

RESEARCH REPORT

EYE GAZE AND PRODUCTION ACCURACY PREDICT ENGLISH L2 SPEAKERS' MORPHOSYNTACTIC LEARNING

Kim McDonough
Concordia University

Pavel Trofimovich
Concordia University

Phung Dao
Concordia University

Alexandre Dion
Concordia University

This study investigated the relationship between second language (L2) speakers' success in learning a new morphosyntactic pattern and characteristics of one-on-one learning activities, including opportunities to comprehend and produce the target pattern, receive feedback from an interlocutor, and attend to the meaning of the pattern through self- and interlocutor-initiated eye-gaze behaviors. L2 English students ($N = 48$) were exposed to the transitive construction in

We would like to thank the research assistants who helped with data collection, transcription, and coding: Abigail Sherby, Stella Carolina Stella, and Lauren Strachan. This research was supported by grants from the Social Sciences and Humanities Council of Canada (435-2015-1206) and the Canada Research Chairs program (950-221304).

Correspondence concerning this article should be addressed to Kim McDonough, 1455 de Maisonneuve Blvd W., Education Department, FG 6-151, Montreal, QC H3G 1M8 Canada. E-mail: kim.mcdonough@concordia.ca

Esperanto (e.g., *filino mordas pomon* [SVO] or *pomon mordas filino* [OVS] “girl bites apple”) through comprehension and production activities with an interlocutor, receiving feedback in the form of recasts for their Esperanto errors. The L2 speakers’ interpretation and production of Esperanto transitives were then tested using known and novel lexical items. The results indicated that OVS test performance was predicted by the duration of self-initiated eye gaze to images illustrating the OVS pattern during the comprehension learning activity and by accurate production of OVS sentences during the production learning activity. The findings suggest important roles for eye-gaze behavior and production opportunities in L2 pattern learning.

Construction learning studies have reported considerable variability in second language (L2) speakers’ success at learning novel morpho-syntactic patterns following brief exposure to meaning-based, comprehension activities (Fulga & McDonough, 2016; McDonough & Fulga, 2015; McDonough & Trofimovich, 2013, 2015, 2016; Nakamura, 2012; Year & Gordon, 2009). For example, previous studies that investigated L2 speakers’ ability to learn the key morphological (accusative *-n* case marking) and syntactic (variable word order: SVO or OVS) features of the Esperanto transitive construction (as in *bubalo pelas kapron* [SVO] and *kapron pelas bubalo* [OVS] “buffalo chases goat”) reported an average success rate of only 23% (Fulga & McDonough, 2016; McDonough & Fulga, 2015; McDonough & Trofimovich, 2016). The participants in these studies consistently misinterpreted OVS transitives as being SVO transitives, likely because they treated the first noun as the agent, which can be attributed to both general processing strategies (Ferreira, 2003) and the influence of word-order cues in the participants’ first languages (L1s) (MacWhinney, 2012; VanPatten, 1996, 2004).

Another potential reason that these L2 speakers experienced difficulty acquiring the novel morphosyntactic pattern in Esperanto may have been due to the nature of the learning activities. In these studies, the learning phase consisted of noninteractive, forced-choice comprehension activities administered to groups of participants through timed presentation slides, with participants required to listen to Esperanto transitives in both SVO and OVS word orders and select the image that corresponded to their meaning. The activities did not enhance the salience of the accusative case marking (e.g., by stressing the suffix), provide any feedback if participants selected the incorrect pictures, or offer any production opportunities. Furthermore, the presentation of images and aural sentences was strictly timed, giving participants little time (approximately 10 seconds) to view the images and select the appropriate image.

Consequently, participants were under considerable time pressure and received little additional evidence that could have helped them abandon a familiar word-order cue (SVO) in favor of a key morphological cue (accusative *-n* suffix). This difficulty with extracting novel linguistic cues from positive evidence alone supports linguistic arguments about the insufficiency of positive evidence in L2 learning (e.g., VanPatten, 2007; White, 2007) and illustrates such (negative) cognitive learning phenomena as overshadowing and blocking (Ellis et al., 2014; MacWhinney, 2012).

In light of the difficulty that L2 speakers have shown with pattern learning through these types of listening activities, an interesting question is whether one-on-one learning activities might create a more acquisition-rich environment. In contrast to a timed, group presentation format using prerecorded aural sentences, an individual session with an interlocutor might be a more effective way to learn a novel pattern. During one-on-one, individualized learning comprehension activities, for example, learners can request repetition if they do not understand an Esperanto sentence, receive the correct answer if they misunderstand an OVS transitive sentence, and take as much time as they need to look at the visual images. Production activities with an interlocutor give learners opportunities to produce the target structures, and receive feedback if they produce nontargetlike forms. All these features of more individualized learning activities could help learners recognize that accusative *-n* case marking is an important cue for interpreting Esperanto transitives.

Carrying out activities with an interlocutor may also facilitate pattern learning by creating conditions for joint attention, which is the human capacity to coordinate attention with a social partner (Moore & Dunham, 1995). Joint attention occurs during conversation when interlocutors coordinate their behavior by initiating and responding to visual cues, such as facial expressions, gestures, and eye gaze. Among visual cues, a speaker's eye gaze has been shown to have the most consistent impact on listeners' responses (Bavelas, Coates, & Johnson, 2002). For instance, an interlocutor's initiation of joint attention through eye gaze is associated with requests for repair (Rossano, Brown, & Levinson, 2009), while mutual eye gaze (i.e., when a speaker looks to a listener who returns eye gaze) is associated with listener responses (Bavelas et al., 2002). In the context of one-on-one learning, for example, besides providing feedback and eliciting repair, an interlocutor can use visual cues to orient L2 speakers to images that illustrate key form-meaning relationships, which may help them interpret Esperanto transitives accurately. Furthermore, the L2 speakers could use eye gaze as a sign that they are attending to these key form-meaning relationships. Just as mutual eye gaze and L2 speaker eye gaze duration have been shown to predict targetlike responses to recasts (McDonough, Crowther,

Kielstra, & Trofimovich, 2015), both L2 speaker and interlocutor eye gaze behavior may also play a role in morphosyntactic pattern learning.

In sum, due to the challenges that L2 speakers face when adopting a morphological cue for sentence interpretation, it is possible that the noninteractive listening activities provided in previous experiments were insufficient for promoting learning. Therefore, the current study explored whether one-on-one learning activities with an interlocutor facilitate novel morphosyntactic pattern learning. More specifically, the study examined which specific characteristics of carrying out learning activities with an interlocutor—namely, opportunities to comprehend and produce the target pattern, receive recasts, and orient to the correct visual image through self- and interlocutor-initiated eye-gaze behaviors—are predictive of novel pattern learning. The research question was: *What aspects of performance during one-on-one learning activities are associated with L2 speakers' morphosyntactic pattern learning?*

METHOD

Participants

The participants were 48 L2 English students (22 women, 26 men) studying degree programs at an English-medium university in Canada. They ranged in age from 18 to 52 years with a mean of 25.6 years ($SD = 6.6$) and had resided in Canada for a mean of 33 months ($SD = 44.6$). They reported speaking numerous L1s, including Farsi (10), Mandarin (8), French (8), Portuguese (5), Spanish (4), Vietnamese (3), Russian (3), Telugu (2), Hindi and Urdu (2), and Arabic, Punjabi, and Japanese (1 each). The students reported having studied English for a mean of 9.5 years ($SD = 4.8$). Based on the sample size ($N = 48$), the number of predictor variables selected for entry into the regression model ($N = 3$), the alpha level (.05), and an anticipated medium effect size, power was estimated at .70.

Design

An associational design was used to explore the relationship between characteristics of the learning activities and L2 speakers' morphosyntactic pattern learning. The outcome variable was pattern learning, which was operationalized as the ability to comprehend and produce Esperanto OVS transitives, which have been tested in previous construction learning studies that explored L2 speakers' ability to adopt morphological cues for sentence interpretation. As mentioned previously, Esperanto transitives

are characterized by the presence of the suffix *-n* to mark all nouns as functioning as an object along with variable word order. For example, the sentence *girl bites apple* can be expressed as either *filino mordas pomon* (SVO) or *pomon mordas filino* (OVS). Due to speakers' tendency to interpret the first noun as the agent, which makes SVO sentences a highly familiar target, the outcome variable focused more narrowly on their accuracy at identifying and producing OVS sentences. To reduce the number of regression analyses performed on the data, a single outcome variable was created by summing the OVS accuracy scores from the comprehension and production tests.

The initial set of predictor variables reflected various aspects of the participants' performance during the learning activities, and included (a) accurate comprehension of OVS sentences, (b) production of accurate OVS sentences, (c) OVS recasts received, (d) self-initiated eye gaze to OVS pictures during sentence comprehension, (e) self-initiated eye gaze to OVS pictures during sentence production, and (f) other-initiated eye gaze to OVS pictures during recasts. The first three predictors involved various aspects of L2 speakers' performance during one-on-one learning, reflecting key linguistic characteristics of the learning session, such as comprehension and production accuracy and the number of feedback episodes received. The decision to include self- and other-initiated eye-gaze behaviors as predictors of OVS performance was motivated by the possibility that L2 speakers' own looking behavior could be associated with their detection of the key morphosyntactic feature of the Esperanto transitives (accusative *-n* suffix) and that interlocutor's eye gaze could cue L2 speakers—by way of joint attention—to attend to this feature. To increase variability in the participants' eye gaze, the interlocutor initiated attention to the images while speaking Esperanto sentences and giving recasts through either a single visual cue (head turn only) or two visual cues (head turn plus pointing). Participants were randomly assigned to carry out the learning activities with an interlocutor who provided either single ($n = 21$) or dual ($n = 27$) cues toward the visual images.¹

Materials

Learning Activities. The learning phase included both comprehension and production picture-based activities. The comprehension activity consisted of 32 sentences (16 SVO, 16 OVS) created from two verbs (*mordas* [bite], *batas* [hit]) and seven nouns (*pomo* [apple], *pilko* [ball], *knabo* [boy], *filino* [girl], *bubalo* [buffalo], *pordego* [gate], *kapro* [goat]). The sentences were grouped into four sets, with the first set containing human subjects and inanimate objects ($n = 8$), while the second set had

animal subjects and inanimate objects ($n = 8$). The third set had either human or animal subjects paired with inanimate objects ($n = 8$), and the fourth set presented inanimate subjects paired with either human or animal objects ($n = 8$). The nouns occurred equally in subject and object position, and the verbs were used an equal number of times. Each sentence was paired with two images printed side by side on poster boards (labeled as images A and B), with one image correctly depicting the meaning of each sentence. The distracter image differed from the correct image in terms of a single lexical item, with the contrasting element counterbalanced across the subjects, verbs, and objects to ensure that the participants had opportunities to orient to both the word order and form of subjects and objects in Esperanto transitives. For example, the distracter picture for the sentence *pomo batas bubalon* (SVO, “apple hits buffalo”) showed an image of an apple hitting a goat. The picture description activity consisted of eight pairs of images printed side by side on poster boards. The images showed the same nouns and verbs used in the comprehension learning items. However, each poster board had images of fully reversible events. For instance, one poster board showed an image of an apple hitting a gate (image A) while the second image was a gate hitting an apple (image B). This use of target and distracter images is comparable to the use of visual materials in processing instruction (VanPatten, 1996), where images are often chosen to cue aspects of L2 morphology and syntax distinguishing meaning, such as the use of pronominal clitics to indicate objects.

Immediate Comprehension Test. The immediate comprehension test contained eight previously unheard sentences (four SVO, four OVS) created from the same seven nouns and two verbs. Each noun occurred two or three times as subjects and objects, and each verb was used four times. Animate nouns occurred as both the subject and object for four sentences, with the remaining four sentences containing animate subjects with inanimate objects and inanimate subjects with inanimate objects. Each sentence was paired with two images printed side by side on poster boards with one image correctly depicting the meaning of each sentence (also labeled images A and B). The pictures in the immediate test phase presented images of fully reversible events (e.g., a picture showing a ball hitting a girl alongside an image of a girl hitting a ball). Thus, it was impossible to identify the correct image based on lexical knowledge, as both images depicted the same two nouns and verb. Consequently, participants required knowledge of the morphosyntactic features of the Esperanto transitive construction to select the correct picture.

Generalization Test. The generalization comprehension test included comprehension and production activities that targeted a new set of

lexical items to determine if participants were able to generalize the morphosyntactic features of Esperanto to novel words. The sentences and images were created from two verbs (*pedbatas* [kick], *tiras* [pull]), and six nouns (*zebro* [zebra], *makropo* [kangaroo], *doktoro* [doctor], *gardenisto* [gardener], *cervo* [deer], and *cevalo* [horse]). For the comprehension activity, 12 sentences were created (six SVO, six OVS), with each noun occurring four times and each verb used six times. All nouns were animate, so every sentence had an animate subject and an animate object. Each sentence was paired with two images printed side by side on poster boards, labeled as images A and B. As in the immediate test, the images for each sentence were fully reversible to ensure that the participants relied on morphosyntactic features rather than lexical items to identify the correct picture. For example, the sentence *makropo tiras gardeniston* (SVO, “kangaroo pulls gardener”) was paired with an image of a kangaroo pulling a gardener and a gardener pulling a kangaroo. The production activity consisted of 12 sets of reversible images created from the same lexical items. The reliability value (Cronbach’s α) of all comprehension test items was .65.

PROCEDURE

The participants carried out the experimental activities during an approximately 90-minute individual session with a researcher and a research assistant (RA) trained by the researchers. The researcher was responsible for explaining the project, administering forms, and monitoring the eye-tracking equipment, whereas the RA carried out the Esperanto activities. The faceLAB 5 eye-tracking system was used to capture the eye movements of the RA and the participant, who were seated at a table opposite each other with four cameras positioned on two stereo heads in the middle of the table, so that two cameras tracked and recorded the eye gaze and movement of each person. Two Logitech webcams were placed next to the four cameras to record the scene, which included the head and torso of the RA and participant. Together these cameras integrated the eye movement and field-of-vision data, specifically where in the scene (as depicted visually by a green dot in the field of vision) each person looked. For the participants, the cameras were calibrated to determine if they were looking at the RA or the target images. For the RA, only their general gaze to the participant was tracked. The cameras were connected to two synchronized DELL Latitude E5520 laptops recording both interlocutors’ eye gaze and movement.

After filling out consent and background information forms (15 minutes), the participants completed a brief calibration process for the eye-tracking equipment (15 minutes). They then interacted with an RA to learn the initial Esperanto vocabulary (10 minutes) and then carried

out the experimental activities across the learning (15 minutes), immediate test (5 minutes), and generalization test phases (13 minutes). Prior to undertaking the sentence trials, the RA checked to make sure that the participant knew the meaning of all the vocabulary items. During all experimental phases, the RA held the poster boards with the images by the top left and right corners at chest height so the participant could see them easily above the eye-tracking equipment and keep their eye gaze within the calibration field. For all comprehension activities, the RA produced each sentence twice before the participant stated whether the corresponding picture was image A or B, but the participant could request repetition if needed. For the production activities, the RA indicated which image the participant should describe.

During the learning phase only, the RA provided feedback if the participants selected the incorrect image or produced an ungrammatical sentence. When a participant selected the wrong picture or produced an ungrammatical sentence, the RA initiated eye gaze with the participant and turned her head to look at the correct image while repeating the correct sentence or recasting the participant's incorrect Esperanto sentences. For all recasts, the RAs produced complete Esperanto transitives without any additional stress or rising intonation, and paused so that the participant could repeat the recast. Any errors in English that the participants produced while carrying out the Esperanto tasks were ignored. For participants assigned to the dual cue condition, the RA also tapped an index finger at the top corner of the correct image while looking at it. No feedback was given during immediate and generalization tests. Following the Esperanto activities, each participant completed an exit interview (10 minutes) to answer questions regarding their perceptions about the eye-tracking equipment, the RA's feedback, and Esperanto grammar, and to generate Esperanto sentences using four nouns and two verbs not included in the experimental materials. All verbal interaction was audio-recorded using a Sony portable digital recorder.

ANALYSIS

The audio-recordings were transcribed and verified by the researchers and RAs. To create a single outcome variable, the OVS immediate and generalization comprehension tests ($k = 10$) were summed and added to the number of correct OVS sentences produced during the generalization production test (max = 12).² For the picture identification items, the participants' initial responses were coded for accuracy (i.e., selection of the correct image). For the picture description items, the participants' initial sentence was coded for its form (i.e., SVO or OVS) and accuracy (i.e., correct or ungrammatical), and the number and type of

recasts (SVO or OVS) received during the learning phase were summed. The picture identification and description coding was carried out independently by two researchers, and there were no disagreements about the form or accuracy of Esperanto transitives.

The eye-gaze data were coded using the Captiv video-analysis program, which allowed the researchers to view videos from both interlocutors simultaneously. For each picture identification and picture description item, the total duration of self-initiated looks to the correct image was summed, with a look defined as having a minimum length of 180 milliseconds, which is the lowest mean fixation duration (180 to 275 milliseconds) for a visual search (Rayner & Castelano, 2007). For picture identification items, the length of eye gaze was calculated from the time participants looked at the correct picture while or after hearing the sentence until the moment they gave an answer. For the picture description items, the eye-gaze duration was computed from when participants were told which picture to describe until they produced a sentence. Other-initiated looks to the correct image were also coded for looks and total duration. When a participant received recasts, the total duration of other-initiated looks to the correct image was summed from when the RA completed the visual cue and began speaking until they finished delivering the recast. To account for variation in the number and type of transitives produced by the participants and recasts provided by the RAs, mean eye-gaze duration was computed per participant for all OVS items. Interrater reliability for the eye-gaze data (i.e., number of looks and duration) was calculated for a subset (25%) of the data. Cronbach's alpha was .92 for the number of looks and .96 for duration.

RESULTS

The participants' Esperanto OVS test scores and their performance during the interactive learning activities are summarized in Table 1. In terms of the outcome variable, the participants accurately identified the meaning of OVS transitives in the immediate and generalization tests combined at a mean rate of 37%. In terms of their production during the generalization test, they produced 12% accurate OVS transitives. In general, the participants had relatively low test accuracy for OVS transitives in both comprehension and production, which is not unexpected due to the challenges that L2 speakers face when learning new morphosyntactic patterns. During the learning phase, when lexical knowledge could help participants identify the correct picture, their accuracy was much higher (94%), and their self-initiated eye gaze to the correct image was approximately 3.5 seconds. When producing Esperanto transitives during the production learning task, the participants produced

Table 1. Descriptive statistics for all variables

Variable	<i>M</i>	<i>SD</i>	95% CI
Esperanto tests			
Comprehension (<i>k</i> = 10)	3.71 (37%)	3.38	[2.73, 4.69]
Production (<i>k</i> = 12)	1.40 (12%)	2.70	[0.61, 2.18]
Total (<i>max</i> = 22)	5.10 (23%)	5.03	[3.64, 6.56]
Esperanto learning phase			
Comprehension accuracy (<i>k</i> = 16)	15.06 (94%)	1.14	[8.77, 10.65]
Comprehension mean eye gaze	3.58	0.74	[3.36, 3.80]
Production accuracy (<i>k</i> = 8)	1.29 (16%)	1.56	[0.84, 1.74]
Production mean eye gaze	7.90	2.97	[6.98, 8.83]
Recasts received	2.46	1.58	[2.00, 2.92]
Recasts mean eye gaze	2.27	1.33	[1.87, 2.67]

few accurate OVS sentences (16%), and spent nearly 8 seconds on average looking at the images while producing those utterances. In response to their ungrammatical sentences, the participants received a mean of 2.46 OVS recasts ($SD = 1.58$).³ While receiving recasts, the participants' other-initiated eye gaze to the correct picture was approximately 2 seconds.

To address the research question, which asked about the predictors of L2 morphosyntactic pattern learning, a correlation analysis was carried out first to identify which variables were most closely associated with OVS test accuracy.⁴ To be included in the regression model, a predictor variable's correlation coefficient had to reach $\pm .25$, as this value is considered a benchmark for weak associations in L2 research (Plonsky & Oswald, 2014). As shown in Table 2, three predictor variables met the inclusion criterion: self-initiated eye gaze during the comprehension learning task, production accuracy, and other-initiated eye gaze during recasts.

The three predictor variables were entered into a regression model using forced entry. The model was statistically significant, $F(3, 44) = 15.84$,

Table 2. Summary of Pearson correlations

Variable	2	3	4	5	6	7
1. Test accuracy	.21	.37	.65	-.07	-.12	-.32
2. Comprehension accuracy	—	.25	.28	.02	-.12	.03
3. Comprehension eye gaze		—	.15	.37	.30	.07
4. Production accuracy			—	-.14	-.12	-.43
5. Production eye gaze				—	.22	.11
6. Recasts					—	.30
7. Recast eye gaze						—

$p = .01$, and accounted for a total of 52% of the variance in Esperanto test performance ($R^2 = .52$, adjusted $R^2 = .49$). As shown in Table 3, the only significant predictors in the model were comprehension eye-gaze duration and production accuracy. As indicated by positive beta values, longer self-initiated eye-gaze duration to OVS items during the comprehension learning task and greater production of accurate OVS sentences during the production learning task were associated with more accurate Esperanto test performance. While holding all other variables constant, a 1-second increase in eye gaze to the correct OVS image and production of one additional OVS transitive led to increases in test scores of 2.01 and 1.76 points, respectively. In terms of assumptions and model fit, there was no multicollinearity among the predictor variables (correlations of .43 or lower and no tolerance statistics below .2), only two cases (4%) had standardized residuals greater ± 2 , no cases had Cook's distance values greater than one, the Durbin-Watson value was 2.05, and the plot of the predicted and standardized residuals showed no signs of heteroscedasticity.

The significant, positive relationship between OVS production accuracy and test performance is relatively straightforward in that participants who were able to produce correct OVS sentences during the learning activities scored higher on the subsequent tests. However, the relationship between self-initiated eye gaze during the comprehension learning activities and test performance is less transparent. To obtain a more nuanced view of that relationship, a subset of the data from the participants with the highest and lowest test scores were analyzed to identify potential patterns in their eye-gaze behavior. Unlike the low scoring participants ($n = 15$), who had test scores of 0 or 1, the higher scorers ($n = 13$) had test scores ranging from 8 to 22 (which was the maximum score possible). A possible explanation for their divergent test performance would be that the low scorers simply misunderstood the OVS picture identification items during the learning phase. However, both sets of participants had equally high accuracy rates during the comprehension learning activity (89–98% across groups and items).

An alternate explanation is that the high scorers were able to detect the key features of the Esperanto transitive construction during the learning activities, whereas the low scorers continued to rely on a more

Table 3. Summary of regression model

	<i>B</i>	<i>SE B</i>	95% CI	<i>b</i>	<i>t</i>	<i>p</i>
Comprehension eye gaze	2.07	0.73	[0.60, 3.54]	0.31	2.84	0.01
Production accuracy	1.76	0.38	[1.00, 2.52]	0.55	4.68	0.01
Recast eye gaze	-0.55	0.42	[-1.39, 0.28]	-0.15	-1.33	0.19
Constant	-3.40	2.65				

familiar word-order cue (SVO) when identifying the pictures. Assuming that longer eye-gaze durations would be reflective of at least some aspects of L2 speakers' morphosyntactic learning, such as their detection of the $-n$ suffix as a key feature, these participants' mean eye-gaze duration for the OVS items across the four learning sets was compared to determine whether eye gaze could shed light on the differences in their learning outcomes. As shown in Table 4, the high test scorers had longer mean eye-gaze duration (in seconds) to the OVS items than the low test scorers for every set. Independent-samples t tests (equal variance assumed) using an adjusted alpha level of .01 (.05/4) indicated that the high test scorers looked at the OVS pictures significantly longer during the final set. Based on these comparisons, the high scorers' longer self-initiated eye-gaze to the OVS items, particularly during Set 4, suggests that they may have begun to detect the underlying pattern of Esperanto transitives during the learning phase. In contrast, the low scorers spent less time looking at the OVS pictures, especially in Set 4, which suggests that they may have oriented to the distracter pictures whose actions corresponded with an incorrect, agent-first (SVO) interpretation of the OVS sentences.

An interesting question raised by the two-predictor regression model (comprehension eye gaze and production accuracy) is whether participants with one behavior also engaged in the other. In other words, it could be that the participants who had longer eye gaze during the comprehension items also produced accurate OVS sentences during the subsequent production task. However, as previously shown in Table 2, the correlation coefficient between eye-gaze duration during the comprehension learning task and production accuracy was small (.15), which suggests little relationship between the variables. To investigate this relationship further, a post hoc comparison of the OVS eye-gaze duration for the participants who produced no OVS transitives ($n = 22$, $M = 3.52$, $SD = .87$) with the participants who produced at least two OVS transitives ($n = 17$, $M = 3.67$, $SD = .74$) revealed no statistically significant difference in their eye-gaze duration, $t(37) = .55$, $p = .59$, $d = .13$.

Table 4. Eye gaze during comprehension learning sets by high and low test scorers

Set	High test scorers ($n = 15$)		Low test scorers ($n = 13$)		Comparison		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
Set 1	2.47	.72	2.15	.95	1.01	.32	.51
Set 2	2.46	.48	2.28	.53	.92	.37	.36
Set 3	2.61	.60	2.17	.45	2.17	.04	.83
Set 4	2.26	.42	1.77	.50	2.47	.01	1.06

Taken together, the regression model and the post hoc analyses suggest that there were two independent routes to accurate test performance: one route through self-initiated eye gaze to OVS items during the comprehension learning activity and a second route through production accuracy.

DISCUSSION

The current study investigated the relationship between L2 speakers' success in learning a new morphosyntactic pattern (Esperanto transitive construction featuring OVS word order and *-n* accusative case marking) and several characteristics of one-on-one learning activities, which included opportunities to comprehend and produce the target pattern, receive feedback from an interlocutor, and orient to the meaning of the pattern through self- and interlocutor-initiated eye-gaze behaviors. Test performance with the target OVS pattern was predicted by two separate variables—self-initiated eye-gaze duration while viewing the images illustrating OVS sentences during the comprehension learning activity and accuracy of OVS picture descriptions during the production learning activity. Compared to the findings of prior research investigating the learning of the Esperanto transitive construction (e.g., Fulga & McDonough, 2016; McDonough & Trofimovich, 2013) and other novel structures (e.g., Nakamura, 2012; Year & Gordon, 2009), these results suggest that using individualized learning activities may be no more effective than group presentation of timed listening activities for promoting pattern learning. Compared to the Esperanto learning activities used in previous studies, which provided prerecorded audio materials, paired with relevant target and distracter images and delivered through timed PowerPoint slide presentations, the 37% comprehension test performance in this study was noticeably lower than the 54–63% accuracy rates for instruction that included explicit (deductive) information about the pattern (McDonough & Trofimovich, 2013) or the 54–59% accuracy rates for learners whose L1s have differential object marking related to definiteness (Fulga & McDonough, 2016).

Despite the overall low test performance, longer self-initiated eye-gaze duration to OVS items during the comprehension learning activity was a significant predictor of test scores. In essence, the participants who spent more time looking at the images that illustrated the meaning of OVS learning items were more accurate with the subsequent test items than participants with shorter eye-gaze durations. During the last set of comprehension learning items, high test scorers looked at the target OVS images on average for nearly 1 second longer than low test performers (Set 4 in Table 4). Self-initiated looking behavior thus distinguished those who succeeded and who generally failed in morphosyntactic learning.

Comments from the exit interview confirmed that participants with longer looks to the images in Set 4 could articulate the key features of the Esperanto transitive or generate OVS transitives with completely new nouns and verbs during the interview, even if they had not produced any OVS sentences during the production learning activity (e.g., *the subject and the object of the action, you can move that around, and it does not matter . . . what matters is that there's the -on*). In contrast, none of the participants with short looks, who similarly failed to produce any OVS transitives during the production learning task, were able to articulate any of Esperanto's morphosyntactic features or generate OVS sentences during the interview (e.g., *the same grammar order as in English*).

The positive association between self-initiated eye-gaze duration and morphosyntactic learning in this study is generally consistent with research showing the positive role of visual cues, such as gestures and pointing, in various forms of learning such as the learning of L2 lexis and pronunciation (e.g., Gullberg, Roberts, & Dimroth, 2012; Kelly & Lee, 2012). However, when the interlocutor initiated joint attention using either a single (head turn) or dual (head turn with pointing) visual cue, there was no difference in participants' orientation to the images of OVS transitives. The two-gesture condition led to no differences for any learning outcomes, test performance, or eye-gaze behaviors (see note 1). Furthermore, other-initiated eye gaze to the target images following the interlocutor's recasts was not associated with test performance rates (see Table 2). Rather, what mattered for test performance was the L2 speakers' self-initiated looking behavior during the learning items. This confirms the finding of the previous studies of eye-gaze behavior during interactive tasks (McDonough et al., 2015), which reported that L2 speakers' production of targetlike responses to recasts was predicted by the length of their self-initiated eye gaze while reformulating, rather than their interlocutor's eye-gaze duration. Taken together, the findings suggest that self-initiated eye gaze may help shed light on when learners make use of learning opportunities. However, in the absence of further evidence as to potential reasons for their self-initiated eye-gaze behaviors, the link between eye-gaze duration and test performance can be interpreted only descriptively: Longer eye-gaze durations illustrate that learning happens but do not directly explain why and how it occurs.

The other significant aspect of interaction positively linked to test performance was greater production accuracy with the OVS items during the production learning activity. This finding highlights production practice as an important route to test accuracy in morphosyntactic pattern learning. This finding is compatible with structural priming research showing that it is the production (rather than comprehension) component of priming tasks that is associated with learners' subsequent production accuracy rates (McDonough & Chaikitmongkol, 2010;

McDonough & Kim, 2009). It appears, then, that the production opportunities provided by interactive activities might drive at least some learning of target morphosyntactic features, likely because generation of utterances requires learners to attend more extensively to the relationship between form and meaning. Comments from the exit interview confirmed that all the participants who produced at least three accurate OVS transitives during the production learning activity could articulate the key features of Esperanto transitives or generate new OVS sentences during the interview (e.g., *makropon piedbatas cervo*, or you can say *cervo piedbatas makropon* [deer kick kangaroo] *they are the same*). This learning function of output activities might potentially be interpreted within views of production practice as reflecting a tuning of the language production system as a result of experience (Ferreira & Bock, 2006).

Another finding was that the number of individual recasts received by participants in response to their production of inaccurate OVS transitives was not related to their test performance. This finding is unsurprising given that recasts of morphosyntactic errors, compared to those targeting aspects of lexis and pronunciation, are often misinterpreted (Mackey, Gass, & McDonough, 2000). Moreover, recasting was overall infrequent and potentially ambiguous as it targeted both SVO and OVS utterances, with each participant receiving on average only around two OVS recasts. Finally, most participants were unaware they had received feedback during activities. Based on the exit interview comments, when asked if the interlocutor provided feedback, many participants indicated that no feedback had been given, while others mentioned positive feedback (*nice, good work, smiling*), pronunciation (*sometimes you correct my pronunciation*), and word choice (*I said the wrong word and you corrected it*). Only one participant mentioned receiving feedback about the accusative *-n* suffix. This finding provides further evidence that L2 speakers may not recognize or be able to articulate the corrective function of recasts, which have been the focus of considerable debate (e.g., Goo & Mackey, 2013; Lyster & Ranta, 2013). The recasts provided in this study may not have encouraged the students to tune to the critical cue (accusative case marking) necessary for accurate sentence production. Other forms of feedback, such as explicit correction, might fare better in enabling learners to become aware of key morphosyntactic cues and produce target structures.

Limitations and Future Research

As an initial step in exploring the effectiveness of individualized activities at promoting novel morphosyntactic pattern learning by L2 speakers,

the findings highlight several avenues for future research that may help overcome some of the current methodological limitations. First, this dataset was based on a relatively small sample, which resulted in lower statistical power (.70) for detecting potentially meaningful relationships. Although it surpassed the mean power level of .57 reported in applied linguistic research (Plonsky, 2013) and .56 in interaction research (Plonsky & Gass, 2011), it did not reach the recommended .80 level (Cohen, 1992). As a consequence of low power, it is possible that predictor variables with a true relationship to the outcome variable remain undetected. Second, this analysis targeted the initial stages of morphosyntactic learning during which students had opportunities to make key form-meaning mappings through brief exposure to the target language (15 minutes). As a result, it is not clear whether the same characteristics of learning that are relevant during this initial stage would also play a role during more extensive learning activities or affect extension or proceduralization. With respect to feedback, future research should explore whether feedback that more explicitly highlights the targeted feature (e.g., by emphasizing the *-n* accusative suffix through stress and syllable lengthening) impacts test performance. Finally, although the participants carried out learning activities with an interlocutor, their ability to engage in the free communication of meaning associated with the interactionist approach to L2 acquisition (e.g., Gass, 2003; Long, 1996; Mackey, 2012) was negatively impacted by their limited knowledge of Esperanto. Future studies involving L2 speakers who are not true beginners would allow for the implementation of more communicative activities. To clarify how features of interactive activities contribute to learning, especially those that transcend verbal behavior, future research is needed to explore the role of self- and other-initiated eye gaze in enabling learners to extract key features of novel morphosyntactic patterns.

Received 22 April 2016

Accepted 15 September 2016

Final Version Received 20 September 2016

NOTES

1. The data were checked to ensure that there were no significant differences in the performance of participants who received the single or dual cue for any variable. Comments during the debrief interview also indicated no differences, with three participants in each condition mentioning pointing when asked explicitly about the researchers' body language, despite the fact that the single cue group did not receive any pointing gestures. In sum, the comparison suggests that any variation in the participants' success at learning the novel pattern was not due to the type of cue used to draw their attention to the correct pictures.

2. Pearson correlations indicated that there were significant relationships among all three test scores with *r* values ranging from .30 to .57. For exploratory purposes, separate

regression models of each test score were calculated, and the same two predictor variables revealed p values ranging from .01 to .12 in all three models.

3. In addition to the OVS recasts, the participants also received a mean of 3.17 SVO recasts ($SD = 1.75$).

4. Because prior research with the Esperanto transitive construction has shown differences in test performance based on the participants' L1 (Fulga & McDonough, 2016), we checked to ensure that there were no differences in the test scores due to L1 features (no case marking, restrictive case marking, differential object marking based on either animacy or definiteness), $F(3,44) = .08$, $p = .97$, $\eta_p^2 = .01$.

REFERENCES

- Bavelas, J., Coates, L., & Johnson, T. (2002). Listener responses as collaborative process: The role of gaze. *Journal of Communication*, 52, 566–580.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & Sagarra, N. (2014). An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics*, 35, 547–579.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164–203.
- Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, 21, 1011–1029.
- Fulga, A., & McDonough, K. (2016). The impact of L1 background and visual information on the effectiveness of low variability input. *Applied Psycholinguistics*, 37, 265–283.
- Gass, S. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 224–255). Oxford, UK: Blackwell.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35, 127–165.
- Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *IRAL*, 50, 239–276.
- Kelly, S. D., & Lee, A. L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, 27, 793–807.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of language acquisition: Second language acquisition* (Vol. 2, pp. 413–468). San Diego, CA: Academic Press.
- Lyster, R., & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition*, 35, 167–184.
- Mackey, A. (2012). *Input, interaction and corrective feedback in L2 learning*. Oxford, UK: Oxford University Press.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497.
- MacWhinney, B. (2012). The logic of the unified model. In S. M. Gass & A. Mackey (Eds.), *Routledge handbook of second language acquisition* (pp. 211–227). New York, NY: Routledge.
- McDonough, K., & Chaikitmongkol, W. (2010). Collaborative syntactic priming activities and EFL learners' production of *wh*-questions. *Canadian Modern Language Review*, 66, 817–841.
- McDonough, K., & Fulga, A. (2015). The detection and primed production of novel constructions. *Language Learning*, 65, 353–384.
- McDonough, K., & Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of *wh*-questions. *The Modern Language Journal*, 93, 386–398.
- McDonough, K., & Trofimovich, P. (2013). Learning a novel pattern through balanced and skewed input. *Bilingualism: Language and Cognition*, 16, 654–662.
- McDonough, K., & Trofimovich, P. (2015). Structural priming and the acquisition of novel form-meaning mappings. In S. Eskildsen & T. Cardierno (Eds.), *Usage-based perspectives on second language learning* (pp. 105–123). Berlin: Mouton De Gruyter.
- McDonough, K., & Trofimovich, P. (2016). The role of statistical learning and working memory in L2 speakers' pattern learning. *The Modern Language Journal*, 100, 428–445.

- McDonough, K., Crowther, D., Kielstra, P., & Trofimovich, P. (2015). Exploring the potential role of eye-gaze in eliciting English L2 speakers' responses to recasts. *Second Language Research, 31*, 563–575.
- Moore, D., & Dunham, P. J. (1995). *Joint attention: Its origin and role in development*. Hillsdale, NJ: Lawrence Erlbaum.
- Nakamura, D. (2012). Input skewedness, consistency, and order of frequent verbs in frequency-driven second language construction learning: A replication and extension of Casenhiser and Goldberg (2005) to adult second language acquisition. *International Review of Applied Linguistics, 50*, 31–67.
- Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition, 35*, 655–687.
- Plonsky, L., & Gass, S. (2011). Quantitative research methods, study quality, and outcomes: The case of interaction research. *Language Learning, 61*, 325–366.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effects sizes in L2 research. *Language Learning, 64*, 878–912.
- Rayner, K., & Castelano, M. (2007). Eye movements. *Scholarpedia, 2*, 3649.
- Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 187–249). Cambridge, UK: Cambridge University Press.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Westport, CT: Ablex.
- VanPatten, B. (Ed.) (2004). *Processing instruction: Theory, research, and commentary*. Mahwah, NJ: Lawrence Erlbaum.
- VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115–135). Mahwah, NJ: Lawrence Erlbaum.
- White, L. (2007). Linguistic theory, universal grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 37–55). Mahwah, NJ: Lawrence Erlbaum.
- Year, J., & Gordon, P. (2009). Korean speakers' acquisition of the English ditransitive construction: The role of verb prototype, input distribution, and frequency. *The Modern Language Journal, 93*, 399–417.