The Influence of the First Language on the Cognition of Temporal Properties of Motion Events in the Second Language

BA Thesis English Language and Culture, Utrecht University
Programme Language Matters
Mojca Kloos
Student number: 6155111
Supervisor: Dr. Nynke de Haas
Co-reader: Dr. Elena Tribushinina
February 2020
6030 words

Abstract

The question of whether a language has grammaticised means of expressing aspectual distinctions in conceptual categories is used in investigating linguistic relativity in speakers of different languages. This study focuses on the extent to which advanced learners of a second language are affected by the concepts set in their first language when lexicalising temporal properties of motion events. The analysis, which covers native speakers of Dutch who are advanced learners of English, involves data from a linguistic experiment in which the participants described video clips in both their first and their second language. The current study investigated the endpoint behaviour of this population, and it was found that there is no difference in the encoding of endpoints between the first and the second language. This finding may indicate that the speakers still employ L1 principles in their L2, which could hold implications for second language education. Previous studies, such as von Stutterheim and Carroll (2006), have also found that advanced learners of a second language rely on the principles of their L1 when lexicalising temporal properties of motion events. These findings were consolidated by the results of the present study. This provides evidence for the framework on linguistic relativity as designed by Slobin (1996), i.e. the thinking-before-speaking *hypothesis, as in the present study, native speakers of Dutch are more attentive to the maximal temporal viewing frame, as they are when speaking their native language, even when speaking a second language.*

Table of contents

1. Introduction

1.1 Linguistic Relativity. The idea of language influencing or limiting the thoughts of an individual has occupied linguists since Whorf (1940/1956) formulated the principle of linguistic relativity, which he explains as "users of markedly different grammars [being] pointed by their grammars toward different types of observations and different evaluations of externally similar acts of observation" (p. 221). According to this hypothesis, speakers are limited to certain observations or thought processes depending on the language that they speak. Alternatively, a weaker version of Whorf's hypothesis was developed by Slobin (1996), who formulates it as a *thinking-before-speaking* hypothesis. He states that effects of language can only be seen during the process of language use, where speakers need to attend more to the aspects of events that are mandatory in their language than those that are not. In Slobin's hypothesis, speakers are not limited to specific observations, but they are more attentive to those concepts that are more salient in their native language. The underlying question behind these two hypotheses, i.e. whether speakers of different languages think in different manners, has led to linguists investigating this phenomenon in different fields, e.g. Athanasopoulos (2009) on the domain of colour, Gennari, Sloman, Malt and Fitch (2001) on the cognition of spatial properties of motion events; Athanasopoulos (2006) on the influence of grammatical number marking in a speaker's language on their cognition; Von Stutterheim and Carroll (2006), von Stutterheim, Andermann, Carroll, Flecken and Schmiedtová (2012), Bylund and Jarvis (2011) and Bylund, Athanasopoulos and Oostendorp (2013) on the cognition of temporal properties of motion events.

Speakers of varying languages differ in the way they lexicalise properties of motion events when they consider certain properties of these events. According to Bylund, Athanasopoulos and Oostendorp (2013), a motion event is "a situation involving physical

displacement, whereby an entity occupies different spatial locations at different points in time" (p. 930).

    1.2 Temporal properties of motion events. A distinction can be made between aspect and non-aspect languages, with regard to temporal properties of motion events. In an aspect language (such as English), aspectual distinctions are grammaticalised, and speakers focus on the ongoingness of an event (which is called the intermediate temporal viewing frame). Research has shown that speakers of aspect languages demonstrate a tendency to leave out the endpoint when describing a motion event (e.g. *two women are walking*). In contrast, non-aspect languages (such as German) have no grammatical means to express aspect and speakers generally look at a motion event through a maximal temporal viewing frame, meaning that they do not just zoom in on the event itself; speakers of these languages show a tendency to encode endpoints and the focus on boundedness of the given event (e.g. *two women are walking **to a house***) (Bylund, Athanasopoulos & Oostendorp, 2013).

    A great deal of research has been carried out in the domain of temporal properties of motion events. Firstly, regarding monolinguals, Von Stutterheim and Carroll (2006) found that endpoints were encoded in 25.2 percent of the descriptions given by monolingual English speakers versus 76.4 percent in monolingual German speakers. This pattern was also found in von Stutterheim et al. (2012), who investigated a difference in the encoding of endpoints by speakers of languages in an 'aspect group' (i.e. English, Spanish, Russian and Arabic) and a 'non-aspect group' (i.e. Dutch, German and Czech). The 'aspect group' encoded endpoints in about sixty percent of their descriptions of motion events, while the 'non-aspect' group did so in about forty percent. In this study, the results further display the tendency of speakers of non-aspect languages to mention endpoints at a significantly higher frequency than speakers of aspect languages do. Flecken, von Stutterheim and Carroll (2014) found that speakers of German encode endpoints more often that speakers of Arabic (an aspect language), which

demonstrates the tendency of speakers of a non-aspect language (here: German) to encode endpoints in their lexicalisation of motion events. The consensus that arises from these studies is thus that monolingual speakers of a non-aspect language encode significantly more endpoints in their lexicalisation of motion events than speakers of an aspect language.

Regarding the effect of a second language on conceptualisation in a speaker's first language, Bylund and Jarvis (2011) show that Spanish-Swedish bilinguals, whose first language was Spanish, encode endpoints of motion events more often than their monolingual Spanish peers. Spanish is an aspect language, whereas Swedish is a non-aspect language, and the authors hypothesise that there is a connection between grammatical aspect and event conceptualisation; the results of the study indicate that the bilinguals were affected by the focus on boundedness that is present in Swedish. These results may be explained by the idea that individuals experience cognitive restructuring as a result of learning a second language, i.e. that their second language is affected by, but frequently still different from their first language. Furthermore, according to Bylund, Athanasopoulos and Oostendorp (2013), the patterns demonstrated by the participants in their descriptions of motion events in Afrikaans (a non-aspect language) align more with Swedish (also a non-aspect language) than English (an aspect language). The participants were native speakers of Afrikaans, but since English is the language of education in South Africa, they were seen as advanced learners of this language. In this study, a key finding was that the more frequently a participant spoke English, the more their endpoint behaviour aligned with English monolinguals. However, studies have also found evidence that the languages an individual speaks may not fully determine the manner in which he/she lexicalises a certain concept. For example, Sharpen (2016) argues that the parameters that are set in a speaker's native language affect their performance with regard to motion events in their L2; they show L1 conceptual transfer in their second language. These findings are in line with Slobin's *thinking-before-speaking*

hypothesis. Von Stutterheim and Carroll (2006) demonstrate further evidence for this hypothesis, as they found that German advanced learners of English omit endpoints in situations where these can easily be inferred (for example, *a boy jumping off a table **onto the ground***), but not in those situations where they had to infer the endpoints themselves (e.g. *two women walking **to a house***) which is in line with the pattern found in German monolingual speakers. In this study, English advanced learners of German did not mention endpoints more often than English monolinguals, which provides evidence that the principles of the L1 are still present in the speakers' productions in their L2.

The studies discussed here give rise to the question whether it is a universal phenomenon that speakers who are fluent in both an aspect and a non-aspect language undergo cognitive restructuring in their L2 compared to their first language as a result of learning the L2. Bylund and Athanasopoulos (2015) argue that the extent to which verbal behaviour of L2 speakers approximates that of native speakers depends on concept similarities between the two languages, i.e. the similarity of two related concepts in the two languages, conceptual prototypicality (here, how motion events are conceptualised in a given language) and factors such as length of exposure, age of acquisition and frequency of L2 use. These factors will not be tested in the present study, though they may constitute a basis for future research in the domain of linguistic relativity.

1.3 Typology of motion events in Dutch and English. Dutch and English differ in their typological patterns with regard to the temporal properties of motion events: Dutch is generally regarded as a non-aspect language, in which speakers tend to encode endpoints (von Stutterheim et al., 2012). However, van Beek, Flecken and Starren (2013) argue that Dutch may not be a true non-aspect language, because although it has no grammatical means of marking ongoingness, speakers are able to denote it by using lexical means such as *aan het (V-inf) zijn* ('to be at the (V-inf)'). Von Stutterheim, Carroll and Klein (2009) contend that

structures like this one are in the process of grammaticalisation, and it could therefore be possible that native speakers of Dutch employ a more intermediate temporal viewing frame than speakers of a true non-aspect language, such as German. However, I decided to classify Dutch as a non-aspect language for the purposes of this study, since no empirical studies have been conducted on whether Dutch is in the process of becoming an aspect language. Moreover, since previous research on Dutch within the domain of temporal properties of motion events (e.g. von Stutterheim et al., 2012) classified this language as a non-aspect language, I will do so as well. In contrast to Dutch, English is an aspect language, in which speakers can employ grammatical means to describe aspect, and they tend to leave out the endpoint of a motion (von Stutterheim et al., 2012). Therefore, although Dutch may not be seen as a true non-aspect language, the ways in which speakers of English and Dutch lexicalise aspect differ and possibly the manner in which they regard motion events as well, so these speakers lend themselves well to investigation on whether the typology of a speaker's first language remains dominant in the acquisition of a second language. As there presently is no research on this topic regarding advanced Dutch learners of English in the domain of temporal properties of motion events, only on advanced German learners of English (von Stutterheim and Carroll, 2006), and Dutch monolinguals (von Stutterheim et al., 2012), the current paper will investigate the possible influence that Dutch as a first language may have on the cognition of temporal properties of motion events in English as a second language. This will be executed by observing the endpoint behaviour of native speakers of Dutch who are proficient in English and analysing these data in order to examine whether the participants' descriptions of motion events in their L2 differ from their native language.

The present experiment will be executed in the same manner as von Stutterheim and Carroll (2006), who scrutinised the influence of L1 German on L2 English with regard to the cognition of motion events. These authors found that German speakers mention endpoint

significantly less often when speaking L2 English than when speaking L1 German (in 36.7 versus 76.4 percent of the test items). However, they state that the L1 German participants only left out the endpoint when this could easily be inferred from the motion and that, even at an advanced level, learners of a second language rely on their first language with regards to the lexicalisation of motion events. They further argue that identifying and implementing L2 principles is a recurring issue in teaching and learning a second language, as it is difficult for learners of a second language to identify and activate the target principles of reporting motion events, and they may still draw on their L1 principles to do so.

Since German and Dutch are comparable regarding their typology of motion events, as they are both non-aspect Germanic languages, the same pattern may be expected when this study is conducted with participants whose first language is Dutch, meaning that they will encode an endpoint more often when speaking their first language, namely Dutch, than when speaking their second language, which is English. This may then indicate that speakers undergo cognitive restructuring when lexicalising motion events in a second language that differs in typology from their first language, which might have implications for L2-teaching, as it may make a speaker sound more native-like if this restructuring is considered in the learning curriculum.

2. Research Question

How does Dutch as a first language affect the lexicalisation of goal-oriented motion events in English as a second language, and is the pattern found in Dutch advanced learners of English similar to the pattern found in German advanced learners of English, as investigated in von Stutterheim and Carroll (2006)?

3. Hypotheses

3.1 Hypothesis 1. Native speakers of Dutch will express a lower endpoint preference when they lexicalise goal-oriented motion events in English (i.e. events where a possible endpoint is

not reached), than when they do so in Dutch, as they are influenced by the pattern of focus on the ongoingness that English native speakers have. However, this effect will be limited, as, according to von Stutterheim and Carroll (2006), learners of a second language still rely on the typology of their first language when lexicalising a motion event.

3.2 Hypothesis 2. In the control items, where the endpoint of a motion is reached, no difference between the two languages is expected. In these videos, the endpoint is visibly reached, so regardless of whether a participant speaks English or Dutch, they should mention endpoints equally as often in both languages.

3.3 Hypothesis 3. The pattern of lexicalisation of goal-oriented motion events found in Dutch advanced learners of English will be similar to the pattern found in German advanced learners of English: Dutch learners will lexicalise endpoints in around 76 percent of their descriptions of motion events in Dutch, and around 36 percent in English, which are the numbers found in von Stutterheim and Carroll (2006).

3.4 Hypothesis 4. The exposure to English that a participant undergoes will have a positive relation to endpoint encoding in the L2, to the extent that more exposure correlates with more native-like patterns of encoding of endpoints, i.e. that fewer endpoints are mentioned in English than in Dutch.

4. Methodology

4.1 Participants. The participant group consisted of 20 native speakers of Dutch, from comparable socio-cultural backgrounds (i.e. students enrolled in a programme at a university), aged between 20 and 23 (with a mean of 21.5). All participants were given a questionnaire about their linguistic background with regard to English (see appendix A); their average self-awarded score concerning their proficiency in English was 7.5 out of 10. They all received lessons in English around the age of 10 (*groep 6* 'group 6' in Dutch primary schools) until the last year of secondary school, when they were around the age of eighteen. In the

questionnaire, the participants were asked to indicate how many hours a day they were exposed to the English language. This was divided into four main categories: Listening, watching (tv shows, series, etc.), reading and gaming. See table 1 for the mean exposure of the participants to English in the four categories, as well as their total mean exposure. As is evident from the table, the participants showed the greatest variance in listening, and the smallest in gaming.

Table 1

Mean Exposure to the English Language, in Hours per Day, in Four Categories and the Total Daily Exposure

	Listening	Watching	Reading	Gaming	Total
M	1.75	2.38	1.40	0.25	5.93
SD	1.08	0.95	0.93	0.53	2.38

If a speaker indicated that they were proficient in a language other than Dutch or English, they were excluded from the study, in order to eliminate the possible interference of another language on the participants' lexicalisations in Dutch and English. With each participant, I checked whether the majority of their studies was done in English. If this was the case, they were also excluded from the study, as they may have had a higher level of English than the other participants. Students of linguistics were not allowed to participate in the experiment, as they may have had metalinguistic awareness that could have influenced the results.

4.2 Materials. The materials consisted of forty clips, each clip being about six seconds long. The clips were developed by the team of Christiane von Stutterheim, and kindly lent to me for the purposes of this study. The study after which the present paper is modelled, i.e. von Stutterheim and Carroll (2006), included more test and filler items, but their team only sent me twenty items that served as critical/control items and twenty that served as fillers,

which is why this number of items was used in this study. After each clip, a black screen with a white focus point was shown, which lasted eight seconds. During this interval, the participants were asked to verbalise the information they gathered from the clip. Each clip depicted an everyday situation. Ten critical items were combined with twenty filler items and ten control items. The critical items depicted a scene in which a figure (an animal, person or a vehicle) moved along a trajectory, e.g. down a road or an alley. In these items, only the initial and/or intermediate phase of the movement was shown, so the scene ended before the object reached the endpoint. However, a possible endpoint was depicted in all cases (e.g. a van driving along a road, with a village visible in the background). The control items depicted scenes in which an endpoint was reached, e.g. a woman walking into a shop. See appendix B for a description of the critical and control stimuli. The filler items showed both static scenes, e.g. a bicycle resting against a tree, and dynamic scenes, e.g. a woman making a necklace.

The participants described half of the items in Dutch, the other half in English. Two videos were created, video A containing critical items 1-5 and control stimuli 1-5 and ten filler items; video B contained critical items 6-10 and control items 6-10 combined with ten filler items. Each video contained twenty video clips in total. Critical and control items were pseudo-randomly alternated with an equal number of filler items: One control item would be followed by one filler item, but when a control item was followed by a critical item, two fillers would follow. No more than two control or critical items were shown after one another. The videos were presented to the participants in different orders, and they were asked to describe the clips in either Dutch or English, depending on which group they were assigned to. This was done to minimise the biasing effect that the order in which the videos were presented could possibly have. Group 1 described the video A in Dutch and the second in English; group 2 did the same but the participants in this group were presented with video B before video A. Group 3 described video A in English and video B in Dutch; group 4 did so

too but these participants were presented with video B first. Before starting the experiment, the participants underwent a training phase. The videos in this phase consisted of filler items, for example a man shredding a piece of paper.

  4.3 Procedure. The present study follows the procedure as designed by von Stutterheim and Carroll (2006), in order to ensure that the results from this experiment and that conducted by von Stutterheim and Carroll could be compared. In that study, the instructions for the participants were not explicitly mentioned, but dr. von Stutterheim confirmed that they were the same as in von Stutterheim et al. (2012), in which they were written in full.

  Each participant was presented with the same instructions. The English instructions were the same as those in von Stutterheim et al. (2012, p. 856), but I changed the number of items from sixty to twenty per language; I translated these instructions into Dutch myself for the purposes of the present study (see appendix C). The general instructions at the start of the experiment were presented in Dutch to each participant. All participants underwent an initial training phase, in which they were asked to describe three video clips in English and three in Dutch. For the training phase, they received instruction in Dutch for the first three practice videos and afterwards, they received instruction in English for the next three videos. The participants were allowed to ask questions during this phase. Whether the participant first received instructions in English or Dutch on the subsequent experiment depended on the group they were assigned to. If a participant was assigned to a list in which they had to begin with describing the clips in English, the instructions were first given in this language, in order to ensure that they had activated it and understood that they were to answer in English. When they had finished the first half of the experiment, the same instructions were given, but translated into Dutch (or vice versa if the participant was assigned to a different list). The responses were recorded and transcribed. Each session lasted about twenty minutes. After

completing the task, they filled in a questionnaire concerning their language background (see appendix A).

4.4 Analysis. The data were transcribed and subsequently coded for references to endpoints. The analysis was done in a similar manner to von Stutterheim et al. (2012), since the analysis of von Stutterheim and Carroll (2006) was not explicitly stated. The percentage of endpoints mentioned was analysed and compared between the two conditions (i.e. critical and control). No difference was made in how the endpoint was referenced, so if a participant did so by means of verbal morphology or by using adjuncts, these types of encoding were both counted as an endpoint (e.g. *entering X* and *driving towards X*). The number of endpoints that were encoded was counted and a mean percentage was calculated per language (Dutch or English) and condition (critical or control).

5. Results

Table 2 depicts the percentage of endpoints mentioned by the participants in the different conditions. As is evident from this table, the participants encoded higher percentages of endpoints in both the critical and the control condition in Dutch, but this is merely a small discrepancy.

Table 2
Mean Percentage of Endpoints Mentioned in Critical and Control Items, Lexicalised in Dutch and English.
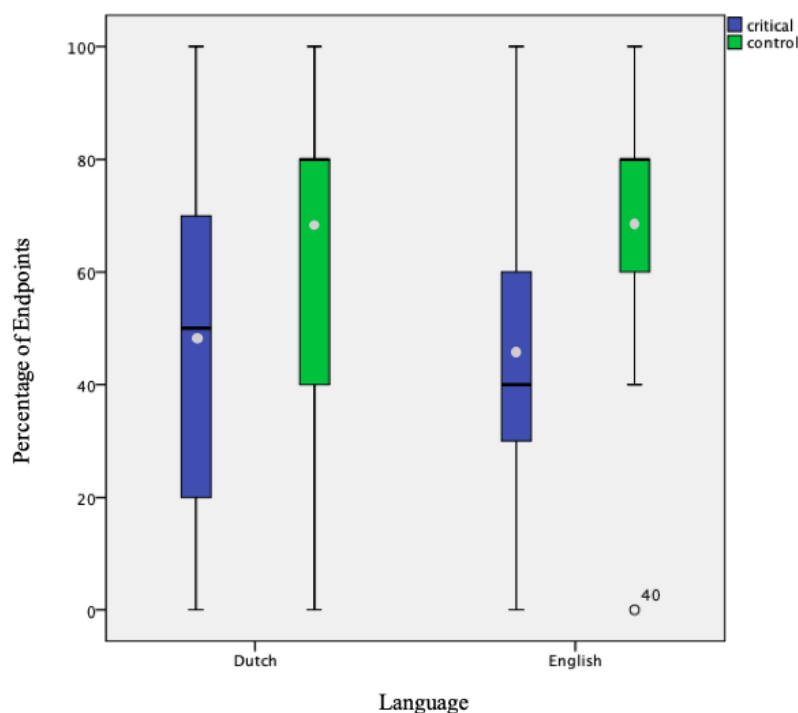
	Dutch		English	
	Critical	Control	Critical	Control
M	48.00	67.00	44.00	69.00
SD	31.39	28.49	27.22	22.92

In order to test the hypotheses of the present study, a Repeated Measures ANOVA was conducted, in which condition (i.e. critical vs. control) and language (i.e. Dutch vs. English) were included as between-subjects effects. The test revealed no main effect of language ($F =$

(1, 19) = 0.041, *p* = 0.843). As a result, the first hypothesis, i.e. that Dutch advanced learners

of English would encode more endpoints when speaking Dutch, is rejected. Contrastingly, a

main effect of condition was found when conducting this test (*F* (1, 19) = 22.65, *p* < 0.005).

The second hypothesis, i.e. that the participants would not demonstrate a difference in the

encoding of endpoints in the control items, can therefore be accepted. Between the two

variables, no significant interaction was found (*F* (1, 19) = 0.774, *p* = 0.039). This lack of

interaction may suggest that the participants encoded more endpoints in the control condition

than in this critical condition, regardless of which language they speak.

Figure 1
Percentage of Endpoints Mentioned in the Critical and Control Condition (Mean Indicated by Small Dots), Lexicalised in Dutch and English

The third hypothesis, i.e. that Dutch advanced learners of English would demonstrate

the same pattern in their descriptions of motion events as German advanced learners of

English, as scrutinised by von Stutterheim and Carroll (2006), is also rejected, as the

percentages of encoding of endpoints does not coincide between the present study and von

Stutterheim and Carroll (2006).

The relation between the mean daily exposure to the English language and endpoint behaviour of the participants was studied by examining the correlations between these two variables. Weak positive correlations were found between exposure and the percentage of endpoints mentioned by the participants. This could indicate that the more a participant was exposed to English, the more endpoints they encoded. However, since this correlation was not significant (r (18) = .131, p = 0.292), the fourth hypothesis of the present study, i.e. that more exposure to the English language would correlate with the encoding of fewer endpoints in English, is rejected.

Surprisingly, figure 1 demonstrates that the variation in the data about the control items is skewed, as the median coincides with the 75th percentile. This means that the dispersion among the smaller values of the data set (i.e. the lower percentages of endpoints mentioned) is greater than among the larger values. This pattern can be identified in both Dutch and English. However, the range is smaller for the English control items than for the Dutch control items (i.e. 40-100% vs. 0-100%), which means that the distribution of the English control items was smaller than that of the Dutch control items. One outlier can be seen in the box plot: This participant did not encode any endpoints at all. The box of the Dutch critical items shows that fifty percent of the data (between the first and the third quartile) is more widely dispersed around the median (namely from 20% to 70%) than the critical items lexicalised in English (namely from 30% to 60%). Figure 1 thus demonstrates that, although no significant difference was found between the mean percentages of endpoints encoded in both conditions and languages, a difference in the distribution of these variables can be seen in the box plots.

6. Discussion

The current study found that Dutch advanced learners of English do not encode endpoints more in their first language than they do in their second language. This result may be explained by a variety of factors. Firstly, the results of this study revealed that advanced Dutch learners of English encode endpoints in their lexicalisations of motion events at an equal percentage in their first and their second language, in contrast to the first hypothesis of the present study, which stated that these speakers would encode more endpoints in Dutch. I hypothesised that the learners would be influenced by the intermediate temporal viewing frame on motion events that speakers of English have, as opposed to their native maximal temporal viewing frame. The results may be explained by the fact that, although the learners were advanced in English, they may still have to infer some psycholinguistic aspects of the English language, including a viewing frame through which to regard motion events that differs from their first language, as Dutch speakers tend to employ a holistic viewing frame, as an effect of the lack of grammatical aspect in Dutch. When English is taught in Dutch primary and secondary schools, the emphasis is mostly on being able to express oneself comprehensibly in a second language, rather than being able to employ the same psycholinguistic structures as native speakers, so Dutch learners of English need to infer this from their own experience, but the chances that they achieve this may be small.

Secondly, another factor that might explain the results of this study is the argument made by van Beek, Flecken and Starren (2013), namely that Dutch is not truly a non-aspect language, because, though it does not have a true grammaticised means of aspectually marking ongoingness, it can be denoted by using constructions such as *aan het (V-inf) zijn* ('to be at the (V-inf)'). Von Stutterheim, Carroll and Klein (2009) contend that structures like this one are in the process of grammaticalisation, and it is therefore possible that speakers of Dutch start to employ a more intermediate temporal viewing frame, in which they focus more

on the movement than on the image as a whole, when processing temporal information about motion events. This might then explain why no difference between the English and Dutch descriptions of the videos was found in this study. This is further supported by the result that von Stutterheim and Carroll (2006) found, namely that German native speakers encoded endpoints in 76.4 percent of their descriptions of videos depicting goal-oriented motion events. If Dutch were truly a non-aspect language, such as German, the percentage of endpoints mentioned should have been around the same number, which is not the case (in the present study, the participants only mentioned endpoints in 48 percent of the cases). In von Stutterheim and Carroll (2006), the English monolingual speakers encoded endpoints in 25.2 percent of the cases, so the participants tested in this study are in between the English and the German norm. Moreover, when comparing the box plots in the present study (see figure 1) to those in von Stutterheim et al. (2012), the mean relative frequency (about 45 percent) of endpoints mentioned found with monolingual English speakers in von Stutterheim et al. (2012) coincides with the mean frequency of endpoints mentioned in Dutch in the current study (the relative frequencies in von Stutterheim et al. (2012) were compared to the data in the present study by converting the numbers from a scale of 0-1 to a scale of 0-100). It could then be argued that, because the patterns found in Dutch do not coincide with those associated with a true non-aspect language, the present study may provide evidence that Dutch may be moving towards becoming a language in which the aspect of a verb can be grammaticised and speakers employ an intermediate temporal viewing frame as a result.

Moreover, some participants used the adverb *voorbij* ('past') when describing the videos in Dutch, see the example in (1):

(1) *Er   loopt   een   vrouw   voorbij*

There   walk-3SG   a   woman   past

'A woman walks by'

However, they did not employ a similar construction in their English descriptions. Using this adverb prevents the possible encoding of an endpoint, which they might have done if they had employed another type of description. Since this kind of structure was not expressed in the English descriptions, this may have led to a more minor difference between the percentage of encoding of endpoints between the two languages. Additionally, one participant did not encode any endpoints in their descriptions of the videos, either because they understood the instructions too literally ('describe the motion only'), or they employed the temporal viewing frame in both Dutch and English. Since I do not know which is the case, and moreover, because excluding this participant from the study would not result in a statistically significant difference between the English and the Dutch descriptions, I decided to include these data in my analysis.

Contrary to the expectation to find the same pattern as von Stutterheim and Carroll (2006), namely that learners of English mention endpoints less often in English than in their first language (in the study conducted by von Stutterheim and Carroll (2006), this was German), the present study found that Dutch learners of English demonstrated an equal percentage of endpoints encoded in their first and second language. In von Stutterheim and Carroll (2006), advanced learners were also asked to produce descriptions of videos depicting motion events. The difference between this previous study and the present one could have been caused by the level of English that the participants had, and the conventions of teaching in Germany. No information, other than the fact that the participants were advanced learners of English, was given about them. Therefore, I cannot state with certainty that the population employed in von Stutterheim and Carroll (2006) was comparable to the one employed in this study. If German schools put more emphasis on producing utterances containing native-like underlying psycholinguistic structures, it is reasonable to expect that the difference in the percentage of encoded endpoints between German and English will be higher, as German

learners of English may have acquired the L2 principles at a higher level than Dutch learners. Additionally, the participants in von Stutterheim and Carroll (2006) may have been exposed to English for a longer time, which may be another explanatory factor for the contrast between the two studies. Moreover, as von Stutterheim and Carroll (2006) state, L1 principles remain dominant in learners' productions in a second language, so although the German learners in the study encoded more endpoints in German than English, they did not fully move toward the English target norm. This phenomenon may have occurred with the participants in this study as well, but the principles of Dutch may have remained more dominant, either because the participants were less proficient, or they may have had less experience with the temporal viewing frame of English native speakers than the German learners.

Moreover, in von Stutterheim and Carroll (2006), it was found that German learners of English move more towards the target norm than English learners of German when encoding endpoints in their lexicalisation of motion events, which may indicate that having a first language that is a non-aspect language has a larger effect on a second language that is an aspect language than vice versa. Since the present study did not find an effect of the first, non-aspect language, on the second language, which was an aspect language, future research may investigate English learners of Dutch, in order to scrutinise whether speakers in this population demonstrate a different pattern than that found in the current study. Von Stutterheim et al. (2012) found that English advanced learners of German do not demonstrate a difference in the encoding of endpoints between their first and their second language, which indicates that an aspect language may have the same effect on a non-aspect language as vice versa, when investigating advanced learners of these languages.

The findings from the current study are in accordance with Slobin's *thinking-before-speaking* hypothesis, since it can be argued that the participants in the present study are more attentive to those concepts that are more salient in their native language, namely Dutch, even

when speaking a second language, in this case English. The participants may have employed the same temporal viewing frame that they use in their L1 when lexicalising motion events in their L2, which could explain why the same percentage of endpoints that were encoded was found between the two languages.

Lastly, Bylund and Athanasopoulos (2015) state that the extent to which L2 speakers' verbal behaviour approximates the verbal behaviour of native speakers depends on concept similarities between the two languages and other factors such as frequency of L2 use, length of exposure and age of acquisition. Future research may vary these factors when investigating the influence of L1 Dutch on L2 English concerning temporal properties of motion events, as more exposure to English, an earlier age of acquisition and more frequent use of the language might cause Dutch learners of English to encode fewer endpoints in their L2, because they may have encountered the English native temporal viewing frame more often. Furthermore, English and Dutch are different with regard to the viewing frames that speakers of these languages typically employ (intermediate vs. maximal). Research regarding language pairs in which these viewing frames are more similar, for example German and Dutch, may identify whether the variation in the encoding of endpoints is a result of the first language being conceptually different than the second language, or whether it is an effect of speaking a second language in general.

7. Conclusion and Implications

The analysis of the data resulted in the finding of no significant difference in the encoding of endpoints in Dutch and English. It can therefore be concluded that Dutch as a first language affects the lexicalisation of temporal properties of motion events in English to such an extent that advanced learners still employ the Dutch viewing frame when speaking English. Although the participants acquired the grammatical means for expressing aspect in English,

they might not have obtained the temporal viewing frame that native speakers employ, which would have caused them to encode endpoints to a lesser extent.

This finding, though preliminary, holds implications for the teaching of a second language; English in particular. The participants in the present study had completed the whole English language curriculum typical for Dutch pupils, so this study provides evidence for the idea that this curriculum does not enable students to acquire the native English viewing frame. If a learner wishes to produce more native-like utterances, it is crucial that they understand not only the grammatical structures and vocabulary, but also the underlying psycholinguistic aspects of a given language. This can be achieved by focussing on these underlying principles at an earlier stage in the teaching curriculum. More research in this domain is required, then, as linguistic relativity may not only be present in speakers of English and Dutch, and not only in the domain of motion events, but in all languages and probably in all linguistic aspects. As von Stutterheim and Carrol (2006) argue, identifying and implementing L2 structures remains an issue in learning and teaching a second language, so future studies on linguistic relativity in learners of L2 learners are recommended if second language education is to be improved.

References

Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition, 9,* 89-96. doi:10.1017/S1366728905002397

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition, 12*, 83-95. doi:10.1017/S136672890800300X

van Beek, G., Flecken, M., & Starren, M. (2013). Aspectual perspective taking in event construal in L1 and L2 Dutch. *International Review of Applied Linguistics in Language Teaching, 51,* 199-227. doi:10.1015/iral-2013-0009

Bylund, E., & Athanasopoulos, P. (2015). Introduction: Cognition, motion events, and SLA. *The Modern Language Journal, 99,* 1-13. doi:10.1111/modl.12175

Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion event cognition and grammatical aspect: Evidence from Afrikaans. *Linguistics, 51,* 929-955. doi:10.1515/ling-2013-0033

Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualisation. *Bilingualism: Language and Cognition, 14,* 47-59. doi:10.1017/S1366728910000180

Flecken, M., Von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: findings from a cross-linguistic non-verbal recognition task. *Language and Cognition, 6,* 45-78. doi:10.1017/langcog.2013.2

Gennari, S.P., Sloman, S.A., Malt, B.C., & Tecumseh Fitch, T. (2001). Motion events in language and cognition. *Cognition, 83,* 49-79. doi:10.1016/S0010-0277(01)00166-4

Navarro, S., & Nicoladis, E. (2005). Describing motion events in adult L2 Spanish narratives. In D. Eddington (Ed.). *Selected proceedings of the 6th conference on the acquisition of*

Spanish and Portuguese as first and second languages (pp. 102-107). Somerville, MA: Cascadilla Proceedings Project.

Nikolayeva, O., Buz, E., Liu, L., Watts, A., & Jaeger, T. F. (2015). *Web based tutorial on experimental research design.* Retrieved from https://www.hlp.rochester.edu/resources/BCS152-Tutorial/

Sharpen, R. (2016). L1 conceptual transfer in the acquisition of L2 motion events in Spanish and English. *Open Linguistics, 2*, 235-252. doi:10.1515/opli_2016-0011

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.) *Rethinking linguistic relativity* (pp.70-96). Cambridge: Cambridge University Press.

von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics, 50,* 833-867. doi:10.1515/ling-2012-0026

von Stutterheim, C., & Carroll, M. (2006). The impact of grammatical temporal categories on ultimate attainment in L2 learning. In H. Byrnes, H. Weger-Guntharp & K. A. Sprang (Eds.), *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment* (pp. 40-53). Washington, DC: Georgetown University Press.

von Stutterheim, C., Carroll, M., & Klein, W. (2009). New perspectives in analyzing aspectual distinctions across languages. In W. Klein & P. Li (Eds.), *The expression of time* (pp. 163-184). Berlin, Germany: De Gruyter.

Talmy, L. (2000). *Toward a cognitive semantics.* Cambridge, MA: MIT Press.

Whorf, B. L. (1940/1956). Linguistics as an exact science. In J. B. Carroll (Ed.), *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (pp. 220-232). Cambridge, MA: MIT Press.

Appendix A: Questionnaire language proficiency

Beantwoord alle onderstaande vragen.

Leeftijd: _____

Geslacht: _____

Opleiding: _____

1. Welke taal (of talen) spreek je thuis?

_____

_____

2. Wanneer kreeg je voor het eerst lessen Engels? (Bijvoorbeeld: groep 6, eerste klas etc.)

_____

3. Deze vraag gaat over hoeveel je wordt blootgesteld aan de Engelse taal.

Hoeveel uur **per dag** besteed je gemiddeld aan de volgende activiteiten? Vul het schema in.

Je kunt ook 0 uur of 0,5 uur invullen.

Activiteit:	**Uur per dag:**
Luisteren naar Engelse muziek/audioboeken/podcasts	
Kijken van Engelse video's op YouTube	
Kijken van Engelse tv-programma's	
Kijken van Engelse series/films (op Netflix etc.)	
Lezen van Engelse artikelen/boeken/blogs	
Spelen van games in het Engels	

4. Hoe goed vind je dat je Engels spreekt op een schaal van 1 tot 10? Omcirkel het cijfer.

Heel slecht   1   2   3   4   5   6   7   8   9   10   Bijna native

Appendix B: Descriptions of critical stimuli and control stimuli

Critical stimuli: Endpoint not reached

Video clip	Description
1	A woman and a child walking (towards a playground)
2	A woman riding a horse (towards a barrier)
3	Two women walking (towards a house)
4	A man walking across a street (towards a car)
5	A van driving (towards a village)
6	A woman walking across a parking lot (towards a car)
7	A dog running through a meadow (towards a fence)
8	A boy walking (towards a playground)
9	A woman walking down an alley (towards a barrier)
10	A woman walking (towards a café)

Control stimuli: Endpoint reached

Video clip	Description
1	A train driving into a tunnel
2	A woman riding a horse into a stable
3	A man walking into a building
4	A cat walking into a room
5	A woman walking into a shop
6	A man cycling into a forest
7	A man cycling into a gate
8	A woman running into a train station
9	A van driving into a gate
10	A dog running into a building

Appendix C: Instructions for participants

We beginnen met een oefenvideo. Je ziet zometeen twee keer een set van 3 videoclips. Bij de ene set krijg je instructie in het Nederlands, bij de andere in het Engels. Antwoord alsjeblieft in de taal waarin de instructie is gegeven. Je mag tijdens de oefenvideo vragen stellen als er iets onduidelijk is. ('We will start with a practice video. You will see two sets of 3 video clips. With one set, you will receive the instruction in Dutch, with the other one in English. Please answer in the language in which the instruction is given. You may ask questions during the practice video when something is unclear.')

De eerste drie video's: Tussen de clips zie je een zwart scherm met een wit focuspunt, richt je aandacht alsjeblieft op dit punt. Jouw taak is om antwoord te geven op de vraag 'wat gebeurt er', je mag beginnen met praten zodra je herkent wat er gebeurt in de clip. Het is niet nodig om de clip in detail te beschrijven (bijv. 'De lucht is blauw'). Concentreer je alsjeblieft alleen op de beweging.

The next three videos: Between the clips you will see a black screen with a white focus point. Please focus on this point. Your task is to tell 'what is happening', and you may begin as soon as you recognise what is happening in the clip. It is not necessary to describe the video clips in detail (e.g. 'the sky is blue'). Please focus on the event only.

Dat was de oefenvideo. Je gaat nu twee keer een set van 20 videoclips zien. Bij de ene set krijg je instructie in het Nederlands, bij de andere in het Engels. Antwoord alsjeblieft in de taal waarin de instructie wordt gegeven. ('That was the practice video. You will now see two sets of 20 video clips. With one set, you will receive the instruction in Dutch, with the other in English. Please answer in the language in which the instruction is given.

Je ziet zometeen een set van 20 videoclips, waarin alledaagse dingen worden afgebeeld die op geen enkele manier met elkaar verbonden zijn. Voordat elke clip begint zie je een zwart scherm met een wit focuspunt. Richt je aandacht alsjeblieft op dit punt. Jouw taak is om antwoord te geven op de vraag "wat gebeurt er?' en je mag beginnen zodra je herkent wat er gebeurt in de clip. Het is niet nodig om de videoclips in detail te beschrijven (bijv. 'de lucht is blauw'). Concentreer je alsjeblieft alleen op de beweging.

You will see a set of 20 video clips showing everyday events which are not in any way connected to each other. Before each clip starts, a blank screen with a white focus point will appear. Please focus on this point. Your task is to tell "what is happening", and you may begin as soon as you recognize what is happening in the clip. It is not necessary to describe the video clips in detail (e.g., 'the sky is blue'). Please focus on the event only.