Language and thought in a multilingual context: The case of isiXhosa*

EMANUEL BYLUND
Stockholm University, Sweden & Stellenbosch University, South Africa
PANOS ATHANASOPOULOS
University of Reading, UK

Situated within the grammatical aspect approach to motion event cognition, this study takes a first step in investigating language and thought in functional multilinguals by studying L1 isiXhosa speakers living in South Africa. IsiXhosa being a non-aspect language, the study investigates how the knowledge and use of additional languages with grammatical aspect influence cognition of endpoint-oriented motion events among L1 isiXhosa speakers. Results from a triads-matching task show that participants who often used aspect languages and had greater exposure to English in primary education were less prone to rely on endpoints when categorising motion events.

Keywords: linguistic relativity, motion events, grammatical aspect, multilingualism, isiXhosa

1. Introduction

During the past two decades, the study of linguistic relativity, or the question of whether speakers of different languages think differently (Whorf, 1956), has experienced a remarkable surge. Owing to refined definitions – and subsequent operationalisations – of the notions of LANGUAGE and THOUGHT, researchers within this paradigm have been able to break down the big question of whether language influences thought into the more manageable task of examining which linguistic categories influence which cognitive processes under which conditions. A growing body of evidence shows that crosslinguistic differences in the semantic partitioning of reality may sometimes give rise to crosslinguistic differences in the ways that speakers perceive, remember, sort, and categorise, for example, objects and substances (Imai & Gentner, 1997; Lucy, 1992), colour (Athanasopoulos, Dering, Wiggett, Kuipers & Thierry, 2010; Davidoff, Davies & Roberson, 1999), motion (Athanasopoulos & Bylund, 2013; Papafragou & Selimis, 2010), time (e.g., Boroditsky, Fuhrman & McCormick, 2011; Miles, Tan, Noble, Lumsden & Macrae, 2011), and space (e.g., Haun, Rapold, Janzen & Levinson, 2011).

More recently, scholars within the fields of second language acquisition and bilingualism have started to systematically explore the relationship between language and thought in individuals with knowledge of more than one language (Jarvis & Pavlenko, 2008; Pavlenko, 1999, 2005). This development is reflected in the numerous edited volumes and special journal issues in the past few years dedicated to language and thought in second language (L2) speakers (Cook & Bassetti, 2011; Han & Cadierno, 2010; Jarvis, 2011; Jarvis & Pavlenko, 2008; Pavlenko, 2011b). Adopting a bilingual perspective of language and thought can be considered a natural extension of the linguistic relativity principle: if speakers of different languages exhibit different cognitive behaviour, how do speakers who have functional knowledge of more than one language behave? Findings from current research on language and thought in bilinguals suggest that the degree to which language-specific cognitive behaviour is maintained or separated in the bilingual mind relates to factors such as language proficiency, length of cultural immersion, age of acquisition onset, and frequency of language use (e.g., Athanasopoulos, 2007, 2009; Athanasopoulos et al., 2010; Boroditsky, 2001; Bylund, Athanasopoulos & Oostendorp, 2013).

A common trait of this line of research is the operationalisation of speakers of more than one language as L2 speakers or foreign language learners, living in typically monolingual contexts (e.g., Boroditsky, 2001). It is important to keep in mind, however, that these speakers and the contexts in which they live represent only one instance of the different profiles that speakers of more than one language might exhibit. As a consequence of this bias, very little is known about language and

Address for correspondence:
Emanuel Bylund, Centre for Research on Bilingualism, Department of Swedish and Multilingualism, Stockholm University, 106 91 Stockholm, Sweden
manne.bylund@biling.su.se

thought in speakers of more than two languages (i.e., MULTILINGUAL SPEAKERS) or in contexts where more than two languages are used for communicative purposes (i.e., MULTILINGUAL SETTINGS). In 2005, Aneta Pavlenko made the observation that research into multilingualism and thought is "a lacuna . . . still waiting to be filled" (Pawlenko, 2005, p. 447). Close to a decade later, this remark still stands: currently, there is to the best of our knowledge no published study that experimentally addresses language and thought in multilingual speakers living in multilingual contexts. In our view, at least three arguments can be posited in favour of extending language and thought research to the domain of multilingualism. The first relates to the fact that there is at present a scarcity of evidence on how language-specific cognitive patterns develop in the multilingual mind. The second argument concerns the observation that multilingualism represents a common linguistic situation worldwide (e.g., Aronin & Singleton, 2012).[1] There is, in other words, ecological validity to be gained from studying language and thought through a multilingual lens. The third argument relates to context-boundedness in research, and concerns the link between the context in which research is carried out and the methodological and epistemological watermarks that the context leaves on the research practice (e.g., Henrich, Heine & Norenzayan, 2010). Put more concretely, studying multilingual contexts and speakers could allow us to examine how relativistic methods and theories that have developed out of largely monolingual contexts translate to multilingual contexts.

Situated within the grammatical aspect approach to motion event cognition and conceptualisation (Athanasopoulos & Bylund, 2013; Flecken, 2011; von Stutterheim, Andermann, Carroll, Flecken & Schmiedtová, 2012), this paper constitutes a first attempt to empirically approach the question of multilingualism and thought by studying motion event cognition in isiXhosa-speakers in multilingual South Africa. Throughout the article, the terms THOUGHT (as in "language and thought") and LINGUISTIC RELATIVITY will be used to denote crosslinguistic differences in NON-VERBAL BEHAVIOUR, that is, cognitive processes such as categorisation, sorting, memory, and categorical perception that do not take place in relation to overt speech production or comprehension (see Lucy, 1997).[2]

It is precisely these non-verbal behaviours defined in the modern version of the linguistic relativity framework (Lucy, 1997) that this brief research note will focus on.

2. Background

2.1 Bilingualism and linguistic relativity

In spite of the remarks made by von Humboldt in the 19th century about thought processes in bilinguals (Pavlenko, 2011a, p. 11), it is only recently that bilingualism and thought has started to become an integrated area of research (with the exception of a few pioneering studies, e.g., Brown & Lenneberg, 1954). Findings from recent studies usually show that individuals who speak two languages with contrasting semantic structure differ from monolingual speakers in a variety of perceptual domains. For example, in the domain of objects, Athanasopoulos (2007) explored categorisation preferences among L1 Japanese – L2 English bilinguals (see also Cook, Bassetti, Kasai, Sasaki & Takahashi, 2006). Due to grammatical differences in mass and count nouns, Japanese speakers base their object similarity judgements on material or substance, whereas English speakers are likely to base theirs on the shape of the objects (Imai & Gentner, 1997). The results showed that the bilinguals were more prone to matching objects on the basis of shape rather than material, thus approximating English preferences. This behaviour was modulated by English language proficiency, such that those with higher proficiency were more likely to behave like English native speakers.

Studies on grammatical gender have shown that additional language learning may influence the perception of object characteristics (e.g., Forbes, Poulin-Dubois, Rivero & Sera, 2008). Using the voice attribution paradigm, these studies have asked participants to assign male or female voices to a set of objects, showing that learning a new language with a distinct grammatical gender system influences voice attribution. For example, Kurinski and Sera (2011) found that English learners of Spanish were more prone to assign male/female voices in accordance with the Spanish grammatical gender of the objects, even when the task was carried out in English.

---

[1] It should be noted, however, that many estimations of worldwide bilingualism and multilingualism often use the dichotomy monolingual and "non-monolingual". For this reason, claims about exact numbers regarding monolingual and bilingual vis-à-vis multilingual speakers and contexts will remain uncertain.

[2] Admittedly, some of these behaviours (e.g., sorting) may involve inner speech. Such behaviours may be said to be VERBALLY MEDIATED. It is important to keep in mind, however, that there is still an important difference between non-verbal tasks that are verbally mediated (e.g., sorting) and verbal tasks per se (e.g., picture description):

whereas the latter are concerned with eliciting and studying overt speech production (where some lexical or grammatical feature is the dependent variable), and explicitly instruct participants to engage in some form of verbalisation, the former are concerned with measuring the non-verbal outcome (where the dependent variable is some sort of sorting or similarity judgement of something other than words or sentences, such as pictures, videos, objects, colour stimuli, etc.) of a higher-level cognitive process that might have relied on inner speech, and participants are explicitly instructed to engage in some sort of non-verbal behaviour avoiding linguistic cues that might bias responses.

In the domain of time, Boroditsky (2001) investigated the cognitive consequences of spatio-temporal metaphors in English and Mandarin (however, for criticisms see January & Kako, 2007). In English, temporal succession is typically conveyed through horizontal metaphors (e.g., "before", "after"), whereas in Mandarin there is a possibility to express succession through vertical metaphors (e.g., "above", "below"). It was found that L1 Mandarin – L2 English bilinguals (living in the US) with later ages of L2 acquisition were, compared with early bilinguals, more prone to rely on vertical spatial cues when determining the temporal succession of two events. Along a similar line, Miles et al. (2011) investigated time conception in Mandarin–English bilinguals. This study found that the bilingual participants made faster judgements than English monolinguals when relying on vertical cues to determine temporal succession.

In the domain of colour, Athanasopoulos (2009) examined categorical perception in L1 Greek – L2 English bilinguals, taking as a starting point the Greek obligatory lexical distinction between light blue and dark blue. Results demonstrated that the longer the bilinguals had spent in the UK, the more likely they were to exhibit weakened categorical distinction between light and dark blue colour stimuli in a similarity judgement task (see also Athanasopoulos et al., 2010). These results were extended by Athanasopoulos, Damjanovic, Krajciova and Sasaki (2011) in another study on colour perception in L1 Japanese – L2 English bilinguals, which showed that frequency of language use determined the degree to which proficient bilinguals attend to their native colour categorical distinctions (for an early investigation into bilingual colour cognition, see Brown & Lenneberg, 1954).

The domain of motion has to-date remained under-researched with regard to bilingual non-verbal behaviour. Using as a point of departure the finding that speakers of languages without grammatical aspect are more prone to attend to event endpoints than are speakers of aspect languages (Athanasopoulos & Bylund, 2013; Flecken, 2011; von Stutterheim et al., 2012), Bylund et al. (2013) examined event categorisation in L1 speakers of Afrikaans (non-aspect language) who had English (aspect language) as an L2. Results from a triads-matching task showed that those who used English more often were less prone to rely on the reaching of endpoints while categorising the motion events, thus exhibiting a behaviour closer to that of English speakers. Age of acquisition and self-reported proficiency with English did not exert any effect on this behaviour.

2.2 IsiXhosa and other languages in South Africa

The 1996 South African Constitution states that Afrikaans, English, isiNdebele, isiXhosa, isiZulu, Sepedi, Sesotho, Setswana, siSwati, Tshivenda, and Xitsonga are the official languages of the Republic of South Africa.

The provincial governments must promote, regulate and monitor the use of these languages, and use AT LEAST two of them. South Africa also recognises a number of languages that are either historical indigenous minority languages (e.g., the Khoi, Nama, and San languages) or languages brought to the country through immigration, indentured labour, or the slave trade (e.g., German, Greek, Gujarati, Hindi, Portuguese, and Tamil), which the Constitution states must be promoted (for further reading, see Mesthrie, 2002). In this diversity of languages, English is the dominant symbolic resource on the linguistic market – despite being the L1 to less than 10% of the population (Setati, 2005).

The high status of English is also reflected in South African education: although policy documents stipulate that pupils have the right to choose their language of instruction, and that schools should work towards multilingualism, English is the dominant medium of instruction. In schools whose pupils are predominantly L1 speakers of a Bantu language (e.g., isiXhosa), English is typically introduced as a medium of instruction in Grade 4, replacing the previous language of instruction. The extent to which English replaces the former language of instruction, however, varies depending on the school and the teacher. Code-switching between English and the former language of instruction is common in the classroom (Ncoko, Osman & Cockcroft, 2000).

In the Western Cape Province, which is the research site of the current study, Afrikaans, English, and isiXhosa are the main languages. In this region, isiXhosa is the "first language spoken at home" by 24.3% of the population.[3] The corresponding numbers for Afrikaans and English are 49.7% and 20.2%, respectively (Census, 2012). Other home languages in the region include isiZulu, Sesotho, Setswana, and siSwati. On a national level, isiXhosa is South Africa's second largest language, with more than eight million speakers (Census, 2012).

IsiXhosa is a Southern Bantu language, belonging to the Nguni group (code S41 in Guthrie's 1948 classification). The aspectual distinctions of perfectivity and imperfectivity are not grammaticalised in isiXhosa, and must be expressed lexically. It should be noted that there is an alternation between so-called "long forms" and "short forms" in the past, present, and future tenses through the verbal infix -*ya*-, which at one point was interpreted by grammarians as a distinction between continuous and non-continuous aspect (MacLaren, 1936). However, more recent analyses show that this is not the case: rather, -*ya*- is a morphosyntactically motivated infix that appears after the subject concord in the verb phrase depending on the transitivity and negation of the

---

[3] This phrasing was used by the South African 2011 census, and should be understood primarily in functional terms (i.e., "mostly spoken"). It is unspecified with regard to order of acquisition and proficiency.

predication, as well as on the presence of other concords. This infix may also be used for emphatic purposes (Du Plessis, 1978; Hobson, 1999). The long form/short form distinction is also found in isiZulu, another language of the Nguni group. However, similar to isiXhosa, the alternation between these forms relates to morphosyntactic, not aspectual, properties (Buell, 2005).

Aspectual categories are found in siSwati, Sesotho, and Setswana. In siSwati (also part of the Nguni group), there is a subcategory of imperfective aspect. According to Taljaard, Khumalo and Bosch (1991) and Ziervogel and Mabuza (1976), the infix -*sa*- is used in siSwati to denote progressivity. This interpretation is, however, challenged by Nichols (2012), who contends that -*sa*- is more appropriately analysed as a marker of persistivity. Persistive aspect is semantically close to progressive aspect, with the difference that it conveys the continuation of a situation throughout two (as opposed to one) distinct temporal intervals, typically translated into English with the temporal adverbial *still* (e.g., *He is still eating*) (Nurse, 2008). In Sesotho, a language belonging to the Sotho-Tswana group, the morpheme -*sa*- denotes progressive aspect, and may occur in past, present, and future tenses (Doke & Mofokeng, 1985; Motsei, 2010). Sesotho has a highly complex aspectual system, with around 15 aspectual categories, according to Nurse (2008). A similar pattern is found in Setswana, another Sotho-Tswana language, which also has a progressive marker (-*sa*-), along with markers of perfect and habitual aspect in the past, present, and future tenses.

The two Germanic languages spoken in the Western Cape, Afrikaans and English, also differ in their aspectual marking. Whereas Afrikaans has no grammatical means to denote contrasts of imperfectivity/perfectivity (Bylund et al., 2013), English conveys progressive aspect through the periphrastic construction *be* + VERB-*ing* (e.g., Comrie, 1976).

To summarise, English, Sesotho, Setswana, siSwati, and isiZulu have at their disposal systematic, grammatical morphemes to denote the unfolding phase of an event without reference to its temporal boundaries, such as the English progressive form *She is playing the piano* (see next section for a discussion of temporal viewing frames). Afrikaans and isiXhosa, on the other hand, lack such means.[4]

2.3 Aims and scope of the present study

The overall aim of the current study is to examine motion event cognition in L1 speakers of isiXhosa living in multilingual South Africa. Specifically, the study seeks to investigate how the isiXhosa speakers' linguistic trajectories affect their motion categorisation patterns.

Motion event cognition is assessed through a triads-matching paradigm, in which participants are asked to pair video clips depicting goal-oriented motion events (Athanasopoulos & Bylund, 2013; Bylund et al., 2013). The study is situated within the theoretical framework of the grammatical aspect approach to motion (see von Stutterheim et al., 2012). According to this framework, the category of aspect, or more specifically, grammatical aspectual markers denoting the ongoing phase of an event without reference to its left or right boundaries, give saliency to the internal temporal constituency of events during the process of conceptualisation.[5] Speakers of aspect languages are, as a consequence, sensitised toward ongoingness (i.e., a continuous aspectual viewpoint) and more prone to taking an immediate viewing frame of an unfolding event whereby the possible event endpoint is excluded. The reason why speakers of languages without grammatical aspect do not exhibit the same behaviour is because they are not pointed by their grammars to pay attention to the ongoing phase of events. Specifically, in the absence of grammatical markers of ongoingness, speakers of non-aspect languages are more inclined to construe events from a maximal temporal viewing frame; that is, they adopt a holistic event perspective in which event endpoints are included (Langacker, 2008).

Against this background, the specific research question pursued in the current study is as follows:

> How does the knowledge and use of additional languages with grammatical aspect influence cognition of endpoint-oriented motion events among L1 isiXhosa speakers?

The scope of the study is confined to exploring the variation in event categorisation behaviour WITHIN the isiXhosa participant group and explaining this variation through the participants' linguistic backgrounds. No comparison group of monolingual isiXhosa speakers is included, since it would be close to impossible to find such individuals who match the participant group in socio-educational background and age.

3. Method

3.1 Participants

Fifty speakers in their mid-twenties with isiXhosa as L1 participated in the study. These individuals were students at a university in the Western Cape, where English was

---

[4] For further reading on the aspectual systems of these languages, the reader is referred to the cited references. It goes beyond the scope of the paper to delve further into the tense and aspect systems of Bantu or Germanic languages.

[5] Henceforth, we will use the phrases "aspect languages" and "grammatical aspect" to refer to the specific typological feature of a grammatical marker that denotes the ongoing phase of an event.

Table 1. *Information on the languages spoken by the participants. Standard deviations in parentheses.*

Language	Spoken by % of participants	Proficiency (1–5)	Age of acquisition	Use (%)
Afrikaans	46	2.2 (0.8)	11.7 (5.0)	8.7 (6.7)
English	100	3.8 (0.7)	7.7 (2.1)	33.2 (13.8)
IsiXhosa	100	4.4 (0.8)	from birth	51.1 (16.3)
IsiZulu	50	3.4 (1.2)	12.7 (4.1)	12.4 (9.2)
Sesotho	26	2.3 (1.0)	13.1 (6.0)	16.1 (16.8)
Setswana	12	2.0 (1.0)	16.0 (4.9)	8.6 (7.4)
SiSwati	4	2.5 (2.1)	18.0 (7.1)	16.0 (12.2)

Notes: Proficiency scale: 1 = "basic", 5 = "excellent". Use is an estimation of the percentage with which each language was used on a weekly basis.

the medium of instruction. All participants had, in other words, knowledge of English. English was also present to different degrees in the participants' education up to university level. On a five-point scale where 1 was "Only English" and 5 "Only isiXhosa", participants reported that the distribution between these languages in primary school averaged at 4.0 (*SD* = 1.4). In secondary school, English was on average used half of the time (2.4, *SD* = 1.5). In addition to English, the participants spoke other languages, such as Afrikaans, isiZulu, Sesotho, Setswana, and siSwati. Each participant spoke on average 3.4 languages (*SD* = 1.0). Table 1 presents information on the languages spoken by the participants.

All participants except two reported having started learning English via formal instruction in school. Out of those who spoke Afrikaans, 52% had learnt this language in school, whereas the rest had learnt it through interaction with Afrikaans native speakers. In the case of the Bantu languages, naturalistic contexts represented the prevailing learning situation: all of those who spoke isiZulu, Setswana and siSwati had learnt these languages through interaction. Out of those who spoke Sesotho, 92.3% had learnt this language naturalistically and 7.7% formally.

3.2 Materials

Data on event cognition were elicited by means of a memory-based triads-matching task. This task, which was the same as the one used in Athanasopoulos and Bylund (2013) and Bylund et al. (2013), was designed in the following way. Thirty-one video clips from the stimulus pool of the research group of Christiane von Stutterheim and associates at Heidelberg University were used in all permissible combinations to create 19 triads.[6] Each triad consisted of a target and two alternates. The target clip was a scene with an intermediate degree of goal orientation. One alternate, the so-called [–endpoint] alternate, was a scene with a low level of goal orientation, that is, an entity

moving along a trajectory without an obvious endpoint in sight (e.g., a person cycling along a road). The other alternate, the so-called [+endpoint] alternate, was a scene with a high level of goal orientation. In these scenes, the moving entity actually reached an endpoint (e.g., a person cycling into a garage). In each triad, manner and direction of motion and number of agents were controlled for. The clips had also been checked for visual similarity to ensure that the [+endpoint] and [–endpoint] alternates were equally visually similar to the target clips (for details, see Athanasopoulos & Bylund, 2013). These measures were taken to minimise the possibility that the participants made their similarity judgements on features other than goal-orientation.

Participants were also given a short linguistic background questionnaire in isiXhosa.

3.3 Procedure

The participants were tested individually in a quiet room at university by a native speaker of isiXhosa. They were informed that they would see video clips arranged in triads on the computer screen, where clip A would appear first, then clip B, and finally clip X (the target). Participants were instructed to indicate whether they thought clip X was more similar to clip A or more similar to clip B. The 19 triads were presented twice in a counterbalanced ABX format, in which half of the time the [–endpoint] alternate appeared first (clip A) and half of the time it appeared second (clip B), and vice versa for the [+endpoint] alternate. The sequence of the clips in each triad was thus as follows: clip A played, followed by clip B, followed by clip X. Participants were instructed to give their responses only after they had watched clip X in its entirety. Clips A, B, and X played immediately after one another, with no pause in between.

4. Results

The participants' categorisation preferences of endpoint-oriented motion were calculated based on the number of

---

[6] We are very grateful to Christiane von Stutterheim and Barbara Schmiedtová for making their clips available to us.

Table 2. *Pearson correlation matrix of independent variables.*

	Age of acquisition of aspect languages	English in secondary school	English in primary school	Proficiency in aspect languages
Use of aspect languages	.08	.14	−.09	.41
Age of acquisition of aspect languages		−.07	−.06	.34
English in secondary school			.28	−.19
English in primary school				.31

times they matched the target clip (X) with the [+endpoint] alternate. This score was subsequently converted into a percentage that indicated each participant's endpoint preference. On average, the participants matched the target clip with the [+endpoint] alternate in 30.9% ($SD = 12.8$) of the cases.

The next step in the analysis consisted of exploring to what extent the participants' linguistic backgrounds could explain the variation attested in their endpoint preferences. To this end, the following independent variables were taken into account:

(i) Age of acquisition of aspect languages. The ages of acquisition onset of each of the aspect languages spoken by the participants were used to create an average age at which they started to acquire an aspect language. This age was 9.0 (SD 2.9) years.[7]

(ii) Amount of use of aspect languages. This variable was created by totalling each participant's amount of use of each of the aspect languages that he or she spoke. On average, the participants used aspect languages 38.7% ($SD = 14.8$) of the time.

(iii) Proficiency with aspect languages. Here, an average was calculated based on the participants' self-reported proficiencies with each aspect language they spoke.

(iV) Degree of exposure to isiXhosa – as opposed to English – in primary school, which was on average 4.0 ($SD = 1.4$) (1 = "Only English"; 5 = "Only isiXhosa", see Section 3.1 above).

(v) Degree of exposure to isiXhosa in secondary school, which was 2.4 ($SD = 1.5$).

As a first step in the analysis, the independent variables were entered into a Pearson correlation matrix (Table 2).

An inspection shows that only weak to moderate correlations were present between the independent variables (Cohen, 1988), giving a first indication that multicollinearity was insignificant in the data ($rs < .80$, see Meyers, Gamst & Guarino, 2006). The independent variables were consequently entered into a multiple regression analysis, in which endpoint preference was the dependent variable. The overall regression was statistically significant, $F(5,44) = 4.203, p = .003, MSE = 17.736, R^2 = .323$. This means that approximately one-third of the variation in the participants' endpoint preference could be explained by the factors taken into account, whereas the remainder is unaccounted for. The residuals of the dependent variable were normally distributed ($W = .984, p = .745$). In Table 3, beta weights, significance values, and collinearity diagnostics are presented. Frequency of use of aspect languages and Use of English in primary school turned out to be the only variables that exerted a significant effect on endpoint categorisation preferences, such that the more the participants used aspect languages and the more they had been exposed to English in primary school, the less prone they were to match the target scene with an endpoint alternate. Mean age of acquisition of aspect languages did not reach statistical significance, but stayed at the trend level. The Variance Inflation Factor (VIF) and Tolerance values confirmed the initial observation that the multicollinearity in the sample was inconsequential (VIFs < 10 and Tolerances > .20, see Menard, 2001).[8]

5. Discussion

A central finding in the present study is that motion event categorisation is influenced by the frequency of use of languages with grammatical aspect. This finding dovetails with previous research on bilingual motion event cognition (Bylund et al., 2013), which found that the frequency of use of English (aspect language) affected

---

[7] An anonymous reviewer suggests that the age by which the participants started acquiring their first aspect language (as opposed to the average age of acquisition of all of their aspect languages) might predict endpoint preferences. However, a Pearson correlation showed no significant relationship between such age and endpoint behaviour, $r = .01, p > .1$.

[8] Note that even though these are common rules of thumb for interpreting VIF-values, some statisticians suggest that separate VIF-values above 2 or a mean VIF above 1 constitute an even stricter criterion for detecting multicollinearity (Field, 2007).

Table 3. *Beta weights, significance values, and multicollinearity diagnostics for multiple regression with endpoint preference as dependent variable.*

Independent variables	Standardised coefficient (β)	t	p	Tolerance	VIF
Use of aspect languages	−.322	−2.278	.028	.770	1.299
Age of acquisition of aspect languages	.233	1.760	.085	.876	1.141
English in secondary school	−.003	−.020	.984	.854	1.172
English in primary school	−.358	−2.662	.011	.851	1.175
Proficiency in aspect languages	.010	.065	.949	.643	1.555

VIF = Variance Inflation Factor

Afrikaans speakers' endpoint categorisation preferences. Moving beyond the realm of motion events, the effects of frequency of language use on bilingual cognition has also been documented in the study of colour categorisation (Athanasopoulos et al., 2011). However, the difference between the current finding and those reported in the literature is the number of languages included in the variable "frequency of use". Whereas previous studies have defined this variable as "frequency of use of language X", the novelty of the current finding is that frequency of use refers to a constellation of multiple languages with a common typological trait (i.e., grammatical aspect). This is, to the best of our knowledge, the first time such an effect has been documented.

This study shows, then, that the more often the participants used languages that grammatically encode the ongoing phase of events, the less endpoint-oriented they were in their event categorisations. Notably, since isiXhosa does not have grammatical aspect, it is the use of languages other than the participants' L1 that drives this tendency. The attested behaviour is thus classified as conceptual transfer, which is the process whereby language-specific conceptual distinctions acquired in one language transfer to another language (Jarvis, 2011). More specifically, the behaviour may be labelled as cognitive restructuring (Jarvis & Pavlenko, 2008; Pavlenko, 1999), in the sense that the use of additional languages with a contrasting grammatical category has restructured the speaker's cognitive behaviour. Following Cognitive Grammar, the specific mechanisms that operate here are assumed to be entrenchment and routinisation. According to Langacker (2008), the recurrent use of a grammatical construction leads to an entrenchment of the conceptual content represented by that construction. As a function of this entrenchment, the conceptual content becomes part of a cognitive routine, which is manifested in specific ways of construing and categorising reality (Jarvis, 2011). Applying this interpretation to the current findings, frequent use of aspect languages can be said to have led to an entrenchment and routinisation of immediate temporal viewing frames that zoom in on the ongoing phase of

events (Bylund & Jarvis, 2011). As a consequence, the participants who used aspect languages frequently were more likely to exhibit categorisation preferences based on a perspective in which the events are construed as ongoing.

Another striking finding of the current study is the effect of medium of instruction in primary school on event categorisation. Daller, Treffers-Daller and Furman (2011) found that the degree to which Turkish–German bilinguals adhered to Turkish monolingual and German monolingual patterns of event construal was related to the extent to which they had attended German and/or Turkish schools. We are, however, not aware of any previous study that has documented – or even taken into account – the effects of medium of instruction on non-verbal cognition. The current finding thus suggests that those participants who had greater exposure to English were less likely to exhibit high endpoint preferences. English being an aspect language, these effects may be interpreted in a similar way as the effects of use of aspect languages discussed above. That is, frequent exposure to a language with grammaticalised immediate temporal viewing frames increases the chances that such frames will become a salient cognitive routine in the categorisation of events. What is noteworthy about this variable, though, is its diachronicity, as it refers to a situation that took place when the participants – who are now adults – were between six and 12 years of age. Given that this is some 15–20 years prior to testing, it is intriguing that language of instruction has exerted such a long-term effect on the participants' cognitive behaviour. One possible explanation for this lasting imprint is that the intense exposure took place during a period in life when there is heightened susceptibility to language exposure, that is, during the sensitive (or critical) period. Even though different interpretations exist, several empirical accounts posit a terminus for the sensitive period at around 12 years of age (e.g., Abrahamsson & Hyltenstam, 2008; Bylund, Abrahamsson & Hyltenstam, 2012; Lenneberg, 1967; Yeni-Komshian, Flege & Liu, 2000). It is thus possible that the effects of a high degree of English exposure in

school between ages six and 12 can be explained against the notion of the sensitive period.

An interpretative frame based on the sensitive period may also explain why another language experience in the participants' past, namely degree of exposure to English in secondary school, did not influence event categorisation. Since in this case the participants were past the alleged sensitive period, they did not have a similar degree of susceptibility to language exposure. Taken together, these findings suggest that language experiences from the past may give rise to differential effects, depending on their timing.[9] It should be noted, however, that our interpretation does not necessarily suggest maturational constraints have a direct effect on cognitive behaviour, as the maturational effects could also be seen as indirect: high exposure to an L2 at an early age entails a higher probability of attaining advanced proficiency with that language, which in turn may influence cognitive restructuring.

The results also showed that average age of acquisition of aspect languages did not quite significantly influence event categorisation preferences. A possible reason for this is the context in which learning took place. Research has found that learning context is a decisive predictor for whether age of acquisition effects can be expected or not. Lenneberg's (1967, p. 176) formulation of the Critical Period Hypothesis specifies that the predictions concern "automatic acquisition" through "mere exposure", and subsequent research has shown that age effects on ultimate attainment are rarely found in foreign language contexts (Muñoz, 2006). The distinction between formal and naturalistic learning contexts is not clear-cut in the current study; for example, even though most of the participants started learning English formally at school, they were gradually immersed in this language as it took on the role of medium of instruction. This blurs the distinction between the two types of contexts and consequently makes it difficult to provide a clear-cut measure of when the naturalistic learning started, and to what extent it may be labelled "naturalistic". These circumstances might thus have reduced the salience of age effects. It is therefore difficult to posit that the current findings confirm or disconfirm previous findings on the role of age of acquisition in cognitive restructuring.

No effects of language proficiency were detected in the current study. Previous relativistic research on this matter has yielded mixed findings (see Athanasopoulos & Kasai, 2008; Bylund et al., 2013; Cook et al., 2006). These mixed findings may be attributed to the specific methods used by different studies to measure language proficiency:

whereas some rely on formal tests (Athanasopoulos & Kasai, 2008), others use self-reports (such as the current study). High correlations between self-reported proficiency and proficiency assessed via formal tests have been documented (Marian, Blumenfeld & Kaushanskaya, 2007), but self-reports have an inherent component of subjectivity that may render them unreliable. In addition to the caveats with self-reports, there is also a possibility that global proficiency index (which was reported by the current participants) is too coarse-grained to capture the effects that the relevant linguistic properties (i.e., grammatical aspect) may exert on event categorisation. Evidence in support of this suggestion is provided by a study on the effects of language proficiency on linguistic construal of event endpoints: Bylund (2009) and Bylund and Jarvis (2011) found that whereas general grammatical skills did not influence endpoint encoding, proficiency with aspectual distinctions did. Unfortunately, the current study did not have the resources to design such tests for all the different aspect languages involved. We therefore leave this question open for future inquiry.

6. Conclusion

This study set out to investigate motion event cognition in L1 speakers of isiXhosa living in multilingual South Africa. The study specifically examined the influence of linguistic background on motion categorisation patterns. The analyses revealed that frequency of use of aspect languages and degree of English exposure in formal instruction early in life exert a significant influence on the participants' non-verbal behaviour with endpoints. Future studies on motion event cognition in isiXhosa will benefit from also including verbal behaviour measures to establish the extent to which isiXhosa speakers encode endpoints when describing motion. Ideally, future studies would also investigate monolingual isiXhosa speakers from monolingual settings to establish a baseline against which multilingual behaviour may be compared. It is questionable, however, whether there are any such speakers who match the current participants in terms of age and educational background. An additional task for future inquiry consists of extending the triads-matching paradigm used in the current study. Even though the use of real life scenes affords some ecological validity to the current design, it is difficult to ascertain whether the steps taken (e.g., visual similarity norming) efficiently eliminate EVERY possibility that the participants could have made similarity judgements based on features other than endpoint orientation. One way to avoid this problem would be to use artificial stimuli (e.g., computer animations) where the features of each scene are controlled to the last detail.

Apart from addressing language-specific cognitive patterns in multilinguals, the current study has also

---

[9] It should be kept in mind that this interpretation concerns language experiences from the past, and the extent to which these are relevant to cognition at the time of testing. The interpretation does not make any predictions regarding the effects of current language use on cognition.

raised issues relating to research practice in multilingual contexts. As we have sought to show, the linguistic circumstances and characteristics of the participants in the study are fairly different from what has been reported in previous studies on bilingualism and thought. The fact that the participants spoke at least one additional European language (i.e., English, and sometimes Afrikaans), which was used in their education, as well as other local languages (Bantu thereof), which they mostly had learnt naturalistically, is a situation that is typical to former colonies in Africa (Oostendorp, 2012). An important difference is also that whereas participants in previous studies on bilingualism and thought often have come into contact with their additional language through migration (Athanasopoulos et al., 2010) or foreign language learning (Kurinski & Sera, 2011), the current participants were brought up in an environment where their additional languages were spoken. This, for example, means that variables such as length of residence in the L2 setting are difficult to translate to a multilingual context of this kind.

The fact that the current participants knew multiple languages also renders the notion of language use problematic: whereas this variable has been operationalised in a typically dichotomous way (i.e., EITHER L1 OR L2), the current study has used a procedure that groups languages together on the basis of their typological features. Even though this procedure is still dichotomous (i.e., aspect vs. non-aspect languages), it differs in that it allows for inclusion of more than two languages. There are, however, other methodological aspects in the current study that are less easily addressed. Consider, for instance, the possibility that L1 acquisition may have taken place in a setting where several languages were being spoken and the L1 provided in the input was already influenced by other languages. In such a situation, one needs to ask to what extent an adult speaker's event cognition is influenced by, on the one hand, the additional languages he/she learnt along the way, and on the other, the specific characteristics of the L2-influenced L1 variety that he/she acquired. This question is open to future inquiry.

In our view, several possibilities are afforded by continued explorations into language and thought in multilingual settings similar to the one investigated in this study. At a more specific level, researchers interested in the effects of grammatical aspect on cognition will find it fruitful to delve further into the aspectual systems of Bantu languages, which often exhibit morphological aspectual categories (e.g., performative, subsecutive, narrative) that are under-researched and largely differ from those found in the more commonly researched European languages. At a more general level, pursuing research on multilingual contexts will bring attention to a question that, so far, has clearly remained in the background of language and thought research: How does growing up in a society where numerous, typologically distinct languages are regularly used for communicative purposes influence the separation or integration of language-specific cognitive behaviour?

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30,* 481–509.

Aronin, L., & Singleton, D. (2012). *Multilingualism*. Amsterdam: John Benjamins.

Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes, 22,* 689–699.

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

Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science, 37,* 286–309.

Athanasopoulos, P., Damjanovic, L., Krajciova, A., & Sasaki, M. (2011). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition, 14,* 9–17.

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J.-R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition, 116,* 437–443.

Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied Psycholinguistics, 29,* 105–123.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43,* 1–22.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition, 118,* 123–129.

Brown, R., & Lenneberg, E. H. (1954). A study of language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Buell, L. C. (2005). *Issues in Zulu verbal morphosyntax*. Los Angeles, CA: ProQuest.

Bylund, E. (2009). Effects of age of L2 acquisition on L1 event conceptualization patterns. *Bilingualism: Language and Cognition, 12,* 305–322.

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2012). Does first language maintenance hamper nativelikeness in a second language? *Studies in Second Language Acquisition, 34,* 215–241.

Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion event cognition and grammatical aspect: Evidence from Afrikaans. *Linguistics, 51,* 929–955.

Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition, 14,* 47–59.

Census (2012). *Census in brief*. Pretoria: Statistics South Africa.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Cook, V., & Bassetti, B. (eds.) (2011). *Language and bilingual cognition*. New York: Psychology Press.

Cook, V., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism, 10,* 137–152.

Daller, M. H., Treffers-Daller, J., & Furman, R. (2011). Transfer of conceptualization patterns in bilinguals: The construal of motion events in Turkish and German. *Bilingualism: Language and Cognition, 14,* 95–119.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature, 398,* 203–204.

Doke, C. M., & Mofokeng, S. M. (1985). *Textbook of southern Sotho grammar*. Cape Town: Maskew Miller Longman.

Du Plessis, J. A. (1978). *IsiXhosa*. Goodwood: Oudiovista.

Field, A. (2007). *Discovering statistics using SPSS*. London: Sage.

Flecken, M. (2011). Event conceptualization by early Dutch–German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition, 14,* 61–77.

Forbes, J. N., Poulin-Dubois, D., Rivero, M., & Sera, M. (2008). Grammatical gender affects bilinguals' conceptual gender: Implications for linguistic relativity and decision making. *The Open Applied Linguistics Journal, 1,* 68–76.

Guthrie, M. (1948). *The classification of the Bantu languages*. Oxford: Oxford University Press.

Han, Z., & Cadierno, T. (eds.). (2010). *Linguistic relativity in SLA: Thinking for speaking*. Clevedon: Multilingual Matters.

Haun, D., Rapold, C., Janzen, G., & Levinson, S. (2011). Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. *Cognition, 119,* 70–80.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences, 33,* 61–83

Hobson, C. (1999). Morphological development in the interlanguage of English learners of Xhosa. Ph.D. dissertation, Rhodes University, Grahamstown.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62,* 169–200.

January, D., & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). *Cognition, 104,* 417–426.

Jarvis, S. (2011). Crosslinguistic influence in bilinguals' concepts and conceptualizations. *Bilingualism: Language and Cognition, 14,* 1–8.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Kurinski, E., & Sera, M. D. (2011). Does learning Spanish grammatical gender change English-speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition, 14,* 203–220.

Langacker, R. W. (2008). *Cognitive Grammar: An introduction*. Oxford: Oxford University Press.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lucy, J. A. (1992). *Grammatical categories and cognition*. Cambridge: Cambridge University Press.

Lucy, J. A. (1997). Linguistic relativity. *Annual Review of Anthropology, 26,* 291–312.

MacLaren, J. (1936). *A Xhosa grammar*. Cape Town: Longmans.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50,* 940–967.

Menard, S. (2001). *Applied logistic regression analysis*. London: Sage.

Mesthrie, R. (2002). *Language in South Africa*. Cambridge: Cambridge University Press.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation*. London: Sage.

Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (2011). Can a mind have two time lines? Exploring space–time mapping in Mandarin and English speakers. *Psychonomic Bulletin & Review, 18,* 598–604.

Motsei, A. (2010). The expression of aspect in Sesotho. Ph.D. dissertation, Stellenbosch University.

Muñoz, C. (2006). *Age and the rate of foreign language learning*. Clevedon: Multilingual Matters.

Ncoko, S. O. S., Osman, R., & Cockcroft, K. (2000). Codeswitching among multilingual learners in primary schools in South Africa: An exploratory study. *International Journal of Bilingual Education and Bilingualism, 3,* 225–241.

Nichols, P. (2012). *A morpho-semantic analysis of aspectuality in siSwati: A cognitive approach to the analysis of the alterative, persistive and inceptive aspects in siSwati*. London: Lambert Academic Publishing.

Nurse, D. (2008). *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Oostendorp, M. (2012). New perspectives on cross-linguistic influence: Language and cognition. *Language Teaching, 45,* 389–398. [A comparative book review.]

Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes, 25,* 224–260.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition, 2,* 209–230.

Pavlenko, A. (2005). Bilingualism and thought. In J. F. Kroll & A. M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches,* pp. 433–453. Oxford: Oxford University Press.

Pavlenko, A. (2011a). Introduction. In Pavlenko (ed.), pp. 1–28.

Pavlenko, A. (ed.) (2011b). *Thinking and speaking in two languages*. Clevedon: Multilingual Matters.

Setati, M. (2005). Teaching mathematics in a primary multilingual classroom. *Journal for Research in Mathematics Education, 36,* 447–466.

Taljaard, P. C., Khumalo, J. N., & Bosch, S. E. (1991). *Handbook of Siswati*. Cape Town: J. L. van Schaik.

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.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings (edited by J. B. Carroll)*. Cambridge, MA: MIT Press.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition, 3,* 131–149.

Ziervogel, D., & Mabuza, E. (1976). *A grammar of the Swati language*. Cape Town: Van Schaik.