Original Article

Dominant-while-speaking: How bilingual Norwegian–English children conceptualize goal-oriented motion events

Hanna Andresen [iD]
University of Oslo, Norway

Renate Delucchi Danhier
Barbara Mertins
TU Dortmund University, Germany

Abstract

Aim: This study investigates how simultaneous bilingual Norwegian–English children conceptualize goal-oriented motion events in their two languages, which have different and partly conflicting language-specific conceptualization patterns (the so-called holistic, endpoint-oriented perspective vs. a phasal perspective with focus on ongoingness).

Design: The experiment combined three different methodologies to measure attention to endpoints: elicitation of spoken data, eye-tracking, and a subsequent memory test. A total of 23 bilingual Norwegian–English children participated on a separate day for each language. The comparison groups comprised monolingual Norwegian children ($n = 21$), first language (L1) Norwegian adults ($n = 30$), and L1 English adults ($n = 20$).

Data and analysis: The statistical analysis included calculations of endpoints mentioned and endpoints remembered in each language and for both sessions, and an event-related analysis was conducted to establish the total length of all fixations on the areas of interest (AoIs [the endpoints]), in each language and for both sessions.

Findings/conclusions: The results showed an effect of the language of operation on conceptualization, independent of language dominance, but depending on the situation. In their first session, the bilingual children showed a strong awareness of the prototypical conceptualization pattern in the language they were speaking. In their second session, the children's conceptualization was influenced by their first. Hence, we can talk about a flexible conceptual dominance linked to language, the situation, and to previous experience.

Originality: This study for the first time explores simultaneous bilingual children's conceptualization of goal-oriented motion events in their two languages and with different methodologies.

Corresponding author:
Hanna Andresen, MultiLing, University of Oslo, P. B. 1102, Blindern, Oslo 0316, Norway.
Email: h.s.andresen@iln.uio.no

Significance/implications: This study brings new insights into our understanding of early bilingualism at a conceptual level, with implications for the field, but also for parents, caregivers, and stakeholders, who at times need to be reminded about children's unique capacity for language learning.

Introduction

Psycholinguists have long been interested in whether the conceptualization of bilingual speakers differs from that of monolingual speakers of the languages in question: While there is a consensus that the bilingual mind is not simply the addition of two monolingual systems, there is still no consensus as to how the conceptual system(s) of bilinguals interact. While the lion's share of previous research on conceptualization in bilingual speakers has focused on second language (L2) adults (for overviews, see Pavlenko, 2014; Wang & Wei, 2022), the aim of the current study is to investigate the conceptualization of goal-oriented motion events in simultaneous bilingual Norwegian–English children, who are exposed to different and partly conflicting language-specific conceptualization patterns in their two languages. Adhering to the *thinking-for-speaking* framework (Slobin, 1996), we define *conceptualization* as the cognitive planning processes activated immediately before the actual language production: the segmentation, information selection, and structuring of conceptual content that takes place before the preverbal message is carried over to the formulator (Levelt, 1989; Mertins, 2018; Schmiedtová, 2011; Slobin, 1991, 1996; von Stutterheim & Nüse, 2003).

Previous research has shown that there is a connection between grammatical aspect and event conceptualization, and that speakers of aspect and non-aspect languages prefer different conceptualizations of, for example, event construal (von Stutterheim & Nüse, 2003). More specifically, the obligatory marking of progressive or imperfective aspect has been argued to direct the speaker's attention to ongoingness, resulting in sentences with phasal event decomposition (Carroll et al., 2004; von Stutterheim & Nüse, 2003). Speakers of non-aspect languages, in contrast—languages where the category of aspect is not grammaticalized—more often tend to look for closure or a boundary of the event, hence taking a holistic perspective. For example, when describing a video clip of a motion event where someone is walking in the direction of a car but without arriving there within the clip, speakers of non-aspect languages tended to include the inferable endpoint anyway, assuming that the possible endpoint was the goal of the motion, as in *a man is walking to his car*. Speakers of aspect languages, in contrast, more often defocused the endpoint, as in *a man is walking* (e.g., Schmiedtová et al., 2011; von Stutterheim & Nüse, 2003). Furthermore, studies with eye-tracking confirmed that speakers of non-aspect languages also tended to look more at the endpoints than did speakers of aspect languages. Memory was also affected: Speakers of non-aspect-languages tended to remember the endpoints to a larger extent (Schmiedtová et al., 2011).

While Slobin (1996) originally examined first language (L1) speakers of different languages, several studies have explored bilingual *thinking-for-speaking* (for overviews see Pavlenko, 2014; Wang & Wei, 2022). To acquire a new conceptualization pattern as an adult is generally considered difficult (Schmiedtová, 2011). However, different reconceptualization patterns, such as conceptual transfer from one language to the other, bidirectional transfer, and convergence are all possible

outcomes in adult L2 acquisition (for an overview see Wang & Wei, 2022). Research on bilingual verbalization and conceptualization exists on a great variety of bilingual profiles, varying from advanced adult L2 speakers and adult bilinguals who acquired their L2 in late or early childhood, to children acquiring an L2 in early school years. The question of language-specific conceptualization in simultaneous and/or early bilingual or multilingual children, however, remains underexplored. The aim of the current study is thus to investigate the conceptualization preferences of simultaneous bilingual Norwegian–English children, who have different aspectual systems in their two languages.

Language-specific conceptualization in Norwegian and English

In English, marking the verb for the progressive aspect is obligatory (Comrie, 1976). In Norwegian, the temporal concept of ongoingness is not grammaticized, and the temporal sequence is instead based on bounded events, which in turn leads to a more holistic perspective (Andresen, 2023; Behrens & Ramm, 2003; Carroll et al., 2004; von Stutterheim & Carroll, 2006). While it is possible to express ongoingness in Norwegian by using lexical constructions, these are not compulsory (Kinn, 2018; Kinn et al., 2018; Tonne, 1999, 2001), and hardly used in motion events at all (Behrens et al., 2013). In an experiment on goal-oriented motion events, adult native speakers of Norwegian mentioned significantly more endpoints than adult native speakers of English when describing the same video scenes (Andresen, 2023). Hence, while native speakers of English preferred the phasal perspective, native speakers of Norwegian tended to choose a holistic viewpoint (Andresen, 2023; Behrens & Ramm, 2003). In the present study, we therefore ask how goal-oriented motion events are conceptualized by bilingual children who have acquired both languages since birth: Will the bilingual participants mention more endpoints when describing goal-oriented motion events in Norwegian (non-aspect-language) than in English (aspect-language)? Will the bilingual speakers look at the endpoints (area of interest [AoI]) more/longer (operationalized as dwell time and average fixation duration) when speaking Norwegian than when speaking English? Will the participants remember more endpoints when solving the task in Norwegian than when solving it in English? Will the results from the three methodologies all point in the same direction, or will they be conflicting, or so-called bilingual-specific patterns (cf., Flecken, 2011)? Finally, how will the results look as compared with the comparison groups?

Methodology

Participants

The bilingual children ($n = 23$) were aged 7–8 years (M_{age} 8) and had been spoken to in Norwegian and English from birth (20 of 23) or early infancy, that is, 12–24 months (3 of 23). They all had one parent with L1 Norwegian and one with L1 English, except two participants, who had two English-speaking parents (but with an active Norwegian-integrated life). The participating families were comparable in terms of socioeconomic and educational background. The children were not tested for language proficiency in their two languages before the experiment, but the parents were interviewed through a detailed questionnaire to map the children's language proficiency and daily usage in the two languages. Criteria for participation were that the children were active bilinguals, that both languages were used in the home on a daily or weekly basis, and that the child could hold a conversation in both languages (controlled for in warm-up conversation). The children lived in Oslo, Norway and attended Norwegian schools, but some of them were born in and/or had also resided in English-speaking countries for shorter or longer periods. In all, 14 children were reported

to be dominant in Norwegian, whereas 6 were reported to speak and understand both languages equally well; only 3 children were reported being dominant in English. However, 17 of the families (73.9%) said that their children understood the two languages equally well. Since English is a high-status language in Norway, the children had many opportunities to both speak and hear English outside the family (e.g., neighbors, social media, music, gaming, etc.).

Comparison groups: One should be cautious when comparing bilinguals and monolinguals as to what is being compared and why (e.g., De Houwer, 2022; Ortega, 2014). However, without comparing the bilinguals to some other group, it would be difficult to discuss whether they have adopted either the preferences of one of the languages, one converged system, or whether we see traces of conceptual transfer. We therefore included data from Norwegian L1 and English L1 speakers, respectively, using the L1 speakers as a baseline. The monolingual Norwegian children's group (N21) is the most comparable to the bilingual children's group. The participants were aged 7–8 years (M_{age} 8) and were comparable in terms of socioeconomic and social background. The material, procedure, and location (the lab) were identical (except that the monolingual children were only tested once). We also have data from a Norwegian L1 adult group ($N=30$) and an English L1 adult group ($N=20$).[1] The Norwegian adults were university students aged 18–29 years (17 f, 13 m) and were tested at the University of Oslo in connection with another study (Andresen, 2023; Behrens et al., 2013). Unfortunately, we did not have an English monolingual children's group for comparison, which is why we included the adults. Twenty English participants aged 20–26 years (14 f, 6 m) were tested in England at English universities (also in connection with another study, see Schmiedtová et al., 2011). The test material for the adults was the same as for the children: The scenes for comparison (critical and control scenes) were identical, though there was some variation when it came to the filler items. The procedure was also the same, except that there was no eye-tracking.

Stimuli

The stimulus set consisted of 60 video recordings, depicting different types of daily life situations.[2] The properties were systematically manipulated to investigate the extent to which they may promote or constrain the use of aspectual constructions and attention to possible endpoints in each language. Twelve of the recordings were *critical scenes*: goal-oriented motion events where the potential endpoint was not reached. For example, a woman walking toward a car on a parking lot but not reaching the car before the video clip ended (see Appendix 1, Picture 1). A typical response to this film clip could be either "I see a lady walking" (endpoint not mentioned), or "a woman is walking to her car" (endpoint mentioned). Twelve video clips were *control scenes*: goal-oriented motion events where the endpoint was reached. For example, a cat walking from one room into another room (see Appendix 1, Picture 2). A typical response to this film clip could be either "a cat walking into a room" (endpoint mentioned), or "a cat walking around in the house" (endpoint not mentioned). The control items were used to elicit the verbalization of completed actions, thereby attracting more attention to the endpoints than the critical items, which were more likely to be interpreted as ongoing (imperfective or progressive actions). The rest of the videos (36) were distractor scenes of different activities or situations (e.g., a woman decorating a cake, a boat floating in the water). Each clip lasted 6 seconds, followed by a blank screen for another 6 seconds to give the participants time to finish their verbalization. The clips were controlled for the type of protagonist (person, animal, object, vehicle), and the direction from which the protagonist appears (left or right). All the clips were shown in a pseudo-randomized order, to avoid several critical scenes appearing directly one after the other. While watching the video clips, the participants' eye movements were tracked by an SMI Red 250 mobile at 250 Hz. AoIs were defined covering the intended

endpoints of the critical and control stimuli and an event-related analysis was conducted to establish the total length of all fixations on the AoIs (fixation duration and dwell time on the AoIs). The memory task was administered on paper and comprised 15 color screen shots of the videos. There were nine critical items in which the possible endpoint was removed and six control items where a random object was missing (see Appendix 2, Picture 3). The task was to name as fast as possible, with one or a few words, what had been removed from the image, the hypothesis being that the participants following an endpoint-oriented perspective would remember the endpoints better than the speakers following a phasal perspective focusing on ongoingness (cf., Mertins, 2016; Schmiedtová et al., 2011). The control items were used for general memory performance. The memory test was performed in both sessions, although we expected a memory effect in the second session, since the participants had already completed the test in the previous session and had seen the video clips twice.

Procedure

The bilingual children were tested on a different day for each language and with approximately 2 weeks in between the two sessions (to minimize learning effects, but while the participants would still be at the same linguistic and psychological level). The participants in the E1-group ($n = 9$) were tested the first time in English and the second time in Norwegian, and the participants in the Norwegian first (N1)-group ($n = 14$), the other way around.[3] The experiment was performed by an L1 or L1-like speaker, and on the day of testing all communication took place in the language of testing to try to create a monolingual environment (cf., Grosjean, 1998, 2013). The participants were told that they would see a set of video clips showing different everyday situations that were not connected in any way, and that their task was to watch each clip and answer the question "what is happening?" *(Og du skal bare fortelle hva som skjer/And then you will just tell me what's happening)*. The video experiment lasted 15 minutes. Afterwards, the participants were asked questions concerning their linguistic background and language situation at home with the help of a child-friendly questionnaire, then they did some jumping exercises to loosen up, and finally the memory test as described above. The procedure was the same on both occasions, only differing in the language spoken. All participants received a participation diploma and a small prize afterwards (an eraser, a sticker, etc.).

Coding and data quality

The audio recordings were transcribed, and the verbalizations were coded for explicit mentions of endpoints by three independent coders. An intercoder reliability measure showed a high degree of agreement (Fleiss' $\kappa \approx .937$). Examples of endpoint encoding are phrases where the inferable endpoint (critical scenes) or reached endpoint (control scenes) is mentioned, mostly with prepositional phrases: Someone or something is moving (walking/running/hurrying/driving/on the way, etc.) *to, toward, into, in the direction of* something or somewhere.[4] For example, "a boy is walking *to the playground"* is coded as endpoint mentioned. The memory tests were also annotated and paired to the verbalization results: endpoint mentioned (EPM) versus endpoint remembered (EPR).[5] Regarding the memory results, one of the monolingual and two of the bilingual children skipped some pages when solving the memory tasks. The skipped items were excluded when calculating the percentages of remembered endpoint.

For the eye-gaze analysis, SMI's BeGaze software was used. Because of the relatively small number of participants, the analysis of eye-tracking data was run using average measures across participants, as well as averages over item categories, even though this means many items must be tracked

back to the same participants. This is a common practice in psycholinguistics (Mertins, 2016). Before each session of the experiment, calibration was carried out for each participant. Only the right eye of the participants was tracked (tracking ratio 250 Hz). Since the participants were young children, the maximal allowed deviation was defined at 1.3° in both the X and Y dimensions.[6] The mean deviation for the right eye in the X dimension was 0.26° (*SD* 0.18) and 0.31° (*SD* 0.20) for the Y dimension. The SMI Experiment Center recorded eye movement time-locked. The mean tracking ratio of all recordings was 70.0% (*SD* 14.32), which is fairly good considering the young age of the participants. The tracking ratio would have been higher if the quality measures had reported only the critical scenes, and not for all stimuli including fillers, which were perceived as very boring by the children (standard fillers for this research line showing mostly static scenes where not much happens).

Results

Linguistic results

When analyzed by language alone and independent of the order of testing, the bilingual participants mentioned the endpoints to the same degree in the two languages. A paired Wilcoxon test did not find any significant difference between the two language conditions ($V=89$, $N=23$, $p \approx .14$) in the critical scenes. Since this is a paired comparison, the normality of the difference between the two groups has been calculated, and it is just barely under the alpha level ($W \approx 0.910$, $p \approx .041$). In the control scenes, the children as expected mentioned the endpoints to a higher degree both in Norwegian and English, but with no significant difference between the languages (paired *t*-test: $t \approx 1.60$, $df=22$, $N=23$, $p \approx .12$).

However, when considering the order of testing, differences between the languages appear: In their first session, the participants in the N1-group (speaking Norwegian) on average mention the endpoints in 55.36% ($M=6.43$, $SD=2.74$) of the critical scenes, while the participants in the E1-group (speaking English) on average mention the endpoints in only 27.86% ($M=4.11$, $SD=2.93$) of the critical items. This difference is highly significant, *t*-test: $t=-2.7978$, $p=.01365$. Surprisingly, there are also differences in the control scenes: While the participants in the N1-group mention a high number (80.36%) of endpoints ($M=8.93$, $SD=2.34$), the participants in the E1-group mention significantly fewer (58.16%) ($M=6.67$, $SD=2.69$), *t*-test: $t=-2.5124$, $p=.02502$.

In their second session, the participants in the N1—now called E2-group because they spoke English secondarily—still mentioned the endpoints quite often in the critical scenes (50.60%) ($M=5.57$, $SD=2.93$). For this group of participants, there was no significant difference between their two sessions (i.e., N1 vs. E2), neither in the critical nor the control scenes (see Table 1), indicating that they had a similar conceptualization across sessions. The participants in the E1—now N2-group because they spoke Norwegian secondarily—mentioned slightly more endpoints in their second session, both in the critical scenes (35.19%) ($M=4.44$, $SD=2.35$), and in the control scenes (69.38%) ($M=9.22$, $SD=2.17$) (see Table 1). However, while the increase in endpoint mentioning in the critical scenes from their first to their second session (E1 vs. N2) was not significant, the mentioning of endpoints in the control scenes was significantly higher in their second session as compared with their first (see Table 1), indicating that the participants in the N2-group gave more focus to the endpoints in their second session.

When comparing the results for the control scenes in the second session (E2 vs. N2), the difference between the two groups (which was significant in the first session) is no longer significant (*t*-test, $t=.92223$, $p=.3669$).

Comparison groups: When including the linguistic results from the monolingual Norwegian children and the English and Norwegian adults' groups, the participants can be ordered on a continuum

Table 1. The average proportion of endpoint mentions in the two bilingual groups in their two sessions (critical and control scenes).
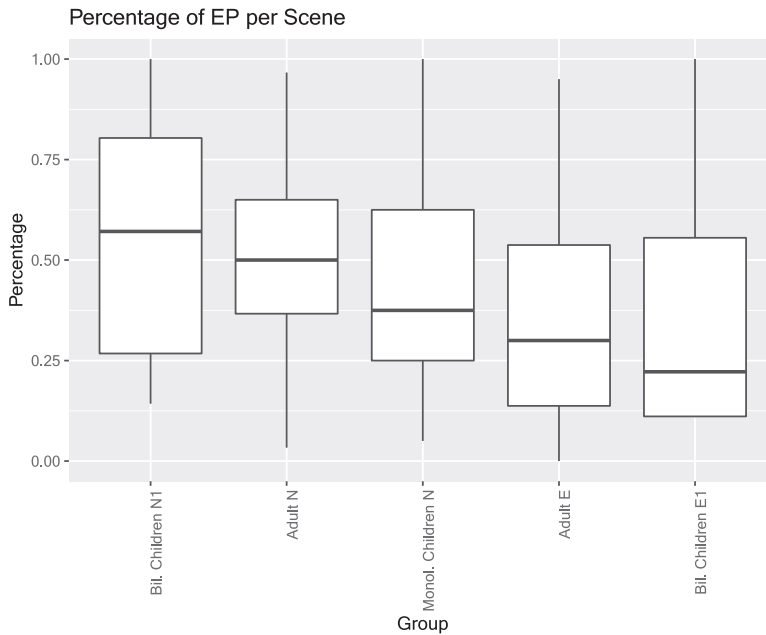
Session	The N1/E2-group		The E1/N2-group	
	1st Norwegian	2nd English	1st English	2nd Norwegian
Critical scenes	55.36%	50.60%	27.86%	35.19%
	t-test, *t*=−.61302, *p*=.5452		*t*-test, *t*=−.66426, *p*=.5161	
Control scenes	80.36 %	77.92%	58.16%	69.38%
	t-test (two-tailed), *t*=−.41711, *p*=.680025		*t*-test (two-tailed), *t*=−2.21832, *p*=.041349	

depending on how probable it is that they will explicitly mention an endpoint while verbalizing the critical scenes, from most likely to mention an endpoint to least likely. In this comparison, the exact same critical scenes are used for all the groups. From the highest to the lowest proportion of endpoints mentioned, the order is as follows: Bilingual children speaking Norwegian, first session (N1): 55.36% > L1 Norwegian adults: 50.3% > Norwegian monolingual children: 42.5% > L1 English adults, 38%: > bilingual children speaking English, first session (E1): 27.86% (see Figure 1).

Interestingly, the two bilingual groups occupy the extremes of the continuum, while the monolingual Norwegian children performed somewhere in the middle, between the two bilingual groups, and between the two groups of adults. When comparing one group to the next on the continuum, there are no significant differences between any adjacent groups. Also, no significant differences were found between the monolingual Norwegian children's group and the two groups of bilingual children, N1-group vs. monolingual (*t*-test [two-tailed], *t*=1.876; *p*=.07166), monolingual group vs. E1 (n. s., *t*-test [two-tailed], *t*=−1.5913, *p*=.1368). However, when comparing the extremes, the N1-group to the E1-group (as shown above), and the adult Norwegian to the adult English, the differences are significant.[7]

Eye-tracking results

The eye-tracking results follow the same pattern as the linguistic results: When comparing all the Norwegian data to all the English data in the critical scenes (independent of the order of testing), there are no significant differences in visual attention linked to language when it comes to fixation duration (*t*-test, *t*=.23689, *p*=.812828) or dwell time (*t*-test, *t*=−.5704, *p*=.56864). However, when separated by order of testing, again, differences between the two groups appear. In their first session, the bilingual participants in the N1-group (speaking Norwegian) had a longer fixation duration in the AoI (*M*=316.50 ms, *SD*=237.83) than the bilingual participants in the E1-group, who were speaking English (*M*=228.9 ms, *SD*=206.6). This difference is significant (*t*-test, *t*=3.0714, *p*=.002381), indicating that the bilingual participants in the N1-group, while speaking Norwegian, looked longer to the endpoints than the participants in the E1-group, who were performing the task in English. Again, the eye-tracking results align with the linguistic results, showing that the participants in the N1-group, while speaking Norwegian, both mentioned more and looked more to the endpoints than the participants in the E1-group, who gave less attention to the endpoints, both verbally and visually. There was no significant difference in dwell time (*t*-test, *t*=−1.3676, *p*=.1728). In the control items, no significant difference was found between the two groups, either in fixation duration (*t*-test, *t*=−1.1406, *p*=.2556) or in dwell time (*t*-test, *t*=−.97468, *p*=.3306) (for all values, see Table 3 in Appendix 3).

Percentage of EP per Scene

Figure 1. Boxplot of proportion of endpoints mentioned in the critical scenes, first session, all comparison groups.

In Session 2, the participants in the N1/E2-group, who were now speaking English, still had a significantly longer fixation duration in the critical scenes than the participants in the E1/N2-group, who were now speaking Norwegian (*t*-test, *t*=2.1711, *p*=.03103) (see Table 4 in Appendix 4). Also, the dwell time was significantly higher for the E2-group versus the N2-group (*t*-test, *t*=−2.6914, *p*=.007574). In the control scenes, there were no significant differences in dwell time (*t*-test, *t*=−1.6451, *p*=.1014) or fixation duration (*t*-test, *t*=−.082042, *p*=.9347) between the groups. When comparing the participants' second session to their first, both groups had a significantly higher value for fixation duration in the critical scenes, indicating that—independent of language—the participants in both groups looked longer to the endpoints in the critical scenes when performing the task for the second time.

Comparison group: Interestingly, and in line with the linguistic results, the monolingual Norwegian children's eye-tracking values lie between the N1- and the E1-group's values from the first session (cf. Table 2).[8] While there was no significant difference between the N1-group and the monolingual children in this regard, the difference in fixation duration between the monolingual children and the E1-group is significant (*t*=−2.7795, *p* = .005788).

In conclusion, when comparing the bilingual children to the comparison groups, we find the N1-group to the left in the endpoint-continuum (cf. Figure 1), displaying a prototypical Norwegian-like conceptualization with many endpoints mentioned and more visual attention to the endpoints. To the right, the English first (E1)-group shows the opposite, with few endpoint verbalizations and less visual attention to the endpoints, the two bilingual groups hence stretching the continuum to the extremes.

Memory results

While there are significant differences between the N1- and E1-groups in the first session regarding attention to the endpoints both verbally and visually, there are surprisingly no differences found

Table 2. Mean proportion of endpoints mentioned and eye-tracking values for the N1-group, the monolingual Norwegian group, and the E1-group.

	N1	Monolingual Norw.	E1
EP mentioned critical	55%	42.5%	27.8%
EP mentioned control	80%	72%	58%
Fixation duration, critical	316.50 ms	285.3 ms	228.90 ms

in the memory data from the first session, either in the critical scenes ($t=.60248$, $p=.5554$), or in the control scenes ($t=-.024741$, $p=.9806$) (see Table 3, Appendix 3). Similarly, there are no differences between the two groups in the second session, either in the critical scenes ($t=.18541$, $p=.8549$), or in the control scenes ($t=.68534$, $p=.502$). In other words, the participants remembered the endpoints equally well, independent of language in both sessions. However, the participants remembered significantly more endpoints in their second session than in their first, in both languages and both in the critical and control scenes, hence showing a clear memory effect between the two sessions (for all values, see Table 4, Appendix 4).

Discussion/conclusion

Dominant-while-speaking and flexible conceptual dominance

While previous studies have shown evidence of different lexicalization patterns in bilinguals (e.g., Aktan-Erciyes et al., 2020; Engemann, 2022), the novelty of the current study is the evidence from the first session of both different verbalization and different visual allocation linked to the bilingual children's language of operation. When *thinking-for-speaking* in Norwegian, the bilingual participants mentioned more and looked more to the endpoints (holistic perspective), than the bilingual participants speaking English (phasal perspective). Furthermore, and quite surprisingly, the bilingual E1-group (speaking English) was placed at one extreme of the endpoint-continuum (cf., Figure 1 above), giving very little attention to the endpoints. While it is less surprising that the participants in the N1-group followed the conceptualization pattern of the societal language (cf., Daller et al., 2011), their placement at the opposite extreme of the endpoint-continuum is striking. If a higher number of endpoints mentioned can be said to be prototypical of Norwegian, the participants in the N1-group seem to "outperform" both the Norwegian-speaking monolingual children and the L1 Norwegian-speaking adults. In addition, the eye-tracking results indicate that the participants in the N1-group look longer to the AoIs than both the monolingual children and the E1-group when placed on the continuum. In other words, in their first session, the bilingual children in both groups show a strong awareness of the language-specific conceptualization patterns of the language of operation: When preparing to speak and/or while speaking, at a conceptual level the bilingual children become *dominant-while-speaking.* Moreover, they stretch the continuum to the extremes, (over-)performing according to the prototypical, language-specific conceptualization. This might explain why the E1-group mentioned so few endpoints not only in the critical, but also in the control scenes in the first session and recalls Flecken's (2011) finding of an increased use of a conceptualization pattern among her bilingual participants. Such (over-)performance indicates something other than pure inhibition of the other language (cf., Bialystok, 2001). It shows that the bilingual speakers fully *lean into* the language they are speaking and should be interpreted as something more than just linguistic (and conceptual) proficiency or awareness. Rather, we interpret it as a sign of both linguistic and emotional or affective confidence, touching upon the positive

variant of what Pavlenko (2014) calls affective processing, and possibly the result of what De Houwer (2009) calls harmonious bilingualism.

The importance of the language of operation is partly in line with the findings from previous studies (cf., Athanasopoulos et al., 2015). However, it does not account for the fact that the children's conceptualization did not change much 2 weeks later, when the children's language of operation was switched. While we would have thought that 2 weeks between the sessions would be enough to "forget" the videos and their verbalizations, the memory tests clearly showed a strong memory effect between the sessions. Hence, we interpret the linguistic and eye-tracking results from the second session to reflect that the children still remembered the situation, the tasks, and partly the solutions for them. This explains why the children in the N1/E2-group kept the Norwegian conceptualization in the second session, when performing in English: Since they had already seen (and mentioned) the endpoints before, it was difficult to disregard them 2 weeks later. In this case, it is difficult to differentiate between conceptual transfer and memory transfer from the first session. Here, of course, it would be interesting to see whether the use of a different set of critical scenes in the second session would have affected the results, and/or for how long a memory effect would last. In the case of the E1/N2-group, the participants did in fact increase the total number of endpoints mentioned (though not significantly) after having switched from English to Norwegian, and their values for fixation duration did increase (in this case significantly), hence suggesting that the children had an awareness of the Norwegian conceptualization after all. However, the dominant-while-speaking effect was overruled, or at least influenced, by memory. In conclusion, while the bilingual children kept the morpho-syntax apart in the two languages on both occasions, their conceptual dominance could be described as flexible and changing, what we could call a *flexible conceptual dominance.* This, again, corroborates previous claims of the malleability or flexibility of the bilingual mind (cf., Athanasopoulos et al., 2015; Pavlenko, 2011).

Limitations and future prospects

Since we had a relatively small number of participants, especially in the E1-group ($n = 9$), for future studies it would be interesting to see whether and under what conditions a follow-up with a larger number of participants would replicate the findings. Furthermore, it would be crucial to compare the bilingual children not only to monolingual populations, but to other bilingual children with different language combinations and in different language contexts. But perhaps as importantly, while the current study aimed for balanced bilinguals, the question remains as to whether a (relatively) balanced proficiency is required for dominant-while-speaking to apply, or whether other factors, such as the frequency of usage, exposure to the languages, the status of the minority language, and/ or questions of identity play a more significant role. Such a line of enquiry would also concern the opposition between harmonious bilingualism (cf., De Houwer, 2009) and the feeling of outsiderness and language anxiety (cf., Sevinç, 2017; Sevinç & Backus, 2019).

In the present study, the Norwegian–English children were not only (more or less) balanced in terms of language proficiency, but also visibly proud of their two languages and the fact that they were bilingual. Moreover, the finding that the children in the first session were (over-)performing according to the prototypical conceptual pattern was interpreted as a sign of fully *leaning into* the language of operation. Feeling comfortable in one's language(s) would make it easier to *lean* or go *all in* into the language of operation: it might even be a prerequisite. Hence, the feeling of pride or comfort may have an impact on the level of conceptual awareness and/or dominance. As English cannot rightfully be called a minority language in a Norwegian context, a future study combining Norwegian with a different minority language, such as the aspect-language Polish, currently one of the biggest immigrant languages in Norway (Statistisk Sentralbyrå [SSB], 2023), and often associated with low social capital and agency (Obojska, 2020), might give a different outcome. In such a context, what would be the relation between language anxiety and conceptual dominance? Such

questions, concerning the relation between bilingual language-specific conceptualization, language anxiety, feelings of (national) identity, and perceptions of "different selves" lie at the intersection of sociolinguistics and psycholinguistics (see Lanza, 2004; Pavlenko, 2011, 2014), and constitute an exciting venue for future research.

ORCID iD

Hanna Andresen https://orcid.org/0009-0002-2092-5900

Notes

1. The Norwegian adults will not be referred to as monolinguals, as they all speak English to various degrees and often other languages as well, such as French, German, and/or Spanish and/or other languages, but they all grew up basically monolingual until they started learning English in school (normally around age 7–10 years).
2. The experimental design follows the design used in Schmiedtová et al. (2011) and von Stutterheim et al. (2012) quite closely.
3. Throughout, E1 refers to "English first" and N1 to "Norwegian first." Likewise, N2 refers to "Norwegian secondarily" and E2 to "English secondarily." Accordingly, the E1/N2-group consists of the same participants, tested in English first and Norwegian secondarily, with 2 weeks in between. The same holds for the N1/E2-group, but the other way around: Norwegian first and English 2 weeks later.
4. The corresponding Norwegian prepositions are *til, mot, i retning av, inn i, inn til, på vei til*. The coding is based on Schmiedtová (2013). For details on Norwegian, see Andresen (2023).
5. The memory tests were coded for correctly remembered endpoint versus wrong answer and/or no answer: 1 (*correct*), 0 (*incorrect*) and NA (*not answered*). 0 and NA are both counted as not remembered.
6. A total of 24 children were tested, but 1 participant was later excluded from the sample due to a lack of responses and poor tracking ratio in the eye-tracking data.
7. The difference between L1 Norwegian adults (*M* 50.3%) and L1 English adults (*M* 38%) is significant (measured with a paired *t*-test, $t \approx 2.40$, $N = 12$, $p \approx .035$, $D \approx .12$, 95% CI = [0.010, 0.23], $d \approx .40$, 95% CI = [0.041, 0.75]; for details, see Andresen, 2023).
8. In this case, we do not have comparable eye-tracking data for the adults.

References

Aktan-Erciyes, A., Goksun, T., Tekcan, A. I., & Aksu-Koç, A. (2020). Children's thinking for speaking: Bidirectional effects of L1 Turkish and L2 English for motion events. *Linguistic Approaches to Bilingualism*, *11*(5), 669–699. https://doi.org/10.1075/lab.19027.akt

Andresen, H. (2023). With Norwegian eyes. The verbalization of goal-oriented motion events in Norwegian. In B. Mertins, & R. Delucchi (Eds.), *Diversity in cognition. Cogniton: Kognitions- und neurowissenschaftliche Beiträge zur natürlichen Sprachverarbeitung* (pp. 147–181) Peter Lang Verlag.

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., Riches, N., & Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, *26*(4), 518–526. https://doi.org/10.1177/0956797614567509

Behrens, B., Flecken, M., & Carroll, M. (2013). Progressive attraction: On the use and grammaticalization of progressive aspect in Dutch, Norwegian, and German. *Journal of Germanic Linguistics*, *25*(2), 95–136. https://doi.org/10.1017/S1470542713000020

Behrens, B., & Ramm, W. (2003). Ser vi det samme eller styres vi av vår språklige kode? [Do we see the same or are we directed by our language? Report from the Norwegian delegation to the international language project "Konseptualisering og språkspesifikk kunnskap"]. *NOA—Norsk som andrespråk*, *24*, 109–142.

Bialystok, E. (2001). *Bilingualism in development. language, literacy, & cognition*. Cambridge University Press.

Carroll, M., von Stutterheim, C., & Nüse, R. (2004). The language and thought debate: A psycholinguistic approach. In C. Habel, & T. Pechman (Eds.), *Approaches to language production* (pp. 183–218). Mouton de Gruyter.

Comrie, B. (1976). *Aspect. An introduction to the study of verbal aspect and related problems*. Cambridge University Press.

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*, *14*(1), 95–119.

De Houwer, A. (2009). *Bilingual first language acquisition*. Multilingual Matters.

De Houwer, A. (2022). The danger of bilingual-monolingual comparisons in applied psycholinguistic research. *Applied Psycholinguistics*, *44*(3), 343–357. https://doi.org/10.1017/S014271642200042X

Engemann, H. (2022). How (not) to cross a boundary: Crosslinguistic influence in simultaneous bilingual children's event construal. *Bilingualism: Language and Cognition*, *25*, 42–54. https://doi.org/10.1017/S1366728921000298

Flecken, M. (2011). Event conceptualization by early Dutch-German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition*, *14*(1), 61–77. https://doi.org/10.1017/S1366728910000027

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Grosjean, F. (2013). Bilingualism: A short introduction; speech perception and comprehension; speech production (Ch. 1–3). In F. Grosjean, & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–69). Wiley-Blackwell.

Kinn, T. (2018). Pseudocoordination in Norwegian. Degrees of grammaticalization and constructional variants. In E. Coussé, P. Andersson, & J. Olofsson (Eds.), *Grammaticalization meets construction grammar* (pp. 75–106). Benjamins.

Kinn, T., Blensenius, K., & Andersson, P. (2018). Posture, location, and activity in mainland Scandinavian pseudocoordinations. *CogniTextes*, *18*(18), 1–34. https://doi.org/10.4000/cognitextes.1158

Lanza, E. (2004). *Language mixing in infant bilingualism. A sociolinguistic perspective*. Oxford University Press.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Massachusetts Institute of Technology.

Mertins, B. (2016). The use of experimental methods in linguistic research: Avantages, problems and possible pitfalls. In T. Anstatt, C. Clasmeier, & A. Gattnar (Eds.), *Slavic languages in psycholinguistics. Chances and challenges for empirical and experimental research* (pp. 15–33). Tübingen.

Mertins, B. (2018). *Sprache und Kognition: Ereigniskonzeptualisierung im Deutschen und Tschechischen*. De Gruyter. https://doi.org/10.1515/9783110615746

Obojska, M. A. (2020). What's in a name? Identity, indexicality and name-change in an immigrant context. *European Journal of Applied Linguistics*, *8*(2), 333–353. https://doi.org/10.1515/eujal-2020-0004

Ortega, L. (2014). Ways forward for a bi/multilingual turn in SLA. In S. May (Ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education* (pp. 32–53). Routledge.

Pavlenko, A. (2011). Thinking and speaking in two languages: Overview of the field. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (Ch. 8, pp. 237–257). Multilingual Matters.

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

Schmiedtová, B. (2011). Do L2 speakers think in the L1 when speaking in the L2? *Vigo International Journal of Applied Linguistics*, *8*(1), 139–179.

Schmiedtová, B. (2013). Traces of L1 patterns in the event construal of Czech advanced speakers of L2 English and L2 German. *International Review of Applied Linguistics in Language Teaching, 5*1, 87–116. https://doi.org/10.1515/iral-2013-0005

Schmiedtová, B., von Stutterheim, C., & Carroll, M. (2011). Language-specific patterns in event construal of advanced second language speakers. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (Ch. 3, pp. 66–107). Multilingual Matters.

Sevinç, Y. (2017). *Language anxiety in the immigrant context. An interdisciplinary perspective* [Doctoral dissertation]. University of Oslo.

Sevinç, Y., & Backus, A. (2019). Anxiety, language use and linguistic competence in an immigrant context: A vicious circle? *International Journal of Bilingual Education and Bilingualism*, *22*(6), 706–724. https://doi.org/10.1080/13670050.2017.1306021

Slobin, D. I. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics*, *1*(1), 7–25. https://doi.org/10.1075/prag.1.1.01slo

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

Statistisk Sentralbyrå [Central Statistical Office]. (2023). Fakta om innvandring (Facts about immigration). https://www.ssb.no/innvandring-og-innvandrere/faktaside/innvandring

Tonne, I. (1999). A Norwegian progressive marker and the level of grammaticalization. *Languages in Contrast*, *2*(1), 131–159. https://doi.org/10.1075/lic.2.1.08ton

Tonne, I. (2001). *Progressives in Norwegian and the theory of aspectuality* [Doctoral dissertation]. University of Oslo.

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*(4), 833–867. https://doi.org/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. D. Weger-Guntharp, & K. Aprang (Eds.), *Educating for advanced foreign language capacities* (pp. 40–53). Georgetown University Press.

von Stutterheim, C., & Nüse, R. (2003). Processes of conceptualization in language production: Language-specific perspectives and event construal. *Linguistics*, *41*(5), 831–881. https://doi.org/10.1515/ling.2003.028

Wang, Y., & Wei, L. (2022). *Thinking and speaking in a second language*. Cambridge University Press. https://doi.org/10.1017/9781009075053

Author biographies

Hanna Andresen is a PhD student at MultiLing Center for Multilingualism across the Lifespan at the University of Oslo. Her research interests include bilingualism, first and second language acquisition and teaching, language typology, and multilingualism.

Renate Delucchi Danhier is a psycholinguist and holds a PhD from Heidelberg University. She currently teaches linguistics at the TU Dortmund University and is part of the research team of the Psycholinguistics Laboratories. Her main research topics include space cognition, bilingualism, and linguistic relativity.

Barbara Mertins is a full professor at the TU Dortmund University with a focus on experimental linguistics and psycholinguistics. She is the Head of the Psycholinguistics Laboratories at the same university. Her main research topics include language processing of multilingualism, linguistic relativity, and language pathology.

Appendix 1

Picture 1. Example of a critical scene: A woman walking in the direction of a car at the parking lot (screenshot of video clip).

Picture 2. Example of a control scene: A cat walking from one room into another (screen shot of video clip).

Appendix 2

Picture 3. Example of a picture from the memory test, where two women are walking toward a house, but the house is cut off.

Appendix 3

Table 3. Comparison of the results (verbalizations, eye-tracking, and memory values): *the N1-group (Norwegian first) versus the E1-group (English first)*.

Measure	First session	
	NOR 1 (*n* = 14)	ENG 1 (*n* = 9)
Verbalization critical	M = 55.36% Median = 6.43 SD = 2.74 t = -2.7978, p = .01365	M = 27.86% Median = 4.11 SD = 2.93
Verbalization control	M = 80.36% Median = 8.93 SD = 2.34 t = −2.5124, p = .02502	M = 58.16% Median = 6.67 SD = 2.69
Eye-tracking Dwell T. critical	M = 992.4 Median = 892.0 SD = 891.3471 t = −1.3676, p = .1728	M = 738.8 Median = 311.9 SD = 981.038
Eye-tracking Dwell T. control	M = 2737 Median = 2556 SD = 1584.856 t = −.97468, p = .3306	M = 2754 Median = 2748 SD = 1742.239
Eye-tracking Fixation D. critical	M = 316.5, Median = 300.0 SD = 237.834 t = 3.0714, p = .002381	M = 228.9 Median = 206.6 SD = 226.9164
Eye-tracking Fixation D. control	M = 508.9 Median = 436.3 SD = 357.9778 t = −1.1406, p = .2556	M = 573.0, Median = 482.2 SD = 508.6775
Memory critical	M = 49.21 % Median = 0.67 SD = 0.22 t = .60248, p = .5554	M = 44.44 % Median = 0.60 SD = 0.20
Memory control	M = 73.57% Median = 8.93 SD = 2.34 t = −.024741, p = .9806	M = 73.33% Median = 6.67 SD = 2.69

Appendix 4

Table 4. Comparison of the results (verbalizations, eye-tracking, and memory values): *the N2-group (Norwegian second) versus the E2-group (English second).*

Meassure	Second session	
	NOR 2 (*n* = 9)	ENG 2 (*n* = 14)
Verbalization critical	M = 35.19% Median = 4.44 SD = 2.35 t = 1.6684, p = .1142	M = 50.60% Median = 5.57 SD = 2.93
Verbalization control	M = 69.38 % Median = 9.22 SD = 2.17 t = .92223, p = .3669	M = 77.92% Median = 6.07 SD = 2.50
Eye-tracking Dwell T. critical	M = 810.9 Median = 575.8 SD = 827.358 t = −2.6914, p = .007574	M = 1114.8 Median = 875.9 SD = 1038.113
Eye-tracking Dwell T. control	M = 2,702 Median = 2,486 SD = 1,717.699 t = −1.6451, p = .1014	M = 3044 Median = 2,886 SD = 1,630.286
Eye-tracking Fixation D. critical	M = 308.7 Median = 258.9 SD = 303.3281 t = 2.1711, p = .03103	M = 358.9 Median = 304.5 SD = 288.0545
Eye-tracking Fixation D. control	M = 583.0 Median = 462.1 SD = 461.5971 t = −.082042, p = .9347	M = 645.6 Median = 478.2 SD = 603.0053
Memory critical	M = 62.96% Median = 0.72 SD: 0.27 t = .18541, p = .8549	M = 61.51% Median = 0.73 SD = 0.18
Memory control	M = 91.11% Median = 9.22 SD = 2.17 t = .68534, p = .502	M = 86.79% Median = 9.36 SD = 3.05