

RESEARCH ARTICLE

# Comparison of Cartesian and Non-Cartesian Real-Time MRI Sequences at 1.5T to Assess Velar Motion and Velopharyngeal Closure during Speech

Andreia C. Freitas<sup>1,2\*</sup>, Marzena Wylezinska<sup>1a</sup>, Malcolm J. Birch<sup>2</sup>, Steffen E. Petersen<sup>1</sup>, Marc E. Miquel<sup>1,2</sup>

**1** NIHR Cardiovascular Biomedical Research Unit at Barts, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom, **2** Clinical Physics, Barts Health NHS Trust, London, United Kingdom

✉ Current address: Neuroradiological Academic Unit, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, London, United Kingdom

\* [a.freitas@qmul.ac.uk](mailto:a.freitas@qmul.ac.uk)



OPEN ACCESS

**Citation:** Freitas AC, Wylezinska M, Birch MJ, Petersen SE, Miquel ME (2016) Comparison of Cartesian and Non-Cartesian Real-Time MRI Sequences at 1.5T to Assess Velar Motion and Velopharyngeal Closure during Speech. PLoS ONE 11(4): e0153322. doi:10.1371/journal.pone.0153322

**Editor:** Frederic Dick, Birkbeck College, UNITED KINGDOM

**Received:** December 10, 2015

**Accepted:** March 28, 2016

**Published:** April 13, 2016

**Copyright:** © 2016 Freitas et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This project is funded by CLEFT (CLEFT-2013-1) and an Institute of Physics and Engineering in Medicine (IPEM) Innovation and Research Award for computer equipment. MEM is partly (20%) funded by the National Institute for Health Research Cardiovascular Biomedical Research Unit at Barts. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Abstract

Dynamic imaging of the vocal tract using real-time MRI has been an active and growing area of research, having demonstrated great potential to become routinely performed in the clinical evaluation of speech and swallowing disorders. Although many technical advances have been made in regards to acquisition and reconstruction methodologies, there is still no consensus in best practice protocols. This study aims to compare Cartesian and non-Cartesian real-time MRI sequences, regarding image quality and temporal resolution trade-off, for dynamic speech imaging. Five subjects were imaged at 1.5T, while performing normal phonation, in order to assess velar motion and velopharyngeal closure. Data was acquired using both Cartesian and non-Cartesian (spiral and radial) real-time sequences at five different spatial-temporal resolution sets, between 10 fps ( $1.7 \times 1.7 \times 10 \text{ mm}^3$ ) and 25 fps ( $1.5 \times 1.5 \times 10 \text{ mm}^3$ ). Only standard scanning resources provided by the MRI scanner manufacturer were used to ensure easy applicability to clinical evaluation and reproducibility. Data sets were evaluated by comparing measurements of the velar structure, dynamic contrast-to-noise ratio and image quality visual scoring. Results showed that for all proposed sequences, FLASH spiral acquisitions provided higher contrast-to-noise ratio, up to a 170.34% increase at 20 fps, than equivalent bSSFP Cartesian acquisitions for the same spatial-temporal resolution. At higher frame rates (22 and 25 fps), spiral protocols were optimal and provided higher CNR and visual scoring than equivalent radial protocols. Comparison of dynamic imaging at 10 and 22 fps for radial and spiral acquisitions revealed no significant difference in CNR performance, thus indicating that temporal resolution can be doubled without compromising spatial resolution ( $1.9 \times 1.9 \text{ mm}^2$ ) or CNR. In summary, this study suggests that the use of FLASH spiral protocols should be preferred over bSSFP Cartesian for the dynamic imaging of velopharyngeal closure, as it

**Competing Interests:** The authors have declared that no competing interests exist.

allows for an improvement in CNR and overall image quality without compromising spatial-temporal resolution.

## Introduction

Velopharyngeal insufficiency (VPI) is a speech impairment resulting from the incomplete closure between the soft palate (velum) and the posterior and lateral pharyngeal walls, i.e. the velopharyngeal port. As a result, air escapes through the nasal cavity during phonation and patients most commonly present hypernasal speech [1]. Clinical assessment of VPI primarily depends on the speech therapists' perceptual evaluation. Imaging is usually performed using x-ray videofluoroscopy and/or nasendoscopy [2]. X-ray videofluoroscopy provides satisfactory visualization of the hard palate and pharyngeal walls, however, soft tissue contrast (e.g. velum) is relatively poor. To improve contrast, a suspension of barium is usually applied to the vocal tract mucosae [3,4]. However, this renders the procedure more unpleasant which is a major constraint in younger patients [4]. Additionally, repetitive exposure to ionizing radiation is of concern. Nasendoscopy consists of the passage of a fiber-optic scope transnasally into the nasopharynx, providing an en-face view of the velopharyngeal port. The introduction of the scope is rather invasive and requires full patient cooperation. Younger patients or those with a deviated nasal septum usually require local anesthetic to the nostril [1]. Additionally, wide-angle distortions and enlarged adenoids may hamper closure assessment [2].

The limitations of both techniques have strongly supported the use of dynamic MRI in speech research, as summarized in a review of the field [5]. MRI provides tomographic images with improved soft tissue contrast, ideal for vocal tract imaging, in multiple image planes without repositioning the patient. An increased number of studies have used real-time MRI to image the upper airway during speaking, singing and swallowing, both in healthy [6–15] and VPI individuals [4,16–19]. Comparison of the performance of real-time MRI with videofluoroscopy revealed good correspondence when assessing the velopharyngeal closure pattern [4].

Despite the advantages, clinical implementation still faces many challenges. Velopharyngeal closure is characterized by the rapid transition of the velum between rest and elevated position; reported between 50 and 150 ms per cycle [20] depending on speech sample and rate. To reliably capture velar motion, previous studies [9,10] have suggested a temporal resolution of 20 frames per second (fps). Although experiments with VPI patients seem to suggest that all closure events during normal phonation are detected at around 10 fps, lower frame rates lead to increased blurring and missed closure events [21]. Additionally, Sagar et al. [16] reported overestimation of the velopharyngeal gap size and misleading diagnostic evaluation when comparing MRI to videofluoroscopy, due to the low frame rate used (2 fps). In summary, temporal resolution is a key issue when assessing velopharyngeal closure and careful consideration should be given to the acquisition frame rate. However, temporal acceleration is limited by the trade-off in signal-to-noise ratio (SNR), spatial resolution and overall image quality.

Most research on clinical scanners has been focussed on Cartesian acquisitions. Turbo spin echo (TSE) “zoom” techniques with partial Fourier acquisition have been used to achieve between 4–6 fps with spatial resolutions of  $1.5 \times 3.1 \times 6 \text{ mm}^3$  [4] and  $3.9 \times 1.9 \times 6 \text{ mm}^3$  [12]. Rapid gradient-echo sequences, both spoiled (like the fast low angle shot (FLASH) sequence) and steady state free precession (SSFP) Cartesian sequences, have also been frequently used in

dynamic vocal tract imaging [6–8,13,16,22]. These have been most commonly implemented with parallel imaging techniques, both image based (SENSE [23]) and k-space based (GRAPPA [24]), to further improve temporal resolution. Vocal tract configuration during singing and swallowing was imaged at 10 fps ( $1.7 \times 2.7 \times 6 \text{ mm}^3$ ) with a FLASH acquisition and GRAPPA [7,8,22]. Scott et al. [6] investigated velopharyngeal closure at various spatial-temporal resolutions (10 fps,  $1.9 \times 1.9 \times 10 \text{ mm}^3$  to 20 fps,  $2.7 \times 2.7 \times 10 \text{ mm}^3$ ) at both 1.5T and 3T using SSFP sequences and SENSE. Also, Martins et al. investigated velar movement in European Portuguese vowels with a Cartesian FLASH sequence and GRAPPA reconstruction, imaging the vocal tract at 14 fps ( $3.3 \times 1.6 \times 8 \text{ mm}^3$ ) [13]. These studies have successfully covered a wide range of speech tasks and clinical applications with standard scanning resources, however image quality and temporal resolution is still limited. Additionally, non-Cartesian acquisitions and nonstandard reconstruction methods have been suggested to improve spatial-temporal resolution [9–11,25–29].

Narayanan et al. [10,30] imaged the vocal tract using an interleaved spiral acquisition, this allowed an acquired frame rate of about 9 fps ( $2.7 \times 2.7 \times 5 \text{ mm}^3$ ), reconstructed up to 24 fps with a sliding window method [31]. Since each frame is acquired in multiple segments, frame rate can be improved by reconstructing each image with the most recent set of spirals' interleaves. However, no additional information is added to the raw data and temporal fidelity is reliant on native frame rate. Improved native temporal resolution (21 fps,  $1.9 \times 1.9 \times 6.5 \text{ mm}^3$ ) was later reported by Bae et al. [9] using a multi-shot FLASH spiral protocol with regional saturation, where saturation bands are applied to eliminate the signal outside the chosen field-of-view (FOV). Niebergall et al. [11] performed imaging of the vocal tract at 30 fps ( $1.5 \times 1.5 \times 10 \text{ mm}^3$ ) using a radial FLASH acquisition with a nonlinear inversion reconstruction method [32]. Although this allowed for improved image quality at higher undersampling factors than SENSE or conventional gridding reconstruction, reconstruction was intrinsically more complex and time consuming. Lingala et al. suggested an optimized system for the dynamic imaging of the vocal tract using a custom-built upper airway coil, multi-shot spiral sampling and a sparse SENSE constrained reconstruction scheme [28]. Temporal resolutions of 83.3 fps (12 ms) for a single-slice and 27.7 (36 ms) for a three-slice acquisition were achieved, with improved temporal fidelity when compared to a fully sampled gridding reconstruction. Higher frame rates were recently achieved by Fu et al., demonstrating a nominal rate of 100 fps ( $2.2 \times 2.2 \text{ mm}^2$ ) based on a Partial Separability model [29]. Although these methodologies have allowed great improvement in spatial-temporal resolution of speech imaging, they are mostly reliant on off-line reconstruction methodologies and/or non-standard resources, thus hampering immediate translation to clinical evaluation.

In summary, dynamic imaging of the vocal tract with real-time MRI is still an open field of research, and there is much variability in the preferred acquisition methods used by different research groups [33]. Therefore, there is the need for a comparison of different acquisition protocols, regarding image quality and temporal resolution trade-off, which could provide with additional insight to researchers interested in the field and assist with future translation to clinical evaluation. The aim of this study is to compare different real-time sequences, Cartesian and non-Cartesian sampling, for the dynamic imaging of speech, in particular the assessment of velopharyngeal closure. It should be underlined that no direct comparison between the k-space samplings' (Cartesian vs. non-Cartesian) performance is intended from this study, we seek instead to provide a comparison of "best practice" protocols for clinical evaluation. To ensure that any resulting protocol could be easily reproduced and adapted to clinical evaluation, only standard hardware, acquisition and reconstruction algorithms provided by the MRI scanner manufacturer were used in this study.

## Materials and Methods

### Subjects

Five adult individuals (2 males and 3 females, range 34–50 years, median 42 years) were recruited from the staff of our institution for the main study, with a further 2 for preliminary development work. All volunteers gave informed written consent according to ethics approval of NHS research ethics committee (LREC 08/H0701/30). None of the participants had any known speech, language or hearing disabilities.

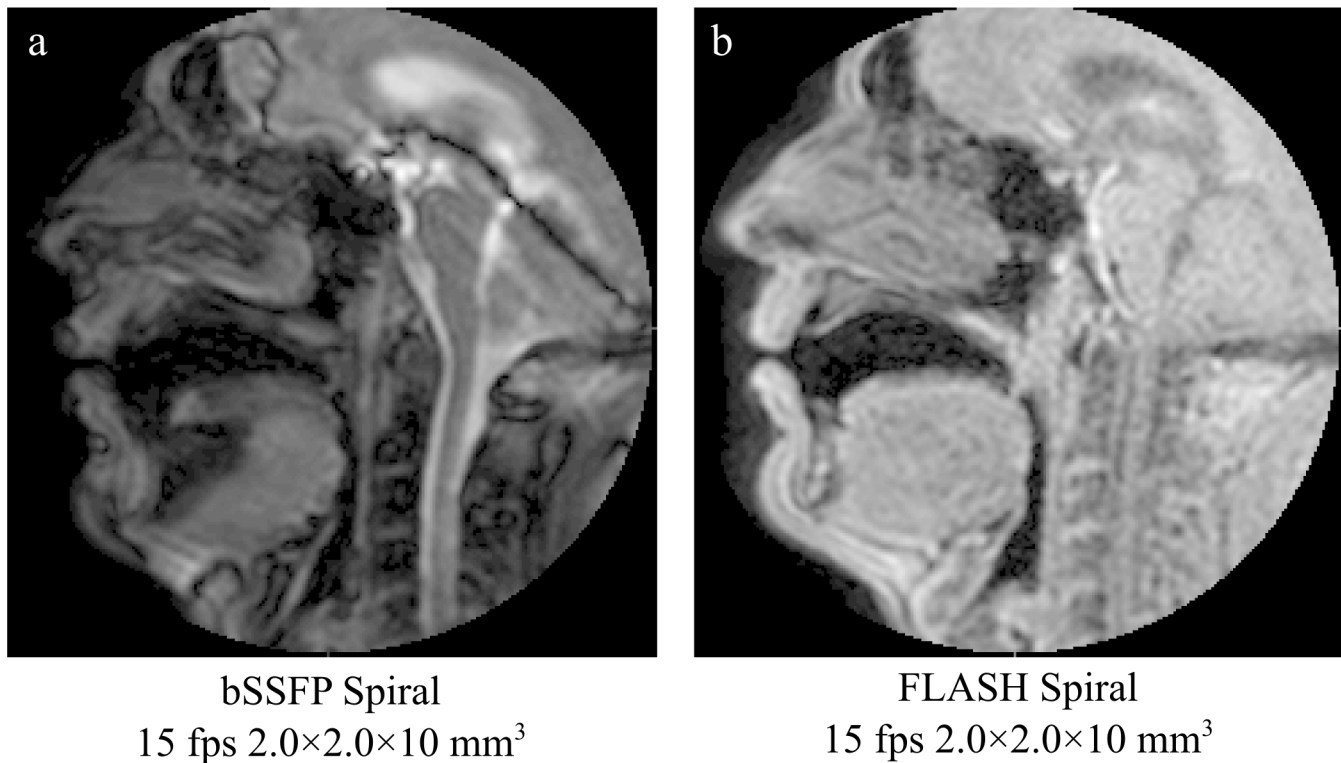
### Speech task and audio recording

Subjects were imaged in the supine position while performing a speech sample consisting of counting (1 to 10), non-sense nasal verbalization (/za-na-za/, /zu-nu-zu/, /ze-ne-ze/) and sustained phonation (/a/ as in 'arm' and /i/ as in 'cheese'). Participants were provided with and asked to repeat the chosen speech sample before entering the MRI examination room. As real-time MRI sequences were used in this study, the speech sample was only produced once per dynamic acquisition, unlike triggered protocols where consecutive repetition of the speech task is necessary [5]. Audio was simultaneously recorded using a fiber-optic MR-compatible microphone (FOMRI II, Optoacoustics, Or Yehuda, Israel), strapped to the coil structure and placed adjacent and parallel to the lips of each subject. A noise-cancelling algorithm was used to reduce the background scanner noise, internal to the microphone system and similar to that described elsewhere [9]. Audio recording was started simultaneously with each MRI acquisition; however, subjects were instructed through the inter-communication system when to initiate phonation in order to allow for the noise cancellation algorithm to adjust, usually between 6 to 10 seconds. Synchronization of the recorded audio and the dynamic image data was possible using a timing trigger signal available from the scanner and recorded as a second channel of the audio signal. Examples of movies generated with synchronized audio can be seen in the supporting information files.

### MRI data acquisition

Images were acquired using a 1.5T Philips Achieva (Philips Healthcare, Best, the Netherlands) R 3.2 MRI scanner (maximum 180 mT/m/ms gradient slew rate and 33 mT/m amplitude) and a 16-channel neurovascular coil. Real-time 2D mid-sagittal images of the head and upper neck were acquired. In order to optimize image quality, the shim volume was centered around the velum.

Preliminary experiments with a phantom and 2 subjects were performed to identify suitable sequences and optimize non-Cartesian acquisition. Cartesian protocols were implemented as described elsewhere [6]. Although Cartesian acquisitions were performed with balanced-SSFP sequences, preliminary testing with bSSFP non-Cartesian sequences revealed dynamic imaging with increased velum blurring and signal void artifacts due to off-resonance effects (Fig 1). Thus, FLASH-like sequences, commonly preferred in speech imaging [5], were used with the non-Cartesian acquisitions. Data acquisition for the main study was performed as follows: Cartesian protocols were implemented with a bSSFP sequence (flip angle 30°, partial Fourier factor of 0.625 and 10 mm slice thickness), while non-Cartesian protocols were implemented with a FLASH sequence (flip angle 10° and 10 mm slice thickness). Non-Cartesian sequences (sequence 1 to 3) were optimized in order to match previously implemented Cartesian protocols [6] in spatial-temporal resolution. Two additional spiral and radial sequences (sequence 4 and 5) were optimized to investigate additional spatial-temporal resolution improvement. By matching protocols in spatial and temporal resolution, a quantitative comparison in image



**Fig 1. Example mid-sagittal images acquired with bSSFP Spiral and FLASH Spiral to demonstrate differences in image quality and velum blurring.**

doi:10.1371/journal.pone.0153322.g001

quality performance (e.g. signal-to-noise ratio, velum distortion and presence of artifacts) between protocols could be performed. Acquisition parameters such as SENSE acceleration, “sliding window” acceleration, number of spiral interleaves, readout time and radial projections were optimized; a detailed description of used acquisition parameters is given in [Table 1](#). Spiral protocols were implemented with 36 interleaves and a readout time of 2 ms. In order to achieve the desired temporal resolution, Cartesian acquisitions were combined with SENSE and non-Cartesian with a “sliding window” reconstruction method.

### Data analysis

Dynamic images were analyzed using measurements of velum thickness and signal homogeneity, dynamic contrast-to-noise ratio (CNR) and image quality visual scoring. SPSS (v.22, IBM, New York) software was used to perform all statistical analyzes. When comparing continuous variables with multiple measurements, as in the case of CNR comparison between sequences and sampling schemes, repeated-measures one-way analysis of variance (ANOVA) was used. This was to test the null hypothesis that the mean of all samples for a certain measurement across sequences were equal, considering a significance level of 0.05. Multiple Bonferroni adjusted paired t-tests were used to identify significant pairs. Paired data sets, such as velum thickness in both velar positions, were compared using a two-tailed paired t-test. Image quality visual scoring was compared using a Kruskal-Wallis test. Significant pairs were identified with multiple pairwise comparisons of the Mann Whitney test and Bonferroni corrected significance level.

**Table 1. Acquisition parameters at 1.5T according to sequence and acquisition sampling scheme.**

Sequence	Spatial-temporal resolution	Acquisition	TE/TR (ms)	FOV (mm <sup>2</sup> )	Sliding window factor	SENSE factor	Radial undersampling
1	1.9×1.9 mm <sup>2</sup> /10fps	Cartesian	1.5/3.0	300×240	-	2.4	-
1	1.9×1.9 mm <sup>2</sup> / 10 fps	Radial	2.3/5.1	180×180	5.0	-	1.57
1	1.9×1.9 mm <sup>2</sup> / 10 fps	Spiral	1.0/5.1	190×190	2.0	-	-
2	2.2×2.2 mm <sup>2</sup> / 15 fps	Cartesian	1.4/2.8	300×240	-	3.0	-
2	2.2×2.2 mm <sup>2</sup> / 15 fps	Radial	2.1/4.7	180×180	6.0	-	1.65
2	2.2×2.2 mm <sup>2</sup> / 15 fps	Spiral	1.0/5.0	190×190	3.0	-	-
3	2.7×2.7 mm <sup>2</sup> / 20 fps	Cartesian	1.2/2.4	300×240	-	3.0	-
3	2.7×2.7 mm <sup>2</sup> / 20 fps	Radial	1.9/4.1	180×180	6.0	-	1.62
3	2.7×2.7 mm <sup>2</sup> / 20 fps	Spiral	1.0/4.8	190×190	4.0	-	-
4	1.9×1.9 mm <sup>2</sup> / 22 fps	Radial	2.3/5.0	170×170	9.0	-	1.71
4	1.9×1.9 mm <sup>2</sup> / 22 fps	Spiral	1.0/5.1	190×190	4.0	-	-
5	1.5×1.5 mm <sup>2</sup> / 25 fps	Radial	2.7/5.9	170×170	16.0	-	1.57
5	1.5×1.5 mm <sup>2</sup> / 25 fps	Spiral	1.0/6.3	190×190	6.0	-	-

Sliding window acceleration factor is defined such as per one fully acquired dynamic scan; a number of sub-dynamic scans defined by the acceleration factor are reconstructed. Therefore, frame rate is increased by the chosen sliding window factor. Radial undersampling was calculated in comparison to a radial aliasing-free sampling case (number of projections times pi/2).

doi:10.1371/journal.pone.0153322.t001

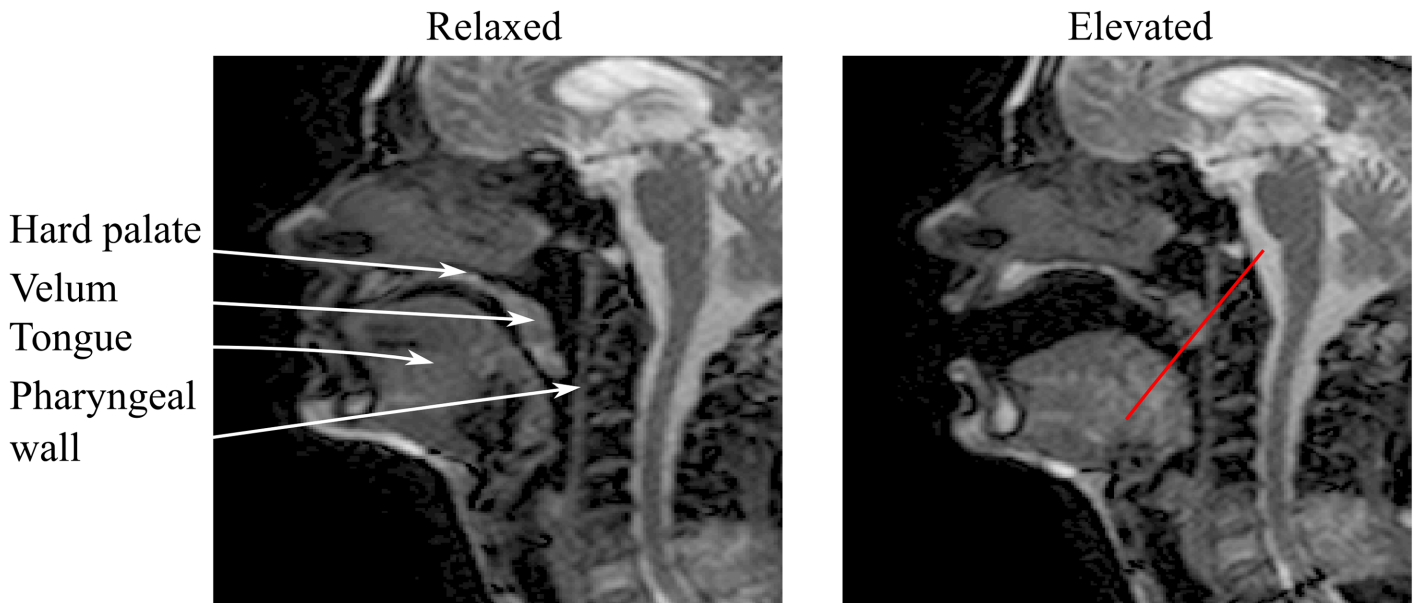
### Velum signal homogeneity and thickness

Measurements of velum thickness and signal homogeneity in the dynamic frames were carried out using OsiriX 6.0.1 32 bit (Pixmeo Sarl, Bernex, Switzerland). Velar measurements in the relaxed position were performed in frames prior to the beginning of phonation (nasal breathing) and in the elevated position, in frames corresponding to the sustained phonation of /a/. Velum thickness was measured as the distance between the velar knee and the velar dimple [34]. Signal homogeneity of the velum was measured as the ratio between the mean and the standard deviation of the signal retrieved from a region-of-interest (ROI) drawn to include the velum structure. This gives an indication of the presence of artifacts or distortion in the selected ROI, as this would result in a decrease in the calculated ratio. Since speech assessment with real-time MRI is yet to be translated into clinical practice, direct correlation between these image quality parameters and clinical relevance still needs to be fully understood. However, low signal homogeneity of the velum could indicate an image where the velum is masked or distorted by artifacts, and thus, clinical assessment of closure could be hampered.

### Intensity-time CNR

CNR measurements were carried out using MATLAB (release 2014b, The MathWorks, Natick, MA). Intensity-time plots were obtained by selecting an intensity profile in each dynamic frame along a reference line (Fig 2b) and stacking profiles from adjacent time frames side-by-side. This allows for a representation of velar motion throughout acquisition, where the horizontal direction is representative of time. CNR was measured in a section of the intensity-time plots, considering two ROIs selected over the velum and adjacent oral cavity, as follows:

$$CNR = \frac{S_{velum} - S_{oral\ cavity}}{\sigma_{oral\ cavity}} \tag{1}$$



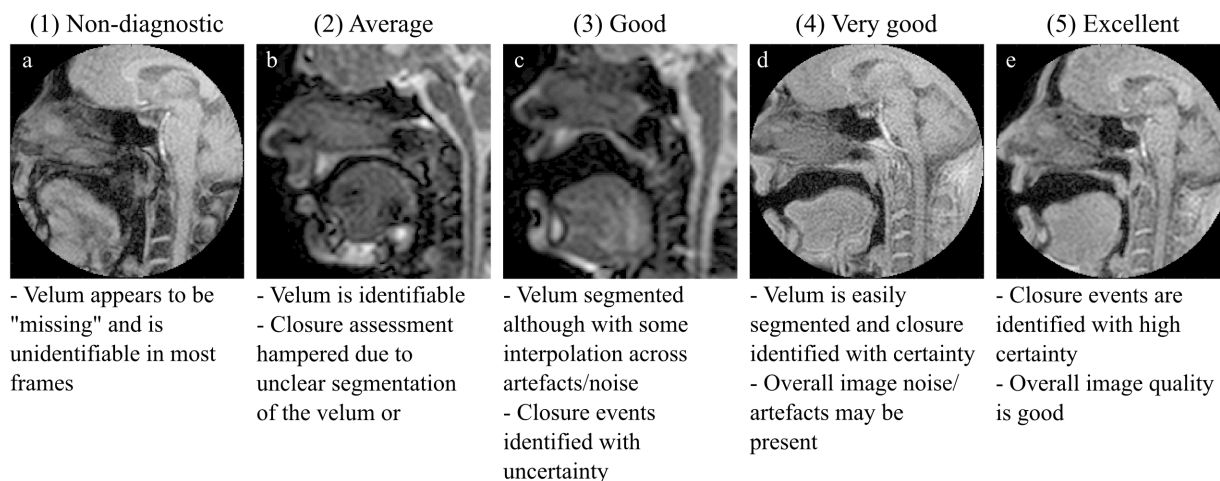
**Fig 2. Example mid-sagittal images to demonstrate the upper vocal tract configuration at the relaxed and elevated velar positions.** Image data acquired in the same subject using Cartesian sequence 1 acquisition protocol. Reference line was selected along the primary direction of motion of the velum to indicate the selected profile when generating the intensity-time plots.

doi:10.1371/journal.pone.0153322.g002

Where  $S_{velum}$  is the mean signal in the ROI drawn in the velum,  $S_{oral\ cavity}$  and  $\sigma_{oral\ cavity}$  are the mean and standard deviation signal in the ROI drawn in the neighboring oral cavity.

### Visual scoring

Image quality of the dynamic data was scored visually. Images were rated blindly and randomly using a five-point scale by two independent observers (imaging physicists) with, 2 and 20 years of MRI experience. For intra-observer reliability data, observer 1 also scored the images a second time, approximately one month after the first scoring. Further details on the chosen scoring scale can be seen in Fig 3.



**Fig 3. Five-point scoring scale from non-diagnostic (a) to excellent (e) image used to visually score image quality.**

doi:10.1371/journal.pone.0153322.g003

## Results

### Velum signal homogeneity and thickness

Example images acquired at both the elevated and relaxed velar positions are shown in [Fig 4](#).

Measurements of velum thickness in millimeters at both velar positions are summarized in [Table 2](#).

Mean velum thickness was determined as  $9.15 \pm 1.51$  mm at the relaxed position and  $11.73 \pm 1.77$  mm at the elevated position ( $p < 0.0005$ ). No significant difference ([Table 2](#) bottom row) in velum thickness was found between sequences 1 to 5 for all sampling schemes at both the relaxed and elevated positions. Additionally, no significant difference ([Table 2](#) right column) in velum thickness was found between sampling schemes (radial vs. spiral vs. Cartesian) at both velar positions.

Signal homogeneity of the velum measured at both velum positions is summarized in [Table 3](#).

It was observed that signal homogeneity of the velum was greater in the relaxed position for all sequences with Cartesian sampling ( $p < 0.05$ ). No significant difference in signal homogeneity was found between the relaxed and elevated positions for all sequences with spiral sampling. ANOVA analysis underlined significant differences between sequences for both velar positions for radial ( $p < 0.05$ , [Table 3](#) bottom row) and spiral ( $p < 0.01$  and  $p < 0.05$ , [Table 3](#) bottom row) acquisitions. Signal homogeneity in the relaxed position was found to be higher for sequence 1 when compared to sequence 5 ( $p < 0.01$ ) for the spiral acquisition. In the elevated position, there were significant differences in signal homogeneity between sequences 1 and 3 ( $p < 0.05$ ) and sequences 3 and 4 ( $p < 0.05$ ) for radial acquisition, and between sequences 2 and 5 ( $p < 0.05$ ) for spiral acquisition. However, no significant difference in signal homogeneity was found between sequences acquired with Cartesian sampling ([Table 3](#) bottom row).

Significant pairs were found between spiral and Cartesian acquisitions for all sequences at both velar positions, and between radial and Cartesian for sequence 3 (relaxed position) and sequences 2 and 3 (elevated position). Comparison of non-Cartesian (spiral vs. radial) acquisitions revealed significant differences in velum signal homogeneity for sequences 1, 2 and 4 in the elevated position.

### Intensity-time CNR

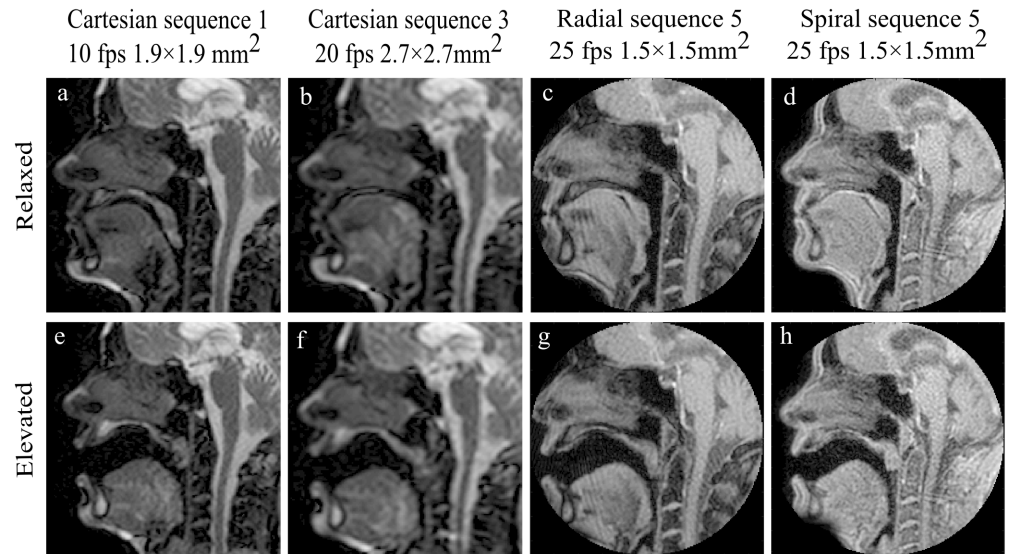
Examples of intensity-time plots for radial and spiral acquisitions are shown in [Fig 5](#). Increased temporal fidelity is noticeable in data acquired with spiral protocols ([Fig 5a, 5c, 5e and 5g](#)) when compared to the otherwise similar radial protocols ([Fig 5b, 5d, 5f and 5h](#)), particularly at higher frame rates. For example at 25 fps, data acquired with spiral sampling ([Fig 5g](#)) presents good distinction of closure events, as all points of contact between the velum and the pharyngeal wall are easily identified. However, intensity-time profiles acquired with the radial protocol showed increased temporal blurring and averaging of consecutive closure events ([Fig 5h](#)).

Intensity-time CNR measurements are summarized in [Table 4](#).

Comparison between sequences showed a CNR increase between sequence 1 and 3 for radial sampling ( $10.21 \pm 1.74$  vs.  $13.27 \pm 1.90$ ), however with borderline significance ( $p = 0.05$ ) and for spiral sampling ( $12.46 \pm 1.31$  vs.  $17.68 \pm 1.51$ ,  $p < 0.0005$ ). However, no significant change in CNR was found between sequences 1 to 3 ( $7.10 \pm 1.87$  vs.  $6.54 \pm 2.71$ ,  $p = 0.93$ ) for Cartesian sampling.

A decrease in CNR was observed for sequences 4 and 5, however with no significant difference ( $p = 1.00$ ) between the two sequences for both radial and spiral acquisitions. In addition, comparison of sequence 4 with sequence 1 for both sampling methods (radial:  $10.21 \pm 1.74$  vs.





**Fig 4. Example mid-sagittal images at elevated and relaxed velum positions acquired with sequence 1 and 3 Cartesian sampling and sequence 5 with radial and spiral acquisitions.**

doi:10.1371/journal.pone.0153322.g004

7.37±1.02,  $p = 1.02$  and spiral: 12.46±1.31 vs. 11.12±0.59,  $p = 0.83$ ) revealed no significant differences in CNR.

For sequences 1 to 3, non-Cartesian acquisitions provided higher CNR than the otherwise similar Cartesian acquisitions. At higher frame rates (sequences 4 and 5), spiral acquisitions provided higher CNR than equivalent radial protocols (refer to Table 4).

### Visual scoring

Intra-observer agreement was good (Cohen’s  $k = 0.59$ ,  $p < 0.0005$ ) with differences in 20 out of 65 analyzed cases and a maximum intra-observer difference of 1 score point. Inter-observer agreement was very good (Cohen’s  $k = 0.74$ ,  $p < 0.0005$ ) with differences between the observers in 13 of the 65 cases and a maximum inter-observer difference of 1 score point. In 11 of these 13 cases, observer 2 scored images higher by 1 point than observer 1. In 12 cases out of the 65,

**Table 2. Mean velum thickness and standard deviation in millimeters (mm) of all subjects measured in the relaxed (nasal breathing) and elevated (sustained phonation of /a/) velar positions.**

Sequence	FLASH Radial		FLASH Spiral		bSSFP Cartesian		p-value
	Relaxed	Elevated	Relaxed	Elevated	Relaxed	Elevated	
1	8.7 (2.2)	11.5 (2.0)	8.7 (1.6)	11.7 (1.3)	9.6 (1.4)	12.3 (2.1)	0.65 <sup>NS</sup> /0.77 <sup>NS</sup>
2	9.1 (2.1)	12.0 (1.6)	8.7 (1.4)	12.3 (0.8)	9.9 (1.6)	12.1 (2.6)	0.59 <sup>NS</sup> /0.95 <sup>NS</sup>
3	9.5 (2.0)	12.2 (2.1)	9.3 (1.9)	12.6 (0.7)	10.0 (0.9)	12.2 (2.5)	0.75 <sup>NS</sup> /0.92 <sup>NS</sup>
4	9.0 (1.4)	11.3 (1.3)	8.5 (1.2)	11.4 (1.3)	-	-	0.56 <sup>NS</sup> /0.86 <sup>NS</sup>
5	9.3 (1.3)	10.5 (2.1)	8.6 (1.0)	10.3 (1.8)	-	-	0.39 <sup>NS</sup> /0.83 <sup>NS</sup>
p-value	0.95 <sup>NS</sup>	0.66 <sup>NS</sup>	0.93 <sup>NS</sup>	0.06 <sup>NS</sup>	0.86 <sup>NS</sup>	0.99 <sup>NS</sup>	

P-values refer to ANOVA analysis of velum thickness between sequences of the same sampling scheme (bottom row) and between sampling schemes within the same spatial-temporal compromise (right column, upper value corresponds to relaxed position and bottom value to elevated position). NS-not significant.

doi:10.1371/journal.pone.0153322.t002

**Table 3. Mean and standard deviation velum signal homogeneity measured from selected frames of the dynamic data at both the relaxed (nasal breathing) and elevated (sustained phonation of /a/) velar positions for all sequences.**

Sequence	FLASH Radial		FLASH Spiral		bSSFP Cartesian		p-value
	Relaxed	Elevated	Relaxed	Elevated	Relaxed	Elevated	
1	4.03(0.40)	2.94(0.27)	4.93(0.95)	4.17(0.84) <sup>c</sup>	3.71(0.51) <sup>a</sup>	2.46(0.64) <sup>a</sup>	<0.005/<0.0005
2	4.50(0.77)	3.27(0.27)	4.63(1.07)	4.29(0.71) <sup>c</sup>	3.71(0.77) <sup>d</sup>	2.20(0.42) <sup>a c</sup>	<0.05/<0.0005
3	4.54(0.41)	3.76(0.55)	4.59(0.77)	3.55(0.77)	3.36(0.43) <sup>a c</sup>	2.29(0.41) <sup>a b</sup>	<0.0005 <0.0005
4	3.66(0.62)	2.84(0.38)	4.17(0.40)	3.66(0.51)	-	-	0.07 <sup>NS</sup> / <sup>&lt;0.01</sup>
5	3.59(0.50)	2.98(0.47)	3.75(0.52)	3.44(0.41)	-	-	0.59 <sup>NS</sup> /0.06 <sup>NS</sup>
p-value	<0.05	<0.05	<0.01	<0.05	0.14 <sup>NS</sup>	0.30 <sup>NS</sup>	

P-values refer to ANOVA analysis of signal homogeneity between sequences of the same sampling scheme (bottom row) and between sampling schemes within the same spatial-temporal compromise (right column, upper value corresponds to relaxed position and bottom value to elevated position).

Post hoc Bonferroni paired t-test was used to identify statically significant pairs, where

<sup>a</sup> p<0.0005 pairwise comparison to spiral acquisition.

<sup>b</sup> p<0.0005 comparison to radial acquisition.

<sup>c</sup> p<0.005 comparison to radial acquisition.

<sup>d</sup> p<0.05 comparison to spiral.

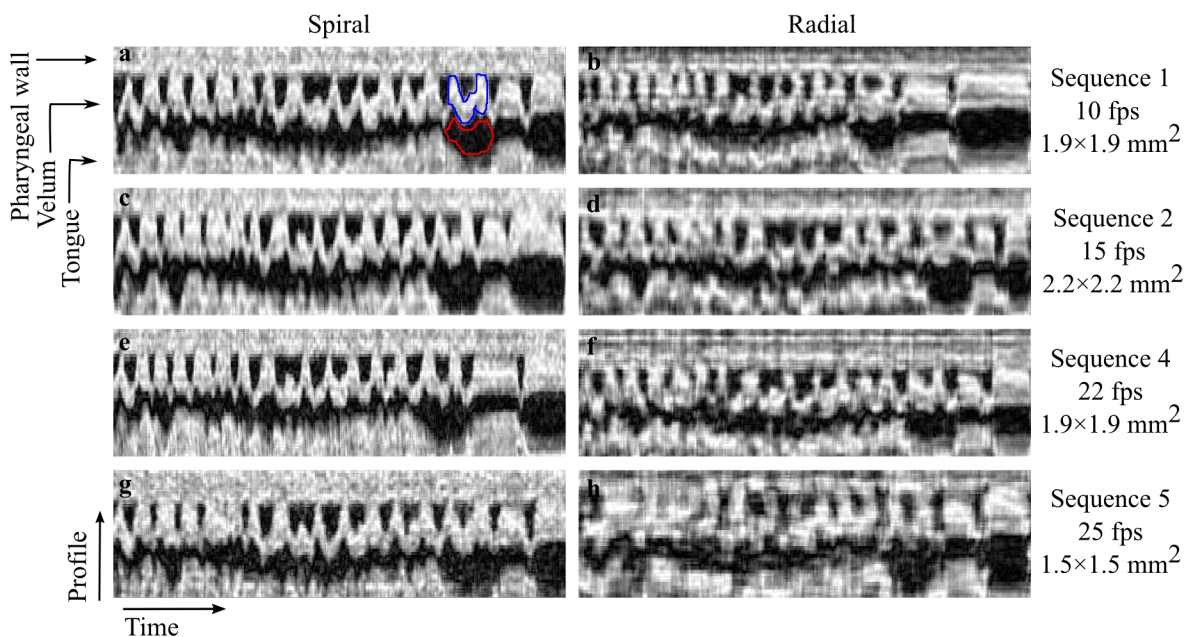
NS—not significant.

doi:10.1371/journal.pone.0153322.t003

correspondent to data acquired with sequence 1 and 5 and spiral sampling, there was complete scoring agreement between the observers.

Histogram representation of visual scoring performed by the two observers is present in Fig 6.

Overall, spiral acquisitions provided superior image quality across sequences 1 to 3 than Cartesian acquisitions (p<0.01). Although visual scoring of data acquired with spiral sequence



**Fig 5. Intensity-time plots for spiral (a,c,e,g) and radial (b,d,f,h) acquisitions at different spatial-temporal resolution sets.** Selected ROIs in the velum (blue) and in the neighboring oral cavity (red) were used to perform CNR measurements. At the highest frame rate of 25 fps (sequence 5), spiral acquisition shows adequate temporal fidelity (g) while radial acquisition shows temporal blurring and averaging of consecutive closure events (h).

doi:10.1371/journal.pone.0153322.g005

**Table 4. Mean and standard deviation CNR measured in a short section of the intensity-time plots.**

Sequence	FLASH Radial	FLASH Spiral	bSSFP Cartesian	p-value
1	10.21 (1.74)	12.46 (1.31)	7.10 (1.87) <sup>a c</sup>	<0.005
2	12.51 (1.92)	13.81 (1.23)	6.67 (2.70) <sup>b d</sup>	<0.0005
3	13.27 (1.90)	17.68 (1.51) <sup>a</sup>	6.54 (2.71) <sup>b d</sup>	<0.0005
4	7.37 (1.02)	11.12 (0.59)	-	<0.0005
5	6.98 (1.09)	9.89 (0.94)	-	<0.005
p-value	<0.0005	<0.0005	0.93 <sup>NS</sup>	

P-values refer to ANOVA analysis of CNR between sequences of the same sampling scheme (bottom row) and between sampling schemes within the same spatial-temporal compromise (right column). Post hoc Bonferroni paired t-test was used to identify statically significant pairs, where

<sup>a</sup> p<0.05 pairwise comparison to radial.

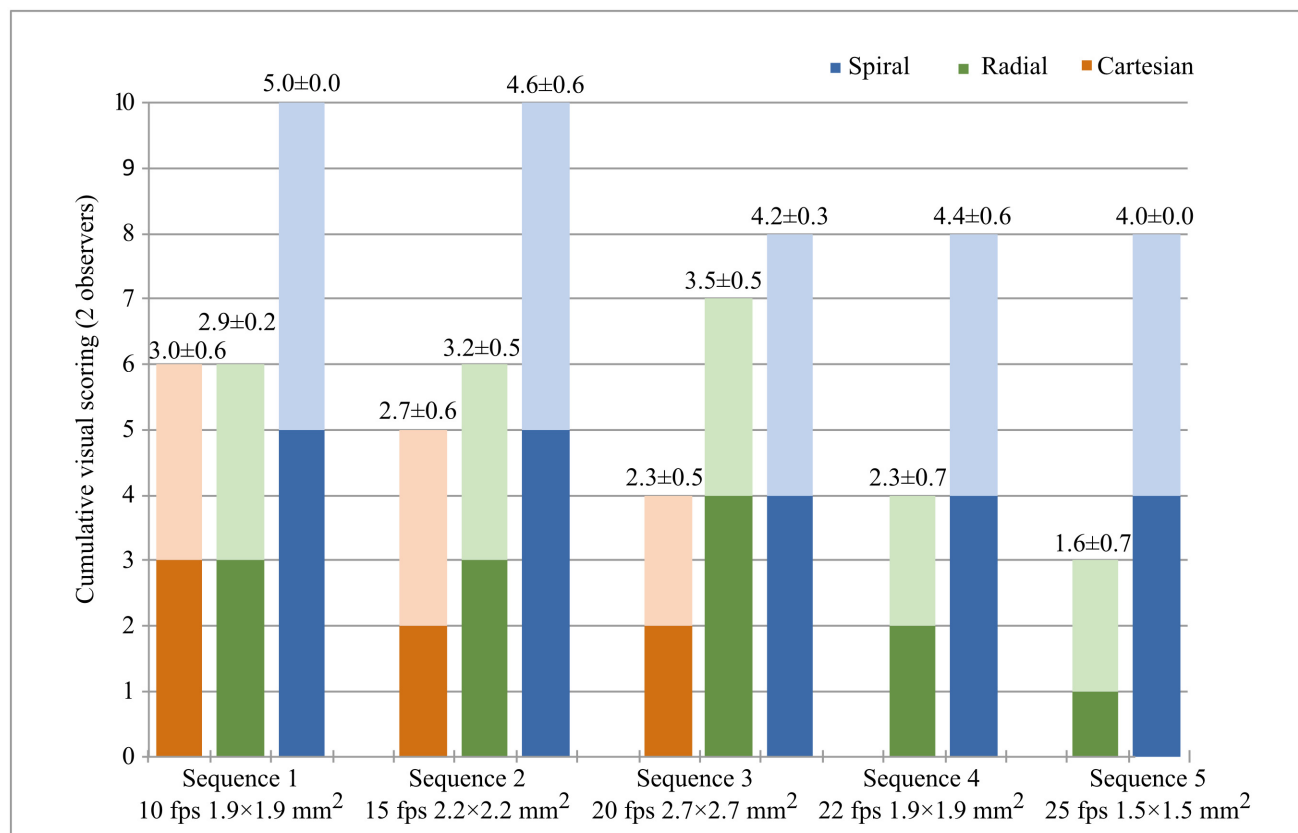
<sup>b</sup> p<0.005 pairwise comparison to radial.

<sup>c</sup> p<0.005 pairwise comparison to spiral.

<sup>d</sup> p<0.0005 pairwise comparison to spiral.

NS—not significant.

doi:10.1371/journal.pone.0153322.t004



**Fig 6. Histogram representation of the image quality cumulative visual scoring performed by 2 independent observers.** Cumulative scoring represented by the sum of each observer independently (maximum scoring of 10) for all sequences and sampling schemes. Mean and standard deviation of visual scoring of both observers is presented numerically on top of each bar plot.

doi:10.1371/journal.pone.0153322.g006

1 presented superior image quality than the equivalent radial acquisition ( $5.0\pm 0.0$  vs.  $2.9\pm 0.2$ ,  $p<0.01$ ), no significant differences were found between the two sampling schemes for sequences 2 and 3. At higher frame rates, visual scoring of spiral data was superior to that of radial acquisitions for both sequence 4 ( $4.4\pm 0.6$  vs.  $2.3\pm 0.7$ ,  $p<0.01$ ) and sequence 5 ( $4.0\pm 0.0$  vs.  $1.6\pm 0.7$ ,  $p<0.01$ ). In total, 12% of the analyzed cases were scored as 'Excellent', all acquired with spiral sampling.

## Discussion

In the present study, we compared Cartesian and non-Cartesian real-time sequences at 1.5T, regarding image quality and temporal resolution trade-off, for dynamic speech imaging.

Previous studies [4,6,12,16] have performed dynamic imaging of velopharyngeal closure with clinically available hardware and Cartesian sequences in both healthy and VPI subjects. However, achieving sufficient temporal resolution to reliably assess VPI, while maintaining satisfactory image quality, is a difficult balance point to achieve and many studies have been limited to low frame rates. Although higher frame rates, 20 fps and upwards, with adequate image quality have been obtained with non-Cartesian acquisitions [9,11,25,28], these protocols are mostly reliant on off-line reconstruction methodologies and/or non-standard equipment, hampering immediate translation to clinical evaluation.

In this study, a non-Cartesian protocol optimized for fast dynamic imaging of the velum and velopharyngeal closure has been proposed. We were able to demonstrate differences in image quality and temporal resolution trade-off between FLASH non-Cartesian (spiral and radial) and bSSFP Cartesian imaging at 1.5T. The optimized non-Cartesian spiral protocol provided improved spatial-temporal resolution (22fps,  $1.9\times 1.9\times 10\text{ mm}^3$  and 25 fps,  $1.5\times 1.5\times 10\text{ mm}^3$ ) in comparison to Cartesian protocols, while still being easily implemented and reproducible with standard resources provided by the MRI scanner manufacturer.

Measured velum thickness was higher in the elevated velum position across all sequences. In agreement with previous literature [6,34], this was an expected result as the posterior-superior elevation of the velum leads to an increase in thickness in the sagittal plane.

Both Cartesian and radial acquisitions showed lower signal homogeneity of the velum in the elevated position across all sequences. As this measurement refers to the homogeneity of the signal intensity within the selected ROI, artifacts present in the velopharyngeal region, which distort and/or mask the velum, strongly reduce the measured parameter. The presence of artifacts in the velopharyngeal region in the elevated position can be seen in Fig 4e. Although both Cartesian and radial data present strong distortion of the velum in the elevated position, and consequently a decrease in the measured homogeneity, artifacts appear to be of different nature. Artifacts present in bSSFP Cartesian data appear to be due to off-resonance effects. This could be explained by that fact that when at rest, all velopharyngeal structures sit in close contact; however, during phonation, the elevation of the velum and separation of the velopharyngeal structures creates a larger area of tissue-air interface and bSSFP sequences are particularly sensitive to susceptibility differences. On the other hand, artifacts in radial data present a spokes-like pattern and are most likely due to radial under-sampling. However, by implementing a FLASH sequence with a spiral sampling, we were able to reduce artifacts or distortion in the region and improve overall signal homogeneity of the velum (Table 3) compared to the other two sampling methods.

Intensity-time CNR performance with non-Cartesian acquisitions was found to be superior to Cartesian acquisitions (Table 4). At higher frame rates, spiral protocols were optimal and provided higher CNR than equivalent radial protocols. Although no significant differences were found between Cartesian sequences 1 to 3, a gradual increase in CNR was found for radial

and spiral acquisitions. As expected with the decrease in pixel size, a decrease in CNR follows for sequences 4 and 5 for both non-Cartesian. However, comparison between sequences 4 and 5 for both spiral and radial acquisitions revealed no significant difference, therefore indicating that improvements in spatial-temporal resolution (from 22 fps,  $1.9 \times 1.9 \times 10 \text{ mm}^3$  to 25 fps,  $1.5 \times 1.5 \times 10 \text{ mm}^3$ ) can be achieved with no significant loss in CNR. In addition, no significant difference in CNR was found between sequences 1 and 4 for both non-Cartesian acquisitions. This suggests that both non-Cartesian acquisitions allow doubling the temporal resolution, from 10 fps to 22 fps, while maintaining spatial resolution ( $1.9 \times 1.9 \text{ mm}^2$ ) and CNR. Although SSFP sequences commonly allow for superior image SNR than spoiled sequences like FLASH [5], this particular measurement of CNR reflects both the intrinsic signal-to-noise ratio and the presence of artifacts in the selected ROIs, i.e. the oral cavity. Thus, although spiral protocols were implemented with a FLASH-like acquisition, due to the reduced presence of artifacts in the velopharyngeal region, the measured CNR was in fact higher than that of bSSFP Cartesian acquisitions. This seems to suggest that the use of the FLASH spiral protocol should be preferred over the bSSFP Cartesian in this particular application, as an improvement in CNR and accurate distinction of the velum boundaries is possible without compromising spatial-temporal resolution.

Qualitative visual scoring provided additional insight on the overall image quality achieved by the different protocols. Overall, FLASH spiral data was scored higher than bSSFP Cartesian across all sequences. Although no significant difference was found between radial and spiral acquisitions at lower temporal resolutions, at higher frame rates of 22 and 25 fps, spiral acquisition was optimal in providing good visual image quality. Radial images, on the other hand, showed increased blurring and spokes-like artifacts, reducing the overall image quality scoring.

Differences between the chosen reconstruction algorithms must also be considered; while SENSE provides a true temporal acceleration of the Cartesian data, the “sliding window” method (non-Cartesian acquisitions) produces an interpolation of the data in time. In this case, since no additional information is being added to the raw data, temporal reliability is still dependent on the acquired frame rate and temporal smoothing is introduced. However, the study intended to compare different “best practice” protocols, i.e. that could be easily reproduced with resources routinely available and still provide adequate imaging. Since spiral acquisitions are intrinsically fast [35] and presented a superior native frame rate (about 6 fps) than radial, lower sliding window acceleration factors were required to achieve the desired frame rates than the equivalent radial acquisitions (Table 1). Radial acquisitions were less optimal with the current clinical setup, as these presented below optimal native frame rate (less than 2 fps) and therefore, blurring of the velum due to motion and averaging of consecutive closure events was present.

Limitations of the present work include the small sample size, as well as increased acoustic noise when using spiral protocols that may require additional care with more sensitive and/or younger subjects. However, noise cancellation and quality of audio was still optimal (supplementary data). Although peripheral nerve stimulation (PNS) clinical warning was displayed for some acquisitions, none of the scanned subjects reported PNS sensation.

## Conclusion

In conclusion, our results suggest that non-Cartesian real-time sequences are a promising tool to improve overall image quality and temporal resolution of dynamic speech imaging. We found that for all proposed sequences, FLASH non-Cartesian (spiral and radial) acquisitions provided higher CNR than bSSFP Cartesian acquisitions, within the same spatial-temporal resolution. With this clinical setup, FLASH spiral sequences were optimal and provided dynamic

imaging with superior CNR, velum signal homogeneity and visual image quality. At temporal resolutions of 22 and 25 fps, spirals showed good temporal reliability of data while radial acquisitions showed increased temporal blurring and were less adequate. It should be underlined that it is possible to further improve image quality and temporal resolution of non-Cartesian acquisitions using alternative reconstruction methods and/or custom equipment and software [9,11,25,28,29]. However, the purpose of this study was to compare and present an easily reproducible protocol to researchers equipped with standard MRI resources and translatable to clinical practice.

## Supporting Information

**S1 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Cartesian 10 fps sequence 1.**

(MP4)

**S2 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Spiral 10 fps sequence 1.**

(MP4)

**S3 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Cartesian 20 fps sequence 3.**

(MP4)

**S4 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Spiral 22 fps sequence 4.**

(MP4)

**S5 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Radial 22 fps sequence 4.**

(MP4)

**S6 File. Example video of rt-MRI of the upper vocal tract during speech acquired with Spiral 25 fps sequence 5.**

(MP4)

## Acknowledgments

We are grateful to Dr. Matthew Clemence, Philips Healthcare Clinical Science, Dr. Redha Boubertakh and Matthieu Ruthven for their contribution and support.

## Author Contributions

Conceived and designed the experiments: ACF MW MJB SEP MEM. Performed the experiments: ACF MW MEM. Analyzed the data: ACF MEM. Wrote the paper: ACF.

## References

1. Sell D, Pereira V. Instrumentation in the analysis of the structure and function of the velopharyngeal mechanism. In: Howard S, Lohmander A, editors. *Cleft Palate Speech: assessment and intervention*. Wiley-Blackwell; 2012. p. 145–62.
2. Havstam C, Lohmander A, Persson C, Dotevall H, Lith A, Lilja J. Evaluation of VPI-assessment with videofluoroscopy and nasoendoscopy. *Br J Plast Surg*. 2005 Oct; 58(7):922–31. PMID: [15922997](#)
3. Henningsson G, Isberg A. Comparison between multiview videofluoroscopy and nasendoscopy of velopharyngeal movements. *Cleft Palate Craniofac J*. 1991 Oct; 28(4):413–7; discussion 417–8. PMID: [1742312](#)

4. Beer AJ, Hellerhoff P, Zimmermann A, Mady K, Sader R, Rummeny EJ, et al. Dynamic near-real-time magnetic resonance imaging for analyzing the velopharyngeal closure in comparison with videofluoroscopy. *J Magn Reson Imaging*. 2004 Nov; 20(5):791–7. PMID: [15503349](#)
5. Scott AD, Wylezinska M, Birch MJ, Miquel ME. Speech MRI: morphology and function. *Phys Med*. Elsevier Ltd; 2014 Sep 28; 30(6):604–18.
6. Scott AD, Boubertakh R, Birch MJ, Miquel ME. Towards clinical assessment of velopharyngeal closure using MRI: evaluation of real-time MRI sequences at 1.5T and 3T. *Br J Radiol*. 2012; 85(1019):1083–92.
7. Echternach M, Sundberg J, Arndt S, Markl M, Schumacher M, Richter B. Vocal tract in female registers—A dynamic real-time MRI study. *J Voice*. Elsevier Ltd; 2010 Mar; 24(2):133–9.
8. Breyer T, Echternach M, Arndt S, Richter B, Speck O, Schumacher M, et al. Dynamic magnetic resonance imaging of swallowing and laryngeal motion using parallel imaging at 3 T. *Magn Reson Imaging*. Elsevier Inc.; 2009 Jan; 27(1):48–54.
9. Bae Y, Kuehn DP, Conway CA, Sutton BP. Real-time magnetic resonance imaging of velopharyngeal activities with simultaneous speech recordings. *Cleft Palate Craniofac J*. 2011 Nov; 48(6):695–707. doi: [10.1597/09-158](#) PMID: [21214321](#)
10. 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(4):1171–776.
11. Niebergall A, Zhang S, Kunay E, Keydana G, Job M, Uecker M, et al. Real time MRI of speaking at a resolution of 33 ms: undersampled radial FLASH with nonlinear inverse reconstruction. *Magn Reson Med*. 2013; 69(2):477–85. doi: [10.1002/mrm.24276](#) PMID: [22498911](#)
12. Demolin D, Hassid S, Metens T, Soquet A. Real-time MRI and articulatory coordination in speech. *C R Biol*. 2002; 325:547–56. PMID: [12161933](#)
13. Martins P, Oliveira C, Silva S. Velar Movement in European Portuguese Nasal Vowels. In: *IberSpeech 2012– VII Jornadas en Tecnología del Habla and III Iberian SLTech Workshop*. Madrid, Spain; 2012. p. 231–40.
14. Wylezinska M, Freitas AC, Birch M, Miquel ME. K-t BLAST/k-t FOCUSS in real time imaging of the soft palate during speech. In: *ISMRM 23rd Scientific Sessions*. 2015. p. 2302.
15. Freitas AC, Wylezinska M, Birch MJ, Petersen SE, Miquel ME. Real-time speech MRI: a comparison of Cartesian and non-Cartesian sequences. In: *ISMRM 23rd Scientific Sessions*. 2015. p. 655.
16. Sagar P, Nimkin K. Feasibility study to assess clinical applications of 3-T cine MRI coupled with synchronous audio recording during speech in evaluation of velopharyngeal insufficiency in children. *Pediatr Radiol*. 2014 Aug 16; 45(2):217–27. doi: [10.1007/s00247-014-3141-7](#) PMID: [25124806](#)
17. Silver AL, Nimkin K, Ashland JE, Ghosh SS, van der Kouwe AJW, Brigger MT, et al. Cine magnetic resonance imaging with simultaneous audio to evaluate pediatric velopharyngeal insufficiency. *Arch Otolaryngol Head Neck Surg*. American Medical Association; 2011 Mar 21; 137(3):258–63.
18. Drissi C, Mitrofanoff M, Talandier C, Falip C, Le Couls V, Adamsbaum C. Feasibility of dynamic MRI for evaluating velopharyngeal insufficiency in children. *Eur Radiol*. 2011 Feb 2; 21(7):1462–9. doi: [10.1007/s00330-011-2069-7](#) PMID: [21287177](#)
19. Kulinna-Cosentini C, Czerny C, Baumann A, Weber M, Sinko K. TrueFisp versus HASTE sequences in 3T cine MRI: Evaluation of image quality during phonation in patients with velopharyngeal insufficiency. *Eur Radiol*. 2015 Nov 28;
20. Kuehn DP. A cineradiographic investigation of velar movement variables in two normals. *Cleft Palate J*. 1976; 13:88–103. PMID: [1062249](#)
21. Miquel ME, Freitas AC, Wylezinska M. Evaluating velopharyngeal closure with real-time MRI. *Pediatr Radiol*. 2014 May; 73(5):1820–32.
22. Echternach M, Sundberg J, Arndt S, Breyer T, Markl M, Schumacher M, et al. Vocal tract and register changes analysed by real-time MRI in male professional singers—a pilot study. *Logoped Phoniatr Vocol*. Informa UK Ltd UK; 2008 Jan 11; 33(2):67–73.
23. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*. 1999 Nov; 42(5):952–62. PMID: [10542355](#)
24. Griswold M, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med*. 2002; 47(6):1202–10. PMID: [12111967](#)
25. Sutton BP, Conway CA, Bae Y, Seethamraju R, Kuehn DP. Faster dynamic imaging of speech with field inhomogeneity corrected spiral fast low angle shot (FLASH) at 3 T. *J Magn Reson Imaging*. 2010 Nov; 32(5):1228–37. doi: [10.1002/jmri.22369](#) PMID: [21031529](#)
26. Burdumy M, Traser L, Richter B, Echternach M, Korvink JG, Hennig J, et al. Acceleration of MRI of the vocal tract provides additional insight into articulator modifications. *J Magn Reson Imaging*. 2015 Oct 3; 42(4):925–35. doi: [10.1002/jmri.24857](#) PMID: [25647755](#)

27. Proctor M, Bresch E, Byrd D, Nayak K, Narayanan S. Paralinguistic mechanisms of production in human “beatboxing”: A real-time magnetic resonance imaging study. *J Acoust Soc Am*. 2013; 133(2):1043–54. doi: [10.1121/1.4773865](https://doi.org/10.1121/1.4773865) PMID: [23363120](https://pubmed.ncbi.nlm.nih.gov/23363120/)
28. Lingala SG, Zhu Y, Kim Y-C, Toutios A, Narayanan S, Nayak KS. A fast and flexible MRI system for the study of dynamic vocal tract shaping. *Magn Reson Med*. 2016 Jan 17;
29. Maojing Fu, Barlaz MS, Shosted RK, Zhi-Pei Liang, Sutton BP. High-resolution dynamic speech imaging with deformation estimation. In: *Engineering in Medicine and Biology Society (EMBC) 37th Annual International Conference of the IEEE*. 2015. p. 1568–71.
30. Ramanarayanan V, Goldstein L, Byrd D, Narayanan S. An investigation of articulatory setting using real-time magnetic resonance imaging. *J Acoust Soc Am*. 2013; 134(1):510–9. doi: [10.1121/1.4807639](https://doi.org/10.1121/1.4807639) PMID: [23862826](https://pubmed.ncbi.nlm.nih.gov/23862826/)
31. Holsinger A, Wright RC, Riederer SJ, Farzaneh F, Grimm RC, Maier JK. Real-time interactive magnetic resonance imaging. *Magn Reson Med*. 1990; 14(3):547–53. PMID: [2355836](https://pubmed.ncbi.nlm.nih.gov/2355836/)
32. Uecker M, Hohage T, Block KT, Frahm J. Image reconstruction by regularized nonlinear inversion—joint estimation of coil sensitivities and image content. *Magn Reson Med*. 2008 Sep; 60(3):674–82. doi: [10.1002/mrm.21691](https://doi.org/10.1002/mrm.21691) PMID: [18683237](https://pubmed.ncbi.nlm.nih.gov/18683237/)
33. Lingala SG, Sutton BP, Miquel ME, Nayak KS. Recommendations for real-time speech MRI. *J Magn Reson Imaging*. 2016 Jan 14; 43(1):28–44. doi: [10.1002/jmri.24997](https://doi.org/10.1002/jmri.24997) PMID: [26174802](https://pubmed.ncbi.nlm.nih.gov/26174802/)
34. Perry JL. Variations in velopharyngeal structures between upright and supine positions using upright magnetic resonance imaging. *Cleft Palate Craniofac J*. 2011 Mar; 48(2):123–33. doi: [10.1597/09-256](https://doi.org/10.1597/09-256) PMID: [20500077](https://pubmed.ncbi.nlm.nih.gov/20500077/)
35. Delattre BMA, Heidemann RM, Crowe LA, Vallée J-P, Hyacinthe J-N. *Spiral demystified*. *Magn Reson Imaging*. Elsevier Inc.; 2010 Jul; 28(6):862–81.