English or Czech Venitive Verbs in Contrast: Deictic or not?

Michaela Martinková (Palacký University Olomouc)

Markéta Janebová (Palacký University Olomouc)

Abstract
This study compares the self-agentive uses of English and Czech venitive verbs, i.e., *come* and intransitive motion verbs prefixed by *při-*, in a bidirectional parallel corpus of English and Czech subtitles. Their mutual correspondence is only 45.2%, and *při*-verbs are translated by *come* more often (61.4%) than *come* by *při*-verbs (34.7%). As factors restraining the use of *při*-verbs we identify their strong association with the arrival perspective and the fact that the Czech venitive verbs are typically perfective. As a result*, při*-verbs are typically avoided in directives involving immediate imperatives as well as in permission seeking speech acts and their rejections, in comitative contexts, and in situations which call for a construal by an imperfective verb. Also mentioned is the role of functional factors, which seem less strong in Czech than in English. In the opposite direction of translation, contexts in which *při*-verbs tend not to be translated by *come* are those in which they describe motion directed at goals other than the speech participants. It is suggested that the details of the contextualization of the English venitive verb need to be revisited, mainly with respect to its narrative uses, and possibly also others, emerging in the context of changes in our understanding of space in the current world.

Keywords: Czech/English; parallel corpus; venitive verb; deixis; contrastive linguistics

1. Introduction
Cross-linguistically, venitive verbs differ in whether they conflate deixis with motion in a verbal root, code deixis on a verbal affix, or are not

deictic;[1] when deictic, they differ in which extensions of the deictic centre they allow (e.g., Gathercole 1978; Lewandowski 2014; Ricca 1993). Slavic deictic verbs have received much less attention (Lewandowski 2014) than English (Fillmore 1971) and Spanish deictic verbs (e.g., Hijazo-Gascón 2017, 2021; Fernández-Couceiro 2001; Calle-Bocanegra 2019), and some linguists take the deictic character of Slavic prefixes on venitive verbs for granted (Malá 2015; Slobin 2004). Furthermore, while the studies were originally more or less based on native-speaker intuitions (Fillmore 1971; Gathercole 1978; Ricca 1993), in recent years linguists have started to use empirical methods such as experimental ones (Matsumoto et al. 2017); experiments have also been used to test the effect of cross-linguistic differences on the acquisition of Spanish (Lewandowski 2014; Hijazo-Gascón 2017, 2021; Calle-Bocanegra 2019).

Our paper contributes to the cross-linguistic study of venitive verbs in two ways: first, by comparing English with Czech, a West Slavic language, whose venitive verbs have so far received very little linguistic attention (for notable exceptions, see Hirschová 2007 and Calle-Bocanegra 2019), and second, by using corpus tools. More specifically, we compare the use of venitive verbs in English (*come*) and Czech (intransitive motion verbs prefixed by *při-*) in the description of motion events as attested in a bidirectional parallel corpus of English and Czech subtitles available through the multilingual corpus InterCorp (Čermák and Rosen 2012). The language of subtitles was selected since, unlike the language of fiction, it has a movie to resort to if the concrete spatial configuration of speech participants is not clear. Our research questions broadly fall into the following sets:

1. How frequent are the self-agentive uses of English and Czech venitive verbs? What is their mutual correspondence (Altenberg 1999), i.e., in what proportion are they translated by each other in the bidirectional parallel corpus?
2. Do cases of non-correspondence point to specific constraints on the use of these verbs and if so, which ones?

---

[1] We use the term 'venitive verb' to cover verbs that code coming motion, but without presupposing their deictic character; in our use of the term, it includes *come* and its dictionary equivalents in other languages. Similarly, by andative verbs, we mean dictionary equivalents of *go*.

3. What are the methodological challenges of using the corpus methodology for a cross-linguistic study of deictic verbs?

The structure of the paper is as follows: section 2 reviews the literature on the factors conditioning the use of venitive verbs cross-linguistically, with a focus on English, Spanish, Slavic languages, and Czech in particular. Section 3 describes the data as well as methods of their sorting, annotation and analysis. Section 4 summarizes the results; conclusions are presented in section 5.

2. Deictic verbs of motion in linguistic literature
According to Fillmore (1971: 38), deictic expressions are those which require contextualization, with 'context defined in such a way as to identify the participants in the communication act, their location in space, and the time during which the communication act is performed'. When it comes to the semantic descriptions of venitive verbs, the details of contextualization to be sorted out are those related to the goal of motion: whether the presence of speech act participants is required at the goal of motion for the verb to be felicitous, which participant(s), and at what time. More specifically, the question is whether the use of the venitive verb only allows the prototypical deictic centre, i.e., the presence of the speaker at the goal of motion at the time of the speech act (coding time), as in Portuguese; or whether it allows an extension of the deictic centre into 'the point or period that is the temporal focus or background for the event or condition being described in the clause' (Fillmore 1971: 52), i.e., the reference time, as in Spanish; or whether it can also be extended to where the addressee is (English or Catalan); or finally, whether the venitive verb can also be used if neither the speaker nor the addressee is at the goal of motion. Lewandowski summarizes this as shown in Table 1. Crucially, the hierarchy of grounds lexicalized in venitive verbs should be interpreted as a universal implicational hierarchy; the venitive verb 'which can take as the Ground a goal of movement situated lower in the hierarchy than the speaker's location at the coding time automatically allows for any other goal which is placed higher in the established hierarchy' (Lewandowski 2014: 46).

Table 1: Hierarchy of grounds lexicalized in venitive verbs (Lewandowski 2014: 46).

Goal	Languages
1. the speaker's location at the coding time	Portuguese, Shibe, …
2. the speaker's location at the reference time	Jacatlec, Spanish, …
3. the addressee's location	Catalan, English, Nepali, Turkish, …
4. another goal of movement	Czech, Polish, Russian, …

In English, *come* allows the presence of the speaker as well as the addressee at the goal of motion; in the latter case Lyons writes about deictic projection, i.e., the speaker 'project[s] himself into the spatiotemporal location of the addressee' (1977: 579). What follows is that many English sentences with *come* are ambiguous with respect to which speech participants are presupposed at the goal of motion and when. According to Fillmore (1971: 10), a speaker ('I') can say a sentence such as (1) at least in the following situations: 'if I am in the office when I say it, if you are in the office when I say it to you, if I was in the office yesterday morning when John came, or if you were in the office yesterday morning when John came':

(1) John *came* to the office yesterday morning.

Certain readings may be ruled out if other deictic expressions appear in the sentence, e.g., deictic spatial adverbs and/or pronominal reference to the speaker or addressee, or if it is contextually given that both the speaker and addressee are together in the same room when talking to each other. For example, (2a) places the speaker at the goal of motion at coding time, but not at reference time, and (2b) places the speaker there at reference time:

(2) a. He came here two hours before I arrived. (Fillmore 1971: 55)
    b. John came to visit me at that place. (Lewandowski 2014: 46)

Crucially, if an extension of the deictic centre (deictic projection) takes place, the venitive and andative verbs may alternate (Lewandowski 2014: 47). Using sentences from Fillmore (1971), Lewandowski shows that alternations between *go* and *come* are possible if the goal is the location of

the speaker (3b) or addressee (3c) at reference time, but not if the goal is the speaker at coding time, or some other goal; in the former case (3a) *go* is not acceptable, in the latter (3d) *come* is typically not acceptable (Lewandowski 2014: 46):

(3) a. He came/*went here two hours before I arrived. (Goal 1)
    b. He'll come/go to the office tomorrow to pick me up. (Goal 2)
    c. She'll come/go there to meet you. (Goal 3)
    d. Tomorrow, I'll go/*come to John's place. (Goal 4)

Fillmore notes that *come* is also acceptable in the description of motion 'toward the location of the home base of either the speaker or the hearer at reference time' (1971: 61) even if neither the speaker nor the hearer are actually present there; this is what we see in (4a–b):

(4) a. He came over to my place last night, but I wasn't home.
    b. I came over to your place last night, but you weren't home.

However, a sentence such as (5a) is only felicitous if 'the addressee is taken to be in the speaker's home at the time the sentence is said, or that the place is also the addressee's home' (Fillmore 1971: 62). Similarly, 'the conditions on "come" do not allow us to say things like [(5b)] when the home is the addressee's alone, and the speaker is not at the addressee's home at coding time or reference time' (Fillmore 1971: 64).

(5) a. I am going to come home.
    b. Are you going to come home?

Apart from descriptions of motion to a home base, *come* can also be used in certain narrative contexts; 'in discourse in which neither speaker nor addressee figures as a character,' *come* marks 'motion toward a place taken as the subject of the narrative, toward the location of the central character at reference time, or toward the place which is the central character's home base at reference time' (Fillmore 1971: 67). According to Fillmore, there can only be one such deictic centre at a time (1971: 67).

Finally, English, like several other languages, also allows the use of *come* in comitative contexts, irrespective of the goal of motion: *come* is felicitous in the descriptions of 'motion at reference time which is in the

company of either the speaker or the addressee' (Fillmore 1971: 66). This condition licenses the acceptability of (6a–c), but not of (6d), 'unless some of the other appropriateness conditions for "come" are satisfied' (Fillmore 1971: 66).

> (6) a. Can I come (along)?
>   b. Would you like to come (along)?
>   c. Can Johnny come (with you)?
>   d. *Can I come (with Fred)?

According to Gathercole (1978: 84), if the presence of the speaker or hearer is overtly expressed in the *with*-phrase, the contrast between *come* and *go* verbs in this context may be neutralized; '[i]n choosing one of the verbs, however, speakers in some languages, and perhaps to some degree in all languages, draw on the features of intimacy, imminency, and closeness' (Gathercole 1978: 84). These components, argues Gathercole, might play a role also in selecting venitive verbs over andative ones in some other contexts where motion is directed at goals other than the speech participants.

The role of factors other than strictly directional is investigated in an experimental study by Matsumoto et al. (2017). The authors compare the use of venitive verbs in three typologically different languages, satellite-framed English, verb-framed Japanese, and equipolently-framed Thai (Slobin 2004) by eliciting descriptions of video clips in which people are moving towards the speaker (the camera view), in a direction off to a side of the speaker, across in front of the speaker, and away from the speaker. The functional factors tested are the presence/absence of interactional behaviour ('whether the moving person greeted or smiled at the speaker while in motion or not'; Matsumoto et al. 2017: 101), and the presence/absence of the speaker's space, operationalized as the speaker's location in an open space versus his/her location inside of a room and on a lower floor; also tested is the aspect of change in visibility. The authors report that 'in English, motion into or onto the speaker's space (i.e., classroom or staircase level) facilitated the use of *come*, with the difference between the open space scene and the staircase scene statistically significant' (Matsumoto et al. 2017: 105). In all three languages, 'motion onto the speaker's level elicited venitive verbs very often, even when the motion is not spatially directed to the speaker,'

including motion away from the speaker (Matsumoto et al. 2017: 108). Similarly, the use of venitive verbs was higher in a scenario described as 'the invisible to visible exiting scene' than in scenarios with no change in visibility,[2] even when 'the motion in the invisible-to-visible exiting is not directed to the speaker at all' (Matsumoto et al. 2017: 110). Finally, the presence of interactional behaviour (the moving person greeting and smiling at the speaker) enhanced the use of venitive verbs (Matsumoto et al. 2017: 110). The study thus shows the importance of functional components in the speaker's choice of a venitive verb over an andative one, at least in the three languages studied.

2.1. Slavic venitive verbs

If deixis is coded on Slavic venitive verbs, then this happens on their prefixes; dictionary equivalents of *come* are typically perfective verbs created by perfectivizing prefixation of simplex imperfective verbs,[3] including the andative ones. Ricca (1993: 82), however, sees a difference between East and West Slavic languages on the one hand (he includes Polish, Czech, Russian, and Ukrainian), which he classifies as non-deictic, and South Slavic languages such as Serbo-Croatian and Slovenian on the other, which he classifies alongside English, i.e., as predominantly deictic. Gathercole (1978: 75), too, puts Serbo-Croatian in the same group as English, i.e., among languages which allow what she calls 'extended deixis' (speaker or addressee at goal at reference time). Her example is given in (7):

(7) Hoćeš li *doći/ići* k meni?
     want.2SG.PRS QM come.INF.PFV/go.INF.IPFV towards me
     'Will you *come* to my place?'

According to the author (1978: 75), the difference between *doći* and *ići* is that *doći* is perfective while *ići* is imperfective: *doći* emphasizes 'arriving

---

[2] The mover was walking out of a small building, 'with the speaker outside the building, looking sideways at the movement, to whom the moving person is visible only outside the building' (Matsumoto et al. 2017: 109).

[3] Aspect is an intrinsic property of Slavic verbs; in the text, it is marked in a superscript, e.g. *přijíždět*[i] and *přijet*[p], where *i* stands for imperfective and *p* for perfective.

at my house' while *ići* 'ongoing movement away from your house […]. When there is a strong interest in arrival at the GOAL, then *ići* becomes unacceptable' (Gathercole 1978: 75).

However, when the speaker wants to profile an ongoing movement rather than the arrival itself, then andative verbs are used also in languages classified by Ricca as non-deictic, namely Czech, Polish, Russian, and Ukrainian, simply on the grounds of being imperfective. Ultimately, argues Ricca (1993: 85), the andative verbs in these languages can be used to describe both motion to a deictic and non-deictic goal, as in the Czech examples in (8a–b):

(8) a. Zrovna k nám *jde*. (Ricca 1993: 85)
      just towards us walk.3SG.PRS.IPFV
      'He *is* just *coming* towards us.'
      b. *Jde* k támhle tomu domu. (Ricca 1993: 85)
      walk.3SG.PRS.IPFV towards there that house
      'He *is going* to that house over there.'

The real difference between South Slavic languages and East/West Slavic is thus whether their andative verbs are manner-neutral verbs, and whether these verbs merge with a potentially deictic prefix to create a full-fledged deictic verb meaning 'to come'. This is what we see in Serbo-Croatian, where '*doći* "to come" has lexicalized from the combination of *do ići*' (Verkerk 2014: 43),[4] or in Bulgarian (Speed 2015: 56). In East and West Slavic, venitive verbs are those with 'the North Slavic allative prefix *pri-*' (Dickey 2010: 96), which productively combines with manner-of-motion verbs giving fully compositional verbs such as the Czech *přijít*[p] 'při-walk' and *přijet*[p] 'při-drive'; the same seems to apply to other West and East Slavic languages. Linguists disagree, however, in whether they classify this prefix as deictic: Slobin (2004: 227) glosses the Russian prefix *pri-* as 'a deictic prefix on a motion verb,' and Malá (2015: 174) its Czech cognate *při-* as indicating 'directed motion towards the deictic centre (the speaker)'.[5] Lewandowski, on the other hand, puts Czech and Russian in

---

[4] According to Filipović (2007: 114), the Serbo-Croatian prefixes which 'habitually express deixis (a strong feature in Serbo-Croatian lexicalization and narrative pattern)' are *do-/ot-* 'toward/away from.'

[5] According to Šlosar (1996: 203), verbs prefixed by *při-* code events directed at contact with something. Daneš at al. (1981: 99) note that motion coding predicates

the same group as Polish, i.e., languages in which 'the use of C[ome] & G[o] is related to other, non-deictic factors' (2014: 44). According to him, 'in Polish […] C is preferred when the speaker wishes to adopt an arrival-oriented perspective, and G, if the motion event is conceptualized from a source-oriented perspective'; the meaning of arrival is also attributed to the Russian prefix *pri-* in Janda et al. (2013).

Crucially, none of the studies involving East/West Slavic languages provide a systematic analysis of verbs prefixed by *pri-* and its cognates with respect to their potential deictic character: the aim of Lewandowski's study is to investigate a potential transfer into Spanish used by Polish learners (2014), and Czech learners of Spanish are targeted in Fernández-Couceiro (2001) and Calle-Bocanegra (2019). According to the former, a frequent source of errors in the use of *venir* 'come' and *ir* 'go' by Czechs is an erroneous identification of *venir* with *přijít*[p] 'při-walk' and *ir* with *jít*[i] 'walk', common in Czech textbooks of Spanish (Fernández-Couceiro 2001: 3–4). The three major cases where Czech differs from Spanish, but also English, seem to be the following:

Spanish *venir* and English *come*, but not Czech *při-* verbs, are common in comitative contexts such as (9). Spanish and English also differ from Czech in cases such as (10), where motion is directed at the speaker: while *venir*/*come* are obligatory, 'Czech may use the verbs *jít* or *jet* "go"' (Calle-Bocanegra 2019: 89); according to the author, this happens in a departure perspective (2019: 106).

(9) a. ¿Quieres *venir* esta noche al cine conmigo? (Calle-Bocanegra 2019, 106)
   want.2SG.PRS come this night to cinema with me
   b. *Půjdeš* se mnou večer do kina?
   walk.2SG.FUT with me evening to cinema
   c. Would you like to *come* with me to the cinema?

---

with this prefix 'imply' the goal of motion, i.e., the goal of motion has to be known to the speaker.

(10)   a. ¿Ves ese robot que *viene* hacia nosotros? (Calle-Bocanegra 2019, 89)
see.2SG.PRS this robot that come.3SG.PRS towards us
b. Vidíš toho robota, co *jede* k nám?
see that robot that go/drive.3SG.PRS.IPFV towards us
c. Can you see the robot that *is coming* towards us?'

Finally, in (11) the Figure moves in a direction other than to the speaker or to the hearer, and the perspective is one of arrival. Spanish only allows *ir*, and in English *go* seems the only choice if neither the speaker nor the addressee are conceptualized as sharing the space with the goal of motion.[6] Czech, in contrast, allows *při*-verbs in such situations:

(11)   a. Honza *fue* a casa de Martin para ayudarle (Fernández Couceiro 2001, 8)
Honza go.3.SG.PST to house of Martin for help.him
b. Honza *přišel* k Martinovi, aby mu pomohl.
Honza při-walk.PTCP.SG.M to Martin so.would him help. PTCP.SG.M
c. Honza *went* to see Martin to give him a hand.

Calle-Bocanegra (2019: 101) concludes that even the most advanced Czech learners 'have not fully acquired the usage of Spanish deictic verbs of movement.' However, the concrete performance on cases of mismatch varied depending on the task used to elicit the data, and erroneous responses occurred even in some of the contexts where Czech and Spanish are in line. This was the case of (12) where venitive verbs are used to describe motion toward the speaker at coding time in an arrival perspective:

(12)   a. ¿Puedes *venir* aquí un momento? (Calle-Bocanegra 2019: 106)
can.2SG.PRS come.INF here one moment
b. Můžeš sem na chvilku *přijít*?
can.2SG.PRS here.to for while při-walk.INF.PFV
c. Can you *come* here for a moment?

---

[6] *Come* then differs from *venir* in allowing the hearer at the goal of motion.

One reason for the lower performance could be that the Czech andative verb is often possible in descriptions of motion towards the speaker (examples (8), (10), but also (12b)). Another is that if the test items in Calle-Bocanegra's study included imperative forms, then the item to be considered for a potential transfer would be *pojd'/te* '*po*-walk.IMP.SG/PL', in which the combination of the prefix *po-* and a motion verb (including the andative *jít^i* 'walk, go') gives rise to what Biskup (2019: 159) calls 'a speaker-oriented futurate meaning'. The existence of these special imperative forms is noted in Ricca (1993: 89), who writes about immediate imperatives of the type *come here*, which he observes in Polish and in Czech, but not in Russian or Ukrainian. As (13), a sentence from the parallel translation corpus InterCorp, demonstrates, *pojd'* corresponds not only to the Spanish, but also to the English venitive verb:

(13)   "Maxi!" zavolal na něj Švejk, "*pojd'* ke mně!"[7]
          "Max!" Švejk yelled at him. "*Come* to me!"
          ¡Max! - gritó Švejk. - ¡*Ven*!

The deictic nature of this usage appears to be confirmed by the fact that the same forms are employed in the case of the comitative imperative, where the nature of the goal is of no importance (Ricca 1993: 89):

(14)   *Pojd'* se mnou do ordinace.
          Come with me into the examining room.
          Ven conmigo a la consulta.

Now, while *přijd'/te* 'při-walk.IMP.SG/PL' cannot be used in comitative contexts, Hirschová (2007: 195) shows that imperatives of *při*-verbs can sometimes be used if the speaker invites the hearer to move towards him/her; however, the competition between *přijd'* [při-walk.IMP.SG] and *pojd'* [po-walk.IMP.SG] in such situations is not further analysed.

    As follows from this review, information about the distribution of Czech venitive verbs is sketchy at best and not based on a systematic analysis of authentic Czech data; this is what we will attempt in the rest of this paper, against the backdrop of English.

---

[7] Examples with no text identifiers are all taken from our bidirectional parallel corpus, i.e., from InterCorp.

3. Data and methods

The use of subtitles in linguistic research is still a matter of debate: 'if film dialogue is a reflection of real dialogue, subtitles are a reflection of a reflection' (Levshina 2017: 336); adherence to strict time and space constraints[8] leads to omissions and reformulations: 'film subtitles contain fewer pause fillers, reformulations and other discourse markers, which are typical of spontaneous discourse produced under real-time constraints' (Levshina 2017: 335), which makes them problematic 'for full-fledged conversational and discourse analyses as a replacement for spoken language' (Levshina 2017: 336). For our study, however, this is not an issue; we target motion verb predicates, i.e., items crucial to move the action forward, and we value the existence of a visual channel to consult if the concrete viewing arrangement is impossible to retrieve from the verbal context alone. We resorted to the Subtitles section of the multilingual parallel corpus InterCorp (version 15),[9] and created a bidirectional parallel corpus of English and Czech subtitles, following the English-Norwegian Parallel Corpus model (Johansson 2007: 11).

Czech being a small language, the size of the corpus had to reflect the size of the Czech source text subtitles aligned with English translations if the components were to be comparable in size. Films wrongly annotated as having Czech as a source language or occurring twice were excluded, and so were the musicals; the result was 448,999 tokens in 60 original Czech texts of subtitles. For the selection of English texts, we used alignment with several languages as a first filter and then manually selected texts with English as language of the original version. Included were films of various genres (404,704 tokens in 41 films) as well as TV series (96,241 tokens in 21 episodes from 20 different TV series).[10] The

---

[8] See Días-Cintas and Remael (2014) for an overview.

[9] The Subtitles section of InterCorp contains data from the database *Open Subtitles*. For more information about this section of InterCorp, see https://wiki.korpus.cz/doku.php/cnk:intercorp:verze15.

[10] According to Díaz-Cintas and Remael (2014: 23–24), the number of characters per one line of subtitles is lower for TV (35–37 characters) than for cinema (40–41, 43 at some film festivals), which reflects, among other factors, the fact that 'the viewer is able to read subtitles more easily and quickly on a cinema than on a television screen'. Changes in format (including that of a cinema version to a DVD) call for subtitle reductions also because 'the cinema illusion is based on a succession of 24 frames per second, whereas the small screen uses 25 frames per

result was a subcorpus of 500,945 tokens, slightly larger than the subcorpus of Czech originals, reflecting the typological difference between English and Czech.

In the next step we retrieved all tokens of *come* and intransitive motion verbs prefixed by *při-*, namely *přijít^P/přicházet^i* [při-walk], *přijet^P/přijíždět^i* [při-drive], *přiletět^P (přilítnout^P)* [při-fly], *přilézt^P* [při-crawl], *přiběhnout^P* [při-run], *připlout^P* [při-sail], *přiblížit^P se* [při-near; come closer] and *přistoupit^P* [při-step; stand by]. Literal motion, which this paper is targeting, was operationalized as self-agentive motion in the sense of Talmy (2000: 28); coding the data for (non)agentivity was the first step in the classification process.

Uses coded as non-agentive and therefore excluded from further analysis are heterogeneous in both languages: subject arguments are inanimate themes (more in English than in Czech), abstract entities, typically times conceptualized as moving objects (both in English and Czech); *come* can be an aspectual marker or it is a part of an idiomatic expression. Phrasal verbs with *come* were classified as agentive if they coded literal motion (e.g., *come in*) but as non-agentive if literal motion was not necessarily involved (*come from* in the sense of 'originate', 'date back to' or 'descend,' *come after* in the sense of 'be after sb'); this also includes intersubjective uses of the imperative *come on*, which in the encouraging sense and in orders to hurry typically correspond to Czech adverbs such as *honem* 'quickly', verbs such as *dělej* 'do.IMP.SG', and elliptic phrases such as *do toho!* 'for it'. If the activity is discouraged, Czech uses phrases such as *no tak* 'oh so', but also *ale jdi* 'but go.IMP.SG', i.e., the speaker asks the addressee (figuratively) to move away.

Non-agentive uses of *při*-verbs further include various idiomatic uses of *přijít^P*, which may (*přijít^P s plánem* 'come up with a plan', *přijít^P vhod* 'come in handy') but need not correspond to *come* (*přijít^P na to* 'find out', *přijít^P o všechno* 'lose everything', *přijít^P k něčemu* 'get something').

Each token of an agentive *při*-verb/*come* was then classified by the type of its correspondence as congruent, divergent, or zero (Johansson 2007: 35). Congruent correspondences were further classified by the concrete translation equivalent (venitive verb, other motion verb, non-

---

second in Europe and 30 in North America' and so 'the pace of the movie must be slightly sped up' (Díaz-Cintas and Remael 2014: 24). Again, given the nature of the items investigated here, we do not expect these differences to affect our analysis.

motion verb); Czech verbs in the translations were also coded for aspect and the presence/absence of a spatial prefix.

The next step was to calculate the mutual correspondence between agentive uses of *come* and *při*-verbs, and then to provide a systematic analysis of the correspondences, with a focus on cases where the venitive verb is avoided in the translation. To decide whether a translation by a verb other than a venitive verb is a result of a translator's choice or whether there is a constraint blocking the use of the venitive verb, we used our native-speaker intuitions; the availability of *come* in the English translations was checked with two native speakers of American English.

4. Results

Both agentive and non-agentive uses of *come* are more frequent than those of the Czech venitive verbs (Figure 1), and this difference is significant at p<0.0001, χ^2=34.5 for the agentive and 345.5 for non-agentive uses.[11] Agentive uses of *při*-verbs are significantly more common than their non-agentive uses (p<0.0001, χ^2=163.6); the difference between agentive and non-agentive uses of *come* is not statistically significant.
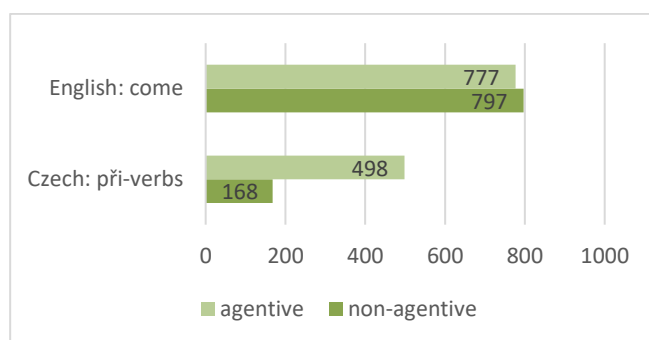

Figure 1: English and Czech venitive verbs in the originals

In their agentive uses (Table 2), the mutual correspondence between *come* and *při*-verbs is 45.2% with a clear translation bias: *come* is translated by *při*-verbs much less often (34.7%) than *při*-verbs by *come* (61.4%).


[11] We used the Corpus Calculator at https://www.korpus.cz/calc/.

Table 2: Mutual correpondence between *come* and *při*-verbs

Mutual correspondence	45.2%
come → *při*-verb	34.7% (270 tokens out of 777)
při-verb → *come*	61.4% (306 out of 498)

The relatively low frequency *při*-verbs in the translation of *come* points to the existence of restrictions on their use; these are analysed below.

4.1 Czech translations of come

Figure 2 shows that in its agentive use, *come* typically has congruent correspondences (89%), i.e., it is translated by verbs, most often indeed by *při*-verbs (35%).
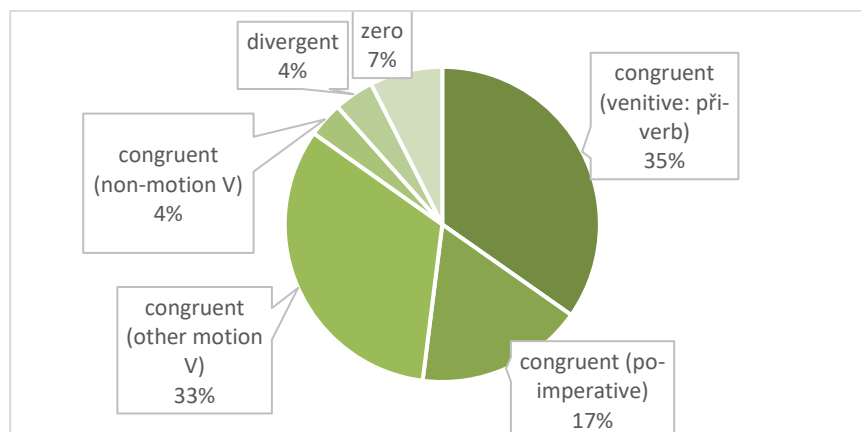
Figure 2: Czech correspondences of agentive *come*

However, *po*-imperatives, typically *pojď/te* 'po-walk.IMP.SG/PL', are also common (17%), and dominant among imperative forms; the imperatives of *při*-verbs are only attested in 12 tokens (*přijď/te* 'při-walk.IMP.SG/PL' and *přijeď* 'při-drive.IMP.SG'). A closer look at the imperatives shows that *při*-verbs do not translate *come* when motion is imminent and the speech participants are communicating face to face, i.e., they are conceptualized as sharing a space. This concerns invitations to *come in*, *come here* in (15), or when the hearer is asked to accompany the speaker (16).

(15)   Hey, *come* here, baby.
         Pojď ke mně, zlato.
         po-walk.IMP.SG towards me dear

(16)   *Come* with me.
         Pojď se mnou
         po-walk.IMP.SG with me

Translations suggest that imperatives of *při*-verbs are used if the speaker and hearer do not share a space at the moment of speech (in (17) they are speaking on the phone), or if the speaker invites the hearer to come back, i.e., it is clear that the hearer will leave the shared space for a while ((18) and (19)).[12]

(17)   *come* and pick me up.
         přijď pro mě.
         při-walk.IMP.SG for me

(18)   *Come* back in ten minutes.
         Přijďte za deset minut.
         při-walk.IMP.PL after ten minutes

(19)   (Now you go back up to school tomorrow,) *come* back
         home at Christmas break
         přijeď domů zase na Vánoce
         při-drive.IMP.SG home again on Christmas

The presence of a shared space and the imminence of motion also blocks the use of a *při*-verb in permission seeking acts such as the one in (20), their rejections, e.g., in (21), and in negative imperatives (22).

(20)   May we *come* in?
         Můžeme *jít* dál?
         can.1PL.PRS walk.INF.IPFV further

---

[12] The imperative form *přistup(te)* 'step.up/towards.IMP.SG/PL' was not found in the translations of *come* and will be discussed in section 3.2.

(21)   No, you can't *come* in.
            Ne, nemůžete *jít* dál.
            no, NEG.can.2PL.PRS walk.INF.IPFV further

(22)   Please, don't *come* in.
            Prosím ne*choď* sem!
            please NEG.walk.IMP.SG here.to

Při-verbs are typically blocked also in situations where motion is in progress or it is habitual. In such contexts, Czech needs an imperfective verb. There are two options: either to use a secondary imperfective form, i.e., one secondarily created for the perfective *při*-verb (e.g., *přicházet^i* for *přijít^p* 'při-walk'), or to resort to an inherently directional simplex imperfective. The data clearly show a preference for the latter; overall, there are only 17 tokens of *přicházet^i* (23), and one of *přijíždět^i* 'při-drive', but 117 tokens of the andative *jít^u/chodit^¬u* 'walk' (in (24) and (25)); and 23 of *jet^u/jezdit^¬u* 'drive':[13]

(23)   We *come* in peace.
            Přicházíme v míru.
            we při-walk.1PL.PRS.IPFV in peace

(24)   All right, I*'m coming*.
            No jo, už *jdu*.
            oh yes already walk.1SG.PRS.IPFV

(25)   That's why they *come* here.
            Proto sem *chodí*.
            for.this here.to walk.3SG.PRS.IPFV.¬U

Secondary imperfectives with *při-* are less common than simplex imperfectives not just because the latter are shorter, but mainly because

---

[13] Czech, like all West/East Slavic languages has a small group of verbs with two imperfective forms, distinguished by the presence/absence of unidirectionality; typically, the unidirectional forms denote on-going motion directed at a single, fixed goal (Mrhačová 1993: 39), while non-unidirectional forms describe repeated/habitual motion. In the data, unidirectional forms have the superscript *u*, non-unidirectional verbs *¬u*.

they construe the arrival as homogeneous, unbounded within the immediate temporal scope and expansible (see Langacker 2008), i.e., there is a focus on the mover getting closer to the goal and the way this is happening. In the situation depicted in (24), what matters is only the fact that the speaker is on his way, and so *přicházet^i* is out; in (25) it is ruled out since habituality is profiled.

The Czech andative verb is sometimes used instead of a *při*-verb also in situations which do not strictly block a construal by a perfective verb, but the translator finds the departure perspective to better suit the situation; this often happens in contexts where the speaker expresses their subjective stance. In (26) the speaker cannot leave everything behind and go see his sister; what matters in (27) is the woman's decision to go rather than her arrival at the destination:

(26) (A: Come home.) B: I can't *come* home right now. (I'm leaving town for a couple of days on business.)
Nemůžu teď *jít* domů.
NEG.can.1SG.PRS now walk.INF.IPFV home

(27) I'm really happy I *came* here.
Jsem ráda, že jsem sem *jela*.
am glad.F that AUX.1SG.PRS here.to drive.PTCP.IPFV.SG.F

In (28), however, the change of perspective involves a loss of semantic information: a man standing below a woman's window phones asking her to come to the window, i.e., to a goal where he (the speaker) is at coding time. His presence at the window is the crucial information he wants to convey, still, the Czech translator decides to conceal it (at least until the moment the women gets to the window and sees him). Instead of using *Přijď k oknu* 'při-walk.IMP.SG to window' or *Pojď k oknu* 'po-walk.IMP.SG to window', which in this clearly intersubjective context would both reveal the speaker's presence under the window in the same way that *come* does, the translator resorts to *běž* 'run.IMP.SG', which, though formally related to *běžet^i* 'run', functions here as an imperative form of the andative *jít^i*:[14]

---

[14] See Martinková (forthcoming) for more on imperatives of the verb *jít^i* 'walk'.

> (28)   Hey, *come* to the window!
>   Ahoj, *běž* k oknu.
>   hi, run.IMP.SG towards window

Simplex imperfectives (in Figure 3 marked as 'simplex Vipfv') attested in the correspondence of *come* include, apart from *jít*ᵘ/*chodit*⁻ᵘ 'walk', *jet*ᵘ/*jezdit*⁻ᵘ 'drive/ride', and *běžet*ᵘ 'run', also *lézt*ⁱ 'crawl' and *letět*ᵘ/*létat*⁻ᵘ 'fly'; overall, they account for 20% of the translations of agentive *come*.
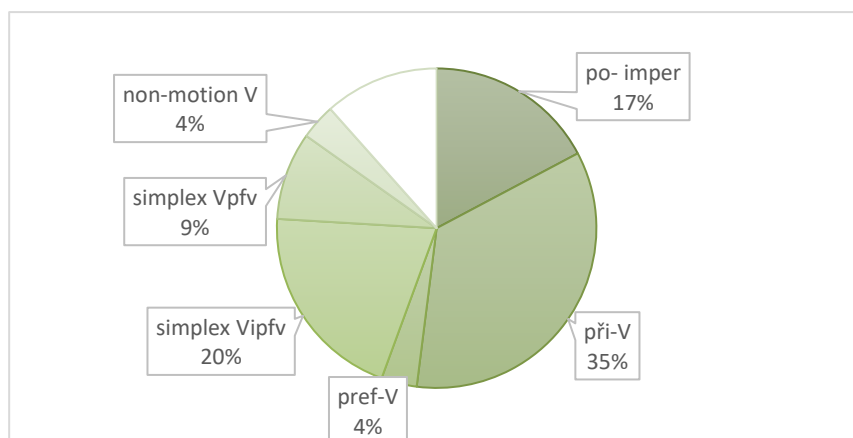

Figure 3: Congruent correspondences of *come* with respect to aspect and prefixation

Furthermore, as Figure 3 shows, translations of agentive *come* also include motion verbs with prefixes other than *při-*/*po-* ('pref-V'). Though these only account for 4% of the translations, they are worth reporting on as they demonstrate that Czech translators may choose to profile other aspects of the path.[15] Typically, this happens if *come* is followed by boundary-crossing satellites such as *out*, *in*, or *through*, whose meanings are then coded on the Czech prefixes. Example (29) illustrates exiting, i.e., a situation where motion implies a change from being invisible to being visible, and (30) entering, i.e., motion into a closed space; both encourage the use of *come* (Matsumoto et al. 2017), but less so of a *při*-verb. Instead,

---

[15] Only one can be profiled at a time, since Czech does not allow cumulating spatial prefixes on a single verb.

the Czech translation makes explicit the boundary crossing by means of the prefixes *vy-* 'out', as in (29), and *v(e)-* 'in', as in (30); (31) exemplifies profiling motion through the ground. Also found are verbs prefixed by *pře-* 'across', *před-* 'in front', and *za-*, literary meaning 'behind', but often used for short visits,[16] as in (32). The prefix *do-* 'to' only occurs once in the path verb *dorazit^p* 'reach'.

(29)   When I *came* out (there were a couple of lads taking some photos.)
         Když jsem vy*šel* ven
         when AUX.1.SG out-walk.PTCP.PFV.SG.M out

(30)   (take off those shoes) before you *come* in
         než ve*jdete* dovnitř.
         before in-walk.2PL.PRS.PFV inside

(31)   when the Sunset Limited *is coming* through at 80mph?
         když tudy pro*jíždí* Sunset Limited rychlostí 128km/h
         when here through-drive.3.SG.PRS.IPFV Sunset Limited speed.INS 128km/h

(32)   (I thought) it might not be a bad idea to *come* by and see how it was going.
         že by nebylo na škodu za*jít* sem se přesvědčit, jak to šlo.
         that   would   NEG.be.PTCP.SG.N   on   shame   behind-walk.INF.PFV   here.to   REFL   convince.INF   how   it   go.PTCP.SG.N

Finally, unprefixed perfective motion verbs in the translation (in Figure 3 marked as 'unprefixed Vpfv') include *stavit^p se* 'stop by' (used as a synonym of *zajít^p*) and *vrátit^p se* 'return'. Non-motion verbs in the translation are infrequent (28 tokens), typically *být^i* 'be' (12 tokens) and verbs of appearance (10 tokens).

---

[16] For the relevant sense of *zajít^p*, Kopečný (1962: 14) writes about the supplementary meaning of temporariness; *Slovník spisovné češtiny pro školu a veřejnost* defines it as 'to go somewhere for a purpose.'

3.2. English translations of při-*verbs*

A look at the English translations of Czech intransitive motion *při*-verbs brings a slightly different picture: 306 out of 498 *při*-verbs are translated by *come* (61.4%), 78 tokens by other motion verbs (16%), and 41 by non-motion verbs (8%). Divergent and zero correspondences represent 8% and 7%, respectively (Figure 4).
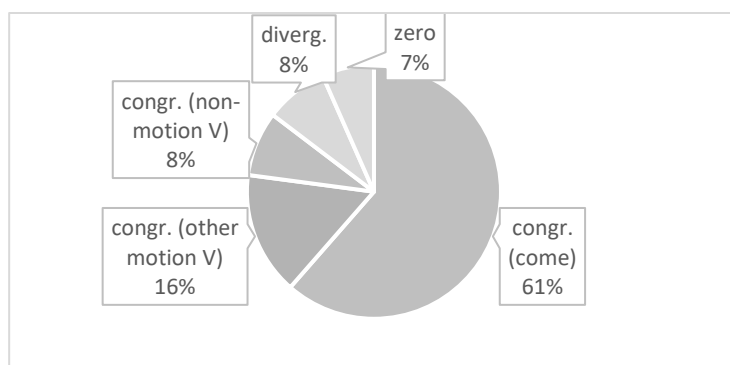

Figure 4: English correspondences of Czech intransitive agentive *při*-verbs

The high frequency of *come* in the translations of *při*-verbs suggests that these verbs are indeed typically used to describe motion to a deictic centre. Imperative forms always prompt the hearer to move in the direction of the speaker, and the verbs *přijď(te)* 'při-walk.IMP.SG/PL' (ten tokens) and *přijeď* 'při-drive.IMP.SG' (one token) are always translated by *come*; coming always involves an arrival from a place different from the speaker's location. Motion is imminent in just two tokens of the imperative form *přistup(te)* 'při-step.IMP.SG/PL'; stepping can hardly be conceptualized as an arrival, and the verb seems to allow deictic uses only when modified by the adverb *blíže* 'nearer'[17] and in rather formal contexts.

Crucially, contexts were found where *come* would not be felicitous in translation of *při*-verbs; according to our native-speaker informants, this is the case of (33), in which a nurse describes a story in which neither she

---

[17] We thank an anonymous reviewer for bringing this to our attention. What also seems relevant is that the stem of *přistoupit*ᵖ 'step up/forward' cannot function independently and fuses with the prefix to create a full-fledged manner/path verb. The verb is rather infrequent in the data: apart from the two imperative forms there are just two tokens in the indicative.

nor the addressee took part; *arrive* is the second most frequent verb after *come* (7.2%) in the translation of *při*-verbs:

(33)   A když *přijel* doctor
           and when při-drive.PTCP.PFV.SG.M doctor
           When the doctor *arrived*

In (34), a mother is talking to her son, who goes to his new job for the first time and she wishes him luck; clearly, she is not and will not be present at the goal of his motion. The verb in the translation is *be*, which is found in the correspondence of *při*-verbs in three cases (6.7%).

(34)   Tak už jdi, at' ne*přijdeš* pozdě!
           so already go.IMP.SG, so.that NEG-při-walk.3SG.PRS.PFV
           late
           Run along or you'll *be* late!

Our English native speakers also reject *come* in situations where the mover visits several places; this is what we see in (35), where the speaker—in response to the question *How do you spend your free time?*—describes his daily routine:

(35)   Ráno *přijdu* do Čáry
           morning při-walk.1SG.PRS.PFV to Čára
           (Quite normally.) In the morning I *go* to non-stop pics,
           (say from 9.30 until noon)

As in the translations of *come*, there are instances where the use of a verb other than *come* in the translation of a *při*-verb is simply the translator's choice. In (36), the speaker refers to a concert which the addressee went to the previous night; the two are talking over the phone. Though English allows the use of *come* for motion to the hearer's location at reference time, the translator opts for *go*:

(36)   Já jsem tam chtěl *přijít*, ale …
           I AUX.1SG.PRS there want.PTCP.SG.M při-walk.INF.PFV,
           but …
           I wanted to *go*, but (it slipped my mind).

5. Conclusions

Our contrastive analysis of agentive uses of English and Czech venitive verbs (*come* and *při*-verbs) in a bidirectional corpus of subtitles reveals a higher frequency of *come* than of *při*-verbs and a relatively low mutual correspondence; in their agentive uses, the verbs are translated by each other in only 45.2% of the cases. Furthermore, there is a translation bias: *při*-verbs are translated by *come* more often (61.4%) than *come* is translated by *při*-verbs (34.7%).

The high percentage of translations by *come* shows that in the majority of cases, *při*-verbs describe motion directed at speech act participants, i.e., they are used as deictic verbs. However, *při*-verbs also allow goals which are not deictic, and the corpus shows this by a lack of *come* in the correspondence of these verbs if motion to goals other than the speech participants is being described. The problem we had to face was that the translators sometimes used verbs other than *come* (*arrive, go* and others) even when motion was directed at the speech act participants. To tell whether *come* is absent in the translation due to a translator's choice, or because it is blocked, was difficult; 'there are no corpora of starred examples: a corpus cannot tell us what is not possible' (Fillmore 1992), and so we had to elicit acceptability judgements from native speakers of (American) English. Their judgements indicate that if the use of *come* is blocked, this typically happens in situations where *při*-verbs describe motion to non-deictic goals. We take this as evidence that *při*-verbs are not limited to deictic contexts.

Given that *při*-verbs are used with a wider range of goals than *come*, their low frequency in the translation of *come* is surprising and suggests the presence of other, strong factors constraining their use. The analysis shows that *při*-verbs often conceptualize the situation as one of arrival, i.e., the goal is viewed as a destination; this is in line with *arrive* being the second most frequent verb in the translation of *při*-verbs. If speakers ask the addressees to move in their direction, *při*-verbs are only used if there is a moment when the speech participants do not share a space. Otherwise, the deictic imperative form *pojď* 'po-walk.IMP.SG' is used (17% of translations of agentive *come*). *Při*-verbs are also avoided in comitative contexts; the imperative *pojď* is either used again, or, alternatively, a simplex imperfective verb is used. Czech simplex imperfectives account for 20% of all translations of agentive *come* and their analysis points to an important role of verbal aspect in the choice between a venitive and

andative verb in Czech: if motion is in progress, the situation must be construed as atelic and a Czech imperfective verb must be used. This constrains the use of *při*-verbs, which are typically perfective counterparts to the imperfective verbs from which they are derived. Rather than use a secondary imperfective, Czech translators often choose to use an inherently directional unprefixed imperfective. Finally, the data also indicate that functional factors such as a shift from invisible to visible and the presence of a bounded space do not trigger the use of *při*-verbs; in such situations, *při*-verbs give way to verbs with the boundary-crossing satellite prefixes such as *vy-* 'out' and *v(e)-* 'in'.

Although choosing the language of subtitles proved fruitful in many respects, there were also cases when the corpus let us down. *Při*-verbs often describe motion directed at speech act participants, but the corpus data cannot tell us whether in the clearly intersubjective contexts (given they can still be construed in an arrival perspective) *při*-verbs implicate the presence of the speaker at the goal of motion as strongly as we think they do. Also, as already mentioned, it was necessary to verify with native speakers of English whether *come* was permissible in the translations of *při*-verbs; especially problematic were cases where *při*-verbs described motion to salient non-deictic goals. This raises the question whether these cases can still be explained by what Fillmore calls the narrative use and what the exact conditions for deictic projection in these contexts are in English. Given the changes in our understanding of space in the current world, we believe that the use of *come*/*go* in these contexts should be revisited, expanded to include computer-mediated communication, and studied experimentally. Also important will be the study of the language of fiction, where the narrative uses of venitive verbs could potentially shed important light on the study of narrative situations and thus be a contribution to the field of corpus-based literary theory. All of this will ultimately be useful to English learners of Czech and Czech learners of English, who will need to learn to switch between the perspectives in order to avoid interference.

References

Altenberg, Bengt. 1999. Adverbial connectors in English and Swedish: Semantic and lexical correspondences. In *Out of corpora. Studies in honour of Stig Johansson*, edited by Hilde Hasselgård and Signe Oksefjell, 249–268. Amsterdam: Rodopi.

Biskup, Petr. 2019. *Prepositions, case and verbal prefixes: The case of Slavic*. Amsterdam: John Benjamins.

Čermák, František, and Alexander Rosen. 2012. The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics* 17(3): 411–427.

Daneš, František, Zdeněk Hlavsa, Anna Jirsová, Eva Macháčková, and Hana Prouzová. 1981. *Větné vzorce v češtině*. Praha: Academia.

Díaz-Cintas, Jorge, and Aline Remael. 2014. *Audiovisual Translation: Subtitling*. London and New York: Routledge.

Dickey, Stephen M. 2010. Common Slavic 'indeterminate' verbs of motion were really manner-of-motion verbs. In *New Approaches to Slavic Verbs of Motion*, edited by Victoria Hasko and Renee Perelmutter, 67–109. Amsterdam: John Benjamins.

Filipović, Luna. 2007. *Talking about motion. A cross-linguistic investigation of lexicalization patterns*. Amsterdam: John Benjamins.

Fernández-Couceiro, Eduardo. 2001. Errores de interferencia en la utilización de los verbos de movimiento. In *El hispanismo en la República Checa II*, edited by Demetrio Estébanez Calderón, 1–11. Praha: Univerzita Karlova v Praze, Filozofická fakulta.

Fillmore, Charles J. 1971. *Santa Cruz Lectures on Deixis*. Berkeley: University of California.

Fillmore, Charles J. 1992. 'Corpus linguistics' or 'Computer-aided armchair linguistics'. In *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4–8 August 1991*, edited by Jan Svartvik, 35–60. Berlin: De Gruyter.

Gathercole, Virginia. 1978. Towards a universal for deictic verbs of motion. *Kansas Working Papers in Linguistics* 3: 72–88.

Hijazo-Gascón, Alberto. 2017. Motion event contrasts in Romance languages: Deixis in Spanish as a second language. In *Motion and space across languages: Theory and applications*, edited by Iraide Ibarretxe-Antuñano, 301–327. Amsterdam: John Benjamins.

Hijazo-Gascón, Alberto. 2021. *Moving across language. Motion events in Spanish as a second language*. Berlin: De Gruyter.

Hirschová, Milada. 2007. Some remarks on spatial deixis in Czech. A semantic-pragmatic approach. *Sborník prací Filozofické fakulty brněnské univerzity. A, Řada jazykovědná* 56: 191–200.

Janda, Laura A., Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nesset, and Svetlana Sokolova. 2013. *Why*

Russian aspectual prefixes aren't empty: prefixes as verb classifiers. Bloomington, IN: Slavica Publishers.

Johansson, Stig. 2007. *Seeing through multilingual corpora. On the use of corpora in contrastive studies.* Amsterdam: John Benjamins.

Kopečný, František. 1962. *Slovesný vid v češtině.* Praha: Nakladatelství Československé Akademie věd.

Langacker, Ronald. 2008. *Cognitive grammar: A basic introduction.* London: Oxford University Press.

Lewandowski, Wojciech. 2014. Deictic verbs: Typology, thinking for speaking and SLA. *SKY Journal of Linguistics* 27: 43–65.

Levshina, Natalia. 2017. Online film subtitles as a corpus: an *n*-gram approach. *Corpora* 12(3): 311–338.

Lyons, John. 1977. *Semantics: Volume 2.* Cambridge: Cambridge University Press.

Malá, Markéta. 2015. Translation counterparts as indicators of the boundaries of units of meaning: A contrastive view of the position of "come V-ing" among the patterns of the verb *come*. In *Cross-linguistic perspectives on verb constructions*, edited by Signe Oksefjell Ebeling and Hilde Hasselgård, 169–191. Newcastle upon Tyne: Cambridge Scholars.

Martinková, Michaela. Forthcoming. *Czech and English motion verbs and constructions in satellite-framing: the case of 'going in'.* Olomouc: Univerzita Palackého.

Matsumoto, Yo, Kimi Akita, and Kiyoko Takahashi. 2017. The functional nature of deictic verbs and the coding patterns of deixis. In *Motion and space across languages: Theory and applications*, edited by Iraide Ibarretxe-Antuñano, 95–122. Amsterdam: John Benjamins.

Mrhačová, Eva. 1993. *Korelativní slovesa pohybu ve slovanských jazycích, zvláště v ruštině.* Ostrava: Sfinga.

Ricca, Davide. 1993. *I verbi deittici di movimento in Europa: una ricerca interlinguistica.* Pavia: La Nuova Italia.

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In *Relating events in narrative: Typological and contextual perspectives*, edited by Swen Strömqvist and Ludo Verhoeven, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.

Slovník spisovné češtiny pro školu a veřejnost (SSČ) [Dictionary of Standard Czech for schools and the public]. 2009. Praha: Academia.

Speed, Traci. 2015. Manner/path typology of Bulgarian motion verbs. *Journal of Slavic Linguistics* 23(1): 51–81.

Šlosar, Dušan. 1996. Slovotvorba. In *Příruční mluvnice češtiny*, edited by Petr Karlík, Marek Nekula, and Zdenka Rusínová, 109–225. Praha: Nakladatelství Lidové noviny.

Talmy, Leonard. 2000. *Toward a cognitive semantics: Typology and process in concept structuring.* Cambridge, MA: The MIT Press.

Verkerk, Annemarie. 2014. The evolutionary dynamics of motion event encoding. PhD dissertation, Max Planck Institute for Psycholinguistics.