



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/104853/>

Version: Accepted Version

Book Section:

Vanek, Norbert (2016) 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In: Howard, Martin and Leclercq, Pascale, (eds.) *Tense-Aspect-Modality in a Second Language: Contemporary perspectives*. Studies in Bilingualism. John Benjamins Publishing Company, Amsterdam, pp. 143-180.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

'It starts to explode.' Phasal segmentation of contextualised events in L2 English

Norbert Vanek

University of York

Abstract

En route to acquiring novel principles of temporal information organisation in the target language, second language (L2) learners exhibit a capacity to build temporal constructions of their own, which are not necessarily fixed in the principles of either their source or their target language system. This study surveys hitherto unattested interlanguage phenomena found in the phasal segmentation patterns of two intermediate-level learner groups with unrelated source languages, and identifies analogies of shared developmental patterns. Film verbalisations and acceptability judgements (AJ) were used to elicit responses from Czech and Hungarian intermediate learners of English, and their analyses yielded a threefold benefit. They generated representative degrees of granularity for each group who experiment with new segmentation techniques. They also showed that the ways in which learners partition events in production (pronounced digression from the target) do not directly replicate patterns in acceptability judgements (closer approximation to the target). And thirdly, overlaps and contrasts between learner and native control speaker preferences for phasal partitioning varied in close relation to specific aspectual properties inherent to the verbs used. The combination of production features and acceptability judgements from L2 groups with distant L1s provides an informative mosaic of how learners at intermediate L2 proficiency strive for an optimal fit when combining available linguistic elements to express specific event phases.

Introduction

Understanding the sensory input from experience or observation of events largely depends on how well we are able to segment the temporally changing flow of situations into meaningful parts. Successful identification of beginnings and endings of event segments is important because it allows us to anticipate the course of development and to plan a fitting response, which depends on whether events do or do not evolve in line with our predictions. Consider the event of a picnic interrupted by rain as an example, in which units such as *unfolding a picnic blanket*, *opening the basket*, *looking up at the sky*, *packing everything up*, *finding shelter*, and so on, represent readily identifiable event segments. Results of behavioural and neuroimaging tests signal that adults (see, for example, Speer *et al.*, 2003) spontaneously partition events into temporal units that are 'reliable, meaningful, and correlated with ecologically relevant features of the action' (Zacks *et al.*, 2001: 651). The fact that we can spontaneously identify appropriate event boundaries in perception, which later help us remember more and learn more proficiently (Zacks & Swallow, 2007), may be crucial for internal cognition. However, it tells us little about links to language, namely about how (much) event segmentation in verbalisation can be modulated by specific language systems. And if it is the case that particular languages do affect speakers' segmentation patterns in different ways, what are the implications for second language learners?

Languages considerably differ in how they encode events and their subunits. Crosslinguistic differences in event segmentation can be attributed to at least two main reasons, language-specific lexicalisation patterns (Talmy, 1985), and availability of syntactic constructions that enable manipulating the 'tightness of packaging' (Bohnenmeyer *et al.*, 2011:47). With respect to contrasts in lexicalisation patterns, it can be argued that the internal divisibility of the event into expressible subunits delineates the basis for segmentation processes. As an illustration, Pawley (1987) reports

that unlike English, the language of Kalam does not provide its speakers with single 'summative' verbs for expressing unit-rich complex events but typically uses serial verb constructions for this purpose (e.g., *hunt* is expressed as *kmn pak dad apl nb okok ad ñbelgpal* '[go forth] > kill game > bring it to camp/home > cook it > eat it > [return home]'; (Pawley, 1987:342-3; square-bracketed units are optional). Despite its intuitive appeal, lexicalisation is important but on its own it is hardly an adequate measure of event segmentation (Givón, 1991) because the lexical level closely interacts with grammatical constraints.

To examine the way in which segmentation principles are linked to grammar, Bohnemeyer and associates propose the notion of *macro-event property* (MEP). In their view, 'an event-denoting construction has the MEP iff it combines only with those time-positional or durational operators that have scope over all subevents it entails' (Bohnemeyer *et al.*, 2011:48) [spelling as in original]. For instance, in *squash the grape flat* or *lick the plate clean*, the subevents expressed by [squash] and [lick] are not available to operators of temporal position or duration at the exclusion of the *flattening, cleaning* subevents, and vice versa; neither are the latter subevents 'temporally separable' from their corresponding first pair parts (e.g., *the chimp squashed the grape flat after the light signal* ≠ *the chimp squashed the grape after the light signal and then it flattened*; or *the dog was licking the plate clean for half an hour* ≠ *the dog was licking the plate for half an hour and then it got clean*). Syntactic properties of this type of tightly-packaged single-verb resultative constructions entailing a state change (unavailable in some languages, e.g., Lao) directly affect segmentation as they exclude the possibility of subevent individuation (unlike e.g., *The chimp squashed the grape and it got flat.*). Besides lexical and syntactic links to segmentation, crosslinguistic differencesⁱ in how native speakers tend to decompose events have also been attested in studies on grammatical aspect, an area of immediate relevance to this work.

Empirical context

Grammatical aspect and event segmentation across L1s

Event segmentation preferences linked to grammatical aspect have been examined in the language production of native speakers of e.g., Swedish and Spanish (Bylund, 2011); and German, English and Arabic (von Stutterheim & Nüse, 2003). These studies employed film-retelling as the elicitation technique. The participants were asked to retell the events in the video while watching it or during the pauses immediately after each short episode. Analyses of variance were conducted for the number of partitions verbalised. Results showed that event segmentation differed across groups in correspondence with availability vs. absence of ongoingness in the aspectual system of the given language. Overall, the observation was that speakers of languages with grammaticalised ongoingness tended to opt for a significantly higher degree of event segmentation than speakers of languages lacking an overt and systematic aspectual marker for ongoingness. Specifically, English, Arabic and Spanish production was found to be significantly more fine-grained and phasally decomposed than that of German and Swedish native speakers.

Empirical support for contrasts in temporal conceptualisation linked to the grammatical devices that languages make available to their speakers comes from a range of angles, using diverse methodological approaches, not exclusively film retellings. For instance, Boroditsky & Trusova (2003) compared frequencies of noticing the difference between ongoing and completed actions and found that Russian-English bilinguals tend to notice such difference more frequently than English monolinguals. What the findings from the surveyed studies in this section collectively signal is that a competent analysis of how we segment events needs to involve a careful consideration of the degree to which aspect is grammaticalised in our languages.

Research on event segmentation in learner varieties

Given the crosslinguistic contrasts attested across first languages, further research has naturally led to examining preferences of segmentation in SLA. In particular, advanced learner groups whose L2 differs from their L1 in terms of event construal patterns has generated strong interest. Focusing on event segmentation in L2 English, von Stutterheim & Lambert (2005) reported that although German learners decomposed events and event phases with target-like frequency, they have not managed to fully reorganise their L1 knowledge about information organisation according to L2 principles. To be more concrete, qualitative analyses showed that German learners (similar to French learners of English L2 tested in the same study) tended to overgeneralise phase segmentation leading to incompatibility with situation types that do not denote opening event chains in the stimulus (e.g., *he starts to wonder* for *he wonders/is wondering*) and thus digressed from target-like uses. In a similar study, Bylund (2011) tested late bilinguals and found that L1 Spanish adults with L2 Swedish resorted to a different event segmentation strategy in their film retellings than is typical of their L1 or TL. L2 learners partitioned temporal information into a set of propositions with significantly higher granularity (i.e. they individuated more event components for expression) than the native Swedish speakers, but their production was not as fine-grained as that of the Spanish natives. It needs to be emphasised that both Bylund (2011) and von Stutterheim and Lambert (2005) tested highly proficient L2 users, but whilst the former study examined L2-immersed naturalistic learners, the latter collected responses from L1-immersed instructed learners. It is noteworthy that *hybrid* or *in-between* performance (Jarvis & Pavlenko, 2008: 247; Pavlenko, 2014: 161-2), i.e. incorporating both source and target language segmentation features, was evident in both studies irrespective of acquisition context. This may signal partial reorganisation in the bilingual system as an effect of competing L1 and L2 segmentation principles. Seen from another perspective, learners whose L2 differed from their L1 in terms of encoding ongoingness remained

influenced by event segmentation patterns typical of their first language despite their advanced formal L2 knowledge. Nevertheless, the key reason for the reported L2 performance was attributed to bilingualism effects (Cook, 2002) rather than to direct L1 transfer.

Empirical research on segmentation patterns at more moderate stages of L2 acquisition, although still embryonic in scale, has generated different types of interesting findings. For instance, Noyau *et al.* (2005) examined the developmental trajectories in temporal segmentation of instructed Polish and Swedish teenage and university learners of L2 French through film retellings and picture descriptions. After 6 months of instruction, learners were able to produce meaningful coherent narratives, in which they typically verbalised the skeletal (macro-)events with a low degree of temporal granularity (i.e. with only a few event partitions in their retellings) but with sufficient clarity to convey the main story line. At 30-38 months of instruction (intermediate level), learners were able to modify temporal resolution from less to more finely-grained micro-events, varying the degrees of specificity in accordance with communicative needs (comparisons of movie-retellings and picture descriptions revealed intra-learner variation). The level of coarser or finer granularity was not quantified in this study, which limits its direct comparability with related research.

In a more quantitatively-driven study, Sugaya & Shirai (2007) examined learners' oral production and acceptability judgements to test the effect of verb types on ongoing and resultative event phase marking. Oral picture descriptions of 59 intermediate German, Russian and English learners of L2 Japanese were used to calculate the frequencies of combining *-te i-(ru)* (an imperfective aspect marker in Japanese, but also used to denote a resultative state) with different event types (activity, accomplishment). The findings of this study need to be interpreted with caution because of highly questionable steps in the analyses as a result of grouping Russian learners together with German learners as having [-progressive] L1s. If we look at the more suitably categorised [+progressive] group, English learners combined the form with ongoing contexts significantly more often than with

resultatives, as the authors predicted. In acceptability judgements, there was a significant effect of verb type, interpreted as 'learners found it more difficult to attach *-te i-(ru)* to accomplishment and semelfactive than to activity verbs [...] regardless of L1' (2007:18).

Combining a production task with Acceptability Judgements (AJs) of ongoing/resultative markers for different verb types is by all means a fruitful method of triangulation leading to richer insights into L2 acquisition patterns of phasal event segmentation. We may expect to find that learners' segmentation patterns in AJs approximate closer to the target than those in a more unplanned oral production task because AJs can be considered less taxing in terms of semantic processing (to understand meaning), noticing (to decide whether a construction is problematic), and reflecting (to figure out what is problematic and possibly why) (Ellis, 2004:256). Another difference relates to the cognitive demand when producing vs. judging pre-constructed alternatives (Bialystok, 1982). What has also remained underexplored is the L2 production and judgements not only of ongoing and resultative but also of inceptive phase marking, and not only of activities and accomplishments but of a fuller range of verb types, including state-like and typically instantaneous achievement-like verbs. Stimulated by empirical evidence that points to intriguing non-standard event decomposition patterns in learner varieties (e.g., von Stutterheim & Lambert, 2005), the present study addresses this gap in L2 temporality research.

The main reason why event segmentation in L2 is seen as a particularly engaging area of research is the idea that the acquisition of information packaging principles may be one of the most challenging steps in L2 learning (Carroll & Lambert, 2003). This is because derivation of such principles involves an intricate network of form-function relations across a number of domains (e.g., morphology, syntax, lexicalisation, discourse structure). Also, these principles represent preferences rather than absolute rules (no rule in the grammar of English suppresses coarse-grained segmentation in favour of highly granular event partitioning). To uncover what underlies segmentation principles at

intermediate stages of L2 development, the present work incorporates two layers of innovation.

Firstly, it uses a time-relational analysis (Klein, 1994, see Central notions section below) to examine phasal segmentation preferences revealed through two different task types. Secondly, it analyses responses of intermediate learners from two distant source languages (SL) en route to the same target language (TL) that differs from both SLs in the ways ongoing phases are encoded. This design allows us to examine which segmentation preferences tend to be shared at the intermediate stage.

Central notions and theoretical grounding

The key concept is *event segmentation*, defined as the process of temporal partitioning of information (Noyau *et al.*, 2005). It displays the way more or less complex situations are divided into smaller units such as subevents, processes or states. Segmentation is seen as a subprocess of conceptualisation, i.e., generation of preverbal messages from the knowledge base (Levelt, 1999). Together with information selection, segmentation can be characterised as a macroplanning process (on the level of 'what to say') (Carroll & von Stutterheim, 2003; Habel & Tappe, 1999). The speaker's options for event segmentation can be located on a continuum between low granularity (presenting a single macro-event or a few events with merged components, e.g., *the campers put up their tent*), and high granularity (a fine-grained resolution in the form of event phases or a series of micro-events, e.g., *the campers unpacked the tent, spread it out, assembled the poles, constructed the frame, attached the tarp, and secured all corners*). The level of granularity is quantified in this study by calculating the number of partitions per episode. *Episode* refers to a 'semantic unit in discourse organisation consisting of a set of related propositions [for which sustained attentional effort] endures until attention is diverted, that is, it is sustained until an episode boundary is

reached' (Tomlin, 1987: 460). In this study, the boundaries of each episode in the video stimulus are clearly indicated by pauses (i.e., just before cutting to a new scene, either the camera pans out or the screen blackens). It cannot be ruled out, however, that participants may have perceived episodic boundaries in a different way, based of cues other than pauses. *Partitions* subsume all propositions that form the main structure of the discourse, comprising events marked as ongoing, completed, and VTAs (i.e., verbs specifically used for temporal/aspectual markings, typically used in verbal constructions in conjunction with another verb to mark properties such as beginning points (*she starts screaming*), continuation (*the boy keeps jumping around*), and endpoints of processes (*he stops playing the guitar*)). In other words, partitions exclude all propositions that do not contribute to temporal movement (Klein & von Stutterheim, 1987), namely non-events (i.e., states, negations, conditions), metacommunicative comments (i.e., propositions marked to express the narrator's viewpoint such as *the camera zooms in*), and staging verbs (*the mother appears; it happens*). The difference between partitions and non-events can also be explained in terms of 'grounding' (Berman & Slobin, 1994), i.e., in marking information as part of the main line (foregrounded partitions) or side structure (backgrounded non-events). Only foregrounded information was relevant for coding partitions in this study.

Another key concept is *Aktionsart*, also referred to as lexical aspect, i.e., the inherent lexical temporal quality associated with verbs. Verbs (and verb phrases) exhibit considerably different properties with regards to their internal temporal features, logical entailments, and combinatory possibilities with temporal modifiers. A classic verb type classification is Vendler's (1957) division of lexico-aspectual categories into states, activities, accomplishments, and achievements (1957:143-160). This four-tiered taxonomy has undergone various modifications, for instance, the aspectual properties were reduced to two-feature oppositions, namely [\pm telic] and [\pm stages] (Rothstein, 2004). This division is of relevance to segmentation as it directly pertains to decomposability of events into

phases, and renders the following classification: (a) states [-stages-telic] (e.g., *believe*); (b) activities [+stages-telic] (e.g., *walk*), (c) accomplishments [+stages+telic] (e.g., *fill an aquarium*), (d) achievements [-stages+telic] (e.g., *land*). Rothstein's (2004) account excludes the combinatory potential of states and achievements with progressive or other phasal marking. She argues that 'achievements do not extend over time but are instantaneous events, and thus stages cannot be distinguished [and] states are non-dynamic, so that every bit is exactly the same as every other bit and therefore no stages can be distinguished' (2004: 12).

Previous research has shown that distinguishing between [\pm telic] and [\pm stages] features in early learner varieties is rarely a straightforward task, since learners often use one verb for many functions (e.g., 'see for see, look, and watch; or search for search and find' (Starren, 2001: 31). The present work addresses this challenge in two ways: (a) by employing a semantically constrained stimulus-based production task plus a contextualised judgement task (see Method section); and (b) by adopting the Basic Time Structure (BTS) (Klein, 1994, 2009) as the analytical framework. Within the BTS, the temporal properties reflected in the lexicon are defined on the basis of how they relate to topic time (TT, i.e., the time about which something is asserted), as shown in Figure 1.

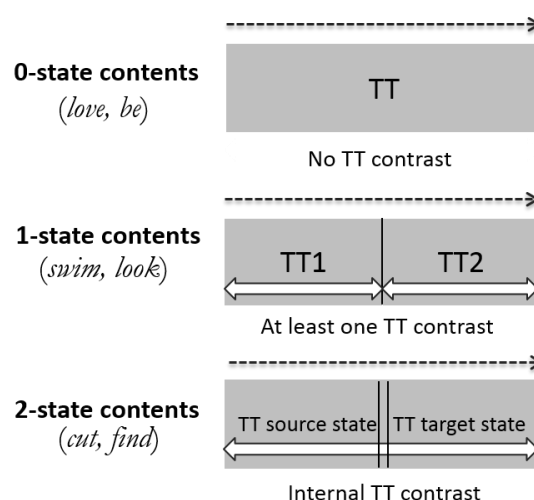


Figure 1. Illustration of the relationships between topic time (TT) and three verb types differing in lexical aspectual properties (0-state contents, 1-state contents, 2-state contents).

0-state contents can be considered atemporal since they typically do not include a TT contrast (i.e., if 0-state contents are linked to a specific TT, it follows that they are automatically linked to any other TT because there is rarely any TT for which the given situation is not true) as in *to be Faroese* or *to love cheesecakes*. Inceptive phase marking in this case induces a change in state, i.e., it modifies the internal temporal constituency that transforms 0-states contents into 1-state contents (e.g., *Olaf started to love cheesecakes*). 1-state contents involve a TT contrast on both sides. This means that a lexical content such as *she is swimming in the sea* is normally limited in time and thus there is a topic time before and also after which she is not swimming. Inceptive phase marking readily combines with 1-state contents because they are inherently dynamic and durative, which also means their TT necessarily extends over some time divisible into subintervals. This does not hold the same way for 2-state contents. 2-state contents include a source state and a target state (e.g., in temporal order, the source state for *to wake up* is *to be asleep* and the target state is *not to be asleep*). For 2-state contents there is a greater variety for positioning topic time. Speakers can locate TT in the source state (just before waking up), it can be in the target state (just after waking up), it can include part of both (partly after waking up), or it can be in part of the target state and the state that follows it (long after waking up) (Klein, 2009: 31-32). Another possibility for 2-state contents is to position TT fully within the phase of transition (while waking up). The lexical content remains identical but what changes is its relation to a topic time, which depends on the internal temporal properties of the individual events that are described.

With respect to temporal partitioning, typically durative 2-state contents (accomplishments) are more hospitable to phasal decomposition (e.g., *the bird starts to build a nest*) than those that typically occupy very brief time spans (achievements, such as *the dog starts to catch the ball*). Nevertheless, the compatibility between characteristically momentary event types and phasal

marking cannot be excluded (e.g., *the company starts to achieve great success*) and its suitability remains a question of perspective. Speakers can defocus temporal boundaries or transition points between one state and the next, and can thus make way for a more granular representation of an achievement-type event with TT located in one of its phases (e.g., *the fireman is starting to extinguish the fire*). If we allowed temporal analysis to hinge purely on [+/-] contrastive features, we would face the risk of missing the fuller temporal picture especially in non-standard form-meaning combinations characterising learner varieties (e.g., Bardovi-Harlig, 1992). By focusing on the location of topic time, the BTS is deemed more flexible than the contrastive feature analysis in also embracing less typical learner-specific configurations.

The third key concept is *grammatical aspect*, which is separate from Aktionsart. It is important to keep the two as distinct layers of analysis because Aktionsart (expresses temporal properties internal to the verb) allows a more language-neutral approach, but grammatical aspect (expresses external temporal properties) operates in very different ways across the source and the target languages which is briefly outlined in the following in relation to the languages of relevance here, namely English, Czech and Hungarian.

Structural contrasts in the source and the target languages

The most pronounced crosslinguistic differences directly linked to segmentation lie in ongoing phase marking exclusively via grammatical means. In brief, English obligatorily uses the *V+ing* form to mark ongoingness across tenses. The regularised imperfective marker provides a steady structural base for English speakers to zoom in on particular event phases and defocus temporal boundaries in relevant contexts (in inceptive phases, e.g., *the passengers are starting to panic*; in

ongoing phases, e.g., *the neighbours are decluttering their garden shed*; as well as in terminative phases, e.g. *the lion is finishing his dinner*).

In contrast, the Czech grammar does not equip its speakers with an equivalent to *V+ing*. Although its lexical phase markers are similar to those in English (e.g., *svět začíná*(IMPERF) *panikařit* 'the world is starting to panic', *fanoušci ztrácí*(IMPERF) *trpělivost* 'the fans are losing their patience'), Czech reserves pure grammatical ongoingness marking only for a small fraction of verbs (Schmiedtová, 2004) using the imperfectivising suffix *-va* (as in *dát/dávat* 'to give/to be giving'). Further differences immediately relevant to the expression of ongoingness are that (a) most aspectual pairs (typically formed from imperfectives via perfectivising prefixation such as *pít/vypít* 'to be drinking/to drink up') are only approximate semantic counterparts, (b) imperfectives are often used for semantic reasons to 'highlight failure to achieve the goal', as in *celý den jsem kupoval*(IMPERF) *kravatu, (ale žádnou jsem nekoupil*(PERF)) 'I spent the whole day buying a tie (but didn't get any)', and (c) an imperfective form may be used to express perfective meaning, as in *tu knihu jsem četl*(IMPERF) *hodně dávno* 'I read that book ages ago' (Short, 2002:481). Further directly relevant explorations have shown that the imperfective form is not used by Czech speakers to encode events as ongoing in contexts where it is used to do so by speakers of Russian (Schmiedtová *et al.*, 2011; 2013), and that the Czech imperfective co-occurs with means that denote event completion, a preference observed in German and Dutch but not in Arabic and English (von Stutterheim *et al.*, 2012). The differences between Czech and English ongoingness marking fuel the assumption of a likely variation when speakers of Czech vs. English express event phases.

Hungarian completely lacks a specialised grammatical aspectual marker for ongoing phases (Csirmaz, 2004; Takács, 2012), and thus simple verb forms oscillate in ambiguity between perfective and imperfective readings (as in *a piac kezd*(AMBIV) *pánikolni* 'the market starts/is starting to panic' *a pszichológus válaszol*(AMBIV) *a kérdésekre* 'the psychologist answers/is

answering the questions'). Disambiguation in these cases can be accomplished with the assistance of durative vs. punctual temporal adverbials (as in *az utcát éppen lomtalanították* 'they were decluttering the street **at that time**' *az utcát rögtön lomtalanították* 'they **instantly** decluttered the street'), gaining an imperfective vs. perfective reading respectively (e.g., Kiefer, 2006; Péter, 2008). With complex verbs (simple verb + one of around forty coverbs such as *ki* 'out(wards)' or *le* 'down(wards)' (Abondolo, 1998:445)), speakers have the possibility to highlight the ongoing phase of an event by locating the coverb postverbally (*amikor a cápa harapott bele a szörfdeszkába* 'when the shark was biting **into** the surfboard') or to express the event as perfective by positioning the coverb before the verb (*amikor a cápa beleharapott a szörfdeszkába* 'when the shark bit into the surfboard'). Viewed on a phase marking continuum, given their structural properties, Hungarian as well as Czech are predicted to gravitate closer to the lower end of phasal granularity, while English is expected to be towards the higher end in the same contexts. Should L1-patterns resist reorganisation in the L2, both Czech and Hungarian proficiency-matched learners would digress from target-like segmentation patterns.

Research questions and hypotheses

The research questions (Qs) and the corresponding hypotheses (Hs) are:

Q1: What event segmentation patterns characterise intermediate learners' language production? Are granularity degrees, grounding preferences, and phasal partitioning in learner production concordant with corresponding source language patterns, target language patterns, or do learners follow different principles?

H1: Intermediate learner varieties represent a developmental stage where temporal reference is typically marked by lexical means (i.e., lexical temporal boundary markers such as *start*, *finish*)

(Klein & Perdue, 1997). Based on this, it is predicted that learners at this stage are sufficiently equipped to modulate granularity depending on communicative needs. Source language patterns are expected to surface as a result of L1-driven principles of organising event information (Carroll & von Stutterheim, 2003) alongside learner-specific interlanguage features (Cook, 2002).

Q2: Are the event segmentation patterns detected in production reflected in learners' acceptability judgements? And further, how does segmentation of inceptive and ongoing event phases vary based on lexical aspectual properties of verbs?

H2: Acceptability judgements are predicted to mirror preferences exhibited in temporal information segmentation in retellings. On the finer level of specific event types, learners can be expected to extend acceptance of inceptive and ongoing phase marking to verbs not typically decomposable into phases (0-states, 2-state instantaneous) as a result of rule overgeneralisation that is characteristic of this developmental stage (Dietrich *et al.*, 1995).

Method

Participants

Film retellings were elicited from 30 learners matched for moderate proficiency (15 L1 Hungarian – L2 English learners recruited from intermediate level classrooms in Budapest, 15 L1 Czech – L2 English intermediate level learners tested in Prague). All participants had completed secondary education and their active use of the TL was 4 hours per week on average in an instructed setting (more details in Table 1).

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

Table 1. Summary of details about participants in the film retelling task and the acceptability judgement task (AJT), including age at the time of testing (mean+range), percentage of using English in a typical day, and the age of onset of learning English (AoL, mean+range)

Task	Group	Mean age	% of English use	Mean AoL
Retelling	CZL1ENL2 (N=15)	26.3 (18-50)	9.1% (0-25%)	17.9 (12-45)
	HUL1ENL2 (N=15)	20.8 (18-53)	7.1% (0-15%)	19.2 (12-35)
	ENL1 (N=15)	21.1 (18-42)	100%	
	CZL1 (N=15)	22.9 (19-50)	-	
	HUL1 (N=15)	27.2 (19-55)	-	
AJT	CZL1ENL2 (N=20)	32.8 (19-46)	4.6% (0-40%)	15.9 (9-37)
	HUL1ENL2 L1(N=20)	22.6 (18-27)	9.9% (0-35%)	10.9 (4-15)
	ENL1 (N=20)	21.9 (18-53)	100%	

All learners were allocated to intermediate level General English classes based on their result in the Oxford Placement Test 2 (Allan 2004) which corresponded to the B1 level on the Common European Framework of Reference (CEFR). Some of the main B1 level descriptors for spoken interaction and production include the learners' ability 'to enter unprepared into conversation on topics that are familiar or pertinent to everyday life; to connect phrases in a simple way in order to describe experiences and events; to narrate a story; to give brief reasoning and explanations' (CEFR 2001:27). All these skills were important to successfully accomplish the set production task.

Acceptability judgements were collected from 40 learners who were not involved in the film retellings, and from 20 monolingual L1 English controls. The learners were matched for language pairs and L2 proficiency with the film retelling group. None of the learners had stayed in an English speaking country for longer than a month, and none of the English controls resided in a non-English speaking country for longer than 6 consecutive months.

Instruments

The first task was a film retelling (example video frames in Appendix B). The story is about a teenage boy at his birthday celebration, which is filled with a rich collection of naturally unfolding and mutually related event sequences. Each part contains consecutive, simultaneous, as well as partially overlapping events, letting the speaker freely decide about the degree of event segmentation. The instructions were to carefully follow the animation divided into six parts by pauses, and to say during the pauses *what happens* in each episode, in a way that a film-maker who has not seen the story could imagine and reproduce the events as accurately as possible. Each participant saw the video only once. After the recording stage, each film retelling was transcribed and coded for partitions, episodes, non-events and metacommunicative comments (following the definitions in Central notions section). Subsequently, granularity indices were calculated for each retelling by dividing episodes by the number of partitions, which were then subjected to between-subject analyses of variance (see Analysis section). A portion of the retellings (>25%) from each group was independently coded for propositional units and partitions by two native speakers of the output language. The reliability check showed a high level of consistency, exceeding 87.3% ($K > .9$) for each of the three pairs of coders.

The second task was an acceptability judgement, for which the Qualtrics® software was used to provide participants with 10 sets of situation descriptions (example set in Appendix A), each consisting of 4 situation types (0-state, 1-state, 2-state cumulative, and 2-state instantaneous), presented in a randomised order. Participants were asked to (a) read and briefly imagine the situation as described, then (b) read the four options under each situation and decide how suitable they found each option to express the given situation. A four-point scale was provided per option (1-perfectly suitable, 2-more or less suitable, 3-somewhat unsuitable, 4-completely unsuitable). Content validation was three-fold. Two English native speakers (applied linguists) independently

rated each situation for event type (0-state, 1-state, 2-state cumulative (CUMUL), 2-state instantaneous (INST)) and imageability (easily imaginable, somewhat problematic, difficult to picture). Only items with 100% intra-rater agreement on both type and imageability were included in the experiment. This set was further reduced during the comprehensibility check with a Czech and a Hungarian intermediate level learner to include only those items that were neither problematic to imagine (self-reported) nor to paraphrase/translate into their respective L1s. To check the reliability of the judgement survey, Cronbach's Alpha coefficients were computed, which showed high degrees of internal consistency ($\alpha=.949$ within Czech learners, $\alpha=.928$ within Hungarian learners, $\alpha=.952$ within English L1).

The status of acceptability judgements in SLA research is controversial because the knowledge type they reflect (implicit vs. explicit) is problematic to discern (Ellis, 2005; Loewen, 2009; Sorace, 1996). This elicitation technique is therefore employed here solely for exploratory purposes, rather than a base on which wide generalisations can be built. The main motive for employing AJs is their potential to serve as a useful cross-task validation tool (Chaudron, 2003) to explore the points at which judgements and production patterns converge. In tandem, AJs can thus usefully complement production data and help accumulate information about the targeted linguistic abilities.

Analysis

Both parametric and nonparametric statistical tests were used, depending on the type of data in each task. For the film retelling task, a parametric one-way between-subject ANOVA proved suitable because the task elicited continuous and normally distributed data. This choice was made after checking each group's granularity degrees for skewness and kurtosis. The deviation of values from symmetry around the mean (CZL1=.31, HUL1=.81, ENL1=.92, CZL2=.12, HUL2=.44) as well as

the distribution around the peak (CZL1=-.58, HUL1=-.35, ENL1=.92, CZL2=-1.18, HUL2=.46)

were found within the standard range of acceptable skewness and kurtosis of ± 2 .

The second task elicited ordinal data, which necessitated the use of nonparametric statistical procedures to test between and within-group variability. In the first step, Kruskal-Wallis tests (including all 3 groups) served to determine whether there was a significant between-group difference in acceptability judgements. Separate Mann-Whitney U tests (two groups per test) were then used to check the extent to which responses in each learner group differed from those in the target group. On the level of within-group differences, Wilcoxon signed-rank tests were run to examine how individual groups' judgements differed for related options (e.g. acceptability of *V+ing* vs. *start V+ing* in responses to the same question).

Results

Production

Table 1 shows the results of two one-way between-subject ANOVAs comparing granularity indices (GIs, i.e., the mean number of partitions per episode) between L1 groups. Analyses of the relationships between segmentation patterns and language group revealed significant differencesⁱⁱ in the levels of granularity between L1 groups [$F(2,41)=3.99, p<0.05$]ⁱⁱⁱ.

Table 2. Comparison of GIs and mean percentages of foregrounded propositions (1-state + 2-state contents) for film verbalisation of L1 and L2 groups (standard deviations in parentheses)

	GI - mean number of partitions per episode	% of foregrounded propositions (1-state + 2-state)
English L1	9.19 (2.58)	57.0%
Czech L1	7.53 (1.65)	58.7%
Hungarian L1	7.98 (2.20)	63.0%
Czech learners	3.22 (0.80)	48.6%
Hungarian learners	3.83 (1.11)	49.7%

Post hoc tests (Tukey HSD for the production data throughout) showed that the average number of events per episode in English retellings ($M=9.19$, $SD=2.58$) significantly differed from the average number of events encoded in the Czech ($M=7.53$, $SD=1.65$) and the Hungarian retellings ($M=7.27$, $SD=1.87$), with no statistical difference found between the latter two groups. This result (reported in Vanek (2012: 149) and imported here to serve as a comparative baseline) indicates that the degree of event partitioning in English L1, for the given type of discourse, is typically higher than that in Czech and Hungarian L1s. The percentages in Table 1 also show the proportion of propositions expressed as part of the foreground (typically 1-state and 2-state contents) in contrast with the remaining propositions forming the side structure (typically 0-state contents). These numbers may give an impression that learners show a general trend to include more background information. However, the between-group differences were not statistically significant. No significant differences in foregrounding were identified between the three L1 groups, suggesting that the processes of segmentation and foregrounding operate independently.

As for granularity in intermediate learner production, the yet nebulous communicative proficiency at this stage of L2 English development (Gyllstad *et al.*, 2014) expectedly gave rise to a less fine-grained event partitioning than that of the respective L1 groups; [$t(16.99)=7.22$, $p<.001$] for Czech learners ($M=3.22$, $SD=0.80$) vs. Czech native speakers ($M=7.53$, $SD=1.65$) and [$t(27)=5.72$, $p<.001$] for Hungarian learners ($M=3.93$, $SD=1.24$) vs. Hungarian native speakers ($M=7.27$, $SD=1.87$).

Table 3. Frequency of inceptive phase markers from the total number of partitions per retelling (means per group and standard deviations in parentheses)

	Inceptive phase markers	% of inceptive phase markers
English L1	4.13/55.13	7.5% (4.6)
Czech L1	3.73/48.00	7.8% (4.4)
Hungarian L1	2.93/43.13	6.8% (3.9)
Czech learners	2.27/19.33	12.5% (10.2)
Hungarian	2.33/23.00	10.0% (9.3)

The following step examines phasal segmentation by quantitatively comparing inceptive phase markers between groups. Table 2 shows the average frequency of inceptive phase markers per mean number of partitions in each group. On the level of group means, both the Hungarian and the Czech learners exceeded the average percentage of inceptive markers used by the TL controls as well as by the corresponding L1 controls. However, when the between-group differences were tested for degrees of variance (one-way between-subject ANOVAs), the frequencies of inceptive phase markers were not found to be significantly different, either between the L2 learners and the English L1 controls [$F(2,42) = 1.26, p=.29$] or between the three L1 groups [$F(2,42) = .096, p=.91$]. The statistical fact that responses to the same stimulus significantly differ in overall granularity degrees but not in the frequency of inceptive phase markers invites further investigation beyond the quantitative level of group means.

To inspect event segmentation contrasts and phasal marking from multiple angles, the next step qualitatively compares representative discourse fragments per group. Examples (1-5) below are all related to the same part of visual input (Appendix B, episode 4.1-4.7).

- (1) Example of event segmentation typical of English L1
 - a. *The mother is **about to start icing** her son's birthday cake.* (INCEPT+1state)
 - b. *She is humming to herself* (1state)
 - c. *as she **starts icing** it using a piping bag.* (INCEPT+1state)

- d. *Suddenly, the blaring sound of the guitar plays out.* (2state.INST)
- e. *The boy **has obviously started to play** it.* (INCEPT+1state)
- f. *She gets startled* (2state.INST)
- g. *so she ruins the icing.* (2state.INST)
- h. *She shouts something* (2state.DUR).
- i. *but Jimmy **carries on playing*** (ONG+1state)
- j. *and plates and crockery **start falling off** of the kitchen shelves.* (INCEPT+1state)
- k. *She exits the kitchen* (2state.INST)
- l. *shouting in fury* (1state)
- m. *and **carries on shouting** back in the living room.* (ONG+1state)

English speakers characteristically inclined to a significantly higher level of segmentation in comparison with the Czech and the Hungarian speakers. Example (1) shows that the finer event granularity in English film retellings was achieved by frequent decomposition of events into inceptive (1a, 1c, 1e, 1j) and ongoing phases (1i, 1m), complemented by specification of circumstances (1b, 1d) as well as causal linkage (1g). Phasal segmentation characteristically coincides with durative event types (1-state or 2-state) and with defocussed right temporal boundaries. Additional information about circumstances and causal links are supportive in conveying specificity in how events progress. This results in high granularity typical of the English discourse. The following is an example from a Hungarian L1 speaker.

(2) Example of event segmentation typical of Hungarian L1

- a. *Az anyuka **ezalatt** a tortáját díszíti a fiúnak,*
In the meantime the mother is decorating her cake for the boy (1state)
- b. *s mikor megszólal az erősítő,*
and when the amplifier plays out (2state.INST)
- c. *el is rontja a torta feliratát.*
she spoils the writing on the cake (2state.INST)

- d. **A nagy hangzavarra minden leesik a konyhában, tányérok s még egy patkány is.**
Upon the sound blast all falls down in the kitchen, plates and even a rat (2state.DUR)
- e. **Azután az anyyuka idegességében berohan a nappaliba**
After that mum whirls into the living room discomposed (2state.DUR)
- f. **és elkezd kiabálni.**
and **begins to shout** (INCEPT+1state)

In comparison with the English production, Hungarian retellings tended to be more coarse-grained. One of the most conspicuous differences from the English segmentation pattern was the frequent explicit linkage of events (high frequency of 2-states) by means of temporal adverbials signalling posteriority, as in examples (2d) and (2e). Since temporal relations between events are often lexically specified in Hungarian, there might be a lesser need for a comprehensive description of circumstances or for a detailed phasal resolution to guide the listener through event progression. As for the Czech retellings, they also differed on the whole from the English L1 in the degree of elaboration and phasal segmentation, in other words, they were also significantly less fine-grained. What they shared with the Hungarian retellings was the relatively frequent occurrence of event links via positional temporal adverbials marking posteriority (e.g., *pak* 'then' in 3f) and simultaneity (e.g., *když v tu chvíli* 'at which point' in 3b).

(3) Example of a segmentation strategy typical of Czech L1 film retellings

- a. *Maminka připravuje v kuchyni dort,*
Mum is preparing a birthday cake in the kitchen (1state)
- b. ***když v tu chvíli začne Jimmy hrát tak hlasitě,***
at which point Jimmy **starts playing** so loudly (INCEPT+1state)
- c. *že padají věci z kredence.*
that things are falling out of the cupboard (1state)
- d. *Maminka se lekne*
Mum gets startled (2state.INST)
- e. *a pokazí nápis na dortu.*
and she ruins the writing on the cake (2state.INST)
- f. ***Pak totálně vynervovaná vyběhne z kuchyně***
Then she storms out of the kitchen totally unnerved (2state.DUR)
- g. *a začne řvát.*

and **starts yelling** (INCEPT+1state)

Supplementary qualitative analyses provide a helpful insight into the nature of difficulties encountered on the way to acquiring TL segmentation principles. Quantitative results might suggest that a distinctly lower granularity in intermediate level learners would coincide with rare attempts to decompose events into phases. Remarkably, this was not the case in either of the learner groups. On the contrary, L2 production in both groups was found to exhibit instances of learners overgeneralising phasal segmentation, as shown in (4) and (5).

(4) Example of overgeneralising phasal segmentation by a Czech intermediate learner

- a. *Her mother **start to preparing** cake for her* (INCEPT+1state)
- b. *but her daughter **start to play** the guitar* (INCEPT+1state)
- c. *and she made mistake on cake with letters.* (2state.INST)
- d. *And guitar was too noisy* (0state)
- e. *that plates **start to falling** to the floor* (INCEPT+1state)
- f. *and her mother **start to shouting*** (INCEPT+1state)

(5) Example of overgeneralising phasal segmentation by a Hungarian intermediate learner

- a. *Yes the mother go to the kitchen* (2state.DUR)
- b. *and there **begin to write** to the birthday cake Happy Birthday* (INCEPT+1state)
- c. *and the child **begin** in room **playing** on guitar.* (INCEPT+1state)
- d. *And it was falling the dishes from the wall* (1state)
- e. *and the mum **begin to shout**.* (INCEPT+1state)

Inchoatives such as *start*, *try* and *begin* in English L1 data typically mark initiation of event chains and are frequently followed by progressive and terminative event phases. By contrast, L2 learners frequently overuse inchoatives, and in addition to initiating event chains, they combine them with any durative verb. The high frequency of inchoatives suggests that the learners have

acquired the means to decompose events into finer-grained phases. However, they often use them in a way that diverts from preferences in the TL. High density of inceptive phase marking results in overinformative temporal partitioning atypical of the target pattern.

Also noteworthy are learner-specific uses of linguistic means to mark duration. To indicate that the event is ongoing, learners were found to employ the temporal adverbials *still* and *again* (as in *the little boy playing still and still*, and in *again and again and he couldn't stop*), instead of more complex verbal structures present in fully-fledged L1s (e.g., *continues to play*, *goes on playing*).

Another remarkable digression from the target pattern was found in learners' atypical uses of inchoatives in combination with typically non-durative verbs. Production in both groups included attempts to employ phasal decomposition of instantaneous events (example from a Czech learner: *it starts to explode*; example from a Hungarian learner: *the boy started catch the ball*). This type of idiosyncratic use may be perceived as a disruption in the temporal information flow that results from stretching the topic time of the source state in typically instantaneous 2-state contents by means of phasal markers. Whether this phenomenon is limited to production or if it surfaces in a different type of performance in which learners can make more extensive use of semantic processing, noticing and reflecting, becomes clearer from the acceptability judgement task.

Acceptability judgements

Figure 2 shows the extent to which learners' acceptability of inceptive and ongoing phase marking for 2-state contents, which typically denote instantaneous events (INST), differs from/overlaps with English L1 controls. Analyses of four 2-state contents are provided to illustrate such variation. Contrasts emerged for the phasal decomposition of [catch], [reach] and [stop]. Results of a Kruskal-Wallis test showed a significant between-group difference ($p=.002$) for ratings in the acceptability of ongoing phase marking '*The dog is catching the ball in his mouth*' for the

situation [imagine you threw a ball to a dog that wants to play]. Mann-Whitney U paired between-group tests showed that the Czech learners' [$U_{CZ}(1,38)=90.5, p=.002$] and also the Hungarian learners' acceptance [$U_{HU}(1,38)=95.0, p=.003$] is significantly stronger than that of L1 controls. The reverse shows for [stop] in 'The truck is stopping' for the situation [imagine you are in a car park and you see a slowing truck], which both learner groups accepted significantly less [$U_{CZ}(1,38)=121.0, p=.020$] [$U_{HU}(1,38)=97.5, p=.003$] than the L1 controls.

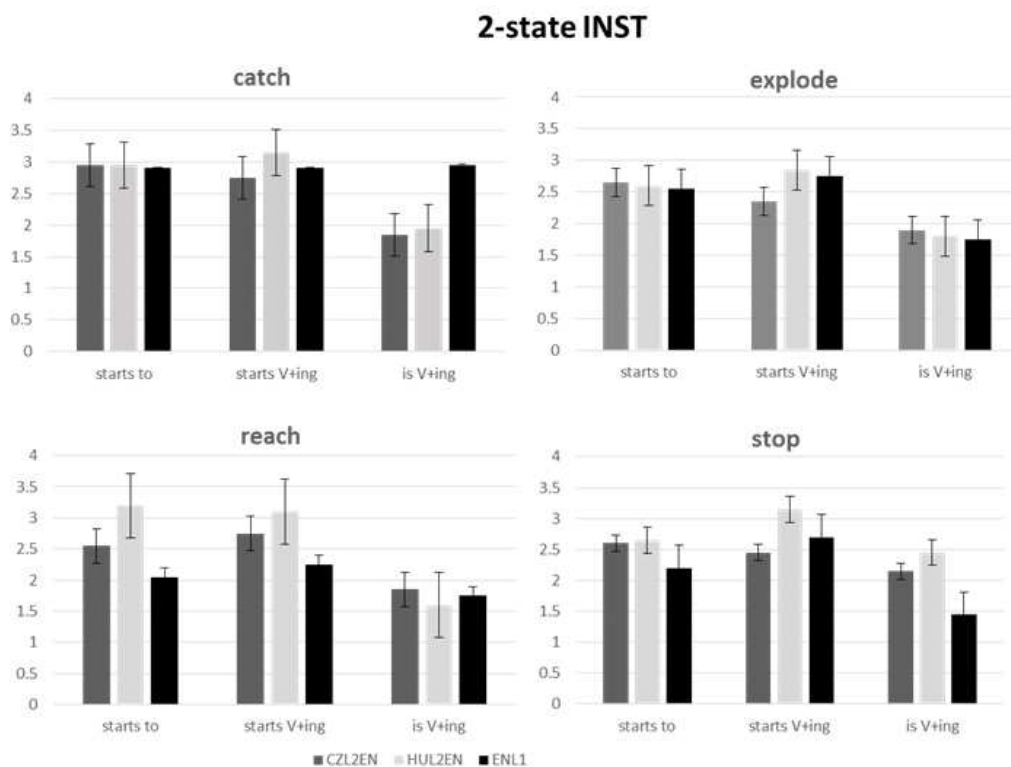


Figure 2. Mean scores for the acceptability judgements of 2-state INST per group (1-Perfectly suitable, 2-More or less suitable, 3-Somewhat unsuitable, 4-Completely unsuitable).

L2 versus L1 differences for judging the compatibility of inceptive phase markers with 2-state INST were limited. Of 10 in total, they were found only for [reach] in the situation [imagine you are on the top of a mountain and you see a group of teenagers as they arrive]. Learners accepted 'the group of teenagers starts reaching the top of the mountain' less than the L1 controls [$U_{CZ}(1,38)=143.5, p=.010$] [$U_{HU}(1,38)=105.0, p=.007$]. Also, 'the group of teenagers starts to reach the top of the

mountain' was accepted less by the learners, [$U_{CZ}(1,38)=156.0, p=.223$] [$U_{HU}(1,38)=76.0, p=.001$], but only the Hungarian learners' responses differed significantly from those of the L1 controls. In most 2-state INSTs, learners' judgements aligned with how (non-)suitable native controls deemed specific situation descriptions (for instance *explode+ing* was found more or less suitable by all three groups while *start to explode/start exploding* was somewhat unsuitable to express the situation [imagine you see the destruction of an old bridge with the use of dynamite]).

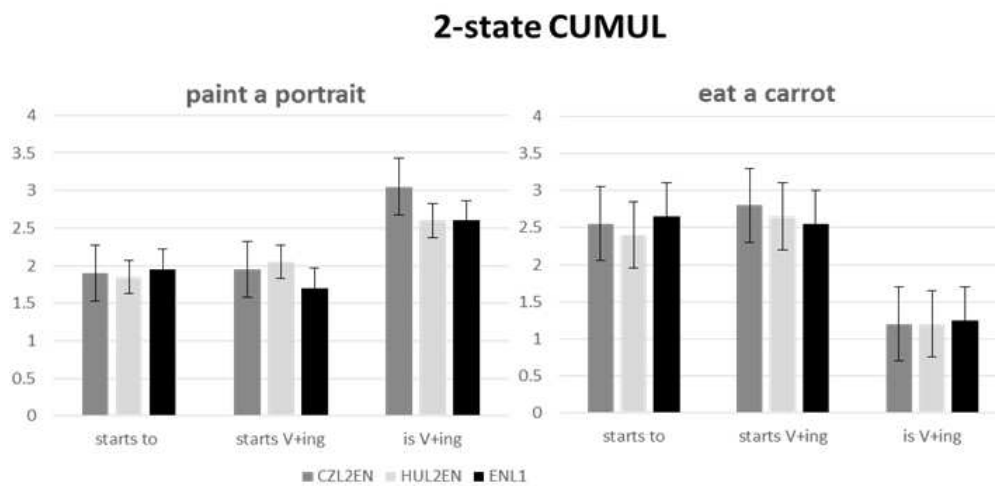


Figure 3. Mean scores for the acceptability judgements of 2-state CUMUL per group.

Remarkably, learners exhibited target-like flexibility in their judgements about phasal segmentation of typically cumulative 2-state contents (CUMUL). For situations with a stronger bias towards the source phase, such as [imagine you are in a square full of tourists where a street artist has a paintbrush in his hand with everything ready for a new a portrait] (Figure 3), learners preferred '*the street artist starts to paint a portrait*' significantly more than '*the street artist is painting a portrait*', with a within-group contrast of [$Z_{CZ}(1,19)=-2.34, p=.018$] for Czech learners and [$Z_{HU}(1,19)=-2.03, p=.043$] for Hungarian learners (Wilcoxon tests for within-group comparisons throughout the acceptability judgement section). For situations with a weaker bias towards the source phase, such as [imagine you are in a zoo and you see a camel with a carrot in its mouth],

learners' preference of '*the camel is eating a carrot*' was significantly stronger than of '*the camel starts to eat a carrot*', with a within-group contrast of [$Z_{CZ}(1,19)=-3.50, p<.001$] for Czech learners and [$Z_{HU}(1,19)=-3.00, p=.003$] for Hungarian learners. The change of preferences in learners' judgements largely mirroring those of native controls suggests a closer approximation to target-like phasal segmentation in cumulative 2-state contents than in typically instantaneous 2-state contents.

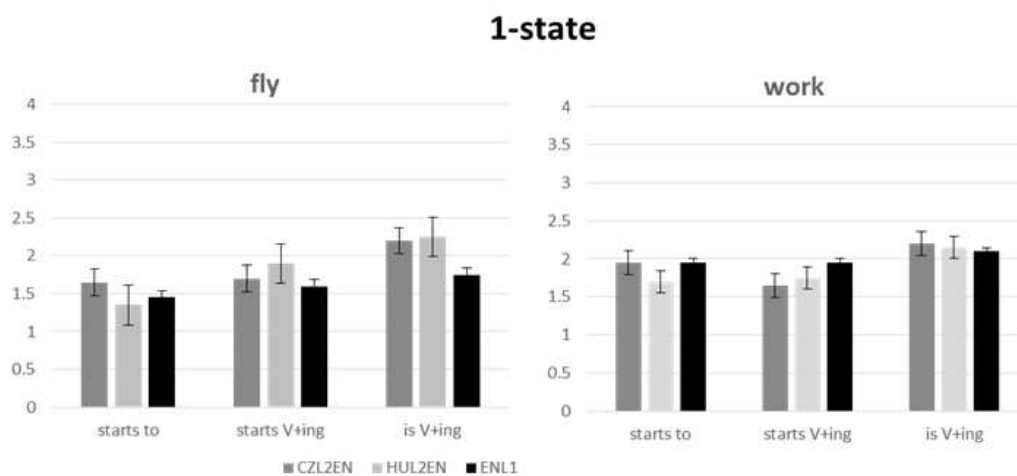


Figure 4. Mean scores for the acceptability judgements of 1-state contents per group.

Preferences for accepting inceptive and ongoing phases in 1-state contents were also largely concordant between the learner groups and the L1 controls (Figure 4). As an illustration of responses to situations for which the topic time is located more explicitly in the post-contrast phase, such as in [imagine you see a building site shortly after the lunch break], between-group differences were marginal when judging '*the builders start to work*' ($p=.529$), '*the builders start working*' ($p=.699$), and also '*the builders are working*' ($p=.890$). Also negligible were the differences for 1-state contents for which the topic time was located more closely to the point of contrast, such as in [imagine you see a nature documentary showing a young eagle that just left the nest for the first time]. Neither the acceptability of '*the young eagle starts to fly*' ($p=.438$) '*the young eagle starts*

flying' ($p=.618$) nor of '*the young eagle is flying*' ($p=.247$) differed significantly between groups.

Similarly minute differences across the three groups were characteristic also for the suitability judgments of other 1-state contents including [panic], [joke], and [play].

The most pronounced variation for the acceptance of ongoing phase marking emerged in situations typically known as 0-state contents (Figure 5). In some situations, learners tended to accept ongoing phase marking of 0-state contents with an added topic time contrast to a greater extent than L1 controls did. For instance, for the situation [imagine three divas enter an elevator which you are in], learners accepted '*the elevator is smelling like a perfume shop*' significantly more, [$U_{CZ}(1,38)=112.0, p=.011$] [$U_{HU}(1,38)=130.0, p=.046$], than the L1 controls. Differences were also found for the acceptance of ongoing phase marking in 0-state contents which did not have an added topic time contrast. For instance, in [imagine you see a swimmer on the diving board], Czech learners accepted '*the swimmer is being ready*' significantly more, [$U_{CZ}(1,38)=105.0, p=.006$], than the L1 controls, with Hungarian learners between the two groups. Interestingly, there were also 0-states for which learners accepted ongoing phase marking significantly less. For instance, in [imagine your friend Jane is at the bus stop at night with no more buses due because it is late], both learner groups accepted '*Jane is having to call a taxi*' significantly less [$U_{CZ}(1,38)=50.0, p<.001$] [$U_{HU}(1,38)=48.0, p<.001$] than the L1 controls.

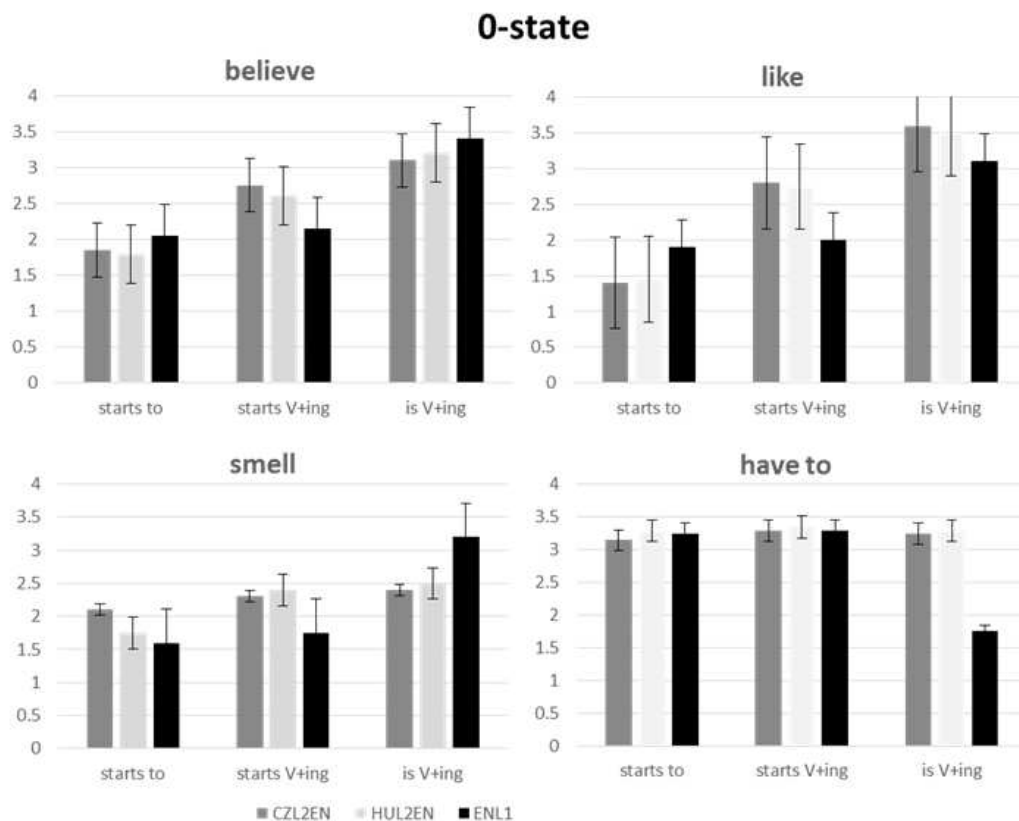


Figure 5. Mean scores for the acceptability judgements of 0-state contents per group.

Acceptance of inceptive phase marking in 0-states also showed some unexpected between-group variation. For instance, for a 0-state content [like] with an added topic time contrast [imagine you have a cat that did not eat fish but now does], both learner groups accepted inceptive phase marking ‘*the cat starts liking fish*’ significantly less [$U_{CZ}(1,38)=116.0, p=.017$] [$U_{HU}(1,38)=120.5, p=.026$], but they accepted ‘*the cat starts to like fish*’ significantly more [$U_{CZ}(1,38)=124.0, p=.023$] [$U_{HU}(1,38)=131.5, p=.042$], than the L1 controls. Judgements of other 0-states showed more target-like flexibility. For instance, for a 0-state content [believe] with an added topic time contrast [imagine your friend Jack is amazed to find circles in his cornfield with a hot meteorite in one of them] (Figure 2), learners, just like native controls, preferred ‘*Jack starts to believe in UFOs*’ significantly more than ‘*Jack is believing in UFOs*’, with a within-group contrast of [$Z_{CZ}(1,19)=-$

3.10, $p=.002$] for Czech learners and [$Z_{HU}(1,19)=-3.06$, $p=.002$] for Hungarian learners. Comparing the four situation types, phasal segmentation of 0-state contents showed the highest degree of between-group (L2-TL) variation in acceptability judgements.

Discussion

Q1: *What event segmentation patterns characterise intermediate learners' language production?*

Are granularity degrees, grounding preferences, and phasal partitioning in learner production concordant with corresponding source language patterns, target language patterns, or do learners follow different principles?

Analyses of temporal event segmentation across learner groups showed that intermediate learner production was characteristically more coarse-grained than that in the corresponding L1s or the TL. However, at this stage learners already showed that they are able to modify granularity degrees in correspondence with different communicative needs. Another closely related finding is that L2 attempts to approximate to TL patterns were often accompanied with non-standard decomposition of events, most typically in the form of overgeneralising inceptive phase marking.

Q2: *Are the event segmentation patterns detected in production reflected in learners' acceptability judgements? How does segmentation of inceptive and ongoing event phases vary based on lexical aspectual properties of verbs?*

Segmentation patterns detected in production were not directly mirrored in learners' acceptability judgements. Learners did not exhibit systematically higher acceptance levels for phasally partitioned events, as their production might suggest. Levels of overlap between learners' and native speakers' acceptability judgements were found to vary depending on lexical aspectual

properties of the verbs in question. Although limits related to their construct validity block wider generalisations, exploration via acceptability judgements yielded a cross-task confirmation of learner-specific segmentation preferences. The following subsections provide interpretations of these results, links to theory and to empirical context.

Learner-specific modification of granularity degrees

In the developmental stage when lexical and discourse-pragmatic means serve to establish the central temporal relations (Klein & Perdue, 1997), learners from both intermediate groups demonstrated that grammatical competence is not a necessary prerequisite to manipulate the levels of event granularity according to communicative needs. Learners showed that they are not only able to achieve a finer-grained event decomposition by means of a clever management of temporal adverbials (e.g., the iterative use of *still* and *again* to indicate extended topic time), but they are also capable of reference to particular event phases with the use of emerging lexical temporal boundary markers (*start, finish*). This is in line with the observations by Dietrich *et al.* (1995: 36) that the pre-grammatical system is very simple compared to the source and target languages but extremely versatile, allowing an easy expression of when something happens or is the case, provided (a) there are enough adverbials, and (b) they are cleverly managed.

Non-standard decomposition of events – overgeneralised inceptive phase marking

Interestingly, overgeneralisation of inceptive phase marking, one of the interlanguage features also present in advanced L2 varieties (von Stutterheim & Lambert, 2005), is already evident at the more intermediate stage in this study^{iv}. Two similarities emerge with von Stutterheim & Lambert's (2005) findings, despite the between-study difference concerning learners' L2 proficiency. The first analogy is that there was an L2 increase in the frequency of inceptive markers compared to TL controls, but in neither study was this increase significant. Second, both studies found learner-specific digressions from the target regarding contexts of use. In the present study, intermediate

level learners were found to overuse inceptive phase marking for any activity-type situation (regardless of whether the outcome was failure or success to achieve the goal), often without further specification of relevant progressive and/or terminative event phases. This segmentation technique naturally coincided with less developed narratives, where learners' production heavily relied on implicit discourse-pragmatic principles to guide the listener in decoding the unspecified event phases. As a yet different kind of digression from the target, some uses of inceptive marking exhibited attempts to phasally decompose instantaneous events, which is atypical for TL as much as for the corresponding L1s. Overall though, it was found that learners at the intermediate stage are able to segment events and to modify granularity levels in discourse in order to comprehensibly convey event development in a fairly complex communicative task.

Discrepancy between acceptability judgements and production

Statistically significant contrasts between learners and L1 controls were found in the acceptability judgements of ongoing phase marking for a number of typically instantaneous 2-state contents. However, inceptive phase marking for this verb type (e.g., *catch*, *explode*) did not render L2-TL contrasts in acceptability judgements, unlike in production. If learners exhibit target-like flexibility in their acceptability judgements, why then did they overgeneralise inceptive phase marking in production? This discrepancy between production and judgements in intermediate L2 learners aligns with related findings in the literature (e.g., Swain, 1985; Trenkic *et al.*, 2014), which signal that structural advances at intermediate level in L2 production can fall behind those in accessing knowledge which the speaker need not formulate. In other words, having access to the right linguistic tools does not necessarily imply their correct usage. In a wider context, the identified production vs. judgement contrasts highlight the need to treat the processes involved in each as separate layers of knowledge. A complementary treatment is required as the scale of differences surfacing in production cannot serve as an adequate prognosis of the scale of differences in less

time-constrained acceptability judgements. And vice versa, the activation of knowledge to decide about individual event phases when a narrative context is absent cannot directly compare to, or predict, the type of knowledge needed for sequential event encoding.

Variation in judgements linked to verb type

Differences in the strength of acceptability emerged in close connection with specific verb types. Learners exhibited target-like preferences for typically cumulative 2-state contents, both in contexts with and without a bias towards the source phase. Acceptability judgements of learners for inceptive and ongoing phase marking in 1-state contents also showed a successful approximation to preferences of the L1 controls, which proved to be the case in situations where the topic time was explicitly located more as well as less closely to the point of contrast. However, a number of 0-state contents showed lower response systematicity without a clear single direction between L2-TL contrasts. Responses to ongoing phase marking in some 0-states with and also without an explicitly highlighted topic time contrast was found significantly more acceptable in L2 than in TL, while in other 0-state contents without an added topic time contrast the learners' acceptance level was significantly lower. Multiple directions in learners' preferences also emerged (to a lesser extent) in the judgements of inceptive phase marking in 0-state contents. The reason this verb type presents more difficulty can be attributed to *associative attentional tuning* (Ellis, 2006; Ellis & Sagarra, 2010). If we consider that 0-states are the only verb type not typically associated with a topic time contrast, the learner's attention is likely to be tuned to their inherent atemporality. It is thus not surprising that a purely external (context-induced) topic time contrast may present an increased challenge compared to other verb types when decisions about phasal segmentation need to be made. The difficulty for judging if phasal segmentation is suitable is further compounded for learners whose L1 grammar either does not encode phasal marking (Hungarian) or it encodes phasal marking only for a small group of verbs (Czech).

L1-specific segmentation and its influence on L2

Language-specificity in event segmentation surfaced as a contrast between the preferences of English speakers, who in production tend to opt for a highly granular event resolution and focus on individual phases, and Czech and Hungarian speakers who instead systematically favour a more coarse-grained resolution. Crosslinguistic variation in segmentation was closely interlinked with aspectual operators in the given L1 system. The tendency to refrain from a detailed decomposition of events into phases corresponds with a less regular or absent specific grammatical ongoingness marker. While a fully grammaticalised concept of ongoingness in English fits well with the susceptibility to express more event phases, irregular marking of ongoingness in Czech and no grammatical marking of ongoingness in Hungarian might direct less attention to the internal temporal constituency of events. These findings are consistent with previous research showing that patterns of event conceptualisation are language-specific and to a great extent grammatically driven (e.g., Bylund, 2011; von Stutterheim & Nüse, 2003).

L1-specific segmentation processes are at least partly accountable for the reduced granularity degrees in learner production. In this sense, there is likely to be some influence of the L1 on event conceptualisation which is resistant to reorganisation in the L2 (Slobin, 1996). However, the findings from both production and acceptability judgements suggest that learners' segmentation patterns tend to cluster more strongly around the shared developmental stage rather than the specific L1 backgrounds. If L2 segmentation depended more on language distance in L1 ongoingness marking (Hungarian > Czech > English), we would expect Hungarian learners' digressions from the English pattern to be more pronounced than those of the Czech learners. While this could be the case at different L2 stages, the present findings with intermediate learners do not point in this direction. The outcomes of the present study are thus limited in how much they can reveal about the modulation of second language production and comprehension in correspondence with structural

similarities between the source and the target language pairs (Ellis, 2006; Tolentino & Tokowicz, 2011).

Limitations and avenues for future inquiry

Epistemological concerns about the use of acceptability judgements in SLA research do not allow interpretations of the current results from task two to extend beyond the exploratory level. The construct validity of AJs remains contentious because the types of knowledge they actually tap into (i.e. what learners know to be a grammatically correct L2 construction vs. what they know about how to use a particular construction in the L2) are not clearly distinguishable (Ellis, 2005; Tremblay, 2005). When AJs are performed, learners may judge a sentence as unacceptable based on its structure rather than on its suitability for the given situation. To at least partly address this problem and channel attention more towards how to use language in spontaneous situations (implicit knowledge) rather than to knowledge about a particular language construction (explicit knowledge), the current AJ task specified in the instructions that 'this is not a language test, and all options are theoretically possible'. Although to a limited extent, in this way, learners could be less strongly guided to assess 'how correct' the constructions are and perhaps would instead focus more on 'how well they fit' in the presented contexts. Future AJs may benefit from including more controlled time limits for task completion (Gutiérrez, 2013) as another way to further increase the predisposition of learners to draw on implicit rather than explicit knowledge. Time constraints were not imposed in the present study, but the software recorded total response times and thus allowed excluding responses that exceeded 40 minutes (i.e. 1 minute for a set of four judgements on average). This filtered out responses with arguably reduced spontaneity and more extensive reflection. Nonetheless, AJs on their own will remain a questionable data type from which linguistic competence can only be inferred. Cross-task validation of AJs with production to see whether and where the two converge (e.g. Chaudron, 2003) is therefore important.

One might also justly ask whether the extent to which two different types of context introduction could have influenced acceptability judgements in different ways. It is indeed likely that context setting without verb mention as in [imagine you threw a ball to a dog that wants to play] for judging '*The dog is catching the ball in his mouth*' sensitises attention to event phases differently than context setting with verb mention such as [imagine you are on the top of a mountain and you see a group of teenagers as they arrive] for judging '*The group of teenagers starts reaching the top of the mountain*'. Analysing responses from differently set contexts together would therefore pose a legitimate concern. To address this issue and avoid conflation of various types of context introduction, every single between-group analysis of AJs in this study was conducted separately per context. This approach naturally reduces generalisability over different contexts even within the same lexico-aspectual verb type, but it ensures response comparability, important for competently assessing the extent to which event segmentation preferences between L1 and L2 groups vary. A fruitful extension in this direction will be to check for possible effects of instruction type via a counterbalanced design in which the same contexts are introduced both with and without verb mention.

Given its exploratory nature, this study naturally raises the question whether the identified production vs. judgement discrepancy is attributable to different participants tested per task rather than to different knowledge types. To verify effect stability and increase result validity, it is clear that extensions in this line of research need to involve within-participant comparisons of segmentation patterns in production and judgements. Another area from which future research on event segmentation would surely profit is extending the tests of interaction between grammatical ongoingness marking and higher order processes such as (non-)verbal memory and categorisation of event phases (Athanasopoulos & Bylund, 2013) to different L2 levels and L1-L2 pairs, using a balanced combination of various verb types. In this way, it will not only be possible to examine a

richer developmental picture but also to better identify the extent to which event segmentation in L2 is L1-mediated.

Conclusion

This study set out to examine which event segmentation patterns characterise the production and acceptability judgements of Czech and Hungarian intermediate learners of L2 English. In particular, it compared granularity degrees, grounding preferences, and phasal partitioning preferred by learners with those typical of corresponding source and target languages. In addition, preferences for inceptive and ongoing phase marking were analysed in close relation with the lexical aspectual properties of four different verb types. Production results showed that learners employ forms that do indeed represent the phasal perspective typical of the TL (i.e., a high frequency of inchoative structures for initiating event chains), but they encounter difficulties in locating them in larger information units (overgeneralisation of inceptive phase marking to any durative verb without elaborating subsequent phases). Other pronounced digressions from the target were manifested in learners' acceptability of phasal marking in verbs without an inherent topic time contrast. These results are interpreted as shared developmental interlanguage phenomena attributable to the increased difficulty in attentional tuning to specific context-modulated aspectual properties. Taken together, the reported findings are an informative addition to the study of event segmentation patterns, enriching our understanding of what principles underlie temporal information organisation in a second language.

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

Acknowledgements

This research was partially supported by the A. H. Lloyd Fund at Christ's College Cambridge and the Centre for Language Learning Research at the University of York. Many thanks to Erzsébet Kulcsár, Aba Szöllősi, Eva Bonaventurová, Radka Šmahelová and Kirstie Wailes-Newson for their assistance with participant recruitment, transcriptions and coding. I am also very grateful to Julie Lawrence and the editors of this volume, Pascale Leclercq and Martin Howard, for their insightful comments on the previous drafts of this article.

References

- Abondolo, D. (1998). Hungarian. In D. Abondolo (ed.), *The Uralic languages* (428–456).
Routledge: London.
- Allan, D. (2004). *Oxford Placement Test 2*. Oxford: Oxford University Press.
- Athanasopoulos, P. & Bylund, E. (2013). Does grammatical aspect affect motion event cognition?
A crosslinguistic comparison of English and Swedish speakers. *Cognitive Science* 37, 286–
309.
- Bardovi-Harlig, K. (1992). The relationship between form and meaning: A cross-sectional study of
tense and aspect in the interlanguage of learners of English as a second language. *Applied
Psycholinguistics* 13, 253–278.
- Berman, R. & Slobin, D. (1994). *Relating events in narrative: A crosslinguistic developmental
study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bialystok, E. (1982). On the relationship between knowing and using forms. *Applied Linguistics*
3, 181–206.

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

- Bohnenmeyer, J., Enfield, N., Essegbey, J. & Kita, S. (2011). The macro-event property: The segmentation of causal chains. In J. Bohnemeyer & E. Pederson (eds.), *Event representation in language and cognition* (43–67). Cambridge: CUP.
- Boroditsky, L. & Trusova, E. (2003). Cross-linguistic differences in the representation of events: Verb aspect and completion in English and Russian. In R. Alterman & D. Kirsch (eds.), *Proceedings of the twenty-fifth annual conference of the Cognitive Science Society* (13–19). Mahwah, NJ: Lawrence Erlbaum.
- Bylund, E. (2011). Ultimate attainment of event segmentation and temporal structuring patterns in speakers of L2 Swedish. *Vigo International Journal of Applied Linguistics* 8, 29–53.
- Carroll, M. & Lambert, M. (2003). Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In C. Dimroth & M. Starren (eds.), *Information structure and the dynamics of language acquisition* (267–287). Amsterdam: Benjamins.
- Carroll, M. & von Stutterheim, C. (2003). Typology and information organisation: Perspective taking and language-specific effects in the construal of events. In A. G. Ramat (ed.), *Typology and Second Language Acquisition* (365–402). Berlin: de Gruyter.
- Chaudron, C. (2003). Data collection in SLA research. In C. J. Doughty & M. H. Long (eds.), *Handbook of Second Language Acquisition* (762–828). Oxford: Blackwell.
- Cook, V. (2002). Background to the L2 user. In V. Cook (ed.), *Portraits of the L2 user* (1–28). Bristol: Multilingual Matters.
- Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: CUP.

Csirmaz, A. (2004). Perfective and imperfective in Hungarian: (Invisible) differences. In S. Blaho, L. Vicente & M. de Vos (eds.), *Proceedings of Console XII*. University of Leiden. Presented at Console XII. Patras, Greece, December 12–14, 2003.

Dietrich, R., Klein, W. & Noyau, C. (1995). *The acquisition of temporality in a second language*. Amsterdam / Philadelphia: Benjamins.

Ellis, N. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27, 164–194.

Ellis, N. & Sagarra, N. (2010). Learned attention effects in L2 temporal reference: The first hour and the next eight semesters. *Language Learning* 60(2), 85–108.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning* 54, 227–275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27, 141–172.

Givón, T. (1991). Serial verbs and the mental reality of 'event'. In E. Traugott (ed.), *Approaches to grammaticalization* (81–127). Amsterdam: Benjamins.

Gutiérrez, X. (2013). The construct validity of grammaticality judgement tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition* 35, 423–449.

Gyllstad, H., Granfeldt, J., Bernardini, P. & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR. *EUROSLA Yearbook* 14, 1–30.

Habel, C. & Tape, H. (1999). Processes of segmentation and linearization in describing events. In R. Klabunde & Ch. von Stutterheim (eds.), *Processes in language production* (117–153). Wiesbaden: Deutscher Universitätsverlag.

Kiefer, F. (2006). *Aspektus és akciómínőség: különös tekintettel a magyar nyelvre* [Aspect and Aktionsart: with particular attention to Hungarian] Budapest: Akadémiai Kiadó

Klein, W. (2009). How time is encoded. In W. Klein & P. Li (eds.), *The expression of time* (39–82). Berlin: de Gruyter.

Klein, W. (1994). *Time in language*. London: Routledge.

Klein, W. & Perdue, C. (1997). The Basic Variety. Or: Couldn't natural languages be much simpler? *Second Language Research* 13(4), 301–347.

Klein, W. & von Stutterheim, C. (1987). Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109, 163–183.

Levelt, W. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (eds.), *The neurocognition of language* (83–120). Oxford: Oxford University Press.

Noyau, C., de Lorenzo, C., Kihlstedt, M., Paprocka, U., Sanz Espinar, G. and Schneider, R. 2005. Two dimensions of the representation of complex event structures: granularity and condensation. Towards a typology of textual production in L1 and L2. In H. Hendriks (ed.), *The structure of learner varieties* (157–201). Berlin: de Gruyter.

Pavlenko, A. (2014). *The bilingual mind and what it tells us about language and thought*. Cambridge: Cambridge University Press.

Pawley, A. (1987). Encoding events in Kalam and English: differing logics for reporting experience. In R. Tomlin (ed.), *Coherence and grounding in discourse* (329–360). Amsterdam: Benjamins.

Péter, M. (2008). Az aspektusról – más aspektusból [On aspect – from a different aspect]. *Magyar Nyelv* 107, 428–452.

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

Rothstein, S. (2004). *Structuring events*. Oxford: Blackwell.

Short, D. (2002). Czech. In B. Comrie & G. Corbett (eds.), *The Slavonic languages* (455–532). London: Routledge.

Schmiedtová, B. (2013). Traces of L1 patterns in the event construal of Czech advanced speakers of L2 English and L2 German. *International Review of Applied Linguistics in Language Teaching* 51(2), 87–116.

Schmiedtová, B., von Stutterheim, Ch. and Carroll, M. (2011). Language-specific patterns in event construal of advanced second language speakers. In A. Pavlenko (ed.), *Thinking and Speaking in Two Languages* (66–107). Clevedon: Multilingual Matters.

Schmiedtová, B. (2004). *At the same time: The expression of simultaneity in learner varieties*. Berlin: de Gruyter.

Slobin, D. (1996). From “thought and language” to “thinking for speaking”. In J. Gumperz and S. Levinson (eds.), *Rethinking linguistic relativity* (70–96). Cambridge: CUP.

Sorace, A. (1996). The use of acceptability judgements in second language acquisition research. In W.C. Ritchie & T.K. Bhatia (eds.), *Handbook of second language acquisition* (375–409). San Diego, CA: Academic Press.

Speer, N., Swallow, K. & Zacks, J. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective and Behavioral Neuroscience* 3, 335-345.

Starren, M. (2001). *The second time: The acquisition of temporality in Dutch and French as a second language*. Utrecht: LOT Dissertation Series 50.

Sugaya, N. & Shirai, Y. (2007). The acquisition of progressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese. *Studies in Second Language Acquisition* 29, 1–38.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and

comprehensible output in its development. In S. Gass & C. Madden (eds.), *Input in second language acquisition* (236–244). Rowley, MA: Newbury House.

Takács, E. (2012). A progresszív jelentéstartalom konstruálásának sajátosságai a magyarban

[Idiosyncrasies of the progressive aspect in Hungarian]. *Magyar Nyelvről* 136, 325–335.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed.),

Language typology and syntactic description. Grammatical categories and the lexicon (57–149). Cambridge: CUP.

Tolentino, L. & Tokowicz, N. (2011). Across languages, space and time: A review of the role of

cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP methods. *Studies in Second Language Acquisition* 33, 91–125.

Tomlin, R. (1987). Linguistic reflections of cognitive events. In R. Tomlin (ed.), *Coherence and*

grounding in discourse (455–479). Amsterdam: Benjamins.

Tremblay, A. (2005). Theoretical and methodological perspectives on the use of grammaticality

judgement tasks in linguistic theory. *Second Language Studies* 24, 129–167.

Trenkic, D., Mirkovic, J. & Altmann, G. (2014). Real-time grammar processing by native and non-

native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition* 17, 237–257.

Vanek, N. (2012). Language-specific perspectives in reference to time in the discourse of Czech,

English and Hungarian speakers. In L. Filipović & K. Jaszczołt (eds.), *Space and time in languages and cultures 1: Linguistic diversity* (135–156). Amsterdam: Benjamins.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review* 66: 143–160.

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

von Stutterheim, Ch., Bouhaous, A., Carroll, M., Sahonenko, N. (2012). Grammaticalized temporal categories, language specificity, and macroplanning in expository texts. *Linguistics* 50, 341–371.

von Stutterheim, Ch. & Lambert, M. (2005). Cross-linguistic analysis of temporal perspectives in text production. In H. Hendriks (ed.), *The structure of learner varieties* (203–230). Berlin: de Gruyter.

von Stutterheim, Ch. & Nüse, R. (2003). Processes of conceptualisation in language production: Language-specific perspectives and event construal. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 41(5), 851–881.

Zacks, J., Braver, T., Sheridan, M., Donaldson, D., Snyder, A., Ollinger, J., Buckner, R. & Raichle, M. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience* 4(66), 651–655.

Zacks, J. & Swallow, K. (2007). Event segmentation. *Current Directions in Psychological Science* 16, 80–84.

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

Appendix A.

A typical acceptability judgement set with four situation types: 0-state (0S), 1-state (1S), 2-state cumulative (2SC), and 2-state instantaneous (2SI)

Completion Instructions: After reading the question, briefly imagine the situation as described. Then, your task is to read the four options under the question and to decide how suitable you find each option to express the given situation. Please fill in all four lines for each question.

(0S) Imagine you see a diver under the water on his way back up to the surface to take a breath. How suitable do you find the following to express this situation?

	Perfectly suitable	More or less suitable	Somewhat unsuitable	Completely unsuitable
a. The diver starts to need oxygen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The diver starts needing oxygen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The diver needs oxygen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The diver is needing oxygen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

(1S) Imagine you see a nature documentary showing a young eagle that just left the nest for the first time. How suitable do you find the following to express this situation?

	Perfectly suitable	More or less suitable	Somewhat unsuitable	Completely unsuitable
a. The young eagle starts to fly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The young eagle starts flying.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The young eagle flies.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The young eagle is flying.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

(2SC) Imagine you are in a square full of tourists where a street artist has a paintbrush in his hand with everything ready for a new a portrait. How suitable do you find the following to express this situation?

	Perfectly suitable	More or less suitable	Somewhat unsuitable	Completely unsuitable
a. The street artist starts to paint a portrait.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The street artist starts painting a portrait.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The street artist paints a portrait.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The street artist is painting a portrait.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

(2SI) Imagine you are in a bath and you see a soap bubble land on the radiator. How suitable do you find the following to express this situation?

	Perfectly suitable	More or less suitable	Somewhat unsuitable	Completely unsuitable
a. The bubble starts to burst.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. The bubble starts bursting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. The bubble bursts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. The bubble is bursting.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please indicate how well you could imagine the described situations:

- It was not a problem to imagine the situations
- I found it complicated to imagine the situations
- The situations were extremely difficult to imagine

This is the authors' copy of 'It starts to explode.' Phasal segmentation of contextualised events in L2 English. In M. Howard & P. Leclercq (Eds.), *Tense-Aspect-Modality in a second language: contemporary perspectives* (pp. 143-180). Amsterdam: John Benjamins. [Studies in Bilingualism, 50] <https://benjamins.com/#catalog/books/sibil.50.06van/details> Please contact the publisher for permission to reuse the material in any form.

Appendix B.

Sample fragments from the elicitation video illustrating event progression from episode 1 (E1.1) to episode 6 (E6). The full video and its event map can be downloaded from the IRIS database, an online repository of L2 data collection materials at <http://www.iris-database.org>.



-
- ⁱ Lexicalisation patterns and grammatical constraints are far from a comprehensive list of factors which impact segmentation. It is fully acknowledged that underlying principles of segmentation can be attributed to event schemas that are at least partly dependent on cultural practice and world knowledge (Carroll & von Stutterheim, 2011: 72).
- ⁱⁱ Throughout the whole study, group differences considered as statistically significant mean that the related p-values were below 0.05.
- ⁱⁱⁱ This calculation reflects the exclusion of one extremely outlying case of event segmentation by a participant from the otherwise relatively consistent group of Czech L1 speakers.
- ^{iv} Unlike the present study, von Stutterheim & Lambert (2005) tested highly advanced L2 learners (an advanced learner group of 6 participants with 3 or 4 years of English at university, and a quasi-bilingual group of 6 participants with 4 years of English at university + 2 years at the highly competitive *École Normale Supérieure de Fontenay*), which naturally disallows any direct comparisons purely on the basis of L2 proficiency. However, the learner-specific similarities in segmentation in these two studies are noteworthy. Both studies found L2 uses of *start* in combination with any activity-type situation (e.g. *he starts scratching*, *she starts to turn*), showing that learners overgeneralised inceptive markers to contexts which were not marked this way by the TL controls.