.

Activity



Gradual-terminative

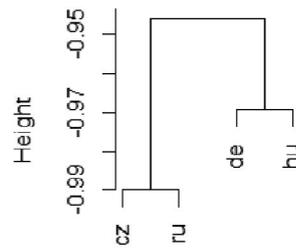
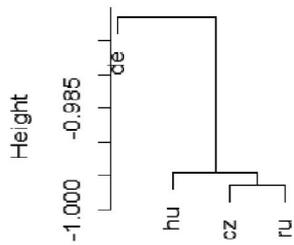


Figure 7: Similarity of the four languages for activity and gradual-terminative verbs.

Relative-stative



Total-terminative

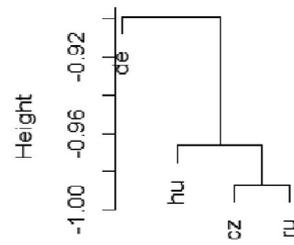


Figure 8: Similarity of the four languages for relative-stative and total-terminative verbs.

- (20) Do you want to build a snowman? (Frozen)
- a. ty xočeš' slepit' snegovika?
 you want.PRS.2SG build.PFV snowman.ACC
 (Russian)
- b. Postavíme sněhuláka?
 build.PFV.1PL snowman.ACC
 (Czech)
- c. Építünk hóembert?
 build.PRS.1PL snowman.ACC
 (Hungarian)
- d. Bauen wir einen Schneemann?
 build.PRS.1PL we a.ACC snowman.ACC
 (German)

This means that although prefixation for gradual-statives is available in both German and Hungarian, the distribution of prefixes rather depends on actional properties and does not correspond to the aspectual distribution of Slavic forms.

Figure 8 illustrates that verbs from the relative-stative and total-terminative classes cluster Hungarian together with Slavic against German. For relative-statives, this can be explained by the fact that German uses verbal prefixes much more frequently for relative-statives than the other three languages (cf. Figure 3 in 5.1).

For total-stative verbs, section 5.1 (cf. Figure 1) showed that the distribution of the prefix in Hungarian follows the distribution of perfectives in Slavic, while there was no tendency for prefixation found in German, which can explain the cluster in Figure 8. An example where Hungarian patterns with Slavic featuring a prefix vs. German using a simple verb is given below:

- (21) Although, I dreamed I was kissed by a troll. (Frozen)
- a. [...] što menja poceloval troll'.
 COMP me.ACC kiss.PFV.PST troll
 (Russian)
- b. [...] že mě políbil troll.
 COMP me.ACC kiss.PFV.PTCP troll
 (Czech)
- c. [...] **meg**-csókolt egy troll.
PFX-kiss.PST.3SG a troll
 (Hungarian)
- d. [...] ein Troll hat mich geküsst.
 a troll AUX.3SG me.ACC kiss.PTCP
 (German)

6 Conclusion

This study addressed the systematicity of the expression of aspectuality in Hungarian compared to Russian, Czech, and German. The first two languages represented the East- and West-Slavic type of aspect respectively, while German functioned as the control language with verbal prefixation available, but without aspect marking.

Based on parallel subtitles, this study could empirically show to what extent Hungarian marks aspectuality by verbal prefixation and in which properties it resembles more Slavic or German behaviour. In addition, the verbs considered were split into four different actional classes to account for the interaction of inherent lexical (actional) properties and aspectual functions.

With respect to actional classes, the raw distributions of the prefix in German and Hungarian suggested that, except for activity verbs, prefixation in Hungarian indeed resembles the distribution of perfective and imperfective forms in Slavic. In German, on the other hand, especially for gradual-statives and total-statives, no such effect could be found. Distance-based similarity measures for the actional classes confirmed that Hungarian clusters with Slavic for relative-statives and total terminatives, while it showed that Hungarian forms a cluster with German not only for activities but also gradual-statives. Although prefixes are available in this class and it is semantically most compatible with both the perfective and imperfective values, the distribution of the prefix in both Hungarian and German is much more dependent on actionality and telicity, which is not the case in Slavic.

Based on random forest models, the importance of the factors annotated (actionality, presence of prefix, presence of suffix, tense, mood, presence of accusative object, negation) was determined. For all the four languages, the accuracy of the model was above 0.89, which means that the factors considered indeed capture the expression of aspect. With respect to the relevance of the factors, Hungarian showed a hybrid behaviour between German and Slavic, which could be confirmed by distance-based similarity measures for the four languages. What grouped it together with German was the fact the main significant factor to predict the aspectual value of a given form was actionality, i. e. an inherent lexical property, which argued against aspect as a grammatical category, systematically expressed and independent of lexical properties. However, actionality was significant in Slavic as well, amongst the presence of the prefix, tense, mood (and the presence of the suffix in Czech). This showed that in Slavic, aspect marking was sensitive to the actional properties of the verb as well. The presence of a prefix was also a highly significant factor in Hungarian, which made it group together with Slavic, whereas the presence of a verbal prefix in German, as expected, did not seem to correlate with aspectual values.

To conclude, this paper showed how the prefixation in Hungarian has to be situated between Slavic and German with respect to aspect marking. It could be shown that prefixation in Hungarian significantly correlates with the expression of aspectuality across the four actional classes. However, lexical properties of the verb, i. e. actionality, still have the greatest influence on the aspectual interpretation of a verb, which argues against a grammatical category of aspect in Hungarian.

References

- Arkadiev, Peter. 2014. Towards an Areal Typology of Prefixal Perfectivization. *Scando-Slavica* 60 (2): 384–405.
- Arkadiev, Peter. 2015. *Areal'naja tipologija prefiksals'nogo perfekтива (na materiale jazykov Evropy i Kavkaza)*. Moskva: Jazyki slavjanskoj kul'tury.
- Baayen, R. Harald; Douglas J. Davidson and Douglas M. Bates. 2008. Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language* 59 (4): 390–412.
- Baayen, R. Harald and Sali A. Tagliamonte. 2012. Models, Forests and Trees of York English: Was/Were Variation as a Case Study for Statistical Practice. *Language Variation and Change* 24 (2): 135–178.
- Breu, Walter. 1994. Interactions between Lexical, Temporal and Aspectual Meanings. *Studies in Language* 18 (1): 23–44.
- Breu, Walter. 2000. Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und Lexikalische Interaktion). In Walter Breu (ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)*, 21–54. Tübingen: Max Niemeyer.
- Csató, Éva Á. 1994. Tense and actionality in Hungarian. In Rolf Thieroff and Joachim Ballweg (eds.), *Tense systems in European Languages*, 231–246. Tübingen: Max Niemeyer.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.
- Dickey, Stephen M. 2000. *Parameters of Slavic Aspect: A Cognitive Approach*. Center for the Study of Language and Information: Stanford, CA.
- Eördögh, Miklós. 1986. Aspect or aspectuality in Hungarian. In Wolfgang Heydrich and Janos S. Petöfi (eds.), *Aspekte der Konnexität und Kohärenz von Texten*, 115–27. Hamburg: Buske.
- É. Kiss, Katalin. 2006. The Function and the Syntax of the Verbal Particle. In Katalin É. Kiss (ed.), *Event Structure and the Left Periphery. Studies on Hungarian*, Vol.68, 17–56. Budapest: Springer.
- Isačenko, Alexandr V. 1960. Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim. Čast'vtoraja: morfologija [Grammatical system in Russian as opposed to Slovak. Part 2: Morphology]. Bratislava: Izdatel'stvo akademii nauk.

- Kiefer, Ferenc. 2006. *Aspektus és akcióminőség: különös tekintettel a magyar nyelvre*. Budapest: Akadémiai Kiadó.
- Ladányi, Mária. 2015. Particle Verbs in Hungarian. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer (eds.), *Word-Formation. An International Handbook of the Languages of Europe*. Vol. 1, 660–672. Berlin/Boston: De Gruyter Mouton.
- Lehmann, Volkmar. 1999. Der Aspekt. In Helmut Jachnow (ed.), *Handbuch der Sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*, 214–242. Wiesbaden: Harrassowitz.
- Lehmann, Volkmar. 2009. Formal-Funktionale Theorie Des Russischen Aspekts. <http://subdomain.verb.slav-verb.org/Aspekttheorie.html> (28 May 2017)
- Levshina, Natalia. 2016. Verbs of Letting in Germanic and Romance: A Quantitative Investigation Based on a Parallel Corpus of Film Subtitles. *Languages in Contrast*. 16 (1): 84–117.
- Lüdeling, Anke. 2001. *On Particle Verbs and Similar Constructions in German*. Dissertations in Linguistics. CSLI: Stanford, CA.
- Piñón, Christopher J. 1995. Around the progressive in Hungarian. In István Kenesei (ed.), *Approaches to Hungarian. Volume Five. Levels and Structures*, 53–92. Szeged: JATE.
- Sasse, Hans-Jürgen. 2001. Recent Activities in the Theory of Aspect: Accomplishments, Achievements, or Just Non-Progressive State? In *Arbeitspapiere des Instituts für Sprachwissenschaft der Universität zu Köln*. Vol. 40.
- Smith, Carlota S. 1997. *The Parameter of Aspect*. Dordrecht, Boston, London: Kluwer.
- Soltész, Katalin J. 1959. *Az Ősi Magyar Igekötők. Meg, El, Ki, Be, Le, Fel*. Budapest: Akadémiai Kiadó.
- Stiebels, Barbara. 1996. *Lexikalische Argumente Und Adjunkte: Zum Semantischen Beitrag von Verbalen Präfixen Und Partikeln*. Berlin: Akademie Verlag.
- Strobl, Carolin; Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin and Achim Zeileis. 2008. Conditional variable importance for ran om forests. *BMC Bioinformatics* 9: 307.
- Tatevosov, Sergej. 2002. The Parameter of Actionality. *Linguistic Typology* 6 (3): 317–401.
- Wiemer, Björn. 2008. Zur Innerslavischen Variation bei der Aspektwahl und der Gewichtung ihrer Faktoren. In Karl Gutschmidt, U. Jekutsch, Sebastian Kempgen and Ludger Udolph (eds.), *Deutsche Beiträge Zum 14. Internationalen Slavistenkongreß, Ohrid 2008*, 383–409. München: Sagner.
- Wiemer, Björn and Ilja A. Seržant. Forthcoming. Diachrony and Typology of Slavic Aspect: What does morphology tell us? In Andrej Malchukov and Walter Bisang (eds.), *Unity and diversity in grammaticalization scenarios*. Berlin: Language Science Press.