Chapter 4

Periphrasis and morphosyntatic mismatch in Czech

Olivier Bonami

Université de Paris, Laboratoire de linguistique formelle, CNRS

Gert Webelhuth

Goethe-Universität Frankfurt a.M.

This paper presents an HPSG analysis of the Czech periphrastic past and conditional at the morphology-syntax interface. After clarifying the status of Czech auxiliaries as words rather than affixes, we discuss the fact that the past tense exemplifies the phenomenon of *zero periphrasis*, where a form of the main verb normally combined with an auxiliary can stand on its own in some paradigm cells. We argue that this is the periphrastic equivalent of zero exponence, and show how the phenomenon can be accommodated within a general theory of periphrasis, where periphrasis is a particular instance of a mismatch between morphology and syntax.

1 Introduction

The term "inflectional periphrasis" denotes a situation where a construction involving two or more words stands in paradigmatic opposition with a single word in the expression of a morphosyntactic contrast. The two Czech examples in (1) illustrate this: where the present indicative of ČEKAT is expressed by the single word *čekáme* in (1a), its past indicative is expressed by the combination of the two words *jsme* and *čekali*, as shown in (1b).

(1) a. Čekáme   na Jardu.
       wait.PRS.1PL for Jarda.ACC.SG
       'We are waiting for Jarda.'

b. Čekali   jsme   na Jardu.
  wait.LF.M.AN.PL be.PRS.1PL for Jarda.ACC.SG
  'We were waiting for Jarda.'

Traditional grammars of European languages treat inflectional periphrases as part of the inflectional paradigm. While this is intuitively satisfactory, capturing that intuition within contemporary lexicalist formal grammar has proven particularly elusive, for reasons outlined with great clarity by Matthews (1991: 219–220): a periphrase is "clearly two words, which obey separate syntactic rules (for example, of agreement). Nevertheless they are taken together as a term in what are otherwise morphological oppositions." Meeting the challenges raised by that observation has been the focus of much attention since the seminal work of Vincent & Börjars (1996) and Ackerman & Webelhuth (1998), including publications such as Sadler & Spencer (2001); Ackerman & Stump (2004); Stump & Hippisley (2011); Brown et al. (2012); Bonami & Webelhuth (2013); Popova & Spencer (2013); Stump & Finkel (2013); Dalrymple (2015); Bonami (2015).

The Czech past indicative presents an additional conceptual challenge for theories of periphrasis. While the expression of the past tense is periphrastic in general, it is not in the third person, where the same form of the main verb is used on its own (2). Two things are remarkable here: the fact that periphrasis is the default while synthesis is the special case, and the apparent finiteness mismatch between what looks like a nonfinite form of the main verb and the finite clause it presumably heads.

(2) Čekali   na Jardu.
     wait.LF.M.AN.PL for Jarda.ACC.SG
     'They were waiting for Jarda.'

The goal of this paper is to show that the approach to periphrasis developed in Bonami (2015) and Bonami et al. (2016) readily accounts for this situation, because it sees periphrasis as a special instance of a more general notion of morphosyntactic mismatch.

Section 2 presents the basic data. Section 3 shows that previous approaches to the Czech facts do not really address the challenges raised by the contrast between (1b) and (2). Section 4 presents the framework and shows how it can be deployed to account for the basic properties of the Czech past tense.

2 The data

2.1 The paradigmatics of the Czech present and past tenses

Table 4.1 shows the positive past subparadigm of a Czech verb.[1]

Table 4.1: Positive past subparadigm of ČEKAT 'wait'

	MAS.ANIM	MAS.INAN	FEM	NEU
1SG	čekal jsem	čekal jsem	čekala jsem	čekalo jsem
2SG	čekal jsi	čekal jsi	čekala jsi	čekalo jsi
3SG	čekal	čekal	čekala	čekalo
1PL	čekali jsme	čekaly jsme	čekaly jsme	čekala jsme
2PL	čekali jste	čekaly jste	čekaly jste	čekala jste
3PL	čekali	čekaly	čekaly	čekala

As can be inferred from the table, all forms of the Czech past tense are based on a form we will call the *l*-form, ending in the suffix *-l*. While it is the historical descendant of a participle, the *l*-form is used only in the formation of the Czech past indicative, and present and past conditional. There are no nonfinite clauses headed by the *l*-participle; passive is also periphrastic, but relies on a different passive participle, as shown by the contrast in (3).

(3) a. Koupil   jsem   knihy.
       buy.LF.M.SG be.PRS.1SG book.ACC.PL
       'I bought books.'
    b. Knihy   byly   koupeny.
       book.NOM.PL be.LF.F.PL buy.PASS.F.PL
       'Books were bought.'

Hence it is misleading to call that form a participle from a synchronic point of view. In addition, when used without an accompanying auxiliary, the *l*-form is the sole exponent of the past. This motivates the fact that traditional grammar calls it the "past form". This term is again a bit misleading, since the *l*-form is

---

[1]For simplicity we do not include polite plural forms such as *čekal jste*, which implement a number mismatch between the main verb and the participle. See https://www.czechency.org/slovnik/VYKÁNÍ (Karlík et al. 2016). These can be integrated straightforwardly in the analysis below by refining the mapping between HEAD and INFL values.

also used in the construction of the conditional periphrases, where it is clearly not an exponent of the past, as we will see below. We will keep on using the morphosyntactically opaque label "*l*-form" and gloss it as "LF".

The *l*-form systematically agrees in gender and number with the subject. Note that in the plural, differences between masculine animate on the one hand, and feminine and masculine inanimate on the other hand, is purely orthographic, as sequences ⟨ly⟩ and ⟨li⟩ note the same phonemic sequence /lɪ/.[2]

In the first and second person past, the *l*-form is obligatorily accompanied by an auxiliary, which we will call the past indicative auxiliary. That auxiliary is homophonous with a present indicative form of the copula BÝT, and exhibits agreement in person and number with the subject. In the third person, by contrast, the *l*-form obligatorily occurs on its own. Despite the existence in Czech of third person forms of the auxiliary, adding such a form to an example such as (2) leads to ungrammaticality.

(4)   * Čekali   jsou.
        wait.LF.M.AN.PL be.3PL

It is worth stressing that, unlike some other Slavic languages, Czech requires the overt presence of a copula in copular constructions in all persons (5). Hence omission of the auxiliary in the past indicative is specific to that (periphrastic) construction.

(5)   a.   Děti   jsou   rády.
            child.NOM.PL be.PRS.3PL happy.F.NOM.PL
            'The kids are happy'
      b. * Děti   rády.
            child.NOM.PL happy.F.PL

Finally, the *l*-form is the locus of expression of negation in the periphrastic past: while negation is expressed as a prefix on the only verb in synthetic forms such as the present or third person past, it is obligatorily expressed on the main verb, and cannot be expressed on the auxiliary, in the periphrastic first and second person.

(6)   Nečekáme   na Jardu.
      NEG.wait.PRS.1PL for Jarda.ACC.SG
      'We are not waiting for Jarda.'

---

[2]In Czech orthography, ⟨y⟩ and ⟨i⟩ note the same short vowel /ɪ/, while ⟨ý⟩ and ⟨í⟩ note the same long vowel /iː/. The ⟨i⟩ vs. ⟨y⟩ contrast indicates presence vs. absence of palatalization for the preceding consonant, for those consonants that are subject to palatalization, which /l/ is not.

(7) Nečekali na Jardu.
    NEG.wait.LF.M.AN.PL for Jarda.ACC.SG
    'They were not waiting for Jarda.'

(8) a. Nečekali jsme na Jardu.
       NEG.wait.LF.M.AN.PL be.PRS.1PL for Jarda.ACC.SG
       'We were not waiting for Jarda.'

    b. * Čekali nejsme na Jardu.

    c. * Nečekali nejsme na Jardu.

We now turn to a brief description of the conditional. As Table 4.2 illustrates, the present conditional is formed by combining a finite form of the conditional auxiliary BY[3] and the *l*-form. As in the past indicative, the auxiliary agrees in person and number, and the *l*-form in number and gender, with the subject. Also as in the past indicative, negation is expressed on the *l*-form.

(9) Nečekali bychom na Jardu.
    NEG.wait.LF.M.AN.PL COND.1PL for Jarda.ACC.SG
    'We would not wait for Jarda.'

Unlike what happens in the past indicative, a form of the auxiliary is obligatorily present in the third person. Hence an *l*-form not accompanied by an auxiliary is unambiguously a past indicative third person form.

Table 4.2: Present conditional subparadigm of ČEKAT 'wait'

	MAS.ANIM	MAS.INAN	FEM	NEU
1SG	čekal bych	čekal bych	čekala bych	čekalo bych
2SG	čekal bys	čekal bys	čekala bys	čekalo bys
3SG	čekal by	čekal by	čekala by	čekalo by
1PL	čekali bychom	čekaly bychom	čekaly bychom	čekala bychom
2PL	čekali byste	čekaly byste	čekaly byste	čekala byste
3PL	čekali by	čekaly by	čekaly by	čekala by

Two older periphrases further illustrate the contrast between the past and the conditional auxiliary.[4] Table 4.3 illustrates the past conditional. This combines

---

[3]The conditional auxiliary is historically a form of the copula BÝT, but is never used as an independent synthetic verb form in contemporary Czech.

[4]We are indebted to Alexandr Rosen for pointing out the relevance of the pluperfect here, and to Olga Nádvorníková for helping us clarify the synchronic status of these periphrases.

the conditional auxiliary with a form homophonous to the *l*-form of the copula, and the *l*-form of the main verb. As expected, the conditional auxiliary agrees in person and number, and both *l*-forms agree in number and gender, with the subject. By contrast, Table 4.4 illustrates the (indicative) pluperfect. In the first and second person, this combines the past auxiliary with a form homophonous to the *l*-form of the copula, and the *l*-form of the main verb. In the third person, just as in the simple past, there is no finite form of the auxiliary, and the apparent *l*-form of the copula is the only auxiliary element.

Table 4.3: Past conditional subparadigm of ČEKAT 'wait'

	MAS.ANIM	MAS.INAN	FEM	NEU
1SG	byl bych čekal	byl bych čekal	byla bych čekala	bylo bych čekalo
2SG	byl bys čekal	byl bys čekal	byla bys čekala	bylo bys čekalo
3SG	byl by čekal	byl by čekal	byla by čekala	bylo by čekalo
1PL	byli bychom čekali	byly bychom čekaly	byly bychom čekaly	byla bychom čekala
2PL	byli byste čekali	byly byste čekaly	byly byste čekaly	byla byste čekala
3PL	byli by čekali	byly by čekaly	byly by čekaly	byla by čekala

Table 4.4: Pluperfect subparadigm of ČEKAT 'wait'

	MAS.ANIM	MAS.INAN	FEM	NEU
1SG	byl jsem čekal	byl jsem čekal	byla jsem čekala	bylo jsem čekalo
2SG	byl jsi čekal	byl jsi čekal	byla jsi čekala	bylo jsi čekalo
3SG	byl čekal	byl čekal	byla čekala	bylo čekalo
1PL	byli jsme čekali	byly jsme čekaly	byly jsme čekaly	byla jsme čekala
2PL	byli jste čekali	byly jste čekaly	byly jste čekaly	byla jste čekala
3PL	byli čekali	byly čekaly	byly čekaly	byla čekala

A possible analysis of the constructions illustrated in Tables 4.3 and 4.4 posits the existence of a general past periphrase combining a finite or *l*-form past auxiliary with an *l*-form of the main verb. Under such an analysis, the past indicative (Table 4.1) relies solely on the general past periphrase with a finite auxiliary; the past conditional (Table 4.3) combines the conditional periphrase with the general past periphrase, using the *l*-form of the past auxiliary; and the pluperfect (Table 4.4) applies the general past periphrase recursively, with both a finite and an *l*-form auxiliary. Note that the *l*-form of the past auxiliary, unlike its finite forms, is not a clitic (see below), and is not dropped in the third person.

More work on earlier stages of the language would be needed to substantiate the feasibility of such compositional analyses of complex periphrases. For present purposes, corpus searches confirm that the forms in Tables 4.3 and 4.4 are clearly no longer in use, and we will not attempt to analyze them further.

To sum up this section, the past indicative contrasts with the present indicative in relying on an *l*-form for the main verb; it contrasts with conditional subparadigms in (i) the use of the past auxiliary in the first and second person, and (ii) the absence of an auxiliary in the third person. Such a distribution can be seen as the periphrastic equivalent of the familiar situation of "zero exponence". Consider the present subparadigm in Table 4.5. Here, the 3SG form contrasts with all other forms in the absence of a suffix following the vowel *-á-*. In the same way, in Table 4.1, the past indicative third person forms contrast with their first and second person equivalents in the absence of a past auxiliary; they likewise contrast with their conditional correspondents in the absence of one or two auxiliaries. By analogy with zero exponence, we will call this phenomenon *zero periphrasis*.

Table 4.5: Present indicative subparadigm of čEKAT 'wait'

	SG	PL
1	čekám	čekáme
2	čekáš	čekáte
3	čeká	čekají

One main goal of the present paper is to account for zero periphrasis in Czech. Before doing so, however, we need to discuss the morphosyntactic status of auxiliaries in this system.

2.2 The morphosyntactic status of the past auxiliary

The Czech past auxiliary is standardly described as a clitic, on the basis of the fact that it is systematically prosodically dependent on an adjacent word. In this context, within a lexicalist framework, it is crucial to establish whether this pretheoretical clitic status is to be analyzed by seeing the auxiliary as a prosodically deficient word, or "true clitic", or as some kind of phrasal affix, inserted by morphology on a word at the edge of some syntactically-defined constituent. In this section we review the evidence on the status of Czech clitics, and draw relevant consequences for the analysis of the past indicative periphrase. We rely mainly on the extensive discussion in Hana (2007), and ignore many complications.

Czech possesses a family of second position clitics. These form a rigidly ordered cluster that cannot be interrupted by any intervening material and consists of the following elements, in the indicated order:[5]

(10)   a.   Past or conditional auxiliary

       b.   Reflexive *se* (reflexivization of direct object) and *si* (reflexivization of indirect object)

       c.   Dative weak pronouns

       d.   Accusative weak pronouns

       e.   Genitive weak pronouns

       f.   Demonstrative *to*

There can be some amount of morphological fusion within the cluster. In particular, the sequence of a 2SG past auxiliary and a reflexive is fused to a portmanteau form, as indicated in (11). In addition to an organization in rigid position classes, this provides limited evidence for the view that the elements in the clitic cluster belong to a single syntactic word, and that the combination of the clitics is governed by morphology rather than syntax.

(11)   a.   jsi   se   > ses
          PST.2SG REFL.ACC   REFL.ACC.PST.2SG

       b.   jsi   si   > sis
          PST.2SG REFL.DAT   REFL.DAT.PST.2SG

In finite clauses, the clitic cluster linearizes after the first major constituent.[6] In most cases, the cluster attaches prosodically to that preceding constituent, as shown in example (12).

(12)   Koupil   =jsem   =je   pro Jardu.
       buy.LF[M.SG] =be.PRS.1SG =ACC.PL for Jarda.ACC.SG
       'I bought them for Jarda.'

However, as discussed by Toman (1996), the clitic cluster attaches to the following, rather than to the preceding constituent whenever a prosodic break needs to

---

[5]We leave aside adverbial clitics such as *už* 'already' and subtler aspects of the distribution of pronominal clitics. Note that *se* and *si* can also be used as part of so-called "inherent reflexive verbs", where they have no referential value and hence no true reflexive function.

[6]See Hana (2007: 98–114) for discussion of situations where the clitic follows what is pretheoretically a partial constituent or a sequence of constituents.

be present after the first constituent. This happens if a parenthetical, e.g. a non-restrictive relative clause modifies that first constituent, or if the first constituent is a clause. Toman's examples are given in (13–14); (15–16) provide parallel examples involving the past auxiliary.

(13) a. Knihy,   které tady vidíte, se=   dnes platí
         book.NOM.PL which here see-2PL REFL= today pay.PRS.3PL
         zlatem.
         gold.INS.SG
         'The books you can see here are paid for with gold today.'
    b. * Knihy, které tady vidíte=se dnes platí zlatem.

(14) a. Poslouchat =ji, by=   ji= asi   nudilo.
         listen   =her would= her= probably bore.
         'It would perhaps bore her$_i$ to listen to her$_j$.'
    b. * Poslouchat =ji =by =ji asi nudilo.
    c. * Poslouchat ji= by= ji= asi nudilo.

(15) Tu   knihu,   která   se mi   moc
     DEM.ACC.SG.F book(F).ACC.SG which.NOM.F.SG REFL 1SG.DAT much
     líbila,   jsem=   koupil   v Praze.
     like.PST.3SG.F be.PRS.1SG= buy.LF.M.SG in Prague.(F).LOC.SG
     'This book, which I like very much, I bought in Prague.'

(16) Že   tam bude,   jsem=   ne-věděl.
     COMP there be.FUT.3SG be.PRS.1SG= NEG-know.LF.M.SG
     'I did not know he would be there.'

Toman's observations provide a strong argument against a phrasal affixation analysis of Czech clitics: if clitics are affixes attached by morphology to the first constituent in the clause, it is predicted that they are always attached to that constituent, as morphology does not normally peer into syntax to decide where affixes should be attached; special mechanisms would need to be introduced to deal with examples (13–16), eliminating much of the appeal of a morphological analysis. On the other hand, this data is compatible with the view according to which clitics are just prosodically deficient words, occurring in a fixed syntactic position and attaching to the preceding or following constituent depending on prosodic properties of the context.

Pointing in the same direction is the observation by Rosen (2001: 210) that some lexical items, including the copula, can be clitic or nonclitic depending on

context. That the copula does not need to be clitic is evident from the fact that it can form a full utterance of its own, as a short answer to a question (second utterance in 17), and can occur in first (first utterance in 17) or third (18) position, unlike, e.g. the past auxiliary.[7] However, that it *can* be a clitic is evident from examples such as (19), cited by Rosen from the Czech National Corpus, where the copula occurs between the first constituent and a pronominal clitic. Since pronominal clitics obligatorily belong to the clitic cluster and the clitic cluster needs to be in second position, the copula has to also be part of the cluster in this and similar examples.

(17) A: Jsou   děti   takové?
          be.PRS.3PL child.NOM.PL such.F.NOM.PL
          'Are kids like that?'

      B: Jsou.   'They are.'
          be.PRS.3PL

(18) Děti   takové   jsou.
      child.NOM.PL such.F.NOM.PL be.PRS.3PL
      'Kids are like that.'

(19) Jedinou   radostí   =jsou   mu   dopisy   z
      only.INS.FEM.SG joy.INS.SG   =be.3PL DAT.MAS.SG letter.NOM.PL from
      domova,   [...]
      home.GEN.SG
      'His only pleasure is the letters from home, [...]'

These facts strongly suggest that Czech clitics are words. If they were affixes, we would need two entirely separate mechanisms to generate the form *jsou*: a lexical entry in (17–18), a rule of morphology in (19). If on the other hand they are words, we just need to assume that clitichood is a property that can be underspecified: some words (e.g. most verbs, strong pronouns) are nonclitics, some (e.g. past and conditional auxiliaries, weak pronouns) are clitics, and some (e.g. finite forms of the copula) can be either.

We thus conclude that Czech clitics cannot be affixes. What remains unresolved at this point is whether each clitic should be considered a separate word, or whether the clitic cluster as a whole should be considered a word. We provided limited evidence for the latter view. However, because such a view raises

---

[7]The past auxiliary can occur clause-initially in questions in colloquial or "Common" Czech, but not in the more formal variety of "Standard" Czech (Hana 2007: 70). In this paper we ignore the complexities of Common Czech.

many issues for a lexicalist formal grammar, and because these issues are largely orthogonal to the analysis of periphrasis, we will not attempt to substantiate it. In the remainder of this paper we thus focus on cases where the only clitic in the clause is the auxiliary, in which case it has to constitute a word. We leave the proper treatment of the clitic cluster for future research.

3 Previous approaches

To the extent that previous approaches to the Czech past indicative within lexicalist formal grammars address the phenomenon of zero periphrasis, they rely on a reductionist approach based on zero auxiliaries.

The most explicit relevant analysis is that of Hana (2007), who assumes a phonologically empty auxiliary (p. 153). Hana takes the past auxiliary to raise all arguments of the *l*-form and combine in a flat structure. This leads to the parallel analyses in Figure 4.1, where sentences in the past indicative first and third person have exactly parallel structures.
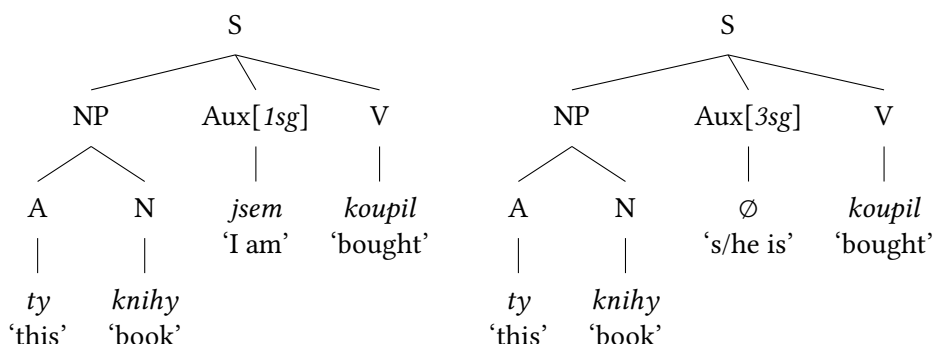

Figure 4.1: Czech auxiliaries according to Hana (2007)

While this is clearly a defendable analysis, it is subject to all the usual arguments against syntactic zero elements (Sag & Fodor 1994; Sag & Wasow 2011). In addition, from the point of view of inflectional morphology, it suffers from the same conceptual defect as all analyses relying on zero morphemes (Matthews 1991; Anderson 1992; Stump 2001; Blevins 2016): instead of modelling directly the fact that Czech grammar efficiently uses the contrast between presence and absence of an auxiliary to encode a morphosyntactic distinction, it treats that situation as a kind of defect of the system, which misleads the analyst (and, presumably, the speaker) into believing that there is nothing where in fact there is

something. Just as in synthetic morphology, it is conceptually more satisfactory to address the descriptive generalization directly.

A different take on the system is proposed by Tseng & Kupść (2007) in the context of a general discussion of Slavic past and conditional auxiliaries. In Polish, there is strong evidence that tense auxiliaries are phrasal affixes. To account for that situation, Kupść & Tseng (2005) propose an analysis along the lines shown in Figure 4.2.
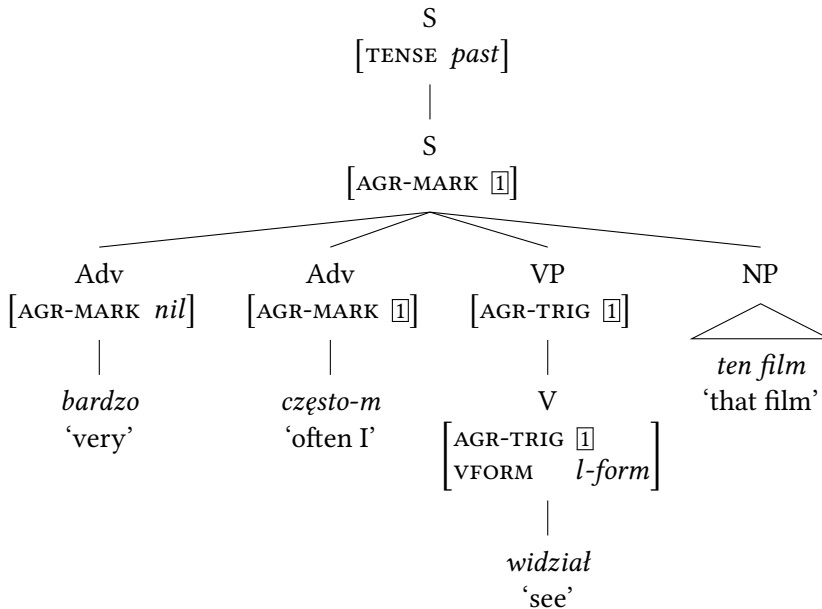

Figure 4.2: Polish auxiliaries as phrasal affixes (adapted from Tseng & Kupść 2007: 269).

The workings of the analysis rely on the two features AGR-TRIG and AGR-MARK. AGR-TRIG is a head feature which transmits the requirement for an agreement marker upwards from the main verb along the head path. At the clause level, the value of that feature is matched with that of the initial constituent's AGR-MARK feature. AGR-MARK itself is a (right) EDGE feature, which transmits information down to the right edge of the relevant subtree to the rightmost word in that tree. At the word level, the value of that feature is interpreted by inflectional morphology, and possibly realized as an affix.

In Polish as in Czech, no form of the auxiliary is used in the third person. Among other desirable features, the analysis in Kupść & Tseng (2005) reduces

this situation of zero periphrasis to a case of zero synthetic exponence: as suggested in Figure 4.3, the syntactic analysis is exactly the same in the third person; it just happens that inflectional morphology provides no exponent for the expression of [ARG-MARK *3sg*].

S
[TENSE *past*]

AdvP
[AGR-MARK [1]]

VP
[AGR-TRIG [1]]

Adv
[AGR-MARK *nil*]

Adv
[AGR-MARK [1]]

V
[AGR-TRIG [1] *3sg*
VFORM   *l-form*]

NP

bardzo
'very'

często
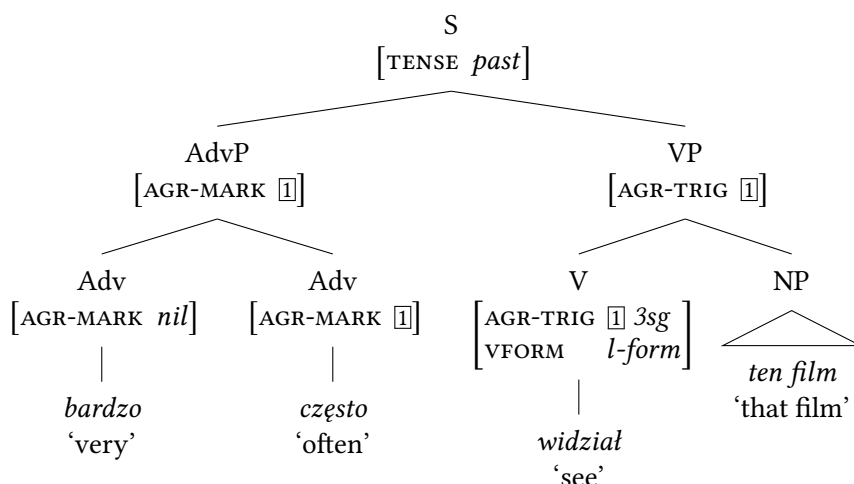'often'

widział
'see'

ten film
'that film'

Figure 4.3: Zero periphrasis in Polish (adapted from Tseng & Kupść 2007: 269).

Tseng & Kupść (2007) suggest that the very same analysis proposed for Polish can be redeployed for Czech. Such an option is untenable, for the reasons we discuss in Section 2.2. Tseng (2009) is aware of this, and provides an extremely rough sketch of an analysis where the Czech copula is a clitic, in the form of the tree reproduced in Figure 4.4. While this tree gives a few hints as to what Tseng has in mind for the first and second person past indicative, with the auxiliary being an adjunct or marker attached to the initial constituent, it is entirely unclear how such an analysis will deal with zero periphrasis, unless a phonologically empty marker is postulated in the third person.

Finally, Petkevič et al. (2015) present a very careful HPSG approach to the formation of past and conditional periphrases in Czech, relying in particular on the idea that, in addition to their individual inflectional category, the auxiliary (called the *surface head*) and the main verb (called the *deep head*) jointly contribute to the construction of an analytic category. There are many similarities between this and Bonami's 2015 use of a distinction between HEAD and INFL features discussed below. However, Petkevič et al.'s approach says nothing on zero periphrasis: the principle regulating the distribution of tense and mood values
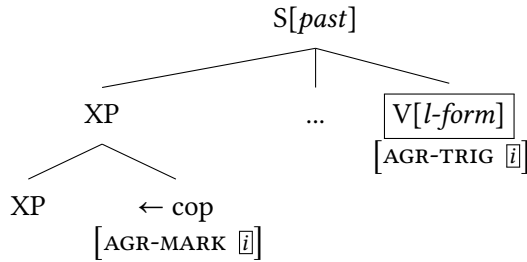
Figure 4.4: Tseng's 2009 sketch of an analysis of the Czech copula

in a periphrase is dependent on the presence of an auxiliary surface head in the syntax. According to Alexandr Rosen (p.c.), the treebank annotation scheme that the paper reports on resolves the issue by positing a third-person past tense form that is homophonous with the *l*-form, but not explicitly related to it. Hence such an approach implicitly treats the similarity between the forms in the first and second person on the one hand and third person on the other hand as synchronically accidental.

We thus conclude that previous literature on Czech and Slavic languages in HPSG and neighboring approaches provides no means of addressing the phenomenon of zero periphrasis.

4 Periphrasis as syntactic exponence

4.1 Main assumptions

In this subsection we outline the general approach to periphrasis that we will rely on in the remainder of this paper, building heavily on Bonami & Webelhuth (2013), Bonami (2015), and Bonami et al. (2016). This relies on three main ideas. First, we adopt an inferential-realizational approach to inflection (Matthews 1972; Zwicky 1985; Anderson 1992; Aronoff 1994; Stump 2001), where inflection and syntax are strictly separated, and the inflectional component deduces the phonological form of words jointly from the lexeme's lexical entry and the morphosyntactic description provided by syntax for that word in the context of a particular utterance. Crysmann & Bonami (2016) and Bonami & Crysmann (2016) present a detailed inferential-realizational approach to inflection within HPSG that is entirely compatible with the proposals discussed here. However, since we will not be discussing matters of synthetic exponence in detail, for present purposes we can simply see inflection as a function f that deduces a phonological form from a SYNSEM object, as indicated in Figure 4.5.

$$word \rightarrow \begin{bmatrix} \text{PHON} & \textbf{\textit{f}}(\boxed{1}) \\ \text{SYNSEM} & \boxed{1} \end{bmatrix}$$

Figure 4.5: Inflection as a function from syntax and semantics to phonology (preliminary version)

Second, we follow Ackerman & Webelhuth (1998); Sadler & Spencer (2001); Ackerman & Stump (2004) in assuming that periphrastic inflection can be seen as an alternative to ordinary (synthetic) inflection, where the combination of the main verb with an auxiliary serves as the exponent of a set of morphosyntactic properties, in the same way as the combination of a stem with an affix may serve as an exponent.

Third, our theory of periphrasis builds on the view that morphosyntactic mismatches in general require a distinction between paradigmatic oppositions as defined by syntax and semantics and their implementation in morphology: although in the canonical situation, the same distinctions made by syntax and semantics are used in morphology, there are various types of situations where morphology makes fewer (syncretism, neutralization), more (overabundance), or different (morphomic distributions, deponency) contrasts than syntax and semantics. This general idea is known under different names in the literature, with important technical differences that do not concern us here directly: Sadler & Spencer (2001) use two disjoint sets of *syntactic* and *morphological* features; Ackerman & Stump (2004) and Stump (2006; 2016) contrast *content paradigms* and *form paradigms*; Bonami & Samvelian (2015) oppose HPSG's SYNSEM attribute, collecting features relevant to syntax and semantics to the exclusion of phonology, to a distinct MORSYN attribute that collects those features that happen to be relevant to inflection. Finally, Bonami (2015); Bonami et al. (2016) make the simplifying assumption that syntactic and semantic contrasts relevant to inflection are coded as HPSG HEAD features, and hence contrast the value of the HEAD feature with that of the INFL feature, which is the direct input to inflection. In this paper we will adopt this final approach, which is sufficient for our purposes.

4.2 Modelling morphosyntactic mismatch

Under such an approach, the input to inflection is the INFL value, which will be identical to the HEAD value in the canonical situation, but may differ from it in grammatically specified ways in particular cases. This proposal is outlined in Figure 4.6, where the dotted line represents the syntax-morphology interface: in simple cases, $\boxed{1}$ and $\boxed{2}$ will be equal, but the grammar will allow for (constrained) mismatches between the two values.

$$word \rightarrow \begin{bmatrix} \text{PHON} & f(\boxed{1}) \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{INFL} & \boxed{1} \end{bmatrix}$$
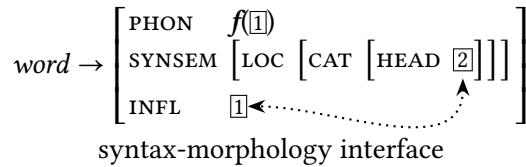
syntax-morphology interface

Figure 4.6: Inflection as a function from syntax and semantics to phonology (final version)

A crucial ingredient of such an approach, then, is a way of licensing limited deviations from identity between HEAD and INFL at the syntax-morphology interface. To this end, Bonami et al. (2016) propose that the grammar contain a set of dedicated interface implicational statements whose antecedent can mention any feature under WORD and whose consequent consists of specifications of feature values within INFL and/or reentrancies between HEAD and INFL. The statement in Figure 4.7 captures the default situation of an absence of mismatch: in the absence of any further specification, HEAD and INFL coincide.[8]

$$[] \Rightarrow \begin{bmatrix} \text{S|L|C|HEAD} & \boxed{1} \\ \text{INFL} & \boxed{1} \end{bmatrix}$$
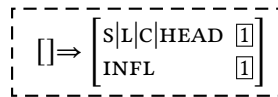
Figure 4.7: Interface statement: default identity between HEAD and INFL

This statement is sufficient to license the correct form in most situations. In particular it is the relevant statement for present forms of the verb in Czech, and contributes to licensing the analysis of the simple sentence in Figure 4.8.

Here we make some explicit assumptions about the feature geometry necessary to capture Czech inflection. As in Sag (2012) and related literature, the feature LID captures lexemic identity – all forms of a lexeme share the same LID value, and no two lexemes have identical LID values. For simplicity we first limit ourselves to the present and past indicative and the *l*-form, see Section 4.4 for an extension to the conditional. This simple subsystem can easily be captured using a single feature VFORM with possible values *l-form, prs, pst*. Our approach can trivially be generalized to the rest of the paradigm using a more elaborate feature geometry. The feature POL governs the inflectional realization of negation as the expression of its *neg* value. Finally, we assume that both finite and nonfinite forms of verbs have a full-fledged AGR value, with gender, number and person features. Implicit here is the hypothesis that rules of morphological exponence encapsulated in the function f relating INFL to HEAD will take care of the fact

---

[8]We display interface statements in dashed boxes, in order to highlight their distinguished status in the grammar. S|L|C abbreviates SYNSEM|LOCAL|CAT.

S $\begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix}$

$$\begin{bmatrix} \text{HEAD} & \boxed{1} & \begin{bmatrix} \textit{verb} \\ \text{LID} & \textit{kupovat-lid} \\ \text{VFORM} & \textit{prs} \\ \text{AGR} & \textit{m.1sg} \\ \text{POL} & \textit{pos} \end{bmatrix} \\ \text{INFL} & \boxed{1} & \begin{bmatrix} \text{LID} & \textit{kupovat-lid} \\ \text{VFORM} & \textit{prs} \\ \text{AGR} & \textit{m.1sg} \\ \text{POL} & \textit{pos} \end{bmatrix} \end{bmatrix}$$ N

kupuju   *knihy*
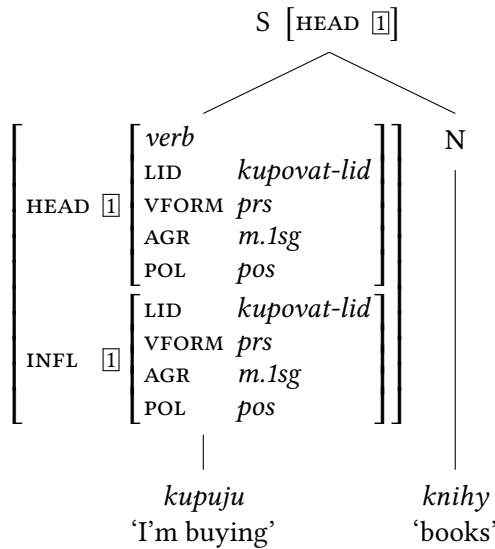'I'm buying'   'books'

Figure 4.8: Analysis of a simple Czech clause in the present tense

that *l*-forms neutralize person distinctions, while finite forms neutralize gender distinctions. Within an inferential-realizational view of inflection (Stump 2001), this simply amounts to having no rule realizing the neutralized category; see Zwicky (1986) for discussion and motivation. An obvious alternative would be to capture neutralizations in the feature system, by complicating the relationship between HEAD and INFL: under such a view, finite and nonfinite forms would have different features under INFL|AGR. Since the two solutions make the same empirical prediction, we adopt the simpler formulation based on morphology proper rather than the morphology-syntax interface.

While the interface statement in Figure 4.8 captures simple cases such as the present, extra statements are necessary to deal with situations of mismatch. For instance, we assume the statement in Figure 4.9 to account for the Czech third person past. What we want to capture here is the fact that the word *čekali* in a sentence such as (2) expresses the past third plural through a form that is not inherently a past form (e.g. it is used in the present conditional) nor a third person form (it is also part of the expression of first and second person plural past). To this end, the statement contrasts the value of VFORM under HEAD with the value of VFORM under INFL: in essence, this states that, to express the past third person, one uses an *l*-form. All other feature values are constrained to be identical under HEAD and INFL. This ensures that the verb will be appropriately inflected for (positive or negative) polarity and for number and gender.

$$
\left[\text{S|L|C|HEAD} \begin{bmatrix} \textit{verb} \\ \text{AGR} \quad \begin{bmatrix} \text{PER} \ \textit{3} \end{bmatrix} \\ \text{VFORM} \ \textit{pst} \end{bmatrix} \right] \Rightarrow \left[\begin{array}{l} \text{S|L|C|HEAD} \begin{bmatrix} \text{LID} \quad \boxed{2} \\ \text{AGR} \quad \boxed{3} \\ \text{POL} \quad \boxed{4} \end{bmatrix} \\ \\ \text{INFL} \begin{bmatrix} \text{LID} \quad \boxed{2} \\ \text{VFORM} \ \textit{l-form} \\ \text{AGR} \quad \boxed{3} \\ \text{POL} \quad \boxed{4} \end{bmatrix} \end{array} \right]
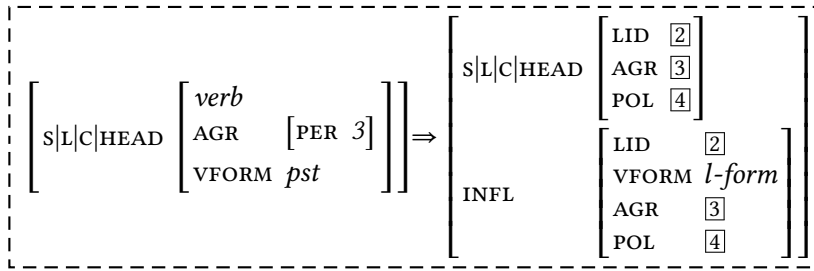$$

Figure 4.9: Interface statement: Third person past indicative

This statement thus licenses forms such as *koupil* in the sentence whose analysis is depicted in Figure 4.10.

$$
\text{S} \ \begin{bmatrix} \text{HEAD} \ \boxed{1} \end{bmatrix}
$$

$$
\left[\begin{array}{l} \text{HEAD} \ \boxed{1} \begin{bmatrix} \textit{verb} \\ \text{LID} \quad \boxed{2} \ \textit{koupit-lid} \\ \text{VFORM} \ \textit{pst} \\ \text{AGR} \quad \boxed{3} \ \textit{m.3sg} \\ \text{POL} \quad \boxed{4} \ \textit{pos} \end{bmatrix} \\ \\ \text{INFL} \begin{bmatrix} \text{LID} \quad \boxed{2} \ \textit{koupit-lid} \\ \text{VFORM} \ \textit{l-form} \\ \text{AGR} \quad \boxed{3} \ \textit{m.3sg} \\ \text{POL} \quad \boxed{4} \ \textit{pos} \end{bmatrix} \end{array} \right] \qquad \text{N}
$$

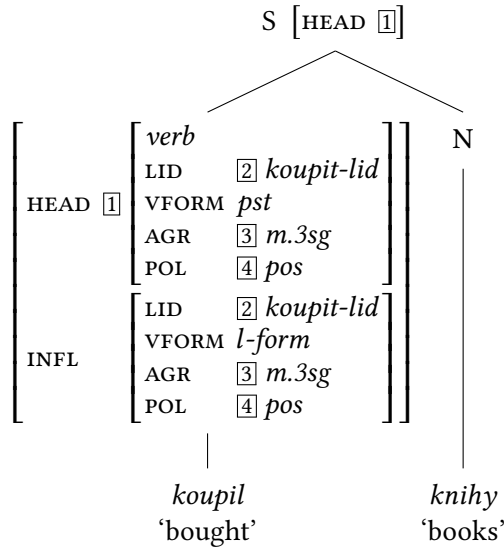|	
koupil	*knihy*
'bought'	'books'

Figure 4.10: Analysis of a simple Czech clause in the third person past

We now have all the ingredients in place to turn to the analysis of periphrastic forms. Figure 4.11 exhibits the lexical entry of the Czech past auxiliary, which embodies a number of assumptions. Following Hana (2007) and Petkevič et al. (2015), we assume that Czech auxiliaries are (surface) heads and raise the arguments of the main verb: both the subject $\boxed{s}$ and the list of non-subject arguments $\boxed{L}$ are raised from the main verb to the auxiliary's ARG-ST list. Following Bonami (2015), we assume that auxiliaries in general have unusual lexical identity. From the point of view of HEAD, they inherit the lexical identity of the main verb,

which they project to phrase level, e.g. for purposes of selection. But from the point of view of inflection, they have their own properties that distinguish them from the main verb. This again can be captured by making use of the HEAD vs. INFL distinction, applied now to the LID feature: note the sharing of LID value ② between the auxiliary's HEAD and that of its *l*-form complement. Finally, the lexical entry also enforces the sharing of HEAD|AGR and HEAD|POL values between auxiliary and main verb, ensuring appropriate inflection on the *l*-form.

$$
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \text{LID} & ② \\ \text{VFORM} & \textit{pst} \\ \text{AGR} & ③ \\ \text{POL} & ④ \end{bmatrix} \\
\text{INFL} & \begin{bmatrix} \text{LID} & \textit{past-aux-lid} \end{bmatrix} \\
\text{ARG-ST} & \left\langle \boxed{s}, \begin{bmatrix} \text{ARG-ST} & \langle \boxed{s} \rangle \oplus \boxed{L} \\ \text{HEAD} & \begin{bmatrix} \text{LID} & ② \\ \text{VFORM} & \textit{l-form} \\ \text{AGR} & ③ \\ \text{POL} & ④ \end{bmatrix} \end{bmatrix} \right\rangle \oplus \boxed{L}
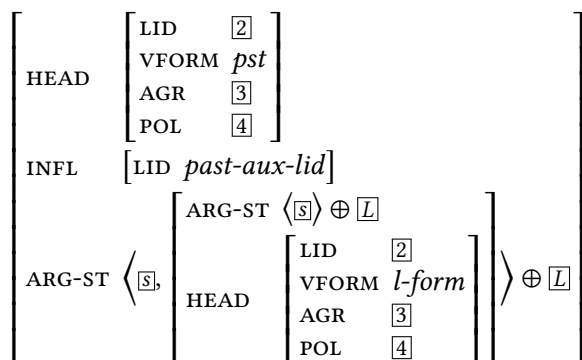\end{bmatrix}
$$

Figure 4.11: Lexical entry for the past auxiliary

It is important to note that neither of the previously stated syntax-morphology interface statements can apply to the auxiliary. The auxiliary is incompatible with both the default statement in Figure 4.7, and the more specific statement in Figure 4.9, since both enforce identity of HEAD|LID and INFL|LID. Thus a third statement, given in Figure 4.12, is necessary.

$$
\begin{bmatrix} \text{S|L|C|HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} & \textit{pst} \end{bmatrix} \end{bmatrix} \Rightarrow
\begin{bmatrix} \text{S|L|C|HEAD} & \begin{bmatrix} \text{AGR} & ③ \end{bmatrix} \\ \text{INFL} & \begin{bmatrix} \text{LID} & \textit{pst-aux-lid} \\ \text{VFORM} & \textit{prs} \\ \text{AGR} & ③ \\ \text{POL} & \textit{pos} \end{bmatrix} \end{bmatrix}
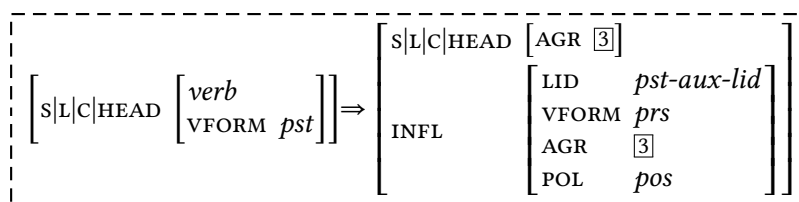$$

Figure 4.12: Interface statement: Third person past indicative

This states that, to inflect a verb in the past, one should use a word form that is the realization of the past auxiliary in the present tense, not inflected for polarity (whether the HEAD|POL value is positive or negative), and with appropriate person and number exponence. The fact that both the lexical entry in Figure 4.11 and the interface statement in Figure 4.12 refer to the INFL|LID value *pst-aux-lid*

ensures that the use of the auxiliary is obligatory to express the past, and that the auxiliary can be used only in the expression of the past (as the only interface statement licensing the use of that auxiliary is restricted to the past).

Figure 4.13 illustrates how the lexical entry for the auxiliary and the interface statement jointly license appropriate analyses for first or second person past indicative sentences. We purposefully choose a negative sentence to highlight the flow of information.
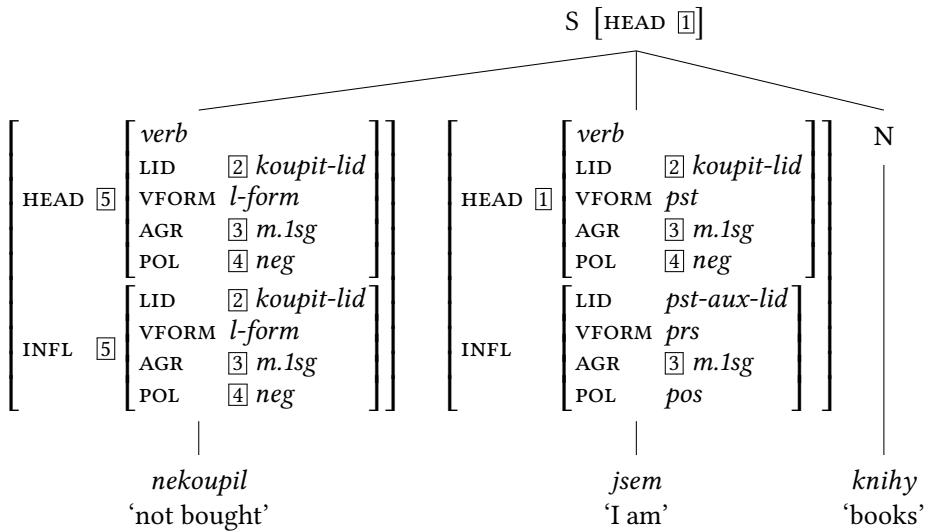
Figure 4.13: Analysis of a simple negative Czech clause in the non-third person past

It is useful to reflect on similarities and differences between the analyses of canonical synthetic inflection (Figure 4.8), mismatching synthetic inflection (Figure 4.10), and periphrastic inflection (Figure 4.13). In all three cases, the head word's HEAD specification is the locus of information relevant to syntax and semantics that gets projected to the phrasal level for purposes of selection and semantic composition. Synthetic and periphrastic past forms have in common a discrepancy between the head word's HEAD specification and its INFL specification, with direct consequences for morphophonology. Thus they both instantiate morphosyntactic mismatch on the head word. What sets the first and second person past apart is the fact that exponence of the phrase's HEAD specification is distributed (Ackerman & Stump 2004) over two words: the main verb realizes polarity, gender and number, the auxiliary realizes person and number, and the

combination of the two, as specified in the auxiliary's lexical entry, holistically realizes tense.

Note that, unlike the auxiliary, the main verb in this construction instantiates canonical morphosyntax: *nekoupil* is an *l*-form of the main verb, both from the point of view of HEAD (i.e., syntax) and INFL (i.e., morphology). This is in contrast with the use of the same word form in the third person past, where an [INFL|VFORM *l-form*] is used as the realization of [INFL|VFORM *pst*].

4.3 Paradigmatic competition

One remaining issue that has not been dealt with is paradigmatic competition between the three inflection strategies at hand: canonical synthetic inflection cannot be used in the past, periphrastic inflection in the past cannot be used in the third person. One way of dealing with this issue would be to add negative stipulations in various places so as to ensure that the three strategies are in complementary distribution. We contend that this is not a satisfactory approach, as it fails to capture the inherently paradigmatic competition between inflection strategies, and the fact that the same types of arbitration mechanisms regulating the choice of a synthetic exponent also regulate the choice between synthesis and periphrasis (Bonami 2015). In the case at hand, specificity seems to be at play: synthesis is the default, preempted by the more specific periphrastic past, which is itself preempted in the third person by the most specific third person past.

To capture this, we follow Stump (2006) in assuming that Pāṇini's principle is active at the syntax-morphology interface, and regulates the use of the most specific inflection strategy wherever more than one strategy is available. Crysmann & Bonami (2016) present an HPSG-compatible formalization of Pāṇini's principle for synthetic inflection defined as a closure operation on the descriptions of rules of exponence. In a nutshell, this assumes that each rule of exponence is a pairing of a description of a morphosyntactic context and an exponence strategy. The closure operation consists in identifying, for each rule R, the set of rules $S = R_1, \ldots, R_n$ whose morphosyntactic context is less specific than that of R, and to strengthen R's morphosyntactic context by the conjunction of the negations of the contexts of all rules in S. Bonami et al. (2016) propose to extend that general modelling strategy to the syntax-morphology interface, through the use of interface statements such as those in Figures 4.7, 4.9, and 4.12. Specifically, they propose the following. The syntax-morphology interface takes the form of a set of conditional statements $S = \{A_1 \Rightarrow C_1, \ldots, A_n \Rightarrow C_n\}$. For each statement $A_i \Rightarrow C_i$, we first find the set of $S_i = \{A_i^1 \Rightarrow C_i^1, \ldots, A_i^k \Rightarrow C_i^k\} \subset S$ of statements

whose antecedent is strictly more specific than A_i. Then each A_i is strengthened with the conjunction of the negations of all A_i^j. As a result, $A_i \Rightarrow C_i$ is replaced by $(A_i \wedge \neg A_i^1 \wedge \cdots \wedge \neg A_i^k) \Rightarrow C_i$, which is mutually exclusive with all the more specific statements in S.

$$
\begin{pmatrix}
[\,] \wedge \\
\neg \begin{bmatrix} \text{HEAD} & [\text{VFORM} \;\; pst] \end{bmatrix} \wedge \\
\neg \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & pst \\ \text{AGR} & 3rd \end{bmatrix} \end{bmatrix}
\end{pmatrix}
\Rightarrow
\begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{INFL} & \boxed{1} \end{bmatrix}
$$

$$
\begin{pmatrix}
\begin{bmatrix} \text{HEAD} & [\text{VFORM} \;\; pst] \end{bmatrix} \wedge \\
\neg \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & pst \\ \text{AGR} & 3rd \end{bmatrix} \end{bmatrix}
\end{pmatrix}
\Rightarrow
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \text{AGR} & \boxed{3} \end{bmatrix} \\
\text{INFL} & \begin{bmatrix} \text{LID} & past\text{-}aux\text{-}lid \\ \text{VFORM} & prs \\ \text{AGR} & \boxed{3} \\ \text{POL} & pos \end{bmatrix}
\end{bmatrix}
$$

$$
\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & pst \\ \text{AGR} & 3rd \end{bmatrix} \end{bmatrix}
\Rightarrow
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \text{LID} & \boxed{2} \\ \text{AGR} & \boxed{3} \\ \text{POL} & \boxed{4} \end{bmatrix} \\
\text{INFL} & \begin{bmatrix} \text{LID} & \boxed{2} \\ \text{VFORM} & l\text{-}form \\ \text{AGR} & \boxed{3} \\ \text{POL} & \boxed{4} \end{bmatrix}
\end{bmatrix}
$$

Figure 4.14: Literal effects of Pāṇinian strengthening

Figure 4.14 shows the literal effects of this process of Pāṇinian strengthening on the set of three interface statements presented respectively in Figures 4.7, 4.9 and 4.12. Figure 4.15 shows equivalent, more readable descriptions. As the reader can check, the net effect of the application of Pāṇini's principle is to end up with appropriately mutually exclusive statements in a principled, rather than stipulative, manner.

We have thus now presented a complete account of the interplay between synthesis and periphrasis in Czech indicative tenses. Crucially for our purposes, this account directly captures the phenomenon of zero periphrasis. First, synthetic and periphrastic past forms have much in common: both are instances of noncanonical morphosyntax, and contrast in this with, e.g. present forms; both rely on an *l*-form of the lexeme being inflected to realize the past. Second, they contrast precisely in that an *l*-form on its own expresses third person, while in combination with an auxiliary it will express first or second person; the use of an auxiliary in the third person is blocked by the existence of a more specific

$$\left[\text{HEAD} \; \neg \left[\text{VFORM} \; pst\right]\right] \Rightarrow \left[\begin{array}{ll} \text{HEAD} & \boxed{1} \\ \text{INFL} & \boxed{1} \end{array}\right]$$

$$\left[\text{HEAD} \left[\begin{array}{ll} \text{VFORM} & pst \\ \text{AGR} & \neg 3rd \end{array}\right]\right] \Rightarrow \left[\begin{array}{ll} \text{HEAD} & \left[\text{AGR} \; \boxed{3}\right] \\ \\ \text{INFL} & \left[\begin{array}{ll} \text{LID} & \textit{past-aux-lid} \\ \text{VFORM} & prs \\ \text{AGR} & \boxed{3} \\ \text{POL} & pos \end{array}\right] \end{array}\right]$$

$$\left[\text{HEAD} \left[\begin{array}{ll} \text{VFORM} & pst \\ \text{AGR} & 3rd \end{array}\right]\right] \Rightarrow \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \text{LID} & \boxed{2} \\ \text{AGR} & \boxed{3} \\ \text{POL} & \boxed{4} \end{array}\right] \\ \\ \text{INFL} & \left[\begin{array}{ll} \text{LID} & \boxed{2} \\ \text{VFORM} & \textit{l-form} \\ \text{AGR} & \boxed{3} \\ \text{POL} & \boxed{4} \end{array}\right] \end{array}\right]$$

Figure 4.15: Simplified effects of Pāṇinian strengthening

strategy. There is no necessity to postulate that the auxiliary is defective, since its third person forms will never be required. This opens the door to capturing the common inflectional makeup between the past auxiliary and the copula by saying that they are distinct lexemes sharing the same PARADIGM IDENTIFIER (Bonami & Crysmann 2018).

4.4 Towards an analysis of the conditional

Having presented an analysis of the Czech past indicative at the morphology-syntax interface, in this final section we briefly present the challenges posed by the analysis of the conditional.

Remember from Section 2 that the Czech conditional comes in two tenses: the present (20a) relies on a finite auxiliary combined with the *l*-form of the main verb, while the past (20b) combines the finite conditional auxiliary also found in the present, a second element identical to the *l*-form of the copula, and the *l*-form of the main verb.

(20) a. Olga   by   koupila   knihy.
          Olga.NOM.SG cond[3SG] buy.LF.F.SG book.ACC.PL
          'Olga would buy books.'

    b. Olga   by   byla   koupila   knihy.
         Olga.ɴᴏᴍ.sɢ cond[3sɢ] be.ʟꜰ.ꜰ.sɢ buy.ʟꜰ.ꜰ.sɢ book.ᴀᴄᴄ.ᴘʟ
         'Olga would have bought books.'

Our analysis extends readily to the present conditional: just adding to the grammar the lexical entry for the conditional auxiliary in Figure 4.16 and the interface statement in Figure 4.17 will license analyses such as that shown in Figure 4.18.

Things are significantly more challenging, both conceptually and technically, for the past conditional. Looking at the examples in (20), it is very tempting to see the past conditional as the compositional combination of two periphrases, one for the expression of the conditional (shared with the present conditional) and one for the expression of the past (shared with the past indicative). Obviously, such an analysis would require modifying the geometry of inflection features to separate expression of tense from that of mood, but that poses no difficulty.

The two real challenges are the following. First, whereas in the indicative, there is no past auxiliary in the present, the past auxiliary is obligatorily realized in the past conditional. Here our general line of analysis provides an appropriate analytic tool: since the third person past indicative requires a dedicated interface statement anyway (see Figure 4.9), we can make that statement specific to indicative mood, while generalizing the statement licensing the past auxiliary (see Figure 4.12) to both moods.

Second, there is a complication with the expression of negation. Both in the past indicative and in the present conditional, negation can only be expressed on the *l*-form, as shown in (21–22). In the past conditional, however, expression of negation is variable, and can occur on either of the two *l*-forms (23), but not both.

(21) a. Nekoupil   jsem   knihy.
         ɴᴇɢ.buy.ʟꜰ[ᴍ.sɢ] be.ᴘʀs.1sɢ book.ᴀᴄᴄ.ᴘʟ
         'I didn't buy books.'
    b. * Koupil nejsem knihy.

(22) a. Nekoupil   bych   knihy.
         ɴᴇɢ.buy.ʟꜰ[ᴍ.sɢ] cond.ᴘʀs.1sɢ book.ᴀᴄᴄ.ᴘʟ
         'I would not buy books.'
    b. * Koupil nebych knihy.

(23) a. Byl   bych   nekoupil   knihy.
         be.ʟꜰ[ᴍ.sɢ] cond.ᴘʀs.1sɢ ɴᴇɢ.buy.ʟꜰ[ᴍ.sɢ] book.ᴀᴄᴄ.ᴘʟ
         'I would not have bought books.'

$$
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \text{LID} & \boxed{1} \\ \text{VFORM} & cond \\ \text{AGR} & \boxed{3} \\ \text{POL} & \boxed{4} \end{bmatrix} \\
\text{INFL} & \begin{bmatrix} \text{LID} & cond\text{-}aux\text{-}lid \end{bmatrix} \\
\text{ARG-ST} & \left\langle \boxed{5}, \begin{bmatrix} \text{ARG-ST} & \langle \boxed{5} \rangle \oplus \boxed{L} \\ \text{HEAD} & \begin{bmatrix} \text{LID} & \boxed{1} \\ \text{VFORM} & l\text{-}form \\ \text{AGE} & \boxed{3} \\ \text{POL} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \oplus \boxed{L}
\end{bmatrix}
$$

Figure 4.16: Lexical entry for the conditional auxiliary

$$
\begin{bmatrix} \text{S|L|C|HEAD} & \begin{bmatrix} verb \\ \text{VFORM} & cond \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{S|L|C|HEAD} & \begin{bmatrix} \text{AGR} & \boxed{3} \end{bmatrix} \\ \text{INFL} & \begin{bmatrix} \text{LID} & cond\text{-}aux\text{-}lid \\ \text{VFORM} & l\text{-}form \\ \text{AGR} & \boxed{3} \\ \text{POL} & pos \end{bmatrix} \end{bmatrix}
$$
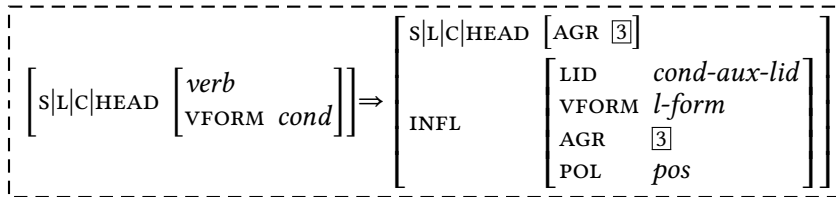
Figure 4.17: Interface statement for the present conditional

$$
\text{S} \begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix}
$$

$$
\begin{bmatrix}
\text{HEAD} \ \boxed{5} & \begin{bmatrix} verb \\ \text{LID} & \boxed{2}\ koupit\text{-}lid \\ \text{VFORM} & l\text{-}form \\ \text{AGR} & \boxed{3}\ m.1sg \\ \text{POL} & \boxed{4}\ pos \end{bmatrix} \\
\text{INFL} \ \boxed{5} & \begin{bmatrix} \text{LID} & \boxed{2}\ koupit\text{-}lid \\ \text{VFORM} & l\text{-}form \\ \text{AGR} & \boxed{3}\ m.1sg \\ \text{POL} & \boxed{4}\ pos \end{bmatrix}
\end{bmatrix}
\quad
\begin{bmatrix}
\text{HEAD} \ \boxed{1} & \begin{bmatrix} verb \\ \text{LID} & \boxed{2}\ koupit\text{-}lid \\ \text{VFORM} & cond \\ \text{AGR} & \boxed{3}\ m.1sg \\ \text{POL} & \boxed{4}\ pos \end{bmatrix} \\
\text{INFL} & \begin{bmatrix} \text{LID} & cond\text{-}aux\text{-}lid \\ \text{VFORM} & prs \\ \text{AGR} & \boxed{3}\ m.1sg \\ \text{POL} & pos \end{bmatrix}
\end{bmatrix}
\quad \text{N}
$$

koupil	*bych*	*knihy*
'bought'	'I could'	'books'

Figure 4.18: Analysis of a simple Czech clause in the present conditional

b. Nebyl bych koupil knihy.
 NEG.be.LF[M.SG] cond.PRS.1SG NEG.buy.LF[M.SG] book.ACC.PL
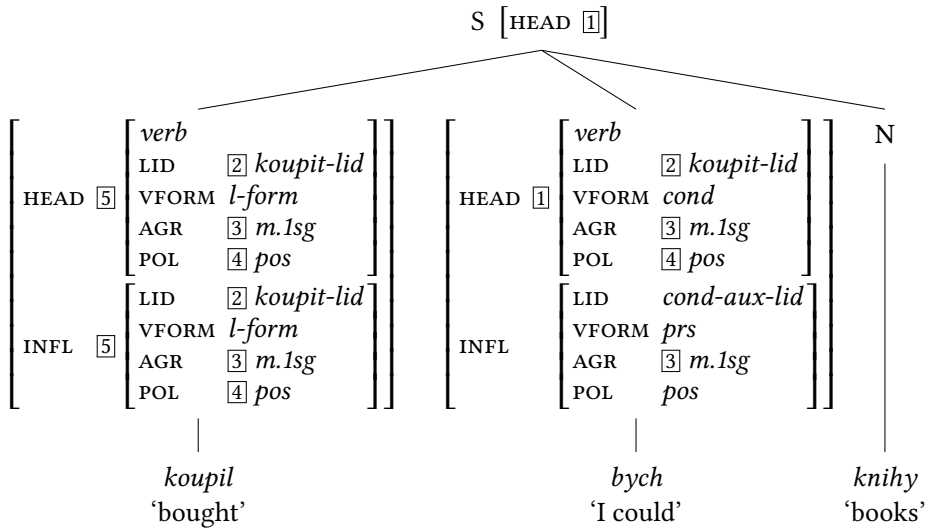 'I would not have bought books.'

c. * nebyl bych nekoupil knihy.

Relevant evidence suggests that both variants in (23) are equally grammatical. Our informants have no consistent preference for one variant over the other, which is unsurprising, given that the past conditional is rarely used in contemporary usage, and felt as archaic. Searches in the Czech National Corpus reported in Table 4.6 suggest that expression of negation on the past auxiliary is preferred when it occurs before the (second position) conditional auxiliary, but that there is no such preference in the opposite order.

Table 4.6: Counts of occurrences of negative conditional forms consisting of three adjacent verbs in the SYN v6 Corpus (Hnátková et al. 2014)

	past > cond.	cond. > past
NEG on past auxiliary	433	372
NEG on main verb	32	307

The existence of such overabundance (Thornton 2012) in the expression of negation presents a significant challenge for the compositional analysis of the past conditional: given what we observe in the past indicative and present conditional, a compositional analysis predicts that negation should be expressible on the main verb only. Evidence from negation thus suggests a holistic analysis of the past conditional periphrase, whereby a single rule of periphrasis licenses a combination of three words, with a dedicated flow of morphosyntactic information. While this is technically feasible, given the vanishing use of this form in contemporary Czech, it might also be defendable that speakers do not have coherent usage, and that two separate competing analyses should be posited. Obviously, more empirical research on the past conditional, its usage in historical stages of the language where it was still frequent, and the conditions of its decay, is necessary to decide which line of analysis is more satisfactory.[9]

[9]In particular, one would want to know more about the historical development of current properties of the past conditional. An appealing scenario would be that the past conditional started out as a more well-behaved combination of periphrases, and over time acquired autonomous properties, such as the unexpected realization of negation on the auxiliary. Future research will have to establish whether that is empirically accurate.

5 Conclusions

Our recent research on periphrasis has emphasized properties that periphrases share, on the one hand, with ordinary syntactic constructions, and on the other hand, with ordinary (synthetic) inflection. In connection with syntax, Bonami & Webelhuth (2013) and Bonami & Samvelian (2015) emphasize the fact that periphrasis builds on the constructional resources available in the language under consideration. In connection with inflection, Bonami (2015) showed that arbitration between synthetic and paradigmatic realization follows the same logic of paradigmatic opposition well documented for arbitration between synthetic strategies; Štichauer (2018) expanded this argument by exhibiting interesting cases of paradigmatic opposition among periphrastic strategies.

In this paper we expanded the set of parallels between synthetic and periphrastic inflection by attending to the phenomenon that we have called "zero periphrasis", by analogy with "zero exponence": this is the situation where the absence of an auxiliary combining with the main lexeme serves as the expression of some morphosyntactic feature. The Czech third person past tense provided a particularly clear example of a phenomenon that is also attested in other languages – see for example Stump & Hippisley (2011) on the past tense in Pamirian languages, or Stump (2013) on the future tense in Sanskrit. To model the phenomenon, we relied on the analytic devices deployed by Bonami et al. (2016) in the analysis of Welsh pseudo-finite constructions. Crucial to the analysis is the observation that ordinary periphrasis is a kind of morphosyntactic mismatch, but not the only possible kind of such a mismatch: another possibility, exemplified in Welsh by the verbs heading *bod* clauses, is that a morphologically nonfinite form of a verb heads a syntactically finite clause. Our analysis of zero periphrasis in Czech is essentially the same: the (finite) third person past is solely realized by a (nonfinite) *l*-form. What is different from the Welsh situation is the fact that the synthetic third person past contrasts with the periphrastic non-third person past. Our analysis states that the same form of the main verb (as expressed by having the same INFL value) can play double duty as the single expression of the past in the third person and as part of a periphrastic expression of the past in the first and second person; this directly captures the nature of zero periphrasis, without any need to postulate empty auxiliaries or other ontologically disputable entities.

Abbreviations

ACC	accusative		M, MAS	masculine
AN, ANIM	animate		N, NEU	neuter
COMP	complementizer		NEG	negative
COND	conditional		NOM	nominative
DAT	dative		PASS	passive
DEM	demonstrative		PL	plural
F,FEM	feminine		POS	positive
GEN	genitive		PRS	present
INAN	inanimate		PST	past
INS	instrumental		REFL	reflexive
LF	*l*-form (or *l*-participle)		SG	singular
LOC	locative			

Acknowledgements

References

Ackerman, Farrell & Gregory T. Stump. 2004. Paradigms and periphrastic expression. In Louisa Sadler & Andrew Spencer (eds.), *Projecting morphology*, 111–157. Stanford: CSLI Publications.

Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates*. Stanford: CSLI Publications.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. 1994. *Morphology by itself*. Cambridge: MIT Press.

Blevins, James P. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press.

Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25. 63–110.

Bonami, Olivier, Robert D. Borsley & Maggie Tallerman. 2016. On pseudo-non-finite clauses in Welsh. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 104–124.

Bonami, Olivier & Berthold Crysmann. 2016. The role of morphology in constraint-based lexicalist grammars. In Andrew Hippisley & Gregory T. Stump (eds.), *Cambridge handbook of morphology*, 609–656. Cambridge: Cambridge University Press.

Bonami, Olivier & Berthold Crysmann. 2018. Lexeme and flexeme in a formal theory of grammar. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology* (Empirically Oriented Theoretical Morphology and Syntax), 175–202. Berlin: Language Science Press. DOI: 10.5281/zenodo.1402520.

Bonami, Olivier & Pollet Samvelian. 2015. The diversity of inflectional periphrasis in Persian. *Journal of Linguistics* 51(2). 327–382.

Bonami, Olivier & Gert Webelhuth. 2013. The phrase-structural diversity of periphrasis: A lexicalist account. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis: The role of syntax and morphology in paradigms*, 141–167. Oxford: Oxford University Press.

Brown, Dunstan, Marina Chumakina, Greville G. Corbett, Gergana Popova & Andrew Spencer. 2012. Defining 'periphrasis': Key notions. *Morphology* 22. 233–275.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52(2). 311–374. DOI: 10.1017/S0022226715000018.

Dalrymple, Mary. 2015. Periphrasis and morphology in LFG. Paper read at the Workshop on Morphology, LFG 15 Conference, Tokyo. http://users.ox.ac.uk/~cpgl0015/periphrasis.pdf (23 February, 2021).

Hana, Jiří. 2007. *Czech clitics in higher order grammar*. Ohio State University. (Doctoral dissertation).

Hnátková, M., M. Křen, P. Procházka & H. Skoumalová. 2014. The SYN-series corpora of written Czech. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, 160–164.

Karlík, Petr, Marek Nekula & Jana Pleskalová (eds.). 2016. *Nový encyklopedický slovník češtiny*. Praha: NLN. https://www.czechency.org (23 February, 2021).

Kupść, Anna & Jesse L. Tseng. 2005. A new HPSG approach to Polish auxiliary constructions. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, 253–273. Stanford: CSLI Publications.

Matthews, Peter H. 1972. *Inflectional morphology. A theoretical study based on aspects of Latin verb conjugation.* Cambridge: Cambridge University Press.

Matthews, Peter H. 1991. *Morphology.* 2nd. Cambridge: Cambridge University Press.

Petkevič, Vladimír, Alexandr Rosen, Hana Skoumalová & Přemysl Vítovec. 2015. Analytic morphology – Merging the paradigmatic and syntagmatic perspective in a treebank. In *Proceedings of the 5th Workshop on Balto-Slavic Natural Language Processing*, 9–16. Hissar, Bulgaria.

Popova, Gergana & Andrew Spencer. 2013. Relatedness in periphrasis: A paradigm-based perspective. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis: The role of syntax and morphology in paradigms*, 191–225. Oxford: Oxford University Press.

Rosen, Alexandr. 2001. *A constraint-based approach to dependency syntax applied to some issues in Czech word order*. Univerzita Karlova v Praze. (Doctoral dissertation).

Sadler, Louisa & Andrew Spencer. 2001. Syntax as an exponent of morphological features. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2000*, 71–96. Dordrecht: Kluwer.

Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford: CSLI Publications.

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In *Proceedings of WCCFL XIII*, 365–384.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, 359–377. Oxford: Wiley-Blackwell.

Štichauer, Pavel. 2018. Lexical splits within periphrasis: Mixed perfective auxiliation systems in Italo-Romance. *Morphology* 28(1). 1–23. DOI: 10.1007/s11525-017-9313-z.

Stump, Gregory T. 2001. *Inflectional morphology. A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge: Cambridge University Press.

Stump, Gregory T. 2006. Heteroclisis and paradigm linkage. *Language* 82(2). 279–322. DOI: 10.1353/lan.2006.0110.

Stump, Gregory T. 2013. Periphrasis in the Sanskrit verb system. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis*, 105–138. British Academy & Oxford University Press.

Stump, Gregory T. 2016. *Inflectional paradigms*. Cambridge: Cambridge University Press.

Stump, Gregory T. & Raphael A. Finkel. 2013. *Morphological typology: From word to paradigm*. Cambridge: Cambridge University Press.

Stump, Gregory T. & Andrew Hippisley. 2011. Valence sensitivity in Pamirian past-tense inflection: A realizational analysis. In Agnes Korn, Geoffrey Haig, Simin Karimi & Pollet Samvelian (eds.), *Topics in Iranian linguistics*, 103–116. Wiesbaden: Ludwig Riechert Verlag.

Thornton, Anna M. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5. 183–207.

Toman, Jindřich. 1996. A note on clitics and prosody. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching second: Second position clitics and related phenomena*, 505–510. Stanford: CSLI Publications.

Tseng, Jesse L. 2009. A formal model of grammaticalization in Slavic past tense constructions. In *Current issues in unity and diversity of languages*, 749–762. Linguistic Society of Korea.

Tseng, Jesse L. & Anna Kupść. 2007. A cross-linguistic approach to Slavic past tense and conditional constructions. In Peter Kosta & Lilia Schürcks (eds.), *Linguistic investigations into formal description of Slavic languages*, 427–439. Berne: Peter Lang.

Vincent, Nigel & Kersti Börjars. 1996. Suppletion and syntactic theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the first LFG conference*. Stanford: CSLI Publications.

Zwicky, Arnold M. 1985. How to describe inflection. In *Proceedings of Berkeley Linguistics Society 11*, 372–386.

Zwicky, Arnold M. 1986. The general case: basic form vs. default form. In *Proceedings of the Berkeley Linguistics Society 12*, 305–314.