

On Non-Grammaticality, "Speaker Ghosting", and the Raison D'être of English Sequence of Tenses (SOT)

By Krasimir Kabakciev*

The sequence of tenses (SOT) phenomenon has so far remained a mystery both across languages and in English, where its observation often triggers sentences contradicting reality. This forces linguists to define it as bizarre or "useless". The analysis of English SOT is based on Bulgarian data that reveal that a huge number of sentences falling into certain schemata, similar to English prototypical examples of SOT observation, are entirely non-grammatical. There exists an intriguing parallelism between certain sentence types in Bulgarian and English: exactly diametrically opposing cases of non-grammaticality. The non-grammaticality of the Bulgarian sentences is explained through a phenomenon called "speaker ghosting". This is partly applied to English data. A reinterpretation of English SOT is proposed: not a "mysterious rule" but, simply, a grammatical mood, and the raison d'être of English backshift mood (SOT) is identified: a crucially important device to prevent the elimination of non-cancelable content.

Keywords: cancelable content, non-grammaticality, noncancelable content, sequence of tenses, speaker ghosting

Overview of SOT

Huge efforts have been spent in linguistics on the description of the sequence of tenses (SOT) phenomenon. According to most current understandings, SOT is idiogenetic and idiosyncratic, manifesting itself in cross-language terms under different disguises (Abusch 1997, 2004, Altshuler 2010, Chung 2007, Comrie 2009, Declerck 1995, Enç 1987, Grønn & von Stechow 2011, Kubota et al. 2009, Ogihara 1994, 2015, Sauerland & Schenner 2007, Shaer 1997, Zagana 2014). In English the phenomenon is referred to as the SOT rule or backshift (back-shift). As shown in many studies (e.g., Declerck 1995), SOT is important because it may be part of some major parameters of tense, as is the case in English. This article is based on presumptions that the essence of English SOT can be revealed through analyses of certain regularities related to the triggering of non-grammaticality and that SOT is similar to phenomena-like aspects, which cannot be understood without appropriate comparisons between languages featuring the relevant types of aspect (verbal and compositional, see Verkuyl 1972, Kabakciev 2000, Bulatovic 2013).

English SOT

Several things are striking in English SOT. First, a grammatical rule ought to be straightforward. English SOT is certainly not. While its outward parameters are not so complex, its raison d'être remains a mystery. Even an agreed-upon term for it is absent in English grammars, authors usually refer to it as backshift (Quirk et al. 1972), backshift/backshifting (Huddleston & Pullum 2002,

*Independent Researcher, Sofia, Bulgaria.

Downing & Locke 2006) and SOT (e.g., Langacker 2008: 302). Grammarians rarely note, let alone challenge, the lack of an explanation of its essence¹. Huddleston and Pullum (2002) call SOT a "backshifted preterite" and "a special use of the preterite", thus somewhat tacitly subsuming it under the tense system. They distinguish between optional and obligatory backshift and, probably for this reason, because of the unclear status of SOT from a purely grammatical point of view, tend to interpret it in pragmatic terms.

Second, a lot of effort has been spent (in journalistic writing, in today's Internet, etc.) on "how to use SOT and when not to use SOT, but common sense instead". If a rule is a grammatical rule proper (cf. agreement of number), one simply uses it and does not ponder in what circumstances "to violate it". Third, a conclusion arises that the lack of an adequate explanation of English SOT amounts to a failure, substantiated by statements in the theoretical literature on SOT such as a "vacuous tense hypothesis", a semantically "null tense", a "tense-deletion" rule (Ogihara 1996), etc. Among these, Enç's (1987: 635-636) concession that SOT is "totally useless" is perhaps the most strongly worded one. So, while the existence of a problem with the definition of SOT is recognized in linguistics, a hefty question remains unanswered: why does English use rather consistently, if not entirely systematically, a rule regarded as "useless"?

SOT across Languages

SOT is assumed to be present in certain languages and absent in others. Among Romance languages, it is claimed to be consistently present in Spanish, Italian and French, as a heritage from Latin. Most Slavic languages, with their simple system of tenses (present, past, future), are generally described as lacking SOT, and the same is taken to be the case for Bulgarian, despite its nine tenses and an extremely complex system of mood. In English, SOT is taken to be well represented, although it is centered on the use of past tense verbs in the matrix clause.

SOT in Bulgarian

Bulgarian grammars do not postulate a SOT rule similar to English or the Romance languages (Maslov 1982, Bulgarian Grammar 1983, Pashov 2013). However, it can be argued that Bulgarian features SOT in the use of certain mood verb forms in dependent clauses. Bulgarian grammars do not employ systematic analyses of non-grammatical sentences², hence neither questions nor conclusions based on non-grammaticality concerning SOT can be encountered in them, despite findings that Bulgarian features non-grammatical sentences related

¹A certain exception is Langacker (2008: 302), who rejects the "traditional analysis" of SOT – but does not propose a new one.

²Bulgarian linguistics was under the influence of Soviet linguistics, in which "Chomskyan approaches" were out of fashion. Partly due to the lack of handling of non-grammaticality in Bulgarian grammars, non-native researchers of Bulgarian often make mistakes in their judgements, assigning correctness to sentences that are non-grammatical.

to SOT (Sauerland & Schenner 2007). As will be shown below, many sentences similar to English prototypical examples of SOT observation (belonging to certain schemata) are non-grammatical.

Non-Grammaticality in Bulgarian Related to English SOT

Consider that the English sentences (1a,1c) below can be said to be prototypical examples of SOT: the preterite must be used in the *that* clause, not the present perfect. With the present perfect these sentences (1b,1d) are normally judged as non-grammatical:

- (1) a. Peter said that he [himself] arrived
 b. *Peter said that he [himself] has arrived
 c. Peter said that Maria arrived
 d. *Peter said that Maria has arrived

Now compare the naturally presumed Bulgarian correspondences of the correct English sentences (1a,1c) – (2a,2b), with verbs in the dependent clauses in the past perfective (Aorist) tense:

- (2) a. *Petar kaza_{perfAorIndic}, che [samiyat toy] pristigna_{perfAorIndic}
 "Peter said that he [himself] arrived"
 b. *Petar kaza_{perfAorIndic}, che Maria pristigna_{perfAorIndic}
 "Peter said that Maria arrived"

Striking Non-Grammaticality

Bulgarian (2a,2b) ought to be the direct correspondences of English (1a,1c). But, actually, they are non-grammatical – and the non-grammaticality is striking. These sentences, like many similar ones, are not "slightly incorrect" or "not quite good, but sometimes possible". They are entirely non-grammatical.

Due to the complexity of the issue, the reason for the non-grammaticality will be revealed gradually. However, it is necessary to identify the features of the verb form *pristigna* "arrived". It is: (i) perfective; (ii) Aorist; (iii) indicative; (iv) witnessed, explicating a specific modal value, despite being indicative; (v) not only witnessed, but also strongly marked as witnessed – which means that, unlike, for example, aspectual values that can be "neutralized" or changed into the opposite value, etc., whoever uses this form categorically states to have personally and directly witnessed the arrival.³

It is worth checking whether the use of the Imperfect instead of the Aorist might not render the sentences grammatical:

³The regularity is valid only for 3rd person forms, singular and plural. But 3rd person is the most important grammatical person, in actual communication it hugely outnumbers 1st and 2nd person verb forms. The "3rd person Bulgarian peculiarity" is mentioned in Sauerland and Schenner (2007: 257).

- (3) a. *Peter kaza_{perfAorIndic}, che pristigashe_{imperfImperfIndic}
 "Peter said that he was arriving"
 b. *Peter kaza_{perfAorIndic}, che Maria pristigashe_{imperfImperfIndic}
 "Peter said that Maria was arriving"

It does not. Note, however, that just like (1a,1c) above are fully correct English sentences, so are the presumed English correspondences, (4a,4b), of the non-grammatical (3a,3b):

- (4) a. Peter said that he was arriving (on time)
 b. Peter said that Maria was arriving (on time)

Thus, if these Bulgarian sentences (2a,2b), on the one hand, and (3a,3b), on the other, are non-grammatical, it begs the question: what will the normal Bulgarian correspondences of English (1a-1c) and (4a-4b) be? There are different options, but it can be assumed that (5a,5b,6a,6b) below are the typical correspondences of English (1a,1c):

- (5) a. Petar kaza_{perfAorIndic}, che pristignal_{perfAorRenarr}
 "Peter said that he arrived"
 b. Petar kaza_{perfAorIndic}, che Maria pristignala_{perfAorRenarr}
 "Peter said that Maria arrived"
- (6) a. Petar kaza_{perfAorIndic}, che e pristignal_{perfPastIndefIndic}
 (literally) "Peter said that he has arrived"
 b. Petar kaza_{perfAorIndic}, che Maria e pristignala_{perfPastIndefIndic}
 (literally) "Peter said that Maria has arrived"

Note that the verb forms in the second clause in (5a,5b) are perfective renarrative in the Aorist (renarrative in Bulgarian grammar being a mood with a regular paradigm), while those in (6a,6b) are in the past indefinite (indicative), broadly corresponding to the English present perfect. Analogously (7a,7b) below are typical correspondences of English (4a,4b) – where the progressive in English requires the use of renarrative imperfectives in the Imperfect:

- (7) a. Petar kaza_{perfAorIndic}, che pristigal_{imperfImperfRenarr} (navreme)
 "Peter said that he was arriving (on time)"
 b. Petar kaza_{perfAorIndic}, che Maria pristigala_{imperfImperfRenarr} (navreme)
 "Peter said that Maria was arriving (on time)"

In this case, and unlike in (5a-5b) and (6a-6b) above (with two correct versions), it is precisely (7a,7b) that are the correct correspondences of English (4a,4b), while Bulgarian (8a,8b) will be the correspondences of English (9a,9b)⁴:

⁴Bearing in mind that (8a,8b) contain a two-way reading in the dependent clause – one progressive, another indefinitely repeated.

- (8) a. Petar kaza_{perfAorIndic}, che e pristigal_{imperfPastIndefIndic} (navreme)
 "Peter said that he was arriving/used to arrive (on time)"
 b. Petar kaza_{perfAorIndic}, che Maria e pristigala_{imperfPastIndefIndic}
 (navreme)
 "Peter said that Maria was arriving/used to arrive (on time)"⁵
- (9) a. Peter said that he used to arrive (on time)
 b. Peter said that Maria used to arrive (on time)

However, let us return to the non-grammaticality of (2a,2b) above and analyze it against the extremely interesting circumstance that the **non-grammatical** complex compound (2a) can certainly be said to be derived from **two perfectly grammatical** simple sentences, (10a,10b), and that the **non-grammatical** complex compound (2b) can also be said to be derived from **two perfectly grammatical** simple sentences, (10a,10c):

- (10) a. Petar kaza_{perfAorIndic}
 "Peter said"
 b. Petar pristigna_{perfAorIndic}
 "Peter arrived"
 c. Maria pristigna_{perfAorIndic}
 "Maria arrived"

Especially noteworthy are several facts. First, (2a,2b) above are non-grammatical not as an exception. All sentences belonging to the *X said that* schema with a perfective Aorist (indicative but witnessed) in the dependent clause are non-grammatical. That is, they are not "non-grammatical but sometimes acceptable", as is the usual case of non-grammaticality; rather, they are absolutely non-grammatical.

A Brief "Philosophical" Detour

A parallel can be made between the system of grammatical sentences in a language and its corresponding system of non-grammatical sentences, on the one hand, and the domains of human health and disease, on the other. As a rule, a healthy person who is not a physician takes one's good health for granted and does not care much about the huge number of diseases that exist and how dangerous many of them are. Similarly with language, it is a communication tool "taken for granted". People use it intuitively, without thinking, and with the conviction that whatever they perceive in communication is (or at least must be) grammatically correct. If sometimes it happens to be incorrect, it is taken to be produced by a non-native speaker, and the intended message is "guessed" and re-structured as correct by the hearer. Or, if it is produced by a native speaker, the mistake is labeled a "slip of the tongue" and the intended message is again "guessed" – by

⁵The verbs in (8a,8b) are imperfective and in the past indefinite (similar to the English present perfect), indicative, non-witnessed – but could be inferential in meaning.

automatically making corrections in the perceived string, filling up omissions, etc. No consideration is given by the receiver of the message (the average native speaker) to the circumstance that, actually, behind the colossal system of correct, grammatical expressions in a language stands a much more colossal system of incorrect, non-grammatical ones.⁶

Therefore, if the average speaker of a language is not to be bothered by problems of non-grammaticality – just like a healthy person is not to be bothered by nosological entities, it is the linguist's task to identify the scope and parameters of non-grammaticality – just like it is the physician's task to draw the line between health and disease on the basis of in-depth knowledge of the parameters of nosological entities. It can be postulated that just like good health is unidentifiable in medical science without prior knowledge of all possible diseases, the identification in linguistic science of the exact mechanisms of language is inseparable from the study and knowledge of non-grammaticality.

Misconceived Bulgarian Examples

It is becoming clear that non-grammaticality in cross-language terms is key to the understanding of SOT. Yet it remains beyond the attention of researchers and, even worse, mistakes are made by non-native investigators. Smirnova (2009: 204) assigns grammaticality to (11):

- (11) Ivan kaza- Ø če Maria be-še bolna
Ivan say-PAST that Maria be-PAST sick
"Ivan said that Maria had been sick" (backward-shifted reading only)

On the basis of this non-grammatical sentence, which is claimed as correct, a generalization is made that Bulgarian does not feature SOT. The author may have been misled by her informants who probably read it in a specific way. The two verbs *kazha* and its English correspondence *say* mean roughly to "report something" or "utter/communicate a message", and it is important to note that a message can be either true or not true. This point is discussed below. But Bulgarian *kazha* and English *say* can, of course, also be used metaphorically, in the sense of "guess and utter the guess", "infer and utter the inference". Note now that while a message can be either true or not true, it is not in the nature of guesses, inferences or reckonings to be true or not true – they are tools for revealing truths. Thus Smirnova's informants probably read (11) as correct in a metaphorical way, as an instance where *kaza* "said" means "uttered a guess/inference/reckoning". In its prototypical meaning, however, (11) is non-grammatical. Here are some other non-grammatical sentences falling into the relevant schema, with verbs in the dependent clause in different tense forms:

⁶It is often forgotten that the enormous set of grammatical sentences in a language is like the visible part of an iceberg, non-grammatical sentences hugely outnumbering grammatical sentences. The average native speaker does not want and does not need to be concerned with non-grammaticality – but a problem arises if linguists are not sufficiently concerned with it.

- (12) a. *Petar kaza_{perfAorIndic}, che Maria beshe pristignala_{perfPastPerfectIndic}
 "Petar said that Maria had arrived"
 b. *Petar kaza_{perfAorIndic}, che Maria shteshe da pristigne_{perfFut-in-PastIndic}
 "Petar said that Maria would arrive"
 c. *Petar kaza_{perfAorIndic}, che Maria shteshe da e ristignala_{perfPerfectFut-in-PastIndic}
 "Petar said that Maria would have arrived"

Note that while Bulgarian (12a,12b,12c) are non-grammatical, (13a,13b,13c) are grammatical:

- (13) a. Petar kaza_{perfAorIndic}, che Maria pristiga_{imperfPresentIndic}
 (literally) "Petar said that Maria is arriving/arrives"
 b. Petar kaza_{perfAorIndic}, che Maria shte pristigne_{perfFutureIndic}
 (literally) "Petar said that Maria will arrive"
 c. Petar kaza_{perfAorIndic}, che Maria shte e pristignala_{perfFuturePerfectIndic}
 (literally) "Petar said that Maria will have arrived"

Note also that the non-grammatical (12a,12b,12c) above contain verbs in the dependent clause in the indicative in the following tenses: past preliminary, broadly identical to the English past perfect (12a), future-in-the-past (12b), and future-in-the-past preliminary, i.e., future-in-the-past perfect (12c). Conversely, the grammatical sentences (13a,13b,13c) contain verbs in the indicative in the following tenses: present (13a); future (13b); and future preliminary (broadly identical to the English future perfect) (13c). The situation thus revealed imposes an obligation for Bulgarian grammars to be reworked along the lines of positing a whole system of SOT rules (complex ones), allowing or disallowing certain combinations of verb forms in the main clause and the dependent clause in the *X said that* schema (and in similar schemata). This is a hugely overwhelming task. Meanwhile, what is needed here is the asking of these unavoidable questions:

- (i) why are sentences such as (12a,12b,12c) non-grammatical?
 (ii) why are sentences such as (13a,13b,13c) grammatical?

An answer, a definitive one, will be given shortly. But before that, let us consider another example of a wrong judgement concerning non-grammaticality. Similarly to Smirnova, Franks (2008: 93) assigns grammaticality to sentence (14):

- (14) Vcera v gradinata Olga sigurno mu dade snimkite
 Yesterday in garden_{DEF} Olga surely him gave pictures_{DEF}
 "Yesterday, in the garden, Olga surely [most probably] gave him the pictures"

(14) is non-grammatical, and the reason is the same as in (11). The Bulgarian native speaker may here again overlook the non-grammaticality due to the

standard (and reasonable) conviction that "any sentence must be good". When a sentence is not well-formed, the native speaker is still reluctant to analyze the reasons for this (apart from frequently being unable to do this).⁷

What is most striking about the different sets of sentences above – grammatical and non-grammatical – is that there obviously exists something in Bulgarian that can be called a "sequence of mood" rule, and it can be labeled a SOT rule in so far as mood forms are tensed. This rule will not be dealt with here, but it broadly amounts to the following. In the *X said that* schema, the preterite form *kaza* "said" – indicative, witnessed – presupposes the use of a renarrative in the dependent clause instead of a witnessed. If we take English (4b) as a starting point and wish to obtain a correct Bulgarian sentence corresponding to it, the equivalent **cannot** be (3b), despite the fact that it appears to be a good candidate; (3b) is non-grammatical. The correct correspondence of (4b) is either (7b) or perhaps (8b):

- (4) b. Peter said that Maria was arriving (on time)
- (3) b. *Peter *kaza*_{perfAorIndic}, *che* Maria *pristigash*_{eimperfImperfIndic}
 "Peter said that Maria was arriving"
- (7) b. Petar *kaza*_{perfAorIndic}, *che* Maria *pristigala*_{imperfImperfRenarr} (navreme)
 "Peter said that Maria was arriving (on time)"
- (8) b. Petar *kaza*_{perfAorIndic}, *che* Maria *e pristigala*_{imperfPastIndef} (navreme)
 (literally) "Peter said that Maria has been arriving (on time)"

The dependent clause cannot contain a witnessed form. The correct form is either renarrative, as in (7b), or past indefinite (corresponding to the English present perfect), as in (8b). Despite being in the indicative mood, it is not witnessed (only Aorist/Imperfect 3rd person forms are) and can be interpreted as either renarrative or dubitative (dubitative being not a separate mood in Bulgarian but a subtype of the renarrative), or both at the same time. If instead of these two correct forms an indicative one is used,⁸ the result is non-grammaticality, (3b). Yet, how can this non-grammaticality be explained? It is an interesting phenomenon for which the term "speaker ghosting" is proposed.

⁷Among the few researchers of Bulgarian discussing non-grammaticality in the *X said that* schema are Sauerland and Schenner (2007), but their analysis is from a different standpoint.

⁸Which, hence, will be non-renarrative: 3rd person Aorist/Imperfect indicative forms are always witnessed and non-renarrative.

Exactly Opposite Phenomena in English and Bulgarian

We should also note another interesting circumstance. While Bulgarian (2b) with a preterite form, indicative and witnessed, is non-grammatical, and (13a) – with a present indicative – is, conversely, fully correct, in English, as a result of SOT, **precisely the opposite is observed**: sentence (1c), with a preterite *arrived*, is correct, while (1c'), with a present (progressive) form *is arriving*, is non-grammatical:

- (2) b. *Petar kaza_{perfAorIndic}, che Maria pristigna_{perfAorIndic}
"Peter said that Maria arrived"
- (13) a. Petar kaza_{perfAorIndic}, che Maria pristiga_{imperfPresentIndic}
(literally) "Peter said that Maria is arriving/arrives"
- (1) c. Peter said that Maria arrived
c'. *Peter said that Maria is arriving

This striking parallelism in two separate (and rather different) languages will be analyzed later. Let me here finally reveal the puzzle of the non-grammaticality of (2b)⁹.

Ghosting

Ghosting is a well-established concept in electronics, meaning "the appearance of more than one image of an entity on a TV/PC screen, when the entity is moving fast".

Speaker Ghosting

"Speaker ghosting" as a linguistic notion is introduced here on the basis of the analysis of the non-grammatical Bulgarian sentences above; it can be said to consist of the following. If we have a sentence such as *Maria arrived* in English, and its producer, the abstract participant in the communication called speaker, knows that Maria arrived because he/she witnessed (saw/heard) Maria's arrival, in Bulgarian the speaker will use the sentence *Maria pristigna* "Maria arrived [witnessed]". But in the non-grammatical sentence (2b), apart from asserting that Maria arrived – through the clause *Maria pristigna* [witnessed], the speaker also asserts Peter's words that Maria arrived. Maria's arrival thus starts to appear problematic.

⁹I asked dozens of native speakers, including my wife, why (2b) is non-grammatical. No success, not the slightest trace of a clue! I myself felt perplexed by the question for months – until I found the answer. The non-grammaticality here is stunning, with its almost full incomprehensibility, and raises the necessity for a better description in linguistics of the possible types of non-grammaticality, especially those resisting explanation.

Why? Because if we rely on the words of a third party that Maria arrived, Maria may not have arrived. What is more, it turns out that the proposition of a sentence like this actually contains not one but two speakers.

Caution! Two Speakers in a Sentence!

In (2b) one speaker is saying that Peter said that Maria arrived. And another speaker is saying that he/she witnessed Maria's arrival. But if one speaker is saying that Peter said that Maria arrived, this opens two possibilities: for Maria to have arrived or not to have arrived. The second of these possibilities is canceled by the second speaker's assertion *Maria pristigna* "Maria arrived [witnessed]". Surely, it is the second speaker, the one who personally saw Maria's arrival, who ought to be trusted by the hearer. Given such a case, the first speaker, the one reporting Peter's words concerning Maria's arrival, ought to be interpreted as an illegitimate speaker, a fake speaker, **a ghost speaker**.

It stands to reason that a sentence cannot freely accommodate two different speakers simultaneously. A sentence cannot, furthermore, be expected to allow these two speakers to contradict each other. For this reason (2b) is conceived of as non-grammatical, because it not only accommodates two speakers, but also they grossly contradict each other:

- (2) b. *Petar kaza_{perfAorIndic}, che Maria pristigna_{perfAorIndic}
"Peter said that Maria arrived"

Conversely, (5b) is fully correct, and it is this sentence that covers the message of (2b):

- (5) b. Petar kaza_{perfAorIndic}, che Maria pristignala_{perfPerfectRenarr}
"Peter said that Maria arrived"

(5b) does not contain two speakers. It contains only one speaker – who asserts Peter's words that Maria arrived, registering them as renarrated through the renarrative *pristignala*, making it clear that Maria's arrival is assumed on the basis of Peter's words only, not on the basis of Peter's words and **simultaneously** on the evidence of another speaker, as in the non-grammatical (2b).

Let us now consider the other non-grammatical sentences given as examples above, to see if the reason for the non-grammaticality is the same. For example, the proper Bulgarian equivalent of English (15) is not the "most naturally presumable" (12b), but is (16):

- (15) Peter said that Maria would arrive
(12) b. *Petar kaza_{perfAorIndic}, che Maria shteshe da pristigne_{perfFut-in-PastIndic}
"Petar said that Maria would arrive"
(16) Petar kaza_{perfAorIndic}, che Maria shtyala da pristigne_{perfFut-in-PastRenarr}
"Petar said that Maria would arrive"

This is because *Petar kaza* "Peter said" generally (though not always, as will be shown below) requires a renarrative form, and sentence (16) features it: *shtyala da pristigne* "would arrive [perfective, future-in-the-past, renarrative]". Note again the existence here of a "sequence of mood" rule, which can also be termed a SOT rule, the verb forms being tensed. Otherwise, of course, it can easily be established that in (12b) the clause *Petar kaza*, containing an Aorist indicative (witnessed) form, means exactly "Peter said", and that the dependent clause *Maria shteshe da pristigne*, containing a future-in-the-past indicative form, means exactly "Maria would arrive". Does this mean that if we combine these two correct sentences into (12b), it will correspond to the correct English sentence (15) above?

No! As the asterisk in (12b) shows, this sentence is not just "incorrect". It is entirely non-grammatical, again a case of speaker ghosting. There are two speakers here: one is asserting Maria's arrival – through the form *shteshe da pristigne* "would arrive [future-in-the-past, indicative]", which envisions the effectuation of a future situation – from the standpoint of the past (although it is not a witnessed form). And the other speaker is saying he/she witnessed Peter's saying that Maria would arrive – through the clause *Petar kaza* [witnessed]. This is a contradiction. If one speaker is saying that Peter said that Maria would arrive, this envisions two different scenarios for the future: for Maria to arrive or not to arrive. Moreover, if another speaker asserts that Maria is certainly arriving – through a future-in-the-past indicative, this means that the second speaker is canceling the scenario for Maria not to arrive, thus contradicting the first speaker. As already postulated, a single sentence in a language cannot normally be expected to "accommodate" two speakers simultaneously, especially when they contradict each other. Hence, (12b) is non-grammatical.

The non-grammaticality of (12c), with a verb in the future preliminary (future perfect) is similar:

- (12) c. **Petar kaza*_{perfAorIndic}, *che Maria shteshe da e pristignala*_{perfPerfectFut-in-PastIndic}
 "Peter said that Maria would have arrived"

The switch from (12b) to (12c) amounts only to the addition of a perfect (tense) marker, which does not change the relevant content underlying the non-grammaticality.

Of course, the following question arises. It was previously established that (2b) is non-grammatical because it manifests speaker ghosting, the feature triggering the non-grammaticality being the marker "witnessed". But if the non-grammaticality of (2b) is related to the marker "witnessed", why are (12b, 12c) also non-grammatical when the marker is absent there? Clearly, the explanation of the non-grammaticality of (12b, 12c) is different, and will be given below.

Recall that Smirnova (2009, 204) assigns grammaticality to (11), and similarly Franks (2008: 93) assigns grammaticality to (14). These sentences are non-grammatical, hence must be preceded by asterisks. The normal sentence corresponding to the one intended in (11) is (11') or (11''), with either *e* "is", a

present tense indicative form, or *bila* "was", a past renarrative of *sam* "be", in the dependent clause:

- (11') Ivan kaza_{perfAorIndic}, che Maria e_{imperfPresIndic} bolna
 (literally) "Ivan said that Maria is sick"
 (11'') Ivan kaza_{perfAorIndic}, che Maria bila_{imperfPastRenarr} bolna
 "Ivan said that Maria was [reportedly] sick"

In (11) the reason for the non-grammaticality is the same, or fairly similar to, that in (2b). The main clause requires a verb in the dependent clause envisioning two possibilities: for Maria to be sick or not (in the past). But the use of *beshe* "was", a witnessed form, cancels the possibility for Maria not to have been sick. In other words, the first speaker expects, and actually requires, the verb in the dependent clause to allow both possibilities. Suddenly, a second speaker appears and reports that he/she actually witnessed Maria's sickness, flouting the requirement of the first speaker. As already stipulated, language cannot allow the simultaneous presence in a single sentence of two speakers contradicting each other¹⁰.

The normal grammatical sentence corresponding to the one intended in Franks' (14) is (14'). There the past indefinite *e dala* "has given" is obligatory (the Bulgarian past indefinite broadly identical to the English present perfect, the difference being that the Bulgarian past indefinite allows past time adverbials):

- (14') Vchera v gradinata Olga sigurno mu e dala_{perfPastIndefIndic} snimkite
 (literally) "Yesterday, in the garden, Olga surely [most probably] has given him the pictures"

It can easily be seen that (14) also exhibits speaker ghosting, due to the simultaneous presence of the adverbial *sigurno* "surely" [most probably] and of *dade* "gave", a witnessed form. The first speaker, producing the string *Vchera v gradinata Olga sigurno* "Yesterday in the garden Olga surely", containing *sigurno* "surely" [most probably], posits a two-way scenario for Olga: to have given the pictures or not to have given the pictures. Suddenly, a second speaker appears and asserts (through the witnessed *dade* "gave") that he/she saw Olga giving the pictures. Having personally witnessed the handover, it is the second speaker who must be trusted. Hence, it remains for the first speaker to be regarded as an illegitimate one, a ghost speaker. Consequently, (14) is non-grammatical.¹¹

¹⁰Regarding *kaza* "said" in (11) as metaphorical, already discussed, this is a possibility, but it can be assigned to (11) only as a peripheral reading in which the sentence will, indeed, be acceptable.

¹¹In fairness to Smirnova and Franks, the identification of non-grammaticality here is an extremely difficult task, far beyond the capabilities of the average native speaker – hence, their misjudgements can be excused. And, as already mentioned, this issue, non-grammaticality adamantly resisting an explanation, must be handled in a better way in linguistic theory.

Non-Grammaticality due to Speaker Ghosting Related to SOT?

Now it remains to explore how the non-grammaticality of the sentences above and the phenomenon of speaker ghosting – which actually constitutes the explanation of the non-grammaticality – relate to the problem of SOT. As far as English SOT is concerned, it is often described as based on two ideas: the "natural order of events" and the causal relations between them. The "natural order of events" corresponds to the "normal perception of events" by humans: moving from the past through the present to the future. Past events occur first, then present, and then future events, not vice versa. The idea of the causal relations between events envisions that if an event happens it may cause a second event, and the second event occurs after the first one. The second event may cause a third event, and this third event will occur after the second one, etc. These two ideas or phenomena, the "natural order of events" and the causal relations between them, are reflected in the technical metaphor called a "time axis". The time axis goes from the left to the right, not from the right to the left (according to convention), and this also reflects the understanding that past situations appear first, then present, then future.

Thus, if we are describing a happening in the past, the reason why it happened will also be normally found in the past, not in the present or in the future. For example, if John sits down, the observer will normally produce (17a), not (17b), although (17b) is, perhaps, not impossible. Note that if we use (17c), the temporal reference of "feeling tired" will change from the real future, which is to the right of the moment of utterance, to the future-in-the-past, which is to the left of the moment of utterance:

- (17) a. John sat down because he felt tired
 b. John sat down because he will feel tired
 c. John sat down because he would feel tired

(17b) may appear as deviating from the "natural order of events". But actually we have the following sequence. John knows he often feels tired. While running, he realizes he may feel tired soon. Therefore, he sits down: the speaker expects John to feel tired in the immediate future. In (17c), conversely, the speaker envisions John's feeling tired after sitting down, but before the moment of utterance, i.e., in the future-in-the past. Hence (17b) and (17c) mean two different things.

In a similar "natural order of events", if it suddenly starts raining and John enters a café, an observer will produce (18a). Language, however, is a subtle tool that allows people to depict happenings in alternative ways, not conforming to the "natural order of events" and SOT. Instead of uttering (18a), one can utter (18b):

- (18) a. John entered the café because it was starting to rain
 b. John entered the café because it is starting to rain

It can be argued that (18b) fails to conform to SOT. But as it can hardly be deemed incorrect, such sentences are labeled "exceptions to SOT" or, as in Huddleston and Pullum's (2002) terms, as optional backshift.

Thus, broadly speaking, there are sentences that reflect the natural order of events and the causal connection between them (e.g., here the rain caused John's entering the café) – but they may appear built according to SOT. Is (18b) really "an exception to SOT"? Let us check using the "speaker ghosting" concept.

There seem to be two speakers in (18b). One is in the past, reporting John's entering the café; another is in the present, reporting the starting of the rain. Perhaps these are not two speakers, it is one speaker transported from the past moment (John's entering the café) to the present (starting to rain). Language does not pose a problem here. There can be no inconsistency *per se* in the introduction of a second (non-contradicting) speaker or in the transportation of a speaker from the past to the present. Compare perfect sentences such as (19), where either there are two speakers (one in the past, one in the present, not contradicting each other) or a single speaker is transported from the location "yesterday" to the location "today":

(19) Yesterday John was sad, today he is happy

Sentence (18b) can be said to conform to both the "natural order of events" and the causal relationship: it starts to rain, and this causes John's entering the café. Still, (18b) is often regarded as simultaneously deviating from SOT and acceptable. And it is then that it is labeled "an exception to SOT". There is a general problem with the delineation of the scope of SOT: it is frequently difficult or impossible to draw a borderline between a "legitimate" and an "illegitimate" deviation.

The explanation usually runs thus: *is starting* (18b) must be backshifted to obtain the "correct sentence" (18a). Such an explanation is not exactly wrong, but it is too general. A correct explanation ought to take into account the following. Imagine the speaker and the hearer in front of a café in a city where it often rains. It is starting to rain now. The speaker and the hearer see John entering the café. The speaker produces (18b). There is nothing wrong, just two consecutive situations: it starts to rain, John enters the café. And although the event order is reversed, it is perfect as it is and semantically equal to (20) – which has the "natural word order" not subject to SOT:

(20) It is starting to rain, John entered the café

This is an aspect of the problem that is usually bypassed in discussions of "exceptions to SOT". (18b) is not an "exception to SOT", it is a meaningful sentence in the setting just described and, simply, there is no reason to apply SOT.

Imagine, however, speaker and hearer again in front of the café. It is starting to rain again. Suppose the speaker recalls that at 5 p.m. yesterday John entered the café. Will the speaker produce (21)?

(21) *John entered the café at 5 p.m. yesterday because it is starting to rain

Of course not. It is starting to rain now, but John entered the café yesterday – because it was starting to rain **then**, not because it is starting to rain now. In this second episode setting – in which the rain and John's entering the café happened yesterday, the sentence is non-grammatical. The corresponding Bulgarian sentence with a present tense form is also non-grammatical, for the same reason:

(21') *Dzhon vlez_{perfAorIndic} v kafeto v 5 sledobed vchera, zashtoto zapochva
imperfPresIndic da vali
(literally) "John entered the café at 5 p.m. yesterday because it is starting to rain"

Must we assume that Bulgarian has a SOT rule, requiring a backshift for sentences like (21')? Perhaps yes, but the backshift is due to the "natural order of events" and the causal relations between them, while in the standard English SOT rule there is something that goes beyond these two factors.

If we return to the Bulgarian data, the overview of the correspondences of English sentences with a verb in the preterite shows that the first thinkable Bulgarian correspondence of a simple sentence with a preterite will always be a sentence in the indicative: either in the Aorist, if a perfective situation is depicted, or in the Imperfect, in an imperfective situation. If we have (22a) in English, the corresponding Bulgarian sentence is (22b); if we have (22c), the corresponding Bulgarian sentence is (10b). On the other hand, if we combine these two perfect sentences (22a,22c) to obtain (22d), the corresponding Bulgarian sentence, built in exactly the same way, is non-grammatical, (2a). The correct Bulgarian sentence corresponding to English (22d) is not (2a) but (22e). And, as already made clear, (22e) corresponds to an English sentence that is non-grammatical, (22f) – at least when referring to a distant past as in (22g).

- (2) a. *Petar kaza_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter said he arrived"
- (10) b. Petar pristigna_{perfAorIndic}
"Peter arrived"
- (22) a. Peter said something
b. Petar kaza_{perfAorIndic} neshto
"Peter said something"

- c. Peter arrived
- d. Peter said he arrived
- e. Petar kaza_{perfAorIndic}, che e pristignal_{perfPastIndefIndic}
(literally) "Peter said that he has arrived"
- f. *Peter said (that) he has arrived
- g. *Last week Peter said he has arrived

Why is English (22f), in the distant past reading, non-grammatical? The standard answer is: "violation of SOT".

Exactly Diametrically Opposing Cases of Non-Grammaticality

In any case, however, clearly what we have here are **diametrically opposing cases of non-grammaticality**. Can this be coincidental? Or is it not, rather, the case that some cross-language (or universal) regularity exists underlying the otherwise different reasons for non-grammaticality in sentences in two different languages? Such a unique circumstance, exactly diametrically opposing cases of non-grammaticality, challenges the linguist to try to identify the reasons.

Let us start with Bulgarian. The reason for the non-grammaticality in (2a) is already explained, it is broadly identical to the non-grammaticality of (2b), speaker ghosting:

- (2) a. *Petar kaza_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter said he arrived"
- b. *Petar kaza_{perfAorIndic}, che Maria pristigna_{perfAorIndic}
"Peter said that Maria arrived"

The first speaker in (2a) asserts that Peter stated that he (himself) arrived. This presents no problem but envisions two possibilities. If Peter is saying that he arrived, then either he has arrived or not. In other words, Peter's statement about his own arrival is not rendered through language as a fact. It is just something that the first speaker is saying: that Peter arrived. Suddenly a second speaker appears, saying that he/she saw Peter's arrival, using *pristigna*, a witnessed form. Obviously, it cannot be the case simultaneously that Peter may have arrived or may not have arrived and that Peter's arrival was witnessed. This is a grave contradiction: a single speaker cannot simultaneously assert that Peter may have arrived or may not have arrived and that Peter did arrive. Hence, there are two speakers in (2a). The existence of two speakers in the same sentence cannot be allowed, especially if they fiercely contradict each other. In any case, Bulgarian (2a) is non-grammatical, and so are innumerable other sentences with a witnessed form in the dependent clause falling into the schema.

Approaching the Essence of SOT

Now two important questions arise. First, if the non-grammaticality of Bulgarian (2a) is already clear, why is it that English (22f), nominally corresponding to Bulgarian (22e), is non-grammatical (in the distant past time reading)?

- (2) a. *Petar kaza_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter said he arrived"
- (22) e. Petar kaza_{perfAorIndic}, che e pristignal_{perfPastIndefIndic}
(literally) "Peter said that he has arrived"
- f. *Peter said (that) he has arrived

Can it be simply because of the flouting of SOT, or are there other reasons? Hence it appears reasonable here to voice the question related to the heart of the issue again. What is the essence of SOT? Second, couldn't the non-grammaticality of English (22f) in the relevant reading be attributed to a similar speaker ghosting as in Bulgarian (2a)?

Speaker Ghosting the Culprit?

It could, perhaps, be argued that the non-grammaticality of (22f) is due to speaker ghosting. However, the reasons will be different from those in (2a). Note that in (2a) both speakers, the legitimate one and the illegitimate one, are deployed in the past. That is, it is not some inconsistency of their time axis location that underlies the non-grammaticality. The inconsistency is in the introduction by the first speaker of two possibilities, for Peter to have arrived or not to have arrived, and in the simultaneous canceling of the second of these possibilities by the second speaker. In English (22f), however, the inconsistency appears to be in the deployment **in the past** of one speaker – who is asserting Peter's arrival as separated from the present (according to the standard definition of the English past simple), and in the simultaneous deployment **in the present**, i.e., on a different time axis segment, of a second speaker. Apart from that, the second speaker is asserting Peter's arrival not as a past situation separated from the present but as a past situation with significance for the present, as per the standard definition of the English perfect.

Arguably, in a single sentence, that ought to be uttered by a single speaker, a past situation separated from the present presupposes a speaker located in the past and observing the situation from there. This does not presuppose a second speaker, located in the present and observing the situation from the moment of utterance. Of course, one and the same situation can, in principle, be observed from the point of view of the past **and simultaneously** from the point of view of the present, as already established. It ought to be observed from the present either by another speaker in a different sentence or by the same speaker in a sentence structured after a redeployment of the speaker from the past into the

present, as in (19), already discussed, in which the speaker is taken to have legitimately changed his/her time axis location:

(19) Yesterday John was sad, today he is happy

(22f) offers two possibilities: (i) two speakers simultaneously deployed at two different time axis locations; (ii) one speaker with a special capacity of attending two different time axis segments simultaneously, one in the past, one in the present. But the irregular deployment of one speaker or two speakers on the time axis does not seem to be the only reason for the non-grammaticality of sentences such as (22f) – or (23a), produced instead of (23b):

- (23) a. *Peter said he will telephone Maria
 b. Peter said he would telephone Maria

Canceling

Recall the non-grammaticality of Bulgarian (2b), consisting in the canceling by the second speaker of one of two possibilities: for Maria to have arrived or not to have arrived. The key word is *canceling*, to be discussed shortly.

Let us have some similar non-grammatical English sentences with a violation of SOT (in the relevant past time reading) such as (23) above: (24a) instead of (24b); (25a) instead of (25b):

- (24) a. *Peter said Maria has arrived
 b. Peter said Maria had arrived
(25) a. *Peter said Maria is playing tennis
 b. Peter said Maria was playing tennis

Why Do Punctuation and the Conjunction Matter so Much?

Note that without the conjunction *that* (24a,25a) – if heard (i.e., not read) – are **not** non-grammatical. Because if heard, (24a) corresponds to (26a) in the written form, and (25a) corresponds to (26b):

- (26) a. Peter said "Maria has arrived"
 b. Peter said "Maria is playing tennis"

Hence all sentences of this type, including (24a,25a), could be regarded as correct. This begs the question: what is it that so easily makes (24a,25a) non-grammatical (not quite acceptable)? The difference in punctuation? The introduction of *that*?

A Comparison with Bulgarian – A Single Sound/Letter Enormously Important

If we make a comparison with Bulgarian, we will see that there exists something negligible at first sight, which eventually turns out to be hugely important. It is the similar impact of a minor morphological element that drastically changes a sentence from a grammatical one to a non-grammatical one, or, vice versa, from a non-grammatical one to a grammatical one. See how the non-grammatical (2a) becomes grammatical (27) after adding the sound/letter *-l*:

- (2) a. *Petar kaza_{perfAorIndic}, che pristigna_{perfAorIndic}
 "Peter said he arrived"
 (27) Petar kaza_{perfAorIndic}, che pristignal_{perfAorRenarr}
 "Peter said he arrived"

Isn't it intriguing that at the very heart of the issue concerning grammaticality/non-grammaticality in two separate languages lies the use of a conjunction instead of quotes (English) or a single sound/letter in the verb (Bulgarian)? Obviously, as regards Bulgarian, this single sound/letter must be the carrier of some extremely significant meaning to be able to transform innumerable non-grammatical sentences into sentences that are grammatical.

What, then, is the meaning of *-l* in Bulgarian, apart from the nominal explanation that it turns a witnessed form into a renarrative? Below is the answer.

Switchover from Non-Cancelable to Cancelable

Adding *-l* to the verb in Bulgarian – in the type of cases under consideration, turns the meaning of the verb, and hence of the VP/clause/sentence, from non-cancelable to cancelable content. "Cancelable" and "non-cancelable" are, of course, well-established concepts in linguistics (pragmatics) – since Grice's (1975) pioneering work.

As already explained, the form *pristigna* "arrived" in (10b) is a witnessed one, used to represent Peter's arrival as a fact:

- (10) b. Petar pristigna_{perfAorIndic}
 "Peter arrived"

It does not matter whether the situation is a fact in the real world or not. It is represented as a fact in linguistic terms only, by means of language, by the speaker, and is called non-cancelable content. The non-cancelability of *pristigna* is easily observed in simple sentences such as (10b): Peter's arrival is taken to be true, a fact, and it is precisely for this reason that the expression (clause) is not allowed in a non-simple sentence such as (2a). The reason is that the *X said that* [content of *that*] schema requires a possibility for Peter not to have arrived (if the content of *said* is "arrived"), and the form *pristigna* "arrived [witnessed]" in (2a) cancels it. Conversely, in (27) the renarrative *pristignal* "arrived" represents cancelable content, something that can be true or not true. For this reason, it is

allowed in the schema, in contrast to *pristigna* "arrived [witnessed]", which is not. It is the content of Peter's words – viz., that he arrived, that is represented in language terms, through the addition of a single phoneme (sound) *-l*, as cancelable. A renarrative form signifies that whatever is described by the speaker was not witnessed and asserted directly by him/her but is reported as somebody else's words. And if something is reported by the speaker as asserted by somebody else, not directly by the speaker, it is by virtue of its very nature either true or not true, and will be represented in language as cancelable content – in one way or another, depending on the particular language.

The analysis thus made shows that in both Bulgarian and English (and surely in many other languages too) a schema such as *X said that* requires that the language expression representing the content of *said* must allow this content to be cancelable. In the dependent clause of Bulgarian sentences such as (27) the content is overtly marked as "cancelable" by the renarrative (hence non-witnessed) form *pristignal* "arrived". This means that if in a certain language the cancelable content, the content of *said* in the *X said that* schema, is not overtly marked as cancelable, there may be a problem. The most obvious problem would be if the content of *said* is overtly marked as non-cancelable. It is precisely this that happens in (2a), where *pristigna* "arrived" is a witnessed form, and is thus non-cancelable. The result? The sentence is conceived of as non-grammatical.

What will happen if the form representing the content of *said* in the *X said that* schema is simultaneously not overtly marked as cancelable **and** not overtly marked as non-cancelable? It can be argued that this is the case in English (1a,1c) – and in all similar sentences.

Let us assume for the moment that, indeed, this is the case. In English sentences like (1a,1c) the form *arrived* is not overtly marked as cancelable or non-cancelable. Yet, since in (1a) *arrived* undoubtedly means that Peter has either arrived or not arrived (because of the impact of *said* in the matrix clause) and, per argumentum a contrario, *Peter arrived* means that Peter arrived, the question is whether:

- (i) the verb form *arrived* in *Peter arrived* is marked as non-cancelable, or that
- (ii) the verb form *arrived* in *Peter said he arrived* is marked as cancelable.

Recall again that in Bulgarian there are two separate forms, each marked for the relevant function: *pristigna* "arrived [witnessed]" is a form marked as non-cancelable through the overt feature "witnessed". Conversely, *pristignal* "arrived [renarrative]" is a form marked as cancelable through the overt feature "renarrative".

Thus, it appears necessary to take for granted that in English *arrived* – in isolation, which means in terms of its various possible manifestations, including sentences of the *X arrived* type – is not marked as non-cancelable, simply because it also appears in sentences like *X said X arrived* or *X said that Y arrived*, where it represents cancelable content, for reasons already clarified.

The Solution? The Opposite One

Thus, ordinary logic and common sense force the assumption that the solution is the opposite one:

- First, the form *arrived* in sentences such as *X arrived* is not marked as non-cancelable;
- Second, it remains for the form *arrived* to be interpreted as marked "cancelable" in sentences such as *X said X arrived* or *X said that Y arrived*.

Indeed, if we take *arrived* to be interpreted as marked "cancelable" in the relevant sentences, it must be due to its participation in a specific schema – because overtly *arrived* is certainly not marked as cancelable. Note that overt marking of mood, aspect and other grammatical values can be effectuated morphologically, as, e.g., in the case of adding the suffix *-l* to the Bulgarian form *pristigna* "arrived [witnessed]" to obtain *pristignal* "arrived [renarrative]", or periphrastically, as in changing the aspectually unmarked English form *arrived* (unmarked for perfectivity or imperfectivity) to the aspectually marked *was arriving*. Note that *arrived* is not marked for perfectivity or imperfectivity because it can be used in perfective (28a) **or** imperfective (28b) sentences (Verkuyl 1993, Kabakciev 2000, 2017):

- (28) a. The train arrived
b. Trains arrived

On the other hand, the form *was arriving* is marked for imperfectivity because it is a progressive and can signify imperfectivity by itself – though this imperfectivity is of a specific type, ongoing actions.

Therefore, it remains to be assumed that the form *arrived* is not marked as non-cancelable content in sentences such as *X arrived* but is marked as cancelable in sentences such as *X said X arrived* or *X said Y arrived* – that is, in certain schemata. There is indeed a problem, in that while the aspectually unmarked *arrived* is contrasted to the aspectually marked *was arriving*, in the case of *X arrived* versus *X said X arrived* there is no overt (morphological) marking. The two forms are the same. Certainly, we can assume that this is a case of grammatical homonymy, an extremely common phenomenon in English. Note also the grammatical homonymy of English *would* that can serve to mark three totally different grammatical values: (i) pure future-in-the-past (29a); (ii) past habitual (29b); (iii) conditional (29c).

- (29) a. That day John would call me in the morning and we would meet in the afternoon
b. John would call his friends whenever he needed money
c. John would tell Maria if he loved her

It can be generalized that grammatical homonymy is such a widespread phenomenon in English that linguists are sometimes prone to overlooking its presence.

Therefore, it can be postulated that there are two separate forms *arrived*: one in general use and unmarked for cancelability/non-cancelability, as in *X arrived*; the other one in specific schemata such as *X said X arrived* or *X said Y arrived* – where *arrived*, although overtly not marked for cancelable content (just like *would* is not marked for "future-in-the-past", "habitual" or "conditional"), is to be interpreted as such, as cancelable (just like *would* is interpreted as "future-in-the-past", "habitual" or "conditional" depending on the sentence used).

Note that in a sentence such as *X said X arrived*, which can be said to be the result of a SOT operation, the form *arrived* does not necessarily and in all cases stand in contrast to *arrives*. (30a) cannot be said to be obtained from (30b) – for different reasons, but especially the fact that these two sentences explicate different aspectual values. (30a) explicates perfectivity in the dependent clause, while (30b) explicates imperfectivity:

- (30) a. Peter said he arrived
b. Peter said he arrives

There are other sentences in which *arrived* can be said to be obtained from *arrives*. Compare (31a), which, after the implementation of SOT, yields (31b) – and the two sentences do not differ in other terms, e.g., in the aspectual value of *arrive/arrived* (imperfective in both cases):

- (31) a. John said "The early morning trains always arrive on time"
b. John said that the early morning trains always **arrived** on time¹²

This means that even though *X said X arrived* is not necessarily obtained from *X said X arrives*, the form *arrived* in *X said X arrived* or *X said Y arrived* can be regarded as a special form manifesting cancelable content. The same can thus be said of all verb forms that appear in schemata such as *X said that*. These are special cancelable content forms – and they may manifest homonymy with forms that are used outside such schemata.

English Backshift Mood

Thus SOT in English can be said to represent a paradigm of opposing forms manifesting cancelable content that more often than not differ overtly (*is playing* > *was playing*). When they do not differ overtly (*had played* > *had played*), these are cases of grammatical homonymy. If instead of "a SOT rule" the phenomenon

¹²An anonymous reviewer expressed doubts as to why *arrived* in (31a,31b) should be treated as imperfective. It is because of the adverbial of non-bounded repetition *always* – which maps its value onto the action itself (Kabakciev 2000: 123-151), rendering it iterative and non-bounded, i.e., imperfective.

is called "mood", the second paradigm can be called "backshift mood", and the first one "indicative". The indicative mood paradigm will contain forms that represent non-cancelable content, while the backshift paradigm will contain cancelable content forms.

Some Comparison with Bulgarian

Of course, in principle, it does not matter much if what is (usually) called SOT rule is labeled backshift mood. The phenomenon is the same. What is to be identified is the essence of the backshift mood (the SOT rule), its *raison d'être*. In an effort to reveal it through a comparison with Bulgarian, recall that *pristigna* "arrived [witnessed]" is non-cancelable content that, on its own, forms perfect sentences such as *X pristigna* "X arrived [witnessed]", in which X's arrival is asserted as a fact. But if *pristigna* is used in a subordinate clause in sentences like (2a,32a,32b), non-grammaticality is triggered. The sentences immediately become non-grammatical, for a reason already explained, speaker ghosting:

- (2) a. *Petar kaza_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter said he arrived"
- (32) a. *Petar misleshe_{imperfImperfIndic}, che pristigna_{perfAorIndic}
(literally) "Peter thought he arrived"
- b. *Petar znaeshe_{imperfImperfIndic}, che pristigna_{perfAorIndic}
(literally) "Peter knew he arrived"

They all manifest the simultaneous presence of two speakers in a single sentence, with the two speakers contradicting each other. The first speaker is saying that Peter said or thought or knew that he (himself) arrived, leaving open the possibility for him not to have arrived, and the second speaker is canceling this possibility by asserting that he/she witnessed Peter's arrival – through the use of the witnessed form *pristigna*. Note specifically the important circumstance that, as can easily be seen in (32a,32b), non-grammaticality in Bulgarian in cases of the type discussed is not restricted to the *X said that* schema. It is also observed with other verbs in the main clause, but not all verbs. Thus, if non-grammaticality in the relevant schema occurs with verbs such as *misleshe* "thought" and *znaeshe* "knew", it does not if verbs such as *vidya* "saw" or *razbra* "realized" are used. Compare (33a,33b) below. These are perfect sentences, the reason being that verbs such as *vidya* "saw" and *razbra* "realized", in contrast to *misleshe* "thought" and *znaeshe* "knew", do not require cancelable content in the dependent clause:

- (33) a. Petar vidya_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter saw he [had] arrived"
- b. Petar razbra_{perfAorIndic}, che pristigna_{perfAorIndic}
"Peter realized he [had] arrived"

Also, note the following difference between (32a,32b) on the one hand, and (33a,33b) on the other. The situations (states) of thinking and knowing in (32a,32b) are preceded by the arrival. Conversely, in (33a,33b) seeing and recognizing the arrival coincide with the arrival.

But if we try to transfer the notion of cancelable content and the general state of affairs thus described for the Bulgarian grammatical system into the English system, exactly the same result can hardly be expected to obtain. Why? Because English features neither witnessed nor renarrative forms. Recall that (25b) is a good sentence that can be decomposed into *Peter said* and *Maria was playing tennis*:

(25) b. Peter said Maria was playing tennis

English *Peter said* entirely matches Bulgarian *Petar kaza*, and English *Maria played tennis* entirely matches Bulgarian *Maria igraeshe tenis*. But while English (25b), built from two correct sentences, is correct, Bulgarian (34a), built again from two correct sentences, is non-grammatical. To become grammatical, (34a) must turn into (34b), by using the renarrative *igraela* "played". And here it is worth noting that (34b) can be said to be derived from (34c):

- (34) a. *Petar kaza_{perfAorIndic}, che Maria igraeshe_{imperfImperfIndic} tenis
 "Peter said Maria played tennis [witnessed]"
 b. Petar kaza_{perfAorIndic}, che Maria igraela_{imperfImperfRenarr} tenis
 "Peter said Maria played tennis"
 c. Petar kaza_{perfAorIndic}, che Maria igrae_{imperfPresIndic} tenis
 (literally) "Peter said Maria plays tennis"

Hence (34b) can be said to manifest the observation of two rules simultaneously, "sequence of tense" and "sequence of mood," because the verb changes not only for tense (present>past) but also for mood (indicative>renarrative).

The issue concerning Bulgarian SOT will not be explored further here, as it requires enormous research. Let us return to the major issue, the *raison d'être* of English SOT.

What is causing the grammaticality and non-grammaticality of sentences associated with SOT in English? Ultimately, what is the reason underlying the existence of SOT in the English tense system?

More on the Absurdities and "Uselessness" of SOT: Is SOT Really Useless?

As is common knowledge, English sentences obtained as a result of SOT often exhibit absurdities. Consider again (25a,25b). Why should (25a) be non-grammatical? Maria is playing tennis right now. The speaker has just heard Peter utter *Maria is playing tennis* and needs to report Peter's words. But instead of producing (25a), the expression perfectly matching the status quo, the speaker

is obliged, at least according to standard grammar, to produce (25b) – "to observe SOT":¹³

- (25) a. *Peter said Maria is playing tennis
b. Peter said Maria was playing tennis

Why? Maria is still playing tennis! Producing *Maria was playing tennis* as part of (25b), the speaker is using an expression contradicting reality. The obvious conclusion to make is that if in English the speaker is forced to build and use expressions totally contradicting reality, this can by no means be coincidental. It cannot be a "bug in the language engine" in the human brain, nor can it be a "whim of human nature". It must be due to a necessity to avoid something extremely important – something even more important than the natural necessity for language to reflect reality. What can this be?

On the "Usefulness" of Backshift Mood (SOT)

As already established, Bulgarian forms such as *igraela* "played [renarrative]" and *igrae* "plays [present, indicative]", as in (34c), and English forms such as *was playing* (in 25b) represent "cancelable content":

- (35) Maria *igraela*_{imperfPresRenarr} tennis
"Maria played tennis"
(34) c. Petar *kaza*_{perfAorIndic}, *che* Maria *igrae*_{imperfPresIndic} tennis
(literally) "Peter said Maria plays tennis"
(25) a. *Peter said Maria is playing tennis
b. Peter said Maria was playing tennis

Suppose the word string *is playing tennis*, containing no backshift, is allowed in an English sentence such as (25a). Language expressions in main clauses such as *X said* or *X thought*, etc., envision two possibilities or scenarios: for the content carried by *said*, *thought*, etc., either to be true or not to be true (i.e., these are possibilities) or to happen or not to happen (these are scenarios, usually for the future). In other words, for the content of *said*, *thought*, etc., basically two possibilities are entailed: true or not true. This means that if a sentence such as (25a) above or (36a,36b) below were allowed in English, the relevant verb expressions (here *is playing tennis*, *plays tennis*, *has played tennis*) could represent "cancelable content" even, and also, when used on their own.

¹³I take it that (25a) is non-grammatical, although it sounds acceptable to some native speakers. According to many grammarians also (e.g., Huddleston & Pullum 2002: 155-156) backshift is sometimes optional, sometimes obligatory. A distinction ought to be made between grammaticality/non-grammaticality as language-system notions, and acceptability/non-acceptability as pragmatic notions.

- (36) a. *Peter thought that Maria plays tennis
b. *Peter knew that Maria has played tennis

This would be inadmissible!

Not only is Backshift Mood (SOT) Useful, its Absence Would Be Catastrophic

If sentences in English such as (37a,37b), etc. were to explicate "cancelable content", this would mean, for example, that a simple sentence such as *Maria plays tennis* would not necessarily mean that Maria plays tennis. It would mean that Maria **either** plays (is playing) tennis **or** does not play (is not playing) tennis.

- (37) a. Maria is playing tennis
b. Maria plays tennis

In a state of affairs of this kind – hypothetical, of course, English (37b) would correspond to Bulgarian (35), a sentence which approximately means "they say Maria plays tennis, but this is questionable; actually, Maria either plays tennis or does not play tennis":

- (35) Maria igraela_{imperfImperfRenarrDubit} tenis
"Maria played/was playing tennis [renarrative, dubitative]"

This would mean that even such a stage could be reached in the language development (hypothetical) in which **any** simple English sentence would carry the meaning that the situation depicted by it is not at all necessarily true. In a state of affairs of this kind (hypothetical), if we are told *It is raining*, we would not believe the speaker. Upon hearing this, we would have to assume that it is **either** raining outside **or** the sun is shining, with no cloud seen in the sky. A state of affairs of this kind in a language is inadmissible, unthinkable.

Bulgarian as a Metalanguage for Revealing the Usefulness of Backshift Mood (SOT)

This hypothetical state of affairs (inadmissible, unthinkable) just described is not easy to imagine in English, when English is used as a metalanguage. On the other hand, Bulgarian offers a convenient platform for revealing it. If someone produces (38a) in Bulgarian, the speaker's words are to be believed. They are not supposed to be questioned or doubted. But if someone produces (38b), it means the following. The speaker is saying that someone said it was raining. Hence, it may be raining – but it may also not be raining. This is because (38b) contains *valyalo* "rained", a renarrative form, which is also dubitative:

- (38) a. Vali_{imperfPresIndic} dazhd
 "It is raining/rains"

- b. Valyalo_{imperfImperfRenarrDubit} dazhd
 "It is raining/rains [reportedly, allegedly ...]"

A renarrative form represents cancelable content by virtue of its nature. It registers an assertion of a third party, and an assertion of a third party is either true or not true, as a default. In addition to this, or perhaps as a result of this, forms such as *valyalo* in (38b) are also dubitative.

If English did not feature backshift mood to mark cancelable content, all normal simple sentences in English would have content that would be cancelable, i.e., subject to doubt. Simply put, a language like this cannot serve its purpose to provide meaningful information between participants in communication.

English Backshift Mood: A Device to Prevent the Elimination of Non-Cancelable Content

Thus, the analysis of the exactly diametrically opposing cases of non-grammaticality, discussed in detail above and exhibited in sentences in two separate languages, Bulgarian and English, leads to the following final generalization.

The reason for the existence of SOT in English is that its systematic observation – even in sentences drastically contradicting reality – prevents the elimination of non-cancelable content found in most verb forms in the indicative mood within the tense system of English. SOT can be called backshift mood, and in a system built in this way, backshift mood will contain forms that represent cancelable content. Its counterpart, the indicative mood, will contain forms that represent non-cancelable content.

English with a "Backshift Mood"

The generalization allows the postulation of the existence of a backshift mood in English. It encompasses all cases of tense verb forms obtained after the observation of SOT. Backshift mood can be described as similar to the conditional mood represented by *would*. However, backshift mood can also be said to have a larger scope. Furthermore, while conditional mood is taken to exist in English despite the fact that it practically has no exponent – in the sense that *would* is used for the explication of other grammatical categories as well (future-in-the-past, habituality), the distinction between backshift mood and indicative mood is reflected in a clear paradigm. Most of the forms in the backshift mood paradigm are different from those of the indicative mood paradigm. Unlike the English conditional mood with *would*, not all English backshift forms manifest grammatical homonymy.

Thus, the reason why English SOT has so far been regarded as "a rule" rather than as a grammatical category (e.g., a mood) appears to lay in the absence of an explanation about the exact meaning of forms obtained as a result of SOT. The analysis above, initiated with Bulgarian data and transferred onto English material,

shows that forms obtained from the SOT (backshift) operation in English represent cancelable content. These backshift forms stand in stark contrast to their indicative mood counterparts – in that the indicative forms represent non-cancelable content.¹⁴

Backshift Mood in English as Cancelable Content

There can hardly be any doubt that English backshift mood forms represent cancelable content. This is because the cancelability of the content of the relevant form is imposed by the use of verbs such as *said*, *think*, etc. in the main clause of the relevant schema. For example, the form *is drinking* in a simple sentence such as (39a) below represents non-cancelable content. If John is drinking beer, he is drinking beer, and he is not drinking vodka or ouzo. It also cannot be the case simultaneously that John is drinking beer and is not drinking beer. If we have a sentence such as (39b), the content in the quotation marks is again non-cancelable content. But if we have a sentence such as (39c), the clause *John was drinking beer*, and especially the form *was drinking* in it, represent cancelable content. This is because the main clause *Maria said* imparts cancelability to the content that was previously signified by "*John is drinking beer*", and was non-cancelable.

- (39) a. John is drinking beer
 b. Maria said "John is drinking beer"
 c. Maria said (that) John was drinking beer

Through the backshift mood, or through SOT – these two ultimately standing for one and the same phenomenon – the English language marks as cancelable the content of subordinate clauses complemented to main clauses such as *X said*, *X thought*, etc., so that it should stand in contrast to the non-cancelable content of the relevant forms in the indicative mood.

¹⁴An anonymous reviewer of this paper raised reasonable objections to the inclusion of backshift in the system of moods in English, arguing that it is so well established that its reconstruction would lead to difficulties. Having taught English for thirty years, I myself do not feel completely at ease to have to propose a reinterpretation of the system of moods. However, certain alternative solutions – backshift as a subtype of tense or as a totally new grammatical category – appear to me much less reasonable, and even more impractical.

English Backshift Mood Forms Cannot Stand Alone

Note that forms in English in the indicative mood, manifesting non-cancelable content, can, of course, be used alone. However, backshift mood forms cannot stand alone. They only appear after main clauses such as *X said, X thought*, etc. On the other hand, backshift forms can by no means be described as some peripheral or exotic phenomenon in English. This is because expressions such as *X said, X thought* and many other similar ones with a following subordinate clause can be said to represent a schema with an extraordinarily high prevalence – in the English language and in many other languages (in both the official style and in everyday communication).

Bulgarian Renarrative Forms Can Stand Alone

In contrast to English, where backshift forms, representing cancelable content, cannot stand alone, renarrative forms in Bulgarian, also representing cancelable content, **can** stand alone. Thus the renarrative with cancelable content *valyalo* "rained [renarrative]", discussed above, can represent a whole sentence entirely on its own:

- (40) Valyalo_{imperfAor/ImperfRenarrDubItInfer}
"It rained/was raining [inferential or renarrative and dubitative]"

(40) has two meanings. One is inferential, which is not an issue here. The other one is renarrative, with cancelable content, which is the one to be considered. (40) can be an answer to a question in the following dialogue. I am asking my wife to ask Peter what is happening outside. My wife asks Peter (41a). Peter's reply is (41b):

- (41) a. Kakav e_{imperfPresentIndic} tozi shum navan?
"What is this noise outside?"
b. Vali_{imperfPresentIndic}
"It is raining"

Then I ask my wife (42), she turns to me and utters (40):

- (42) Kakvo kaza_{perfAorIndic} Petar?
"What did Peter say?"
(40) Valyalo_{imperfAor/ImperfRenarrDubItInfer}
"It rained/was raining [inferential or renarrative and dubitative]"

The form *valyalo* "rained/was raining" [renarrative/dubitative] and the whole sentence (40) renarrate Peter's answer (41b), which is non-cancelable content; (40) is a renarrative and dubitative form – it represents cancelable content and shows that the speaker is not sure whether it is really raining (rains) or not.

Note that (40) above does not imply some mistrust – either my wife's or somebody else's – towards Peter. Peter is a trustworthy person. However, my wife, the speaker, assumes that Peter can be wrong in his assessment (otherwise sincerely made) that it is raining. For example, what Peter heard and thought it was rain may have been noise from cars, not rain. When other people's words are reported, as a general principle, they are treated as either true or not true, they are regarded as cancelable content. Conversely, assertions in English on their own such as *It is raining* or *It rained*, or *John drinks beer*, or *Maria played tennis yesterday* are non-cancelable content.

Conclusion

English SOT/backshift prevents the elimination of non-cancelable content from sentences of a particular type, discussed in detail above. The analysis and its results support a solution that the SOT/backshift phenomenon in English can be regarded as a mood, in spite of the difficulties associated with the necessary restructuring and reinterpretation of the grammatical categories of the verb system. Of course, the idea put forward here to incorporate SOT into the system of moods is just a proposal for a solution. What exactly there is – or must be – in a grammar of a language such as English is decided not by a single researcher at the spur of the moment, but rather by a consensus of the community of linguists and grammarians over time. In any case, however, the idea for SOT/backshift to be regarded as a mood certainly appears much more reasonable in comparison with the following three other possible solutions: (i) to include it somehow in the tense system; (ii) to posit the existence of an entirely new grammatical category; (iii) to dismiss it as a "pragmatic phenomenon".

In Bulgarian, and probably also in other languages, backshift mood (SOT) can be said to be generally unnecessary in the sense in which it is necessary in English. There are separate forms available that represent non-cancelable content, and other forms that represent cancelable content. Still, the Bulgarian language material dealt with here fosters a conclusion that this language actually features both SOT rules and "sequence of mood" rules, with an even higher degree of obligatoriness than English SOT. But this is an issue in need of a lot of future research. What is important is that the inquiry, starting from data in a rather different language, provided an explanation of the *raison d'être* of English SOT: a device to prevent the elimination of non-cancelable content.

Acknowledgements

I am grateful to the anonymous reviewers of this article, who provided valuable comments and suggestions. Hearty thanks, and an apology, to my wife who stoically endured a question thrown at her a thousand times: "Why is **Petar kaza, che pristigna* "Peter said he arrived" a non-grammatical sentence?"

References

- Abusch D (1997) Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20(1): 1-50.
- Abusch D (2004) On the temporal composition of infinitives. In J Gueron, J Lecarme (Eds.), *The syntax of Time*. Cambridge, Mass: MIT Press, pp. 27-53.
- Altschuler D (2010) Aspect in English and Russian flashback discourses. In A Grønn, I Marijanovic (Eds.), *Russian in Contrast, Oslo Studies in Language* 2(1). Oslo, pp. 75-107.
- Bulatovic V (2013) Modern theories of aspect and Serbian EL2 learners. *Belgrade English Language & Literature Studies* V: 65-79.
- Bulgarian Grammar (1983) *Gramatika na savremenniya balgarski knizhoven ezik. Tom 2. Morfologiya* (Grammar of contemporary Bulgarian literary language. Vol. 2. Morphology). Sofia: Bulgarian Academy of Sciences.
- Chung K-S (2007) A two-tiered analysis of sequence of tenses in English. *The New Korean Journal of English Language and Literature* 49(4): 281-304.
- Comrie B (2009) Tense in indirect speech. *Folia Linguistica* 20(3-4): 265-296.
- Declerck R (1995) Is there a relative past tense in English? *Lingua* 97: 1-36.
- Downing A, Locke P (2006) *English Grammar. A University Course*. London, New York: Routledge.
- Enç M (1987) Anchoring conditions for tense. *Linguistic Inquiry* 18: 633-657.
- Franks S (2008) Clitic placement, prosody, and the Bulgarian verbal complex. *Journal of Slavic Linguistics* 16(1): 91-137.
- Grice HP (1975/1989). Logic and conversation. In P Cole, J Morgan (Eds.), *Syntax and semantics III: Speech acts*. New York: Academic Press, pp. 41-58.
- Grønn A, von Stechow A (2010) Complement tense in contrast: the SOT parameter in Russian and English. In A Grønn, I Marijanovic (Eds.), *Russian in Contrast, Oslo Studies in Language* 2(1). Oslo, pp. 109-153.
- Huddleston R, Pullum GK (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Kabakchiev K (2000) *Aspect in English: a "common-sense" view of the interplay between verbal and nominal referents*. Dordrecht: Kluwer.
- Kabakchiev K (2017) *An English Grammar: main stumbling blocks for Bulgarians learning English*. Stuttgart: Mariana Kabakchiev Verlag.
- Kubota Y, Jungmee L, Smirnova A, Tonhauser J (2009) On the cross-linguistic interpretation of embedded tenses. In A Riester, T Solstad (Eds.), *Proceedings of Sinn und Bedeutung* 13. University of Stuttgart, pp. 307-320.
- Langacker RW (2008) *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
- Maslov Y (1982) *Gramatika na balgarskiya ezik (Grammar of the Bulgarian language)*. Sofia: Nauka i izkustvo.
- Ogihara T (1994) Adverbs of quantification and sequence-of-tense phenomena. In M Harvey, L Santelmann (Eds.), *SALT IV*. Ithaca, N.Y.: Cornell University, pp. 251-267.
- Ogihara T (1996) *Tense, Attitudes and Scope*. Dordrecht: Kluwer.
- Ogihara T (2015) Relative tense in relative clauses. *Journal of Japanese Linguistics* 31: 33-56.
- Pashov P (2013) *Balgarska gramatika (Bulgarian grammar)*. Sofia: Hermes.
- Quirk R, Greenbaum S, Leech G, Svartvik J (1972) *A Grammar of Contemporary English*. London: Longman.

- Sauerland U, Schenner M (2007) Embedded evidentials in Bulgarian. In E Puig-Waldmüller (Eds.), *Proceedings of Sinn und Bedeutung 11*. Barcelona: Universitat Pompeu Fabra, pp. 525-539.
- Shaer B (1997) Universal grammar and the parametrization of temporal marking. *CLS (Papers from the Regional Meeting of the Chicago Linguistic Society)* 33: 271-286.
- Smirnova A (2009) A case against "defective" tense in the Bulgarian subjunctive. *FASL 17*. Michigan Slavic Publications, pp. 200-215.
- Verkuyl H (1972) *On the Compositional Nature of the Aspects*. Dordrecht: Reidel.
- Verkuyl H (1993) *A Theory of Aspectuality: the interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press.
- Zagona K (2014) Sequence-of-tense and the features of finite tenses. *Nordlyd* 41(2): 261-272.