Pre-output Language Monitoring in Sign Production

Stephanie K. Riès[1,2], Linda Nadalet[1], Soren Mickelsen[1], Megan Mott[1], Katherine J. Midgley[1], Phillip J. Holcomb[1,2], and Karen Emmorey[1,2]

Abstract

■ A domain-general monitoring mechanism is proposed to be involved in overt speech monitoring. This mechanism is reflected in a medial frontal component, the error negativity (Ne), present in both errors and correct trials (Ne-like wave) but larger in errors than correct trials. In overt speech production, this negativity starts to rise before speech onset and is therefore associated with inner speech monitoring. Here, we investigate whether the same monitoring mechanism is involved in sign language production. Twenty deaf signers (American Sign Language [ASL] dominant) and 16 hearing signers (English dominant) participated in a picture–word interference paradigm in ASL. As in previous studies, ASL naming latencies were measured using the keyboard release time. EEG results revealed a medial frontal negativity peaking within 15 msec after keyboard release

in the deaf signers. This negativity was larger in errors than correct trials, as previously observed in spoken language production. No clear negativity was present in the hearing signers. In addition, the slope of the Ne was correlated with ASL proficiency (measured by the ASL Sentence Repetition Task) across signers. Our results indicate that a similar medial frontal mechanism is engaged in preoutput language monitoring in sign and spoken language production. These results suggest that the monitoring mechanism reflected by the Ne/Ne-like wave is independent of output modality (i.e., spoken or signed) and likely monitors prearticulatory representations of language. Differences between groups may be linked to several factors including differences in language proficiency or more variable lexical access to motor programming latencies for hearing than deaf signers. ■

INTRODUCTION

Healthy adult speakers only err about once every 1000 words under natural speech conditions (Levelt, 1999). The same has been shown in sign language (Hohenberger, Happ, & Leuninger, 2002). Such highly efficient behavior is enabled in part by language monitoring processes, which are responsible for controlling our linguistic production as it is being output. Although the cognitive and neuronal mechanisms underlying speech monitoring have received some attention in the past few years, these mechanisms have been understudied in sign language production.

Various cognitive models of language monitoring have been proposed (Nozari, Dell, & Schwartz, 2011; Postma & Oomen, 2005; Postma, 2000), and all of these models make a distinction between monitoring processes involved before versus after language output. These monitoring processes have been referred to as the inner and outer loops of speech monitoring, respectively. The role of the inner loop is to monitor internal linguistic representations, whereas the outer loop relies on auditory feedback (in overt speech). Differences can be expected between sign and speech monitoring concerning the implementation of the outer loop as the auditory system should not be engaged when signing, and there is

evidence that signers do not rely on visual feedback when monitoring sign production for errors (Emmorey, Bosworth, & Kraljic, 2009; Emmorey, Gertsberg, Korpics, & Wright, 2009; Emmorey, Korpics, & Petronio, 2009). However, it is unclear whether or not differences between the inner loop monitoring mechanisms engaged in sign versus speech production would be observed. The way internal linguistic representations are monitored has been conceptualized in different ways. In particular, speech monitoring models differ in terms of whether the inner loop depends on the language comprehension system (Levelt, Roelofs, & Meyer, 1999) or on the language production system (Nozari et al., 2011; Postma, 2000). They also differ in terms of whether or not a domain-general monitoring mechanism is involved in inner speech monitoring (Acheson & Hagoort, 2014; Riès, Janssen, Dufau, Alario, & Burle, 2011) and whether or not this domain-general monitoring mechanism is conflict-based (Zheng, Roelofs, Farquhar, & Lemhöfer, 2018). Therefore, clarifying whether or not a similar brain mechanism is involved in sign language monitoring before signs are actually produced is a necessary step in furthering the understanding of sign language monitoring. Moreover, finding a similar inner loop mechanism in sign and speech production would be of interest in furthering the understanding of language monitoring more generally, as this would suggest the representations involved are not dependent of language output modality.

[1]San Diego State University, [2]University of California, San Diego

Similar brain regions have been shown to be engaged in sign and in speech production at the single word level but also at the phrase and narrative levels, including left temporal and left inferior frontal regions (Blanco-Elorrieta, Kastner, Emmorey, & Pylkkänen, 2018; Emmorey, Mehta, & Grabowski, 2007; Braun, Guillemin, Hosey, & Varga, 2001). This common neuronal substrate has been argued to underlie common semantic, lexical, and syntactic properties between sign and spoken languages. However, sign and spoken languages differ in several ways, particularly in how they are perceived (visual vs. auditory) and in the modality of output during production (manual and facial movements vs. a phonatory system). Such differences could arguably lead to differences in how sign and spoken languages are monitored. For example, somatosensory feedback in speech monitoring is linked to the movement of speech articulators (Tremblay, Shiller, & Ostry, 2003) but will include manual and facial movements in sign language monitoring (Emmorey, McCullough, Mehta, & Grabowski, 2014; Emmorey, Bosworth, et al., 2009).

Indeed, brain regions known to be associated with speech perception, in particular, the superior temporal cortex, have been found to be sensitive to manipulations affecting speech output monitoring (more specifically the outer loop mentioned above) such as speech distortion or delayed auditory feedback (Tourville, Reilly, & Guenther, 2008; Fu et al., 2006; Hashimoto & Sakai, 2003; McGuire, Silbersweig, & Frith, 1996). These results have been interpreted as supporting the idea that overt language output monitoring occurs through the language comprehension system, as proposed by Levelt (1983).

In sign language, linguistic output is visual and not auditory. Hence, this could imply that visual brain areas could be involved in sign language monitoring, mirroring the involvement of the auditory system in overt speech monitoring. However, the monitoring of sign language production has been proposed to rely more heavily on proprioceptive feedback than visual feedback (Emmorey, Bosworth, et al., 2009). In agreement with this proposal, several studies have reported that parietal regions and not visual regions are more active in sign than in speech production (Emmorey et al., 2007, 2014), including the left supramarginal gyrus and the left superior parietal lobule. Activation of the superior parietal lobule, in particular, has been associated with proprioceptive monitoring during motoric output (Emmorey, Mehta, McCullough, & Grabowski, 2016). Finally, activation of the superior temporal cortex has not been associated with language output monitoring in sign language production, which is likely due to the different modality of output in sign and speech production.

Speech monitoring has also been shown to rely on the activation of medial frontal regions, such as the ACC and the SMA (Christoffels, Formisano, & Schiller, 2007). Activation of the ACC has been shown to be associated with conflict monitoring in and outside language (Piai, Roelofs, Acheson, & Takashima, 2013; Barch, Braver, Sabb, & Noll, 2000; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999).

Therefore, speech monitoring has been proposed to depend not only on the perception of one's own speech through brain mechanisms associated with speech comprehension (see Indefrey, 2011, for a review) but also through the action of a domain-general action monitoring mechanism in medial frontal cortex (Christoffels et al., 2007).

EEG studies of speech monitoring have focused on a component referred to as the error negativity (Ne) or error-related negativity (Zheng et al., 2018; Riès, Fraser, McMahon, & de Zubicaray, 2015; Acheson & Hagoort, 2014; Riès, Xie, Haaland, Dronkers, & Knight, 2013; Riès et al., 2011; Ganushchak & Schiller, 2008a, 2008b; Masaki, Tanaka, Takasawa, & Yamazaki, 2001). This component has a frontocentral distribution (maximal at electrode FCz) and peaks within 100 msec following vocal onset. This component was initially only reported following erroneous utterances and was therefore interpreted as reflecting an error detection mechanism (Masaki et al., 2001). However, this component was more recently also found in correct trials only with a smaller amplitude, suggesting it reflects a monitoring mechanism operating before error detection (Riès et al., 2011). Because of the similar topography, time course, and origin of this component in correct trials and in errors, the component in correct trials has been referred to as the Ne-like wave (Bonini et al., 2014; Roger, Bénar, Vidal, Hasbroucq, & Burle, 2010; Vidal, Burle, Bonnet, Grapperon, & Hasbroucq, 2003). In speech production, this medial frontal monitoring mechanism also starts to be engaged before the onset of verbal responses, suggesting that it reflects the monitoring of inner speech (i.e., the inner loop mentioned above) rather than that of overt speech production (Riès et al., 2011, 2015; Riès, Xie, et al., 2013). Combining a neuropsychological and computational approach, Nozari and colleagues have suggested that accurate speech production relies more heavily on this domain-general monitoring mechanism operating before speech onset than on the speech comprehension-based monitor (Nozari et al., 2011), which would be hosted in the superior temporal cortex. Whether or not a domain-general monitor in medial frontal cortex is also engaged in sign monitoring before signs are produced is unknown.

Several arguments suggest that the medial frontal cortex should be similarly engaged in sign and in spoken language monitoring. One of these arguments is that the representations that this monitoring mechanism operates on are likely to be prearticulatory. Evidence for this proposal comes from the finding that the amplitude of the Ne is modulated by variables that have been tied to stages that precede articulation such as semantic relatedness, lexical frequency, or interference from another language in bilinguals (Riès et al., 2015; Ganushchak & Schiller, 2008a, 2008b, 2009). Such internal representations are likely to be commonly engaged in spoken and sign language production. Another argument is the domain-general nature of the monitoring mechanism

hosted in the medial frontal cortex. Indeed, the Ne and Ne-like waves have been shown to be present in overt speech production, in typing (Pinet & Nozari, 2020; Kalfaoğlu, Stafford, & Milne, 2018), but also in other actions such as in manual button-press tasks (Roger et al., 2010; Burle, Roger, Allain, Vidal, & Hasbroucq, 2008; Vidal et al., 2003; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). The source of the Ne and Ne-like waves has been localized to the medial frontal cortex and, in particular, the ACC (Debener et al., 2005; Dehaene, Posner, & Tucker, 1994) and/or the SMA, as shown through intracranial investigations with depth electrodes inserted in the medial frontal cortex (Bonini et al., 2014). These brain regions are associated with action monitoring generally and are therefore also likely to be engaged in sign language monitoring.

This Study

In this study, we hypothesized that the domain-general monitoring mechanism hosted in the medial frontal cortex and reflected in the Ne and Ne-like wave is similarly engaged during signing and speaking. This study used a picture-naming task and scalp EEG to examine the error (Ne) and error-like (Ne-like) negativities time-locked to the initiation of sign production (as measured through manual key release, as in Emmorey, Petrich, & Gollan, 2013). In particular, we used data gathered during a picture–word interference (PWI) paradigm, which has been shown to elicit more errors than simple picture naming. In the PWI task, used extensively in the field of psycholinguistics, pictures are preceded by or presented with superimposed distractor words (e.g., Bürki, 2017; Roelofs & Piai, 2015, 2017; Piai, Roelofs, Jensen, Schoffelen, & Bonnefond, 2014; Piai, Roelofs, & Schriefers, 2014; Piai et al., 2013; Costa, Alario, & Caramazza, 2005). In the semantic version of the task (used here), the distractor words can be semantically related to the picture (e.g., picture of a dog, distractor word: "cat") or unrelated (e.g., picture of a dog, distractor word: "chair"). Typically in this task, naming the picture takes longer and error rates are higher in the semantically related, compared with the unrelated condition, although the presence of this effect appears to depend on the language input and output modalities (Emmorey, Mott, Meade, Holcomb, & Midgely, under review; Giezen & Emmorey, 2016). Nevertheless, error rates are expected to be higher in this task than in simpler picture naming, which made this paradigm of interest for this study.

We tested both deaf and hearing signers as they named pictures by signing the picture names in American Sign Language (ASL). In addition, we investigated whether or not ASL proficiency, as measured through the ASL Sentence Repetition Task (ASL-SRT), had an effect on the medial frontal monitoring mechanism. Language proficiency has been found to be a potential factor affecting this mechanism, as suggested by several studies (Ganushchak & Schiller, 2009; Sebastian-Gallés, Rodríguez-Fornells, de Diego-Balaguer, & Díaz, 2006). We note, however, that these studies used button-press responses and not overt speech. Therefore, more direct investigations involving overt language production are needed. We had reasons to believe ASL proficiency may be different between the deaf and hearing group because, although the hearing signers participating in these experiments are selected to be highly proficient in ASL, their use of ASL in their everyday lives is typically less frequent than that of deaf signers (see Paludneviciene, Hauser, Daggett, & Kurz, 2012), and hearing signers are also surrounded by spoken English in the environment.

Finding similar components in sign language production would provide strong evidence for the universal nature of inner language monitoring. Indeed, it would suggest that the mechanism reflected by the Ne and Ne-like waves is involved in inner language monitoring irrespective of the language output modality. This would constitute a further argument in support of the idea that the representations monitored by this medial frontal monitoring mechanism are prearticulatory.

METHODS

The data analyzed in this study was initially collected for another study focusing on the effect of the PWI manipulation on ERPs time-locked to stimulus presentation (Emmorey et al., under review). In this study, we focused on the Ne and Ne-like wave time-locked to keyboard release, which reflected the point at which sign production began. There were not enough errors per participant to investigate the effect of the PWI manipulation on the Ne, so we averaged across conditions to increase the number of trials per component.

Participants

A total of 26 deaf signers (15 women, mean age = 34 years, SD = 9 years) and 21 hearing signers (17 women, mean age = 36 years, SD = 10 years) participated in this study. They were recruited through the San Diego area (California) and gave informed consent in accordance with the San Diego State University institutional review board. They received monetary compensation for their time. All had normal or corrected-to-normal vision and no history of neurological impairment. Thirteen deaf participants and five hearing participants were excluded from our analyses because they had less than five error trials remaining after artifact rejection or because they did not follow instructions. Our analyses were therefore conducted on the remaining 11 deaf (eight women, mean age = 35 years, SD = 12 years) and 15 hearing participants (12 women, mean age = 37 years, SD = 12 years). Of these remaining participants, 7 of the 11 deaf participants acquired ASL from birth from their deaf signing families, and 4 acquired ASL in early childhood (before age of 6 years). Of the included 15 hearing

participants, 4 acquired ASL from birth from their deaf signing families and 11 acquired ASL later, at a mean age of 15 years ($SD = 7$ years), 7 were interpreters, and all had been signing for at least 7 years before the experiment (mean = 24 years, $SD = 10$ years). All included participants were right-handed. English proficiency was objectively measured using the Peabody Individual Achievement Test (PIAT) reading comprehension subtest (Markwardt, 1998) and a spelling test from Andrews and Hersch (2010). ASL proficiency was objectively measured using the extended (35 sentence) version of the ASL-SRT (Supalla, Hauser, & Bavelier, 2014). In this task, participants view an ASL sentence and then sign back what was just viewed. Sentence complexity and length increased after each trial. The ASL-SRT task has been shown to differentiate deaf from hearing users of sign language, as well as native from nonnative users (Supalla et al., 2014).

Materials and Design

The stimuli consisted of 200 words representing common nouns and 100 pictures (i.e., line drawings) selected from various sources (Snodgrass & Vanderwart, 1980), presented on a white background. Name agreement in English for the pictures was high (average = 90%, $SD = 14.4$).[1] The average length in letters for the words was 5.05 letters ($SD = 1.87$). The words were presented in capital letters in Arial font (size 60 in height by 30 in width). Fifty of the pictures were presented in an identity condition (e.g., the word "house" followed by a picture of a house), and the other 50 were presented in the semantically related condition (e.g., the word "paper" followed by a picture of scissors). All of the pictures were also presented in an unrelated condition (e.g., the word "ring" followed by a picture of scissors). Therefore, each picture appeared twice (once in a related condition and once in an unrelated condition). Lists were constructed so that they only contained each target item once, half (50 pictures) in the unrelated condition, one fourth or 25 pictures in the semantically related condition, and one fourth or 25 pictures in the identity condition. Lists were counterbalanced across participants so that any target item was presented first in the related condition to half of the participants and first in the unrelated condition to the rest of the participants.

Procedure

The stimuli were presented within a $2° \times 3°$ visual angle at the center of an LCD computer screen at a viewing distance of approximately 150 cm from the participant's eyes. This ensured that participants did not have to make large eye movements to fully perceive the stimuli. The participants were seated in a dimly lit, sound-attenuated room and were asked to hold down the spacebar of a keyboard and only lift their hands when they were ready to produce the sign corresponding to the picture, marking

the naming onset. They were asked to name the pictures as quickly and as accurately as possible and ignore the words. They were each given one practice round that consisted of six trials (these stimuli were not used in the experiment). During the practice, they were instructed to blink during the breaks between stimuli and to minimize facial movements while signing to avoid producing artifacts in the EEG recordings.

Each trial of the experiment began with a fixation cross that was presented in the center of the screen. The cross remained on the screen until the participant placed their hands on the spacebar. The word was then presented for 200 msec and was replaced by the picture that was presented for 2000 msec. Participants were asked to produce the sign corresponding to the picture name as quickly and as accurately as possible, without hesitating. After signing the picture name, the participants were asked to place their hands back on the spacebar. The fixation cross replaced the picture after 2000 msec, and the next trial would start only after the participant placed their hands back on the spacebar. Participants were video-recorded during the experiment so that their sign accuracy could be analyzed off-line. The task was self-paced by use of the space bar. Participants were instructed to rest during the fixation periods before placing their hands back on the keyboard. The whole experiment lasted around 20 min with some variability in time depending on how many breaks each participant took.

EEG Recordings

EEG was continually recorded from a 32-channel tin electrode cap (Electro-Cap International, Inc.; using a 10–20 electrode placement). The EEG signal was amplified by a SynAmpsRT amplifier (Neuroscan-Compumedics), and data were collected by Curry Data Acquisition software at a sampling rate of 500 Hz with a band-pass filter of DC to 100 Hz. To monitor for eye blinks and movements, electrodes were placed under the left eye and on the outer canthus of to the right eye. The reference electrode was placed on the left mastoid, and an electrode was placed on the right mastoid for monitoring differential mastoid activity. Impedances were measured before the experiment started and kept below 2.5 kΩ.

Data Processing

Behavioral Data Processing

RTs were defined as the time separating the picture onset from the release of the spacebar to initiate sign production. The accuracy of sign production was determined off-line by visual inspection of the video recordings from the experiment, and all hesitations were discarded from the analysis. Accuracy and hesitation coding was done by two raters, a deaf native signer and a hearing highly proficient ASL signer. Correct trials were those in which an

accurate sign was produced at the time of keyboard release with no hesitations. Error trials were trials in which the participant produced an off-target sign (e.g., LION instead of TIGER). Trials in which the participant produced an UM sign or where there was a perceptible pause between the keyboard liftoff and the initiation of the sign were excluded from analysis (see Emmorey, Petrich, & Gollan, 2012). Trials in which the participant did not respond were also excluded from the behavioral and EEG analyses.

EEG Processing

After acquisition, vertical eye movements (i.e., eye blinks) were removed using independent component analysis as implemented in EEGLAB (Delorme & Makeig, 2004). Additional artifacts caused by EMG activity associated with facial movements were reduced using a blind source separation algorithm based on canonical correlation analysis (De Clercq, Vergult, Vanrumste, Van Paesschen, & Van Huffel, 2006), previously adapted to speech production (De Vos et al., 2010), and used successfully in previous studies investigating speech monitoring processes (Riès et al., 2011, 2015; Riès, Xie, et al., 2013). Finally, any remaining artifacts were removed through manual inspection in Brain Vision Analyzer (Brain Products). Laplacian transformation (i.e., current source density, estimation), as implemented in Brain Vision Analyzer, was applied to each participant's averages and on the grand averages (as in Riès et al., 2011, 2015; Riès, Janssen, Burle, & Alario, 2013; Riès, Xie, et al., 2013; degree of spline = 3, Legendre polynomial = 15° maximum). We assumed a radius of 10 cm for the sphere representing the head. The resulting unit was $\mu V/cm^2$. Grand averages were created for correct and incorrect trials in both the deaf and hearing groups for the participants with more than five error trials remaining after artifact rejection.

Data Analysis

Proficiency scores for English and ASL as measured by the above-listed tests were compared between groups using two-tailed Student t tests.

Behavioral data were analyzed using linear (for RTs) and generalized mixed-effects models (for accuracy rates). We tested for main effects of Accuracy and Group (deaf vs. hearing) and the interaction between Accuracy × Group on RTs and tested for a main effect of Group on accuracy rates and controlled for random effects of subjects and items. p Values were obtained using Type II analyses of deviance tables providing the Wald χ^2 tests for the fixed effects in the mixed-effects models. For all models, we report the Wald χ^2 values and p values from the analyses of deviance tables as well as raw β estimates (β_{raw}), 95% confidence intervals around these β estimates (CI), standard errors (SE), t values for RTs, and Wald Z and associated p values for accuracy rates.

EEG data were analyzed using two types of EEG measures following methods described in previous studies (Riès, Janssen, et al., 2013; Riès, Xie, et al., 2013; Riès

et al., 2011). The first measure was the slope of the waveforms on a 150-msec time window preceding the key release (the onset of sign production). To find the slope, a linear regression was fitted to the data, and then nonparametric exact Wilcoxon signed-rank tests were used to compare the slopes to 0 for both the errors and correct trials in the deaf and hearing group as the number of error trials was low and the data could not be assumed to be normally distributed. The second measure was peak-to-peak amplitude (i.e., the difference between the amplitude of two consecutive peaks of activity). Peak-to-peak amplitude was calculated by first determining the peak latencies of the Ne and Ne-like wave as well as the preceding positive peak (also referred to as the start of the rise of the negativity) on the by-participant averages. Latencies were measured on smoothed data to minimize the impact of background noise (the length of the smoothing window was 40 msec) and within 100-msec time windows centered around the latency of the peak on grand averages. Then, for each participant, the surface area was calculated between the waveform and the baseline on a 50-msec time window centered around each peak latency as measured in each participant's average waveform. Finally, the difference between the surface areas measured around the Ne or Ne-like wave and around the preceding positivity was considered the peak-to-peak amplitude and is hence independent from the baseline. Again, nonparametric exact Wilcoxon signed-rank one-sided tests (Wilcoxon t tests) were used to compare peak-to-peak amplitudes in errors versus correct trials because the measures were based on few error trials and the normality of the data could not be assumed (as in Riès, Xie, et al., 2013; Riès et al., 2011). The use of one-sided tests was justified as the direction of the difference was expected based on several preceding studies in language (Riès et al., 2011, 2015; Riès, Xie, et al., 2013) and outside language (Vidal et al., 2000, 2003). For each test, we report the W statistic for Wilcoxon signed-rank tests, general Z statistic, associated p value, and effect size r value. In addition, the effect of Group (deaf vs. hearing) on these differences between correct and error trials were tested using an ANOVA.

Finally, we tested for a correlation between the slope of the Ne and Ne-like wave, and the ASL proficiency score as measured with the ASL-SRT, using Spearman's rank correlation coefficient ρ. We report the rho correlation coefficients, S, and associated p values.

All statistical analyses were conducted using R (R Core Team, 2014).

RESULTS

Behavioral Results

Language Proficiency

English. Raw PIAT scores for the deaf group ranged from 42 to 99 ($M = 79$, $SD = 13$), and spelling scores ranged from 62 to 80 ($M = 73$, $SD = 6$). PIAT scores from
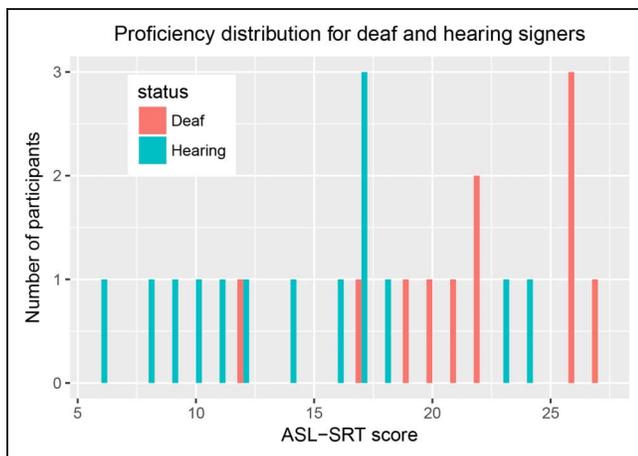
Figure 1. ASL proficiency scores as measured through the ASL-SRT in deaf (blue) and hearing signers (orange).

the hearing group ranged from 67 to 99 ($M = 92$, $SD = 7$), and spelling test scores ranged from 62 to 85 ($M = 78$, $SD = 5$). There was a marginal difference between groups on the spelling test, $t(21.06) = -2.0352$, $p = .055$, and a significant difference between groups on the PIAT, $t(14.56) = -3.016$, $p = .0089$: The hearing participants showed higher performance on these tests of English proficiency than the deaf participants.

ASL. ASL-SRT scores ranged from 12 to 27 ($M = 22$, $SD = 5$) for deaf signers and ranged from 6 to 24 ($M = 14$, $SD = 5$) for hearing signers. An ASL-SRT score was not collected from one of the hearing participants because of her familiarity with the test. There was a significant difference in ASL-SRT scores between the deaf and the hearing signers, $t(22.82) = 3.60$, $p = .0015$: The deaf signers had higher ASL-SRT scores than the hearing signers. Figure 1 provides an illustration of the distribution of the ASL-SRT scores.

RTs and Error Rates

The average RT was 742 msec ($\sigma = 270$ msec) for correct trials and 848 msec ($\sigma = 376$ msec) for errors. There was

a significant main effect of Accuracy on RTs (Wald $\chi^2 = 21.08$, $p < .001$), but no effect of Group (Wald $\chi^2 = 0.71$, $p = .398$), and no interaction between Group × Accuracy (Wald $\chi^2 = 0.40$, $p = .530$). RTs were shorter in correct than in incorrect trials ($\beta_{raw} = -120.20$, CI $[-191.81, -48.59]$, $SE = 36.54$, $t = -3.29$). The median error rate was 5.6% (IQR = 4.6–12.4%), and there was no effect of Group on accuracy rates (Wald $\chi^2 = 0.91$, $p = .341$; $\beta_{raw} = -0.29$, CI $[-0.88, 0.30]$, $SE = 0.30$, $Z = 0.34$, $p = .341$). Mean RT, median error rates, and number of errors are reported in Table 1 (see Tables S1 and S2 for full fixed-effect results from the mixed-effects models[2]). On average, 76% ($\sigma = 17$%) of correct trials and 74% ($\sigma = 20$%) of errors were left after artifact rejection.

EEG Results

Deaf Signers

We observed negativities for correct and incorrect trials. The negativity began to rise on average 209 msec ($SD = 64$ msec) before keyboard release in correct trials and -212 msec ($SD = 60$ msec) in errors. There was no statistical difference between the latency of the positive peak in errors and correct trials, $t(10) < 1$; $W = 30.5$, $Z = -.22$, $p = .852$, $r = .04$. The negativity reached its maximum on average 13 msec ($SD = 31$ msec) before keyboard release in correct trials and 32 msec ($SD = 59$ msec) after keyboard release in errors. The negativity peaked significantly later in errors than correct trials, $t(10) = 3.10$, $p = .011$; $W = 63$, $Z = -2.67$, $p = .005$, $r = .56$. The negativities reached their maximum at frontocentral electrodes, just posterior to Fz and anterior to Cz (see Figure 2; the used montage did not include electrode FCz, which is typically used to study the Ne and Ne-like waves). Slopes measured from -150 to 0 msec were significantly different from 0 in correct trials, $t(10) = -2.31$, $p = .022$; $W = 7$, $Z = -2.60$, $p = .009$, $r = .55$, and incorrect trials, $t(10) = -2.52$, $p = .015$; $W = 7$, $Z = -2.60$, $p = .009$, $r = .55$. The amplitude of the negativity for incorrect trials was significantly larger than for correct trials, $t(10) = 4.03$, $p = .001$; $W = 66$, $Z = -3.49$, $p < .001$, $r = .74$.

Table 1. Mean RTs per Group (Deaf vs. Hearing) for Errors and Correct Trials with Standard Deviations around the Mean (σ), Median Error Rates, Range, and Median (M) Number of Errors per Group with Interquartile Ranges (IQR)

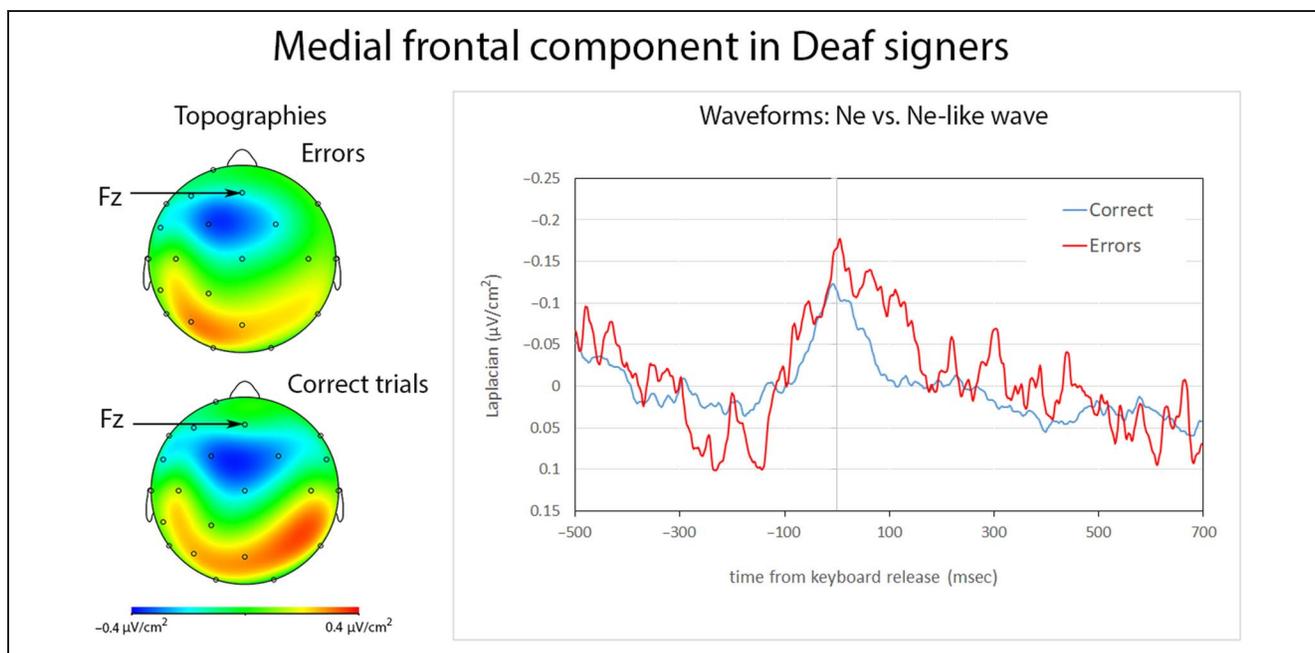|	Deaf		Hearing	
	Correct	*Errors*	*Correct*	*Errors*
Mean RTs	688 msec ($\sigma = 242$ msec)	796 msec ($\sigma = 394$ msec)	781 msec ($\sigma = 290$ msec)	886 msec ($\sigma = 372$ msec)
Median error rates	5.08% (IQR = 4.12–11.81%)		6.70% (IQR = 4.94–11.52%)	
Range and median number of errors	Range = 5–26, $M = 8$, IQR = 7–10		Range = 5–33, $M = 11$, IQR = 9–13	

Figure 2. Ne and Ne-like wave in deaf participants. On the left, topography of Ne at the peak latency. On the right, waveforms for the Ne (in red, errors) and Ne-like waves (in blue, correct trials) time-locked to keyboard release (in msec). The Ne and Ne-like wave reach their maxima around keyboard release time and the Ne is marginally larger than the Ne-like wave.

Hearing Signers

Slopes measured from −150 to 0 msec were significantly different from zero in correct trials, $t(14) = -2.14$, $p = .025$ ($W = 24, Z = -2.06, p = .039, r = .38$). Slopes were not significantly different from zero on the same time window for errors, $t(14) < 1$ ($W = 56, Z = -.53, p = $

.596, $r = .10$). This indicated there was no reliable Ne in errors for the hearing signers at the same recording site as for deaf signers (i.e., Fz). We note, however, that a negativity seemed present at Cz (see Figure 3), although this activity was much smaller than the one reported in deaf participants (see Figure 2; the same scale was used in Figures 2 and 3).
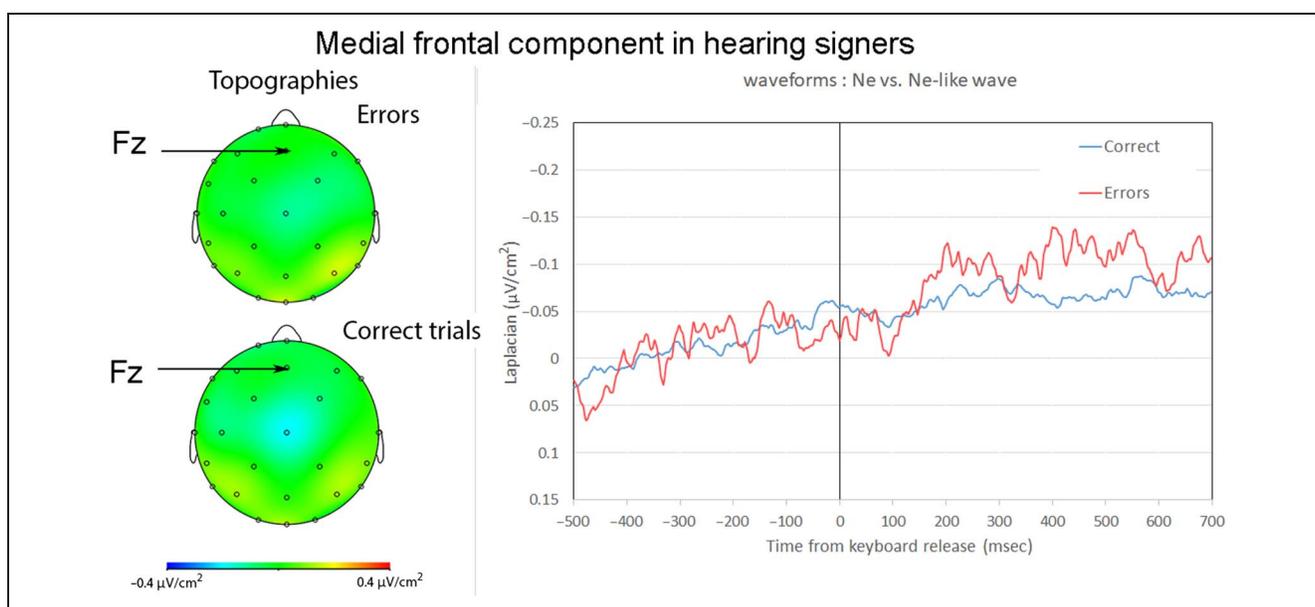

Figure 3. Ne and Ne-like wave in hearing signers. Left: Topography around the peak of the negativity in errors and correct trials. Right: Waveforms for the Ne (in red, errors) and Ne-like waves (in blue, correct trials) time-locked to the keyboard release (in msec). For ease of comparison, the same scales were used as for deaf signers (Figure 2).
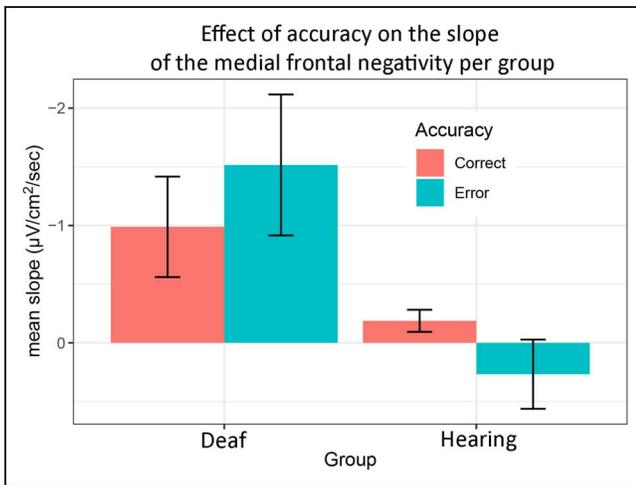
Figure 4. Histogram plotting the mean slope of the Ne and Ne-like wave (with standard deviations from the mean) in deaf and hearing signers.

Deaf versus Hearing Signers and Effect of ASL Proficiency

We tested for effects of Accuracy, Group, and ASL-SRT score on the slope of the medial frontal component (the data of the participant for whom we did not have an ASL-SRT score was excluded). We found a significant effect of Group, $F(1, 21) = 7.14$, $p = .014$, and an interaction between Group × Accuracy, $F(1, 21) = 4.35$, $p = .050$. There was no significant main effect of Accuracy, $F(1, 21) < 1$; ASL-SRT score, $F(1, 21) < 1$; nor any interaction between Group × ASL-SRT score, $F(1, 21) < 1$, and Accuracy × ASL-SRT score, $F(1, 21) = 1.21$, $p = .284$; and no three-way interaction, $F(1, 21) < 1$.

The significant interaction between Group × Accuracy suggests the difference between the slope of the medial

frontal negativity in errors versus correct trials was larger in the deaf than in the hearing signers (Figure 4). When tested independently, we however did not find a significant difference between the slope in errors versus correct trials in deaf signers ($W = 19$, $Z = -1.55$, $p = .120$, $r = .33$), although we note that the amplitudes were significantly different as reported in the previous section. In the hearing group, there was no indication of a difference between errors and correct trials ($W = 76$, $Z = -.08$, $p = .932$, $r = .02$).

Because the effect of hearing status may have been confounded with ASL proficiency, we tested for a correlation between the slope of the Ne and Ne-like wave (as calculated between -150 msec and keyboard release time) and ASL proficiency scores on the ASL-SRT. The slope of the Ne (in errors) was negatively correlated with ASL proficiency scores across deaf and hearing participants ($rho = -0.41$, $S = 3677.3$, $p = .039$), meaning that steeper slopes for the Ne (the steeper the slope, the more negative) were associated with higher ASL proficiency (see Figure 5). No significant correlation was observed between ASL proficiency and the slope of the Ne-like wave (in correct trials, $rho = -0.16$, $S = 3019.3$, $p = .441$).

DISCUSSION

Our results showed that a medial frontal component is present in sign production for both correct responses and errors when deaf signers name pictures in ASL (time-locked to keyboard release). This component is larger in errors than in correct trials. In addition, the slope of the Ne was correlated with proficiency levels in ASL across hearing and deaf signers. The slope was steeper with increasing proficiency, as measured through the ASL-SRT. In hearing signers, this medial frontal

Figure 5. Slopes of the Ne (left) and Ne-like wave (right) versus ASL-SRT scores across participants. Hearing participants are indicated in red and with the letter "H," and deaf participants are indicated in gray with the letter "D."
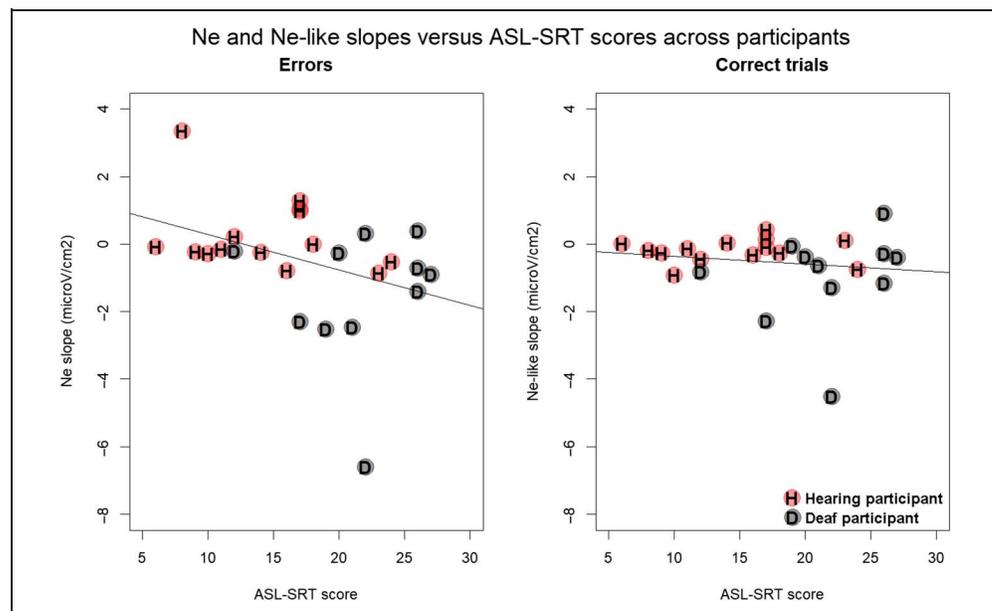
component was not larger in errors than correct trials and was absent in errors at the same recording site as for deaf signers, but was present in correct trials.

Frontomedial Monitoring Mechanism in Sign Language Production

The first important result is that a medial frontal component is present in sign language production in both correct (Ne-like wave) and error trials (Ne). This component had a frontocentral distribution, started to rise before naming onset (as indexed by keyboard release), and peaked just after keyboard release in errors and slightly earlier in correct trials. In addition, it had a larger amplitude in errors than in correct trials. This result suggests that a similar medial frontal mechanism is engaged in pre-output language monitoring in both signed and spoken language production. Indeed, a similar medial frontal negativity was previously reported in overt speech production (Zheng et al., 2018; Riès et al., 2011, 2015; Acheson & Hagoort, 2014; Riès, Xie, et al., 2013). Because this activity starts to rise before vocal onset, it was associated with inner speech monitoring. The topography and time course of this component in sign language production is very similar to that observed in overt speech. In overt speech, the Ne was reported to peak between 30 and 40 msec after vocal onset (Riès, Xie, et al., 2013; Riès et al., 2011), which corresponds to what we observed in this study (mean latency of negative peak: 32 msec after keyboard release). In addition, the preceding positive peak was found to precede the onset of language output (speech and sign) across studies even though there were some differences in the latency of this preceding positive peak: 166 msec (SD = 80 msec) prevocal onset in Riès, Xie, et al. (2013), 46 msec (SD = 52 msec) prevocal onset in Riès et al. (2011), and 212 msec (SD = 60 msec) before keyboard release in this study. One possible explanation for these differences in latencies for the positive peak could be that the task used here was more complex as it required participants to ignore the distractor word and to release the spacebar to start signing. The relatively small difference in amplitude between the Ne and Ne-like wave in this study would be in agreement with this interpretation. Indeed, the difference in amplitude between the Ne and Ne-like wave has been shown to be smaller with increasing task difficulty (leading to higher error rates; e.g., due to decreased response certainty or in time pressure situations; Ganushchak & Schiller, 2006, 2009; Sebastian-Gallés et al., 2006). We note that the RTs were on average longer (742 msec, SD = 270 msec, for correct trials) and that the error rates were higher (median = 5.6%, IQR = 4.6–12.4%) in this study compared with Experiment 2 of Riès et al. (2011), which used simple overt picture naming (mean RT for correct trials = 651 msec, SD = 72 msec; mean error rate = 1.31%, SD = 0.96%). In addition, in this study, the requirement to release the spacebar before signing constitutes an additional step in response

programming that may have caused the increased delay between the latency of the peak of the Ne and that of the preceding negativity. However, the similarity in the latency of the negative peak and the fact that it starts to rise before articulation onset, as well as the similar topographies associated with this component in speech and in sign language production, suggest that the medial frontal component we report here is similar to the medial frontal component reported in overt speech (e.g., Riès, Xie, et al., 2013; Riès et al., 2011). This suggests that this medial frontal component is involved in the inner loop of language output monitoring irrespective of the output modality, which would be in line with the idea that the representations that are monitored by this mechanism are prearticulatory.

In addition to specifying the role of the medial frontal monitoring mechanism in language production, our results also shed light on sign language monitoring more specifically. Indeed, based on the perceptual loop theory of self-monitoring (Levelt, 1983, 1989), previous studies had investigated the role of visual feedback in sign language monitoring. Emmorey, Gertsberg, et al. (2009) found that preventing visual feedback with a blindfold had little impact on sign production (i.e., there is no Lombard effect for sign language). Emmorey, Bosworth, et al. (2009) showed that blurring or completely masking visual feedback did not alter how well novel signs were learned, suggesting that signers do not rely on visual feedback to fine-tune articulation during learning. In fact, production performance of hearing nonsigners was slightly worse with than without visual feedback. This led the authors to suggest that sign language monitoring may rely more heavily on proprioceptive feedback than on visual feedback (see also Emmorey, Korpics, et al., 2009). What the present results suggest is that a medial frontal monitoring mechanism may also be involved in sign language monitoring and that this monitoring mechanism is engaged before proprioceptive feedback is available (i.e., before beginning to move the hand(s) to sign). Evidence for this claim comes from the time course of the Ne and Ne-like waves, which start to rise before sign production onset (i.e., before key release). In addition, Allain, Hasbroucq, Burle, Grapperon, and Vidal (2004) reported Ne and Ne-like waves in a completely deafferented patient. This rare clinical case was tested using a two-response RT task and a go/no-go task and showed the expected Ne and Ne-like wave patterns in both tasks. Our results therefore indicate yet another similarity in the processing of sign and speech production and imply that current models of speech monitoring should be adapted to sign language production (Nozari et al., 2011; Postma & Oomen, 2005; Levelt, 1983).

Difference between Deaf and Hearing Signers

At the group level, no clear Ne was visible in the hearing signers at the same recording site as for deaf signers. Although the hearing signers were highly proficient in

ASL (many worked as interpreters), their scores on the ASL-SRT were significantly lower than for the deaf signers we tested. Therefore, we tested for an effect of proficiency on the slope of the Ne and found that the slope of the Ne was negatively correlated with ASL proficiency scores across deaf and hearing signers, meaning that the higher the ASL-SRT score, the more negative-going the slope of the Ne was. Nevertheless, we note that there was no significant effect of ASL proficiency when tested along with the effect of group. It therefore appeared that group and ASL proficiency were confounded in our study. Consequently, more detailed examinations of the possible effect of ASL proficiency on the medial frontal monitoring mechanism are needed in future studies.

Previous studies investigating the Ne and/or Ne-like wave in overt speech monitoring have been performed in bilinguals (Acheson, Ganushchak, Christoffels, & Hagoort, 2012; Ganushchak & Schiller, 2009; Sebastian-Gallés et al., 2006). In particular, Ganushchak and Schiller (2009) compared German–Dutch bilinguals to Dutch monolingual participants as they performed a phoneme monitoring "go/no-go" task (i.e., they were asked to press a button if the Dutch name of the presented picture contained a specific phoneme) under time pressure versus not. In the time pressure condition, the stimuli were presented for a shorter duration than in the control condition, and this duration was adapted on an individual basis. The results showed differential effects of time pressure on the amplitude of the Ne (referred to as the error-related negativity in their study) as a function of group. German–Dutch bilinguals, who were performing the task in their nonnative language, showed a larger Ne in the time pressure condition than in the control condition, and Dutch monolingual speakers showed the reverse effect. Importantly, the bilingual individuals tested in this study were not balanced bilinguals and had lower proficiency in the language in which they were tested (i.e., Dutch) than in their native language (i.e., German). Although the task we used was very different from the task used in Ganushchak and Schiller (2009), their results suggest that the Ne may be sensitive to language proficiency. Interestingly and similarly to our results, the mean amplitude of the Ne in the control condition, which is more comparable to the setup of our study, appeared to be lower in the German–Dutch bilinguals than in the Dutch monolinguals, although this difference was not directly tested.

Relatedly, Sebastian-Gallés et al. (2006) compared Spanish–Catalan bilinguals who were Spanish versus Catalan dominant in a lexical decision task in Catalan. For Catalan-dominant bilinguals, they observed the expected pattern of a larger negativity in errors than correct trials. However, for Spanish-dominant bilinguals, the amplitude of the negativity was not larger in errors than in correct trials. These results suggest language dominance is an important variable influencing inner speech monitoring abilities. However, we did not test for an effect of language dominance independently from language

proficiency. Typically, English is considered the dominant language for hearing signers because English is the language of schooling and the surrounding community, whereas ASL is considered the dominant language for deaf signers (for discussion, see Emmorey, Giezen, & Gollan, 2016). Interestingly, Sebastian-Gallés et al. (2006) reported a negativity in correct trials (our Ne-like wave) in both groups of bilinguals, which was larger when lexical decision was more difficult (i.e., for nonword trials vs. word trials). This finding is in line with our results as we also found a significant Ne-like wave in the hearing signers, even though the Ne was not statistically reliable at the group level.

Previous reports have also shown modulations of the Ne-like wave outside language as a function of response uncertainty (Pailing & Segalowitz, 2004) and the accuracy of the following trial (Allain, Carbonnell, Falkenstein, Burle, & Vidal, 2004). In particular, the amplitude of the Ne-like wave has been shown to increase with response uncertainty, whereas the amplitude of the Ne has been shown to decrease with response uncertainty (Pailing & Segalowitz, 2004). Hence, one possible interpretation of our results could be that the hearing signers experienced greater response uncertainty compared with deaf signers. This hypothesis would also be in line with the proposal that hearing signers are less aware of their sign errors compared with deaf signers, as suggested by Nicodemus and Emmorey (2015). Another (possibly related) reason for the lack of an Ne in hearing signers could be linked to a time alignment issue with the event used to mark sign production onset, namely the keyboard release. Even though we carefully rejected all trials containing a perceptible pause between the keyboard release time and the onset of the sign (e.g., when the dominant hand reached the target location of the sign; see Caselli, Sehyr, Cohen-Goldberg, and Emmorey (2017) for a detailed description of how sign onsets are determined), naming onset may not have been as closely aligned to keyboard release time in hearing signers as compared with deaf signers. That is, hearing signers may have been more likely to prematurely release the spacebar before they had completely encoded the sign for articulation. This could explain why a later, though not strongly reliable, negativity was observed in the subgroup of proficient hearing signers. Indeed, for these signers, the sign onset itself, occurring after keyboard release, might be a better marker to use for the observation of an Ne. Future studies are needed to clarify this issue.

Conclusion

In summary, our study reports for the first time the presence of a medial frontal negativity associated with inner language output monitoring in sign language production. The presence of this negativity in sign language production strongly suggests that a similar medial frontal mechanism is involved in language monitoring before response initiation irrespective of language output modality and

suggests the representations that are monitored by this mechanism are prearticulatory. In addition, in line with previous studies using phonological monitoring and lexical decision tasks, our results showed that this mechanism was modulated by language proficiency in sign language production, suggesting similar factors affect medial frontal language output monitoring across modalities.

Acknowledgments

Notes

1. ASL name agreement was available for 61 of the stimuli (from an ongoing ASL picture naming study in the Emmorey Lab), and agreement was also high for these stimuli (average = 83.0%, SD = 21.6%).
2. Supplementary material for this paper can be retrieved from https://lbdl.sdsu.edu/wp-content/uploads/2020/02/Supplementary_tables.pdf.

REFERENCES

Acheson, D. J., Ganushchak, L. Y., Christoffels, I. K., & Hagoort, P. (2012). Conflict monitoring in speech production: Physiological evidence from bilingual picture naming. *Brain and Language*, *123*, 131–136.

Acheson, D. J., & Hagoort, P. (2014). Twisting tongues to test for conflict-monitoring in speech production. *Frontiers in Human Neuroscience*, *8*, 206.

Allain, S., Carbonnell, L., Falkenstein, M., Burle, B., & Vidal, F. (2004). The modulation of the Ne-like wave on correct responses foreshadows errors. *Neuroscience Letters*, *372*, 161–166.

Allain, S., Hasbroucq, T., Burle, B., Grapperon, J., & Vidal, F. (2004). Response monitoring without sensory feedback. *Clinical Neurophysiology*, *115*, 2014–2020.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*, 299–318.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoriing of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience*, *12*, 298–309.

Blanco-Elorrieta, E., Kastner, I., Emmorey, K., & Pylkkänen, L. (2018). Shared neural correlates for building phrases in signed and spoken language. *Scientific Reports*, *8*, 5492.

Bonini, F., Burle, B., Liégeois-Chauvel, C., Régis, J., Chauvel, P., & Vidal, F. (2014). Action monitoring and medial frontal cortex: Leading role of supplementary motor area. *Science*, *343*, 888–891.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–181.

Braun, A. R., Guillemin, A., Hosey, L., & Varga, M. (2001). The neural organization of discourse: An H_2 ^{15}O-PET study of narrative production in English and American sign language. *Brain*, *124*, 2028–2044.

Bürki, A. (2017). Electrophysiological characterization of facilitation and interference in the picture–word interference paradigm. *Psychophysiology*, *54*, 1370–1392.

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, *20*, 1637–1655.

Caselli, N. K., Sehyr, Z. S., Cohen-Goldberg, A. M., & Emmorey, K. (2017). ASL-LEX: A lexical database of American sign language. *Behavior Research Methods*, *49*, 784–801.

Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human Brain Mapping*, *28*, 868–879.

Costa, A., Alario, F.-X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture–word interference paradigm. *Psychonomic Bulletin & Review*, *12*, 125–131.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, *25*, 11730–11737.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., & Van Huffel, S. (2006). Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *IEEE Transactions on Biomedical Engineering*, *53*, 2583–2587.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*, 303–305.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.

De Vos, M., Riès, S., Vanderperren, K., Vanrumste, B., Alario, F.-X., Van Huffel, S., et al. (2010). Removal of muscle artifacts from EEG recordings of spoken language production. *Neuroinformatics*, *8*, 135–150.

Emmorey, K., Bosworth, R., & Kraljic, T. (2009). Visual feedback and self-monitoring of sign language. *Journal of Memory and Language*, *61*, 398–411.

Emmorey, K., Gertsberg, N., Korpics, F., & Wright, C. E. (2009). The influence of visual feedback and register changes on sign language production: A kinematic study with deaf signers. *Applied Psycholinguistics*, *30*, 187–203.

Emmorey, K., Giezen, M. R., & Gollan, T. H. (2016). Psycholinguistic, cognitive, and neural implications of bimodal bilingualism. *Bilingualism: Language and Cognition*, *19*, 223–242.

Emmorey, K., Korpics, F., & Petronio, K. (2009). The use of visual feedback during signing: Evidence from signers with impaired vision. *Journal of Deaf Studies and Deaf Education*, *14*, 99–104.

Emmorey, K., McCullough, S., Mehta, S., & Grabowski, T. J. (2014). How sensory-motor systems impact the neural organization for language: Direct contrasts between spoken and signed language. *Frontiers in Psychology*, *5*, 484.

Emmorey, K., Mehta, S., & Grabowski, T. J. (2007). The neural correlates of sign versus word production. *Neuroimage*, *36*, 202–208.

Emmorey, K., Mehta, S., McCullough, S., & Grabowski, T. J. (2016). The neural circuits recruited for the production of signs and fingerspelled words. *Brain and Language*, *160*, 30–41.

Emmorey, K., Mott, M., Meade, G., Holcomb, P. J., & Midgely, K. J. (under review). Bilingual lexical selection without lexical competition: ERP evidence from bimodal bilinguals and picture–word interference. *Language, Cognition and Neuroscience*.

Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2012). Bilingual processing of ASL–English code-blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language*, *67*, 199–210.

Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2013). Bimodal bilingualism and the frequency-lag hypothesis. *Journal of Deaf Studies and Deaf Education*, *18*, 1–11.

Fu, C. H. Y., Vythelingum, G. N., Brammer, M. J., Williams, S. C. R., Amaro, E., Jr., Andrew, C. M., et al. (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, *16*, 969–977.

Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: An ERP study. *Brain Research*, *1125*, 104–115.

Ganushchak, L. Y., & Schiller, N. O. (2008a). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *Neuroimage*, *39*, 395–405.

Ganushchak, L. Y., & Schiller, N. O. (2008b). Brain error-monitoring activity is affected by semantic relatedness: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, *20*, 927–940.

Ganushchak, L. Y., & Schiller, N. O. (2009). Speaking one's second language under time pressure: An ERP study on verbal self-monitoring in German–Dutch bilinguals. *Psychophysiology*, *46*, 410–419.

Giezen, M. R., & Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: Evidence from picture–word interference. *Bilingualism: Language and Cognition*, *19*, 264–276.

Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping*, *20*, 22–28.

Hohenberger, A., Happ, D., & Leuninger, H. (2002). Modality-dependent aspects of sign language production: Evidence from slips of the hands and their repairs in German Sign Language. In R. P. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 112–142). Cambridge, UK: Cambridge University Press.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 255.

Kalfaoğlu, Ç., Stafford, T., & Milne, E. (2018). Frontal theta band oscillations predict error correction and posterror slowing in typing. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 69–88.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Science*, *3*, 223–232.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Markwardt, F. C. (1998). *Peabody Individual Achievement Test–Revised* (Normative update). Circle Pines, MN: American Guidance Service.

Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *NeuroReport*, *12*, 1851–1855.

McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*, 907–917.

Nicodemus, B., & Emmorey, K. (2015). Directionality in ASL–English interpreting: Accuracy and articulation quality in L1 and L2. *Interpreting*, *17*, 145–166.

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, *63*, 1–33.

Pailing, P. E., & Segalowitz, S. J. (2004). The effects of uncertainty in error monitoring on associated ERPs. *Brain and Cognition*, *56*, 215–233.

Paludneviciene, R., Hauser, P. C., Daggett, D., & Kurz, K. B. (2012). Issues and trends in sign language assessment. In D. A. Morere & T. E. Allen (Eds.), *Assessing literacy in deaf individuals: Neurocognitive measurement and predictors* (pp. 191–207). New York: Springer.

Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: Domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, *7*, 832.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.-M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLoS One*, *9*, e88674.

Piai, V., Roelofs, A., & Schriefers, H. (2014). Locus of semantic interference in picture naming: Evidence from dual-task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 147–165.

Pinet, S., & Nozari, N. (2020). Electrophysiological correlates of monitoring in typing with and without visual feedback. *Journal of Cognitive Neuroscience*, *32*, 603–620.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–132.

Postma, A., & Oomen, C. E. (2005). Critical issues in speech monitoring. In R. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 157–166). Hove, UK: Psychology Press.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from www.R-project.org/.

Riès, S. K., Fraser, D., McMahon, K. L., & de Zubicaray, G. I. (2015). Early and late electrophysiological effects of distractor frequency in picture naming: Reconciling input and outpu accounts. *Journal of Cognitive Neuroscience*, *27*, 1936–1947.

Riès, S. K., Janssen, N., Burle, B., & Alario, F.-X. (2013). Response-locked brain dynamics of word production. *PLoS One*, *8*, e58197.

Riès, S. K., Janssen, N., Dufau, S., Alario, F.-X., & Burle, B. (2011). General-purpose monitoring during speech production. *Journal of Cognitive Neuroscience*, *23*, 1419–1436.

Riès, S. K., Xie, K., Haaland, K. Y., Dronkers, N. F., & Knight, R. T. (2013). Role of the lateral prefrontal cortex in speech monitoring. *Frontiers in Human Neuroscience*, *7*, 703.

Roelofs, A., & Piai, V. (2015). Aspects of competition in word production: Reply to Mahon and Navarrete. *Cortex*, *64*, 420–424.

Roelofs, A., & Piai, V. (2017). Distributional analysis of semantic interference in picture naming. *Quarterly Journal of Experimental Psychology*, *70*, 782–792.

Roger, C., Bénar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral cingulate zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage*, *51*, 391–403.

Sebastian-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience*, *18*, 1277–1291.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.

Supalla, T., Hauser, P. C., & Bavelier, D. (2014). Reproducing American sign language sentences: Cognitive scaffolding in working memory. *Frontiers in Psychology*, *5*, 859.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*, 1429–1443.

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, *423*, 866–869.

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: A reexamination of available data. *Biological Psychology*, *64*, 265–282.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the "error negativity" specific to errors? *Biological Psychology*, *51*, 109–128.

Zheng, X., Roelofs, A., Farquhar, J., & Lemhöfer, K. (2018). Monitoring of language selection errors in switching: Not all about conflict. *PLoS One*, *13*, e0200397.