Received: 19 November 2020 | Revised: 9 January 2021 | Accepted: 2 February 2021

DOI: 10.1002/mrm.28746

NOTE

Magnetic Resonance in Medicine

Aliasing artifact reduction in spiral real-time MRI

Ye Tian[1] | **Yongwan Lim**[1] | **Ziwei Zhao**[1] | **Dani Byrd**[2] | **Shrikanth Narayanan**[1,2] | **Krishna S. Nayak**[1]

[1]Ming Hsieh Department of Electrical and Computer Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, California, USA

[2]Department of Linguistics, Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, California, USA

Correspondence
Ye Tian, Ming Hsieh Department of Electrical and Computer Engineering, Viterbi School of Engineering, University of Southern California, 3740 McClintock Ave, EEB 412, Los Angeles, CA 90089-2564, USA.
Email: ytian607@usc.edu

Funding information
National Science Foundation (1514544) and National Institutes of Health (R01DC007124 and R01HL-130494)

Purpose: To mitigate a common artifact in spiral real-time MRI, caused by aliasing of signal outside the desired FOV. This artifact frequently occurs in midsagittal speech real-time MRI.

Methods: Simulations were performed to determine the likely origin of the artifact. Two methods to mitigate the artifact are proposed. The first approach, denoted as "large FOV" (LF), keeps an FOV that is large enough to include the artifact signal source during reconstruction. The second approach, denoted as "estimation-subtraction" (ES), estimates the artifact signal source before subtracting a synthetic signal representing that source in multicoil k-space raw data. Twenty-five midsagittal speech-production real-time MRI data sets were used to evaluate both of the proposed methods. Reconstructions without and with corrections were evaluated by two expert readers using a 5-level Likert scale assessing artifact severity. Reconstruction time was also compared.

Results: The origin of the artifact was found to be a combination of gradient nonlinearity and imperfect anti-aliasing in spiral sampling. The LF and ES methods were both able to substantially reduce the artifact, with an averaged qualitative score improvement of 1.25 and 1.35 Likert levels for LF correction and ES correction, respectively. Average reconstruction time without correction, with LF correction, and with ES correction were 160.69 ± 1.56, 526.43 ± 5.17, and 171.47 ± 1.71 ms/frame.

Conclusion: Both proposed methods were able to reduce the spiral aliasing artifacts, with the ES-reduction method being more effective and more time efficient.

KEYWORDS
gradient nonlinearity, low-pass filter, real-time MRI, speech production, spiral trajectory

1 | INTRODUCTION

Real-time MRI is an established tool to study the dynamics of vocal tract shaping during human speech production. Compared with other imaging modalities, such as CT,[1] electromagnetic articulography[2] and ultrasound,[3] MRI[4-9] involves no ionizing radiation, can provide arbitrary imaging planes, and provides excellent tissue-boundary delineation. Real-time MRI (RT-MRI) for speech production is usually designed to achieve high enough temporal resolution

to reveal fast movement of articulators and constriction events. Spiral sampling trajectory is used commonly due to its high sampling efficiency.[10-12] This, combined with advanced reconstruction, has made it possible to perform single-plane,[13] multiplanar,[14] and even 3D[15,16] visualization of speech articulators. Recently, real-time tagging[17,18] has also made it possible to visualize internal tongue mechanics.

One common artifact, aliasing from outside the desired FOV, occurs in all of these approaches that involve spiral sampling in the midsagittal vocal tract plane. The artifact is a result of gradient nonlinearity, spiral sampling, and subject anatomy. Gradient nonlinearity can cause strong image warping and intensity enhancement,[19,20] especially at locations that are farther away from the magnet iso-center. Furthermore, spiral trajectories rotate, and therefore the readout anti-aliasing low-pass filter does not provide FOV restriction.[21,22] These two effects combine to cause a high signal area outside the FOV. This, convolved with the spiral point spread function, can cause aliasing artifact within the FOV.

In a recent data-collection effort leveraging spiral RT-MRI to study interactions of speech articulators and speech-signal properties across talkers using the MRI protocol,[23] 79% of scans suffered from different levels of the spiral-aliasing artifact. These artifacts can overlap with important articulators, negatively affect the visualization of speech production, and cause problems for automatic postprocessing such as boundary detection.[24,25] Previous solutions have included manually adjusting the imaging plane or adjusting the FOV such that the artifact minimally disrupts the targeted articulators under investigation. This manual approach limits the flexibility of the prescription, does not fully resolve the issue, and is time-consuming. A general solution to address the spiral aliasing artifact is needed.

Some of the streak artifacts that are common in radial sampling[26-30] and quantitative susceptibility mapping[31,32] have a similar origin as the spiral aliasing artifact, arising from a high signal. Previous efforts to solve these artifacts belong to two types. One leverages parallel imaging, using either coil weighting[26,27,29,30] or coil sensitivities[28] to suppress the source of the aliasing. These methods can be limited when the coil array has a limited number of channels and can lead to signal reduction. Another type of artifact reduction performs a two-step reconstruction[31,32]: The high signal is first estimated and subtracted, which then allows the low-signal area to be reconstructed without aliasing. Similar to the two-step method, partial FOV reconstruction has also been proposed to improve image quality[33] or to achieve a higher temporal resolution.[34]

In this study, we propose and evaluate two methods to suppress the spiral aliasing artifact, denoted as the "large FOV" (LF) method and the "estimation-subtraction" (ES) method. The LF method keeps a large FOV including the aliasing source during reconstruction. The ES method estimates the aliasing source signal from temporally combined image, synthesizes its k-space signal, and subtracts this signal in the acquired multicoil k-space. The processed k-space is then used for image reconstruction. We evaluate both methods in the context of speech production RT-MRI, complemented with qualitative evaluation by two expert readers. Reconstruction time without and with both corrections is also compared.

2 | METHODS

2.1 | Speech RT-MRI methods

Data acquisition in this work is based on the speech-production RT-MRI system described in Lingala et al[13] and using a protocol described in Lingala et al.[23] Image acquisition was performed on a 1.5T Signa Excite scanner (GE Healthcare, Waukesha, WI), with a custom eight-channel upper airway receive coil array, and a real-time interactive imaging platform RT-Hawk (Heart Vista, Los Altos, CA)[35] to control the scan. A 13-interleaf spiral-out trajectory supporting a 20×20 cm^2 FOV with a bit-reversed temporal sampling order[36] was used to acquire k-space data. General scan parameters for the 2D spiral gradient-echo sequence were: TR/TE = 6.0/0.8 ms, voxel size = $2.4 \times 2.4 \times 6$ mm^3, receiver bandwidth = $\pm$125 kHz, and flip angle = 15°. The system also records audio synchronized with MRI data, and the details can be found in Lingala et al.[13]

A temporally constrained reconstruction[37] was used to reconstruct the dynamic images. The following cost function was solved using a conjugate gradient algorithm with line search:

$$\text{argmin}_m \|FSm - d\|_2^2 + \lambda_t \left\| \sqrt{\left(\nabla_t m\right)^2 + \varepsilon} \right\|_1 \quad (1)$$

where S is the coil sensitivity map; F is the nonuniform fast Fourier transform (NUFFT); m is the dynamic images to be reconstructed; d is the acquired multicoil k-space data; λ_t is the regularization parameter; ∇_t is the temporal finite difference operator; ε is a small positive constant vector to avoid singularity; and $\|\cdot\|_2^2$ and $\|\cdot\|_1$ are l_2 and l_1 norms, respectively. The algorithm stops at 150 iterations or when the line search finds a step size $<10^{-5}$. Sensitivity maps were estimated from temporally combined coil images using the Walsh method.[38] The regularization parameter was chosen to be $0.08C$ based on previous experience,[13] where C is the highest pixel intensity in the adjoint NUFFT-reconstructed images ($F^T d$). Dynamic images were reconstructed with two spiral arms/frame (12 ms/frame, 83.28 frames per second). Images were processed with 1.25 ×

FOV (105 × 105 matrix size) during the iterative reconstruction, and then cropped to the desired FOV (84 × 84 matrix size).

2.2 | Simulation

To demonstrate the origin of the artifact, we performed simulations using a midsagittal slice of the XCAT phantom[39] covering from the head to midchest. We used Biot-Savart law to simulate coil sensitivities of the eight-channel custom receiver coil. A gradient profile of the Zoom coil (GE Healthcare) in the form of spherical harmonic coefficients was used to approximate the gradient nonlinearity.[40] Fully sampled multicoil spiral k-space data were estimated with the 13-interleaf spiral trajectory, supporting a 20-cm circular FOV. We simulated a dwell time of 1 μs, and then subsampled the k-space to a dwell time of 4 μs, which was the dwell time used in in vivo acquisitions. This allowed for simulation of the readout low-pass filter, and its imperfections.

2.3 | Artifact-reduction methods

We propose two methods to correct for the aliasing artifacts, as illustrated in Figure 1. The LF approach keeps a large FOV including the artifact source during the reconstruction. The ES approach estimates the aliasing source, then subtracts a synthetic signal representing that source from multicoil k-space raw data.

The first step in both methods is to identify the aliasing source (M_1). A 2.5× FOV adjoint NUFFT reconstruction is performed on temporally combined k-space data $d(k,c)$ to get coil images $m(x,c)$. The coil images are root sum-of-squares combined, and the resulting image is smoothed by a 2D Gaussian filter to get $m(x)$. The values of k, x, and c represent the k-space location, spatial location, and coil number. The pixel with the highest intensity outside the $\sqrt{2}\times$ circular FOV ($\mathrm{SI}^{\mathrm{out}}_{\mathrm{max}}$) is identified, and its surrounding pixels with intensities $\geq 0.5 \times \mathrm{SI}^{\mathrm{out}}_{\mathrm{max}}$ compose M_1. The mask is then extended two pixels to smooth the boundary and fill holes. A ratio $r_{\mathrm{out/in}} = \mathrm{SI}^{\mathrm{out}}_{\mathrm{max}}/\mathrm{SI}^{\mathrm{in}}_{\mathrm{max}}$ is calculated to help indicate the level of aliasing artifact, where $\mathrm{SI}^{\mathrm{in}}_{\mathrm{max}}$ represents the highest signal intensity within the $\sqrt{2}\times$ circular FOV.

2.3.1 | Large FOV method

Aliasing is caused by signal leakage into undesired locations. If the aliasing source can be faithfully reconstructed, its aliasing can be reduced. Based on this, we propose to reconstruct the image with a large FOV including the aliasing source. However, reconstructing the image directly with a larger FOV (eg, 2.5× FOV) causes blurring artifact, as the trajectory only supports a circular 1× FOV when fully sampled. Supporting Information Figure S1 illustrates the point-spread-function analysis and an in vivo example of the blurring. To solve this, two masks (M_1 and M_2) are applied during the reconstruction to keep only the signal within the masks, where M_2 is a $\sqrt{2}\times$ circular FOV, to maintain signal within a 1× square FOV when cropped. The masks can constrain the signal leakage and remove the blurring, as illustrated in Supporting Information Figure S1.

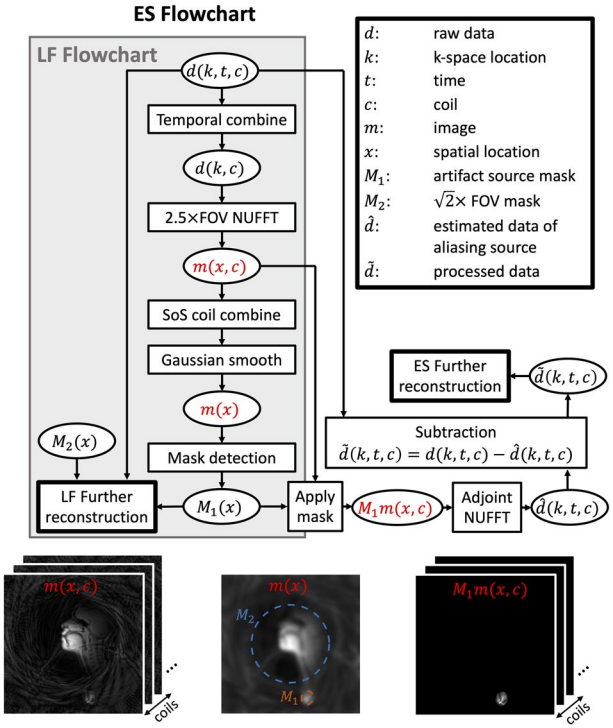

FIGURE 1 Flowchart of the artifact-reduction methods. The raw k-space data were temporally combined and reconstructed to 2.5× FOV to obtain an image for each coil (coil images). Then the coil images were root sum-of-squares (SoS) combined and smoothed with a Gaussian filter. From the smoothed image, a mask (M_1) was automatically detected around the largest pixel intensity outside the $\sqrt{2}\times$ circular FOV by thresholding. The threshold was set to be 0.5× the highest pixel intensity outside the $\sqrt{2}\times$ circular FOV. From there, the large-FOV (LF) reconstruction can be performed by applying two masks, including M_2 ($\sqrt{2}\times$ circular FOV) and M_1, during reconstruction. For the estimation-subtraction (ES) method, M_1 is applied on the coil images, and the masked coil images were used to estimate a k-space for the aliasing source. This estimated k-space was subtracted from the raw k-space for further processing. Abbreviations: NUFFT, nonuniform fast Fourier transform

2.3.2 | Estimation-subtraction method

After the image space-aliasing source $M_1 m(x,c)$ is identified, we estimate the k-space signal of the aliasing source $\hat{d}(k,t,c)$ and subtract it form the acquired k-space $d(k,t,c)$ to

obtain aliasing-free k-space: $\tilde{d}(k,t,c) = d(k,t,c) - \hat{d}(k,t,c)$. The value of $\hat{d}(k,t,c)$ is estimated by applying adjoint NUFFT to $M_1 m(x,c)$. Note that a scale factor was implicitly included in the adjoint NUFFT to match the intensities between $\hat{d}(k,t,c)$ and $d(k,t,c)$. Specifically, the scale factor a was calculated for each data set as $a = min_a \left\| M_1 \sum_t F^T d(k,t,c) - a M_1 \sum_t F^T F \left[m(x,c)_{\times t} \right] \right\|_2^2$ and $\hat{d}(k,t,c) = aF\left[M_1 m(x,c)_{\times t}\right]$, where $\times t$ denotes that the matrix is replicated t times. Further reconstruction is done by replacing k-space data $d(k,t,c)$ in Equation (1) by $\tilde{d}(k,t,c)$.

2.3.3 | Evaluation

We used a data set from a recent speech-study data-collection effort using the protocol[23] to evaluate the aliasing reduction methods. At the time of this study, data for 58 volunteers were prepared for use. Each subject was scanned while producing multiple speech stimuli, and we retrospectively chose the data corresponding to the stimuli provided in the Appendix. The University of Southern California Institutional Review Board approved the study, and consent was obtained from each volunteer. The 58 k-space data sets were reconstructed with the following three methods: no correction, the ES correction, and the LF correction. Initial evaluation of these reconstructions was done by an MRI physicist to identify a threshold of $r_{\text{out/in}}$ for apparent aliasing artifact.

Artifact level was qualitatively evaluated by two speech experts who have 18 years and 28 years of experience in reading speech RT-MRI, respectively. Artifact level was scored using a 5-level Likert scale, in which a score of "1" indicated the least severe artifact, and "5" indicated the most severe artifact. Before undertaking the rating task, two samples selected by an MRI physicist were given to the raters as representative of the two scale endpoints, as defined previously. All videos were evaluated without audio to eliminate potential bias due to audio quality. In the first rating round, 54 data sets *without correction* were rated by both experts (four data sets were excluded by initial evaluation, as explained in section 3). From these results, five data sets were randomly picked from each of the five levels at which the two experts agreed, to balance the rating tasks of the corrected videos. If a certain level did not have enough (≥ 5) data sets, we randomly filled the level with data sets in which at least one expert had given that grade. Three data sets for level 3 were graded 3/2, and two data sets for level 4 were graded 5/4 by expert 1/expert 2. The resulting 25 data sets in three forms (ie, no correction, ES correction, and LF correction) were presented in a second rating task in a randomized order. These 75 videos were scored again by the original raters using the same scale.

The proposed methods were implemented in *MATAB* (2019b; MathWorks, Natick, MA) and executed on a Xeon 2.4-GHz CPU (Intel, Santa Cara, CA) with the iterative reconstruction portion (solving Equation (1)) running on a Tesla P100 GPU (Nvidia, Santa Clara, CA). Reconstruction time without correction, with LF correction, and with ES correction were recorded.

3 | RESULTS

Figure 2 shows the simulation results. Figure 2A illustrates the extracted slice from the XCAT phantom.[39] Figure 2B,C show the root sum-of-squares combined coil-sensitivity map and the phantom image with distortion. Even though the coil has low sensitivity outside the desired FOV, the accumulation of signal due to gradient nonlinearity still produces a strong signal. Figure 2D,E show the limited effectiveness of readout low-pass filter in suppressing the signal outside the FOV, which is consistent with prior reports that document the effects of low-pass filter on spiral sampling.[21,22] Figure 2F shows a representative in vivo acquisition with the spiral aliasing artifact. The simulated artifact closely matches this in vivo artifact, suggesting that the origin of the aliasing is largely due to gradient nonlinearity and spiral sampling. Supporting Information Video S1 shows a video reconstructed with 2.5× FOV and with fully sampled data (13 arms/frame). We observed a pulsatile signal change near the hotspot, which is consistent with cardiac pulsation.

Figure 3 shows a representative example of aliasing reduction results. Two time frames are shown for the three scenarios: without correction, with LF correction, and with ES correction (Figure 3A,B). Time-intensity profiles are shown in Figure 3C. Both correction methods were able to reduce the aliasing artifacts identified by orange arrows, with the LF methods having residual artifacts also identified by orange arrows. Supporting Information Video S2 shows the corresponding video.

The initial evaluation by the MRI physicist found that one data set had no apparent aliasing artifact, although application of the ES and LF methods caused additional artifacts, as shown in Supporting Information Figure S2. This data set had an $r_{\text{out/in}} \leq 0.4$; this threshold also identified three other data sets that had no apparent aliasing, and we excluded these four data sets in the ensuing evaluation.

From the first-round speech-expert evaluation, 46 of 54 datasets were considered to have aliasing artifact by at least one expert (Likert score ≥ 2), resulting in 79% of the acquired data sets (46 of 58) having artifact. Table 1 lists the average scores of the second-round rating. For data sets graded ≥ 2, the average score was improved by 1.1 or 1.4 with LF correction and by 1.4 or 1.3 with ES correction, according to the two raters. The average score for data sets without artifact (score = 1) was 1 for both LF and ES corrections, as scored by rater 1, and was 1.4 for

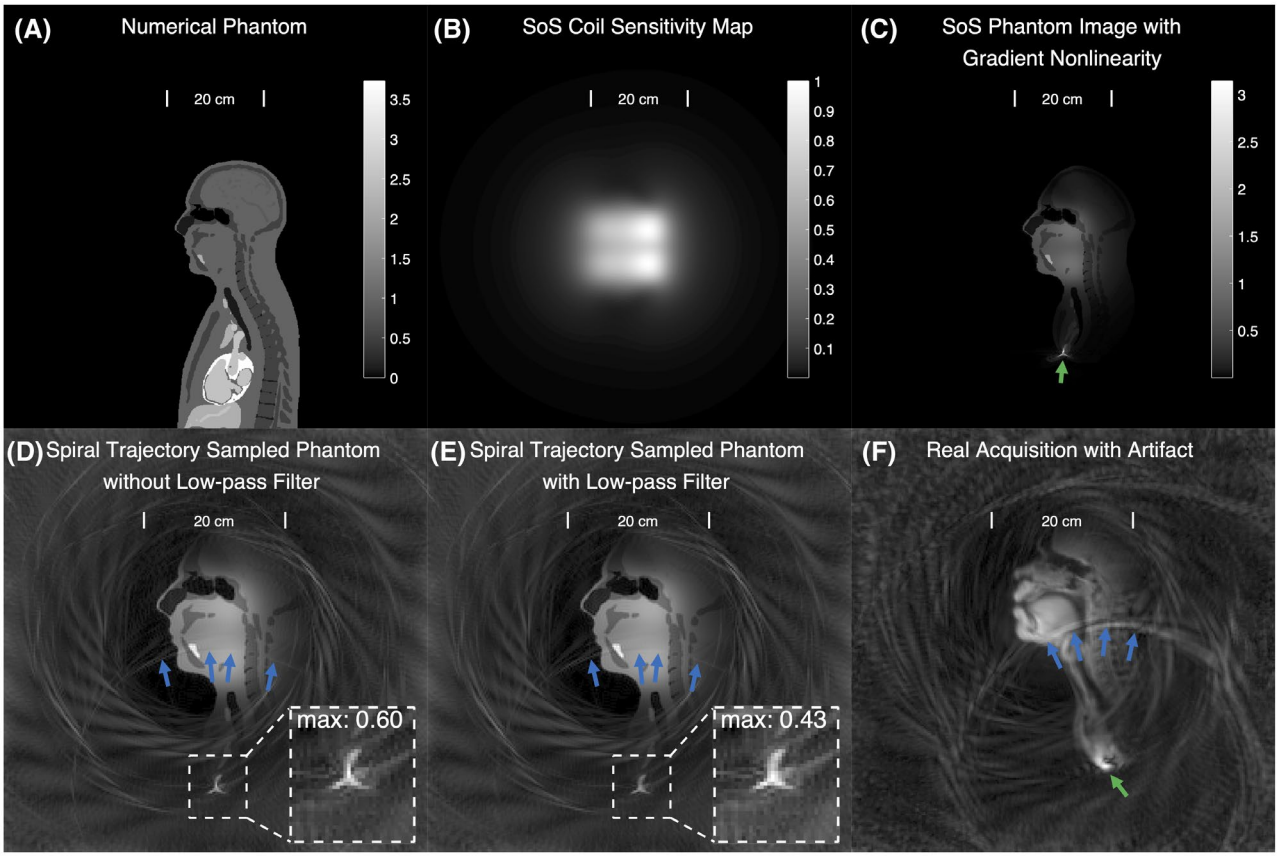

FIGURE 2 Illustration of spiral aliasing artifacts in a numeric phantom and in vivo. A, The numeric phantom is based on XCAT[39] with spoiled gradient echo–recalled contrast, using simulated coil sensitivities (root SoS shown in [B]), and gradient nonlinearity maps (not shown) of the Zoom gradient coil. C, Simulated images with nonlinear gradients illustrate the geometric distortion and the hotspot outside the FOV (green arrow). Simulated spiral reconstructions without (D) and with (E) readout low-pass filter both contain spiral aliasing artifacts appearing within the FOV (blue arrows). This behavior closely matches in vivo data, with a representative example shown in (F)

LF correction and 1.2 for ES correction, as scored by rater 2. Figure 4 shows the second-round evaluation results in Sankey diagrams.

The average reconstruction time was 160.69 ± 1.56, 526.43 ± 5.17, and 171.47 ± 1.71 ms/frame for no correction, LF correction, and ES correction, respectively, as reported in Table 1. Overall, the results suggest that while both methods produce artifact reduction, the ES method is more effective and time efficient than the LF method in suppressing the artifact.

4 | DISCUSSION

We identified the source of a common artifact observed in midsagittal speech RT-MRI with spiral sampling and have demonstrated two effective methods for mitigating this artifact. The source is a spurious signal hotspot caused by gradient nonlinearity and ineffective anti-aliasing filtering during spiral readouts. The two proposed solutions, called LF ("large FOV") and ES ("estimation subtraction"), are both capable of mitigating the issues. The ES method proved to be faster and more effective than the LF method in the context of speech RT-MRI sampled with bit-reversed spiral trajectory.

We used a sagittal slice from the XCAT phantom and the spherical harmonic coefficients of the Zoom coil to demonstrate the spiral aliasing artifact. The spherical harmonic model is an approximation of the actual gradient map, especially at locations that are far away from the magnet isocenter. In practice, we often see the hotspot appears as an oval; however, a triangular hotspot occurred in our simulation. This is likely due to a gradually flattened gradient occurring in practice, but not reproduced by the spherical harmonic model.

We have proposed two methods to reduce the spiral aliasing artifacts for spiral RT-MRI. The LF method requires a larger memory and a longer reconstruction time than the ES method. However, if the spiral trajectory is golden-angle (GA) or pseudo-GA rotated, the supported FOV increases with the number of unique rotations.[41] When the supported FOV is large enough, the LF method can simply be applied without using masks and without introducing blurring. Supporting Information Video S3 shows an example of applying a 2.5× FOV reconstruction on a 34-interleaf pseudo-GA-sampled

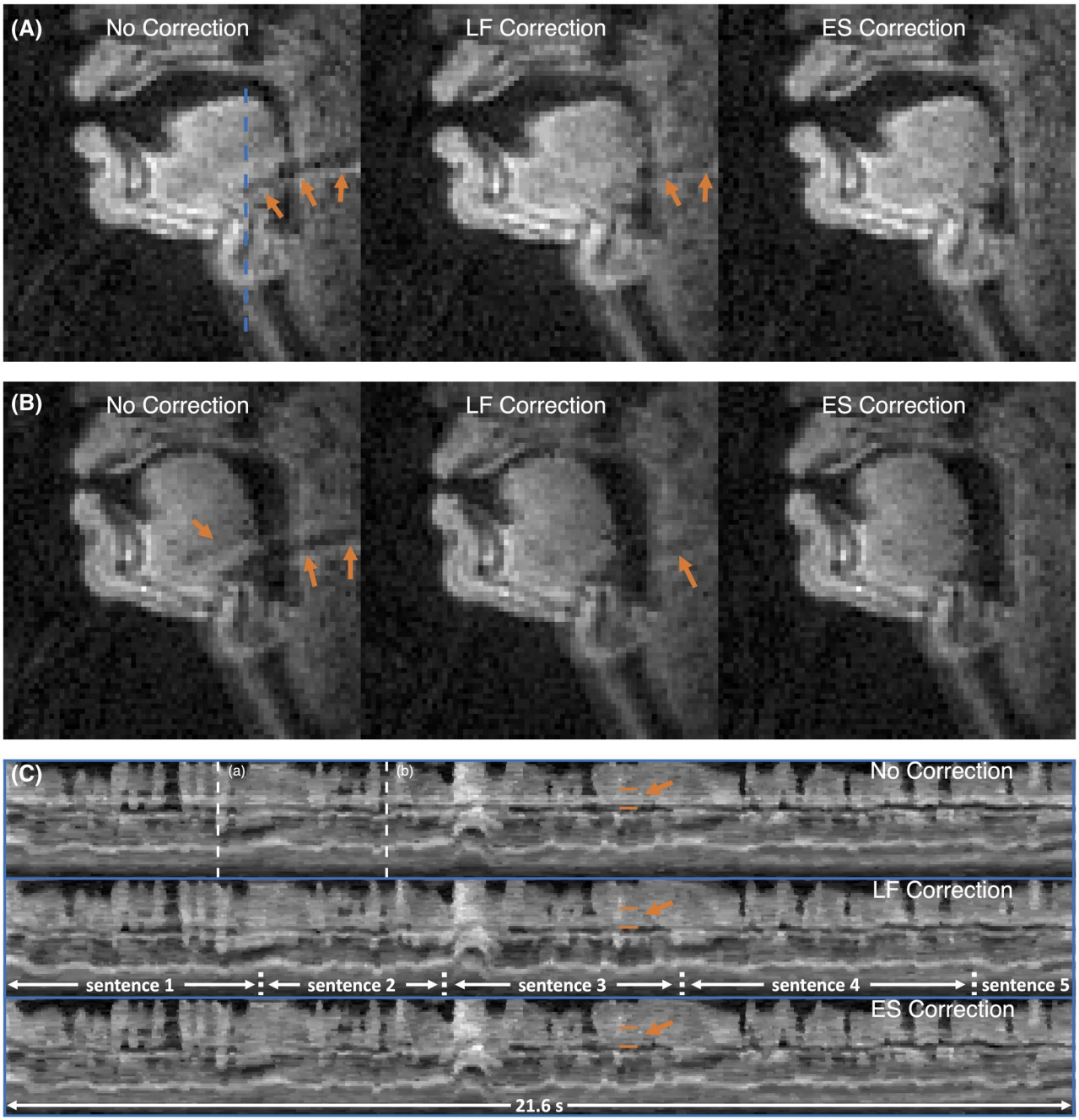

FIGURE 3 Artifact reduction results. This subject was scored as 5/4, 2/2, and 2/2 by expert 1/expert 2 for no correction, LF correction, and ES correction, respectively. A,B, Two time frames from the entire dynamic image series. The orange arrows identify spiral aliasing artifacts passing through the vocal tract. There are remaining mild artifacts in the LF correction, and no apparent artifacts in the ES correction. C, Temporal line profiles of the three methods of the blue line marked in (A). The position of the line profile goes through an area affected by the artifact (identified by orange arrows). The corresponding time frames shown in (A) and (B) are marked with dashed lines. The entire stimuli contain five sentences, with a total duration of 21.6 seconds. Supporting Information Video S2 shows the corresponding video

data set,[14] in which aliasing was reduced without blurring. In speech-production RT-MRI studies in which the speech audio is recorded simultaneously with the image acquisition, an improved noise cancelation is achieved when using a periodic gradient oscillation.[42] This improved noise-cancelation method does not favor GA sampling or pseudo-GA sampling with a large cycle time.

The success of ES correction relies on a faithful reconstruction of the hotspot. In speech RT-MRI, vocal tract motion for speech production occurs in the region(s) of speech articulators and is not directly at the hotspot, which is substantially inferior to the vocal tract in this imaging protocol. Thus, it is possible to temporally combine the dynamic data to obtain a good estimation of the hotspot. Even

TABLE 1 Evaluation results (two raters) and reconstruction time of artifact-reduction methods

	No correction	Method 1: LF	Method 2: ES
Expert 1 w/o artifact	1.0	1.0	1.0
Expert 2 w/o artifact	1.0	1.4	1.2
Expert 1 w/ artifact	3.5	2.4	2.1
Expert 2 w/ artifact	3.3	1.9	2.0
Reconstruction time (ms/frame)	160.69 ± 1.56	526.43 ± 5.17	171.47 ± 1.71

Note: Artifact severity was scored by two expert raters using a 5-level scale, in which a score of "1" indicated the least severe artifact, and "5" indicated the most severe artifact. The listed scores were averaged based on the second rating task results. The "w/o artifact" indicates data sets that have videos without correction rated at "1," and "w/artifact" indicates the rest of the data sets. The reconstruction time was calculated per time frame by averaging recorded time across all 58 data sets.

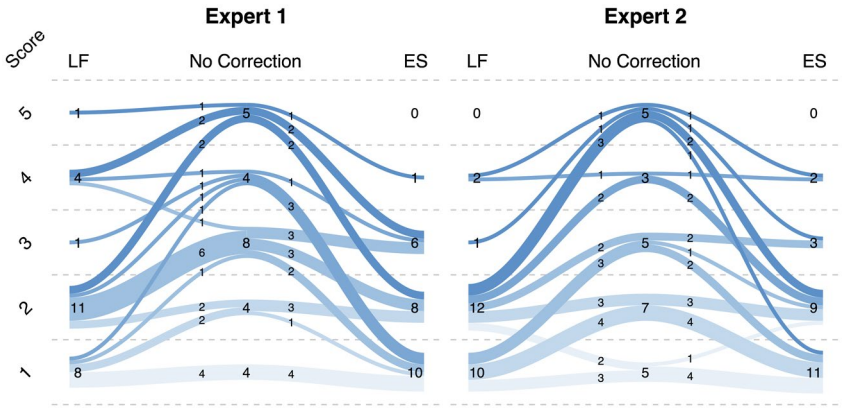

FIGURE 4 Expert qualitative evaluation of the artifact-correction methods. The results are shown in the Sankey diagram separately for the two expert raters. In each diagram, the center column represents reconstructions without correction, and the left and right columns represent reconstructions applied with LF and ES corrections, respectively. The number of data sets graded for each score is shown. The curves have a general trend of decreasing to the left (LF) and right (ES) sides, which means the artifact-reduction methods have decreased the artifact severity

though the prescribed FOV does not support the hotspot, the aliasing from signal within the FOV to the hotspot area is small, leading to a faithful reconstruction of the hotspot. When the hotspot moves, it may be possible to obtain a faithful reconstruction of the hotspot using the sliding-window method (also known as view sharing).[43] If the motion of the hotspot is restricted within a small region, the LF method may be performed using a static mask that includes the whole range of the movement. Both correction methods are not restricted to spiral sampling, and the ES method is not restricted to reducing artifacts arising from outside the FOV.

The proposed methods will require re-evaluation if applied to different body regions, data sets, resolution, and/or FOV settings. Parameters such as the reconstructed FOV when identifying the aliasing source (currently at 2.5× FOV) should be adjusted based on the location of the aliasing source. The automatic detection of the aliasing source may need to be redesigned if there is a major change in the location, shape, and number of sources.

There are several limitations in this study. First, although we found $r_{out/in} \leq 0.4$ can identify aliasing-free data sets, this threshold only selects a limited number of aliasing-free data sets (eg, 4 of 12 in this study). An improved method to identify aliasing-free data sets can further avoid unnecessary application of the aliasing reduction methods. Another limitation is that we have not used quantitative metrics to evaluate the two proposed methods. In fact, we have applied several candidate metrics, including a recent proposed edge sharpness score.[44] However, the unique feature of the artifact (usually a sharp arch across the FOV) may result in improved edge sharpness, leading to biased results. Moreover, qualitative scoring by experts is subject to intrareader and interreader variability. That said, with regard to replication variability, Supporting Information Figure S3 shows the confusion matrixes of the two rating rounds for the no-correction reconstruction with qualitative scores being relatively repeatable, with only one case (one reader out of two, one data set out of 25) in which the two reads differed by more than 1 grade.

5 | CONCLUSIONS

Gradient nonlinearity and spiral sampling can lead to unique aliasing artifact in midsagittal speech RT-MRI. Two proposed artifact reduction methods—a large-FOV method and an estimation-subtraction method—can substantially reduce these artifacts. The ES method is faster and more effective in the context of midsagittal speech RT-MRI with a bit-reversed spiral order. These methods may generally be applied to spiral imaging in which the desired FOV is smaller than the signal-producing region.

ACKNOWLEDGMENT

The authors thank Nam Lee for his help in simulating gradient nonlinearity.

DATA AVAILABILITY STATEMENT

Sample data and reconstruction code are made available on GitHub at https://github.com/usc-mrel/spiral_aliasing_reduction. This includes simulation code that reproduces results shown in Figure 2, and one example data set with reconstruction code that reproduces results shown in Figure 3.

ORCID

Ye Tian https://orcid.org/0000-0002-8559-4404
Yongwan Lim https://orcid.org/0000-0003-0070-0034
Ziwei Zhao https://orcid.org/0000-0003-0281-1141
Dani Byrd https://orcid.org/0000-0003-3319-5871
Shrikanth Narayanan https://orcid.org/0000-0002-1052-6204
Krishna S. Nayak https://orcid.org/0000-0001-5735-3550

REFERENCES

1. Perrier P, Boë LJ, Sock R. Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: modeling the transition with two sets of coefficients. *J Speech Hear Res*. 1992;35:53-67.
2. Perkell JS, Cohen MH, Svirsky MA, Matthies ML, Garabieta I, Jackson MT. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *J Acoust Soc Am*. 1992;92:3078-3096.
3. Stone M, Shawker TH, Talbot TL, Rich AH. Cross-sectional tongue shape during the production of vowels. *J Acoust Soc Am*. 1988;83:1586-1596.
4. Narayanan S, Nayak K, Lee S, Sethy A, Byrd D. An approach to real-time magnetic resonance imaging for speech production. *J Acoust Soc Am*. 2004;115:1771-1776.
5. Demolin D, Hassid S, Metens T, Soquet A. Real-time MRI and articulatory coordination in speech. *C R Biol*. 2002;325:547-556.
6. Honda K, Takemoto H, Kitamura T, Fujita S, Takano S. Exploring human speech production mechanisms by MRI. *IEICE Trans Inf Syst*. 2004;1050-1058.
7. NessAiver MS, Stone M, Parthasarathy V, Kahana Y, Paritsky A. Recording high quality speech during tagged cine-MRI studies using a fiber optic microphone. *J Magn Reson Imaging*. 2006;23:92-97.
8. Ventura SR, Freitas DR, Tavares JM. Application of MRI and biomedical engineering in speech production study. *Comput Methods Biomech Biomed Engin*. 2009;12:671-681.
9. Bresch E, Kim Y, Nayak K, Byrd D, Narayanan S. Seeing speech: capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP]. *IEEE Signal Process Mag*. 2008;25:123-132.
10. Freitas AC, Wylezinska M, Birch MJ, Petersen SE, Miquel ME. Comparison of Cartesian and non-Cartesian real-time MRI sequences at 1.5T to assess velar motion and velopharyngeal closure during speech. *PLoS One*. 2016;11:e0153322.
11. Freitas AC, Ruthven M, Boubertakh R, Miquel ME. Real-time speech MRI: commercial Cartesian and non-Cartesian sequences at 3T and feasibility of offline TGV reconstruction to visualise velopharyngeal motion. *Phys Med*. 2018;46:96-103.
12. Lingala SG, Sutton BP, Miquel ME, Nayak KS. Recommendations for real-time speech MRI. *J Magn Reson Imaging*. 2016;43:28-44.
13. Lingala SG, Zhu Y, Kim YC, Toutios A, Narayanan S, Nayak KS. A fast and flexible MRI system for the study of dynamic vocal tract shaping. *Magn Reson Med*. 2017;77:112-125.
14. Kim YC, Proctor MI, Narayanan SS, Nayak KS. Improved imaging of lingual articulation using real-time multislice MRI. *J Magn Reson Imaging*. 2012;35:943-948.
15. Lim Y, Zhu Y, Lingala SG, Byrd D, Narayanan S, Nayak KS. 3D dynamic MRI of the vocal tract during natural speech. *Magn Reson Med*. 2019;81:1511-1520.
16. Zhao Z, Lim Y, Byrd D, Narayanan S, Nayak K. Improved 3D real-time MRI of speech production. *Magn Reson Med* 2021;85:3182-3195.
17. Chen W, Lee NG, Byrd D, Narayanan S, Nayak KS. Improved real-time tagged MRI using REALTAG. *Magn Reson Med*. 2020;84:838-846.
18. Chen W, Byrd D, Narayanan S, Nayak KS. Intermittently tagged real-time MRI reveals internal tongue motion during speech production. *Magn Reson Med*. 2019;82:600-613.
19. Doran SJ, Charles-Edwards L, Reinsberg SA, Leach MO. A complete distortion correction for MR images. I: gradient warp correction. *Phys Med Biol*. 2005;50:1343-1361.
20. Janke A, Zhao H, Cowin GJ, Galloway GJ, Doddrell DM. Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images. *Magn Reson Med*. 2004;52:115-122.
21. Lee JH, Scott GC, Pauly JM, Nishimura DG. Broadband multicoil imaging using multiple demodulation hardware: a feasibility study. *Magn Reson Med*. 2005;54:669-676.
22. Lazarus C, März M, Weiss P. Correcting the side effects of ADC filtering in MR image reconstruction. *J Math Imaging Vis*. 2020;62:1034-1047.
23. Lingala SG, Toutios A, Töger J, et al. State-of-the-art MRI protocol for comprehensive assessment of vocal tract structure and function. In: Proceedings of Interspeech 2016, San Francisco, California. pp 475-479.
24. Ramanarayanan V, Tilsen S, Proctor M, et al. Analysis of speech production real-time MRI. *Comput Speech Lang*. 2018;52:1-22.

25. Kim J, Kumar N, Lee S, Narayanan SS. Enhanced Airway-Tissue Boundary Segmentation for Real-Time Magnetic Resonance Imaging Data. In: Proceedings of the 10th International Seminar on Speech Production (ISSP), 2014, Cologne, Germany. pp. 222–225.
26. Tian Y, Mendes J, Pedgaonkar A, et al. Feasibility of multiple-view myocardial perfusion MRI using radial simultaneous multi-slice acquisitions. *PLoS One*. 2019;14:e0211738.
27. Feng L, Huang C, Shanbhogue K, Sodickson DK, Chandarana H, Otazo R. RACER-GRASP: respiratory-weighted, aortic contrast enhancement-guided and coil-unstreaking golden-angle radial sparse MRI. *Magn Reson Med*. 2018;80:77-89.
28. Mandava S, Keerthivasan MB, Martin DR, Altbach MI, Bilgin A. Radial streak artifact reduction using phased array beamforming. *Magn Reson Med*. 2019;81:3915-3923.
29. Xue Y, Yu J, Kang HS, Englander S, Rosen MA, Song HK. Automatic coil selection for streak artifact reduction in radial MRI. *Magn Reson Med*. 2012;67:470-476.
30. Holme HC, Frahm J. Sinogram-based coil selection for streak artifact reduction in undersampled radial real-time magnetic resonance imaging. *Quant Imaging Med Surg*. 2016;6:552-556.
31. Li W, Wang N, Yu F, et al. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. *Neuroimage*. 2015;108:111-122.
32. Wei H, Dibb R, Zhou Y, et al. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. *NMR Biomed*. 2015;28:1294-1303.
33. Metze P, Li H, Speidel T, Buckert D, Rottbauer W, Rasche V. Sliding window reduced FOV reconstruction for real-time cardiac imaging. *Z Med Phys*. 2020;30:236-244.
34. Sedarat H, Kerr AB, Pauly JM, Nishimura DG. Partial-FOV reconstruction in dynamic spiral imaging. *Magn Reson Med*. 2000;43:429-439.
35. Santos JM, Wright GA, Pauly JM. Flexible real-time magnetic resonance imaging framework. In: Proceedings of the 26th Annual Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, California, 2004. pp 1048-1051.
36. Kerr AB, Pauly JM, Hu BS, et al. Real-time interactive MRI on a conventional scanner. *Magn Reson Med*. 1997;38:355-367.
37. Adluru G, Awate SP, Tasdizen T, Whitaker RT, Dibella EV. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. *Magn Reson Med*. 2007;57:1027-1036.
38. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array MR imagery. *Magn Reson Med*. 2000;43:682-690.
39. Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BM. 4D XCAT phantom for multimodality imaging research. *Med Phys*. 2010;37:4902-4915.
40. Glover GH, Pelc NJ; General Electric Company. *Method for Correcting Image Distortion Due to Gradient Nonuniformity*. US Patent 4591789A. 1986.
41. Kim YC, Narayanan SS, Nayak KS. Flexible retrospective selection of temporal resolution in real-time speech MRI using a golden-ratio spiral view order. *Magn Reson Med*. 2011;65:1365-1371.
42. Bresch E, Nielsen J, Nayak K, Narayanan S. Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans. *J Acoust Soc Am*. 2006;120:1791-1794.
43. Riederer SJ, Tasciyan T, Farzaneh F, Lee JN, Wright RC, Herfkens RJ. MR fluoroscopy: technical feasibility. *Magn Reson Med*. 1988;8:1-15.
44. Lim Y, Lingala SG, Narayanan SS, Nayak KS. Dynamic off-resonance correction for spiral real-time MRI of speech. *Magn Reson Med*. 2019;81:234-246.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

VIDEO S1 Dynamic illustration of the spiral aliasing artifact. This video was reconstructed to 2.5× FOV using non-uniform fast Fourier transform (NUFFT) with 13 spiral arms per time frame to reduce the impact of regularization. The sagittal imaging plane goes through the main artery and heart of the subject. The aliasing artifact is nearly static; however, a pulsatile signal change close to the hotspot is seen, which is likely due to heartbeat

VIDEO S2 Video corresponding to Figure 3. This example shows the reduced spiral aliasing artifact by both large-FOV (LF) and estimation-subtraction (ES) methods. The video is provided with noise-canceled speech audio

VIDEO S3 Aliasing artifact reduction in pseudo-golden-angle (PGA) sampling. This example shows a 34-interleaved PGA-sampled acquisition reconstructed without correction, with the ES method, and with the LF method. The LF reconstruction was performed directly with 2.5× FOV without any masks. The LF reconstruction has little blurring artifact and reduced aliasing, as a 34-interleaved PGA trajectory supports a large FOV. The horizontal dark line going through the tongue is due to an interleaved multislice acquisition strategy with orthogonal slices. This results in additional saturation of locations where the slices intersect

FIGURE S1 Point spread function (PSF) analysis for the LF reconstruction. Top row shows the PSFs of reconstructions using 1.25× FOV, 2.5× FOV, and 2.5x FOV with a $\sqrt{2}$× circular FOV mask (M_2). Trajectory used to generate these PSFs was a 13-interleaf spiral trajectory supporting an 84 × 84 matrix size, which is the same trajectory as used for in vivo acquisition in this study. The PSFs of 1.25× FOV and 2.5× FOV with mask are similar, although the PSF of 2.5× FOV is wider, as denoted by a black arrow, causing blurring in the reconstructed image. The bottom row shows a time frame of reconstructions with 1.25× FOV, 2.5× FOV, and 2.5× FOV with a $\sqrt{2}$× circular FOV mask (M_2), where the blurring can be seen on the 2.5× FOV reconstruction. Note that this example is used to illustrate the impact of M_2; thus, we picked a data set without aliasing artifact and did not use M_1 during reconstruction

FIGURE S2 Artifacts caused by an oversized mask in ES and LF reconstructions. This figure shows an example of additional artifact caused by the spiral aliasing reduction methods, in which there was no apparent aliasing artifact before correction. A, The automatic masking was able to identify a

hotspot outside the FOV; however, due to its low signal intensity, the mask includes a large area of only aliasing (or signal leakage) from the signal within the FOV. B, When this signal leakage is subtracted from the k-space, as in the ES method, or reconstructed as in the LF method, the signal within the FOV is affected, causing increased artifacts as denoted by yellow arrows. However, only 1 of 58 data sets suffered from additional artifact, and the threshold $r_{out/in} \leq 0.4$ was able to identify it as an aliasing-free data set, preventing it from being processed

FIGURE S3 Confusion matrixes of the two rounds of evaluation. The figure shows confusion matrix data for the two rounds of evaluation in which raters score the no-correction reconstruction videos. The bias is mostly with ±1 grade, with only one instance having ±3 deviation

How to cite this article: Tian Y, Lim Y, Zhao Z, Byrd D, Narayanan S, Nayak KS. Aliasing artifact reduction in spiral real-time MRI. *Magn Reson Med*. 2021;86:916–925. https://doi.org/10.1002/mrm.28746

APPENDIX

Read speech stimuli: "She had your dark suit in greasy wash water all year. Don't ask me to carry an oily rag like that. The girl was thirsty and drank some juice followed by a coke. Your good pants look great however your ripped pants look like a cheap version of a K-mart special. Is that an oil stain on them?" This stimuli is known as the phonetically rich sentences in the linguistics literature.[23]