

UCSF

UC San Francisco Previously Published Works

Title

Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging

Permalink

<https://escholarship.org/uc/item/8g6905bv>

Journal

Journal of Speech Language and Hearing Research, 60(4)

ISSN

1092-4388

Authors

Hagedorn, Christina
Proctor, Michael
Goldstein, Louis
et al.

Publication Date

2017-04-14

DOI

10.1044/2016_jslhr-s-15-0112

Peer reviewed

Research Note

Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging

Christina Hagedorn,^a Michael Proctor,^b Louis Goldstein,^a Stephen M. Wilson,^c Bruce Miller,^d Maria Luisa Gorno-Tempini,^d and Shrikanth S. Narayanan^a

Purpose: Real-time magnetic resonance imaging (MRI) and accompanying analytical methods are shown to capture and quantify salient aspects of apraxic speech, substantiating and expanding upon evidence provided by clinical observation and acoustic and kinematic data. Analysis of apraxic speech errors within a dynamic systems framework is provided and the nature of pathomechanisms of apraxic speech discussed.

Method: One adult male speaker with apraxia of speech was imaged using real-time MRI while producing spontaneous speech, repeated naming tasks, and self-paced repetition of word pairs designed to elicit speech errors. Articulatory data were analyzed, and speech errors were detected using time series reflecting articulatory activity in regions of interest.

Results: Real-time MRI captured two types of apraxic gestural intrusion errors in a word pair repetition task. Gestural intrusion errors in nonrepetitive speech, multiple silent initiation gestures at the onset of speech, and covert (unphonated) articulation of entire monosyllabic words were also captured.

Conclusion: Real-time MRI and accompanying analytical methods capture and quantify many features of apraxic speech that have been previously observed using other modalities while offering high spatial resolution. This patient's apraxia of speech affected the ability to select only the appropriate vocal tract gestures for a target utterance, suppressing others, and to coordinate them in time.

Apraxia of speech (AOS) is a speech motor disorder typically characterized by the presence of distortions and distorted substitutions, articulatory groping and attempts to repair errors, disproportionate difficulty with multisyllabic words and consonant clusters, and prosodic abnormalities, such as reduced rate, increased segment durations, and increased intersegment durations (Duffy, 2013; Ogar, Slama, Dronkers, Amici, & Gorno-Tempini, 2005). AOS is classically defined as a disorder affecting the selection, programming, and execution of speech motor commands specified in a target sequence (Wertz, La Pointe, & Rosenbek, 1984; Ziegler, Aichert, & Staiger, 2012). Although AOS is generally associated with dominant hemisphere pathology, the neural substrates of AOS have yet to be unequivocally identified, and the precise stage at which the speech motor system breaks down in apraxic

patients is still unknown (Ziegler et al., 2012). In general, however, it is agreed that AOS involves failure at a high organizational level of speech: the point at which well-formed phonological representations trigger a sequence of contextually appropriate movements. Some propose that AOS may have its basis in the loss of procedural “memories” of the articulatory movements corresponding with specific gestures or gestural constellations (clusters or syllables; Aichert & Ziegler, 2004; Ziegler, 2008, 2010).

Apraxic speech articulation and articulatory coordination have long been areas of interest among researchers and have been investigated kinematically (Bartle-Meyer, Goozée, Murdoch, & Green, 2009; Itoh, Sasanuma, & Ushijima, 1979) and using electropalatography (EPG; Bartle-Meyer, Murdoch, & Goozée, 2009; Hardcastle, Gibbon, & Jones, 1991; Howard & Varley, 1995). Although these modalities provide useful information about articulatory movement and contact patterns in apraxic speech, they are invasive by nature and do not provide the researcher with a dynamic view of the entire vocal tract during speech.

In this pilot study, we use real-time magnetic resonance imaging (rtMRI) and an analytical method of estimating constriction kinematics based on pixel intensity to investigate aspects of the speech of an apraxic patient that hold both theoretical and clinical importance. Our broad

^aUniversity of Southern California, Los Angeles

^bMacquarie University, North Ryde, New South Wales, Australia

^cUniversity of Arizona, Tucson

^dUniversity of California, San Francisco

Correspondence to Christina Hagedorn, who is now at College of Staten Island, City University of New York: christina.hagedorn@csi.cuny.edu

Editor and Associate Editor: Julie Liss

Received March 23, 2015

Revision received December 19, 2015

Accepted July 15, 2016

https://doi.org/10.1044/2016_JSLHR-S-15-0112

Disclosure: The authors have declared that no competing interests existed at the time of publication.

aims are to utilize rtMRI to shed light on aspects of apraxic speech articulation and to identify and further characterize attributes of apraxic speech that have been previously observed clinically and through analysis of both acoustic and kinematic data. Given that rtMRI provides a complete view of the vocal tract over time, specific attention will be devoted to identifying and investigating the nature of erroneous “double articulations” previously observed using articulometry and EPG data (Bartle-Meyer, Goozée, et al., 2009; Hardcastle, 1985, 1987). Characterizing apraxic speech using articulometry, in conjunction with acoustic measures, is particularly valuable given that different patterns of articulatory miscoordination can lead to a single acoustic percept (Byrd & Harris, 2007; Pouplier & Hardcastle, 2005) and that a speaker’s sub-phonemic articulatory errors can give rise to a listener’s perception of unintended phonemes (Buckingham & Yule, 1987).

To be specific, we investigate (a) gestural coordination in speech errors made by a patient with AOS as compared with typical speakers and (b) covert (silent) articulation in both repetitive and nonrepetitive speech of the apraxic individual that could possibly, under acoustic analysis alone, go unnoticed or be misinterpreted as simple substitution or deletion errors. Errors made in word pair repetition tasks can provide useful insights into a speaker’s ability to coordinate vocal tract gestures appropriately in order to produce linguistically meaningful segments. By observing coordination (and discoordination) patterns in apraxic speech during simple word pair repetition, we are able to probe, given tightly controlled variables (regularly alternating gestural targets), (a) the variability exhibited by apraxic speech and (b) the level(s) at which the system breakdown is likely to occur in AOS. Production of word pairs, such as *top cop*, requires articulatory gestures to be frequency locked in a 1:2 ratio; the tongue gestures associated with /t/ production occur only once for every two instances of the lip gesture associated with /p/ (and similarly for the tongue gestures associated with the /k/). When typical speakers are given the task of repeating these word sequences at a slow rate, they typically are successful in maintaining the appropriate 1:2 frequency-locking ratio between the tongue and lip gestures. As speech rate increases, however, typical speakers tend to produce “gestural intrusions,” coproducing intended onset /t/ gestures with intrusive /k/ gestures and conversely. This can be understood as the system slipping into the intrinsically more simple, stable frequency-locking pattern of 1:1 such that each of the /t/ and /k/ gestures occur once for every coda /p/ gesture produced (Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007; Saltzman & Munhall, 1989). Prior studies using EPG provide indirect evidence that the frequency of these gestural intrusion errors or “misdirected articulatory gestures” is higher in apraxic speech than in typical speech (Edwards & Miller, 1989; Pouplier & Hardcastle, 2005; Sugishita et al., 1987; Washino, Kasai, Uchida, & Takeda, 1981).

Real-time MRI is an ideal modality with which to investigate apraxic speech as it is minimally invasive to the subject and allows for unobstructed viewing of articulatory activity in all parts of the vocal tract over time. Real-time MRI (Bresch, Kim, Nayak, Byrd, & Narayanan, 2008;

Narayanan, Nayak, Lee, Sethy, & Byrd, 2004) is a type of structural imaging specifically designed to reveal the state of the speech articulators at 45-ms intervals (it should not be confused with functional magnetic resonance imaging [fMRI], which examines blood flow in the brain). Because of this, rtMRI aids in the identification and quantification of silent or otherwise hidden speech gestures that may not be detected in the acoustic speech signal that is traditionally used to transcribe disordered speech (Fromkin, 1973). Although other methods of articulometry—for example, electromagnetic articulography (Perkell et al., 1992) or X-ray microbeam (Westbury, Turner, & Dembowski, 1994)—offer high temporal and spatial resolution, they provide information about specific flesh points and may not capture aspects of natural articulation as the receiver coils are known to cause interference with patients’ speech (Katz, Bharadwaj, & Stettler, 2006), which could be expected to have even more serious consequences for an already compromised speaker.

Method

Participant

The participant was a 58-year-old, right-handed gentleman diagnosed with the nonfluent variant of primary progressive aphasia. His first symptoms were sound distortions, phonemic errors, and agrammatism.

In April 2007, at the age of 57, he was diagnosed at the University of California, San Francisco, Memory and Aging Center with the nonfluent variant of primary progressive aphasia, which is defined by the presence of agrammatism and/or AOS (Gorno-Tempini et al., 2011). Although very early in the course of his disease, he showed clear evidence of both agrammatism and AOS (see Table 1). His AOS was rated mild (2 on a 7-point scale), and his

Table 1. Speech, language, and cognitive progression of the patient.

Symptom	April 2007	May 2008	Feb. 2010
Apraxia of speech (MSE, 7)	2	3	4
Dysarthria (MSE, 7)	0	0	3
Fluency (WAB, 10)	9	5	4
Naming (BNT, 15)	14	13	8
Single word comprehension (WAB, 60)	60	59	54
Repetition (WAB, 80)	64	70	63
Expressive morphosyntax (28)	17	7	1
Receptive morphosyntax (55)	45	39	33
Mini Mental Status Exam (30)	27	26	23
CDR	0.5	1	1
CDR sum of box scores	2	4.5	4.5

Note. The expressive morphosyntax test was from Goodglass, Gleason, Bernholtz, and Hyde (1972), and the receptive morphosyntax items comprised 11 subtests of the Curtiss-Yamada Comprehensive Language Evaluation (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004). MSE = motor speech evaluation (Wertz et al., 1984); WAB = Western Aphasia Battery (Kertesz, 1982); BNT = Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983); CDR = clinical dementia rating (Morris, 1993).

aphasia was also mild, with deficits largely restricted to expressive and receptive morphosyntax and impaired repetition (reduced verbal working memory span).

The speech imaging study was carried out in May 2008. At this second time point, the patient's speech and language had declined markedly (see Table 1), and his AOS was now rated mild–moderate (3 on a 7-point scale). His aphasia remained largely restricted to morphosyntax and repetition. A structural brain MRI revealed moderate left-lateralized atrophy (see Figure 1).

A third and final speech and language evaluation was carried out in February 2010, when the participant was 59 years old. Further declines were apparent, and his AOS was now rated moderate (4 on a 7-point scale). His AOS was now accompanied by mild–moderate dysarthria. His aphasia had become more marked, with further declines in morphosyntax as well as production and comprehension of single words.

The participant continued to decline and died in August 2013. An autopsy was carried out. The primary neuropathological diagnosis was corticobasal degeneration with tau-immunoreactive inclusions. The cortical region with the most marked neuronal loss was the left inferior frontal gyrus (pars opercularis). There was also bilateral hippocampal sclerosis, with marked neuron loss in all hippocampal subfields and adjacent entorhinal cortex.

All data presented and analyzed in this study were acquired at the second time point (see Table 1), before the onset of dysarthria, when the participant presented with mild–moderate AOS. The participant gave informed written consent, and the study was approved by the institutional review boards at the University of California, San Francisco, and the University of Southern California.

Procedure

Dynamic imaging of the participant's speech production was acquired using a custom rtMRI protocol (Bresch

et al., 2008; Narayanan et al., 2004). The subject lay supine in the scanner and was able to communicate with experimenters through an intercom system for the duration of the experiment. While lying in the scanner bore, the subject was recorded while producing a diverse corpus of spontaneous speech and three different experiments: (a) isolated word naming, organized in repeated blocks of words; (b) repetition of short phrases; and (c) a self-paced word pair repetition task (see Appendixes A–B).

Stimuli

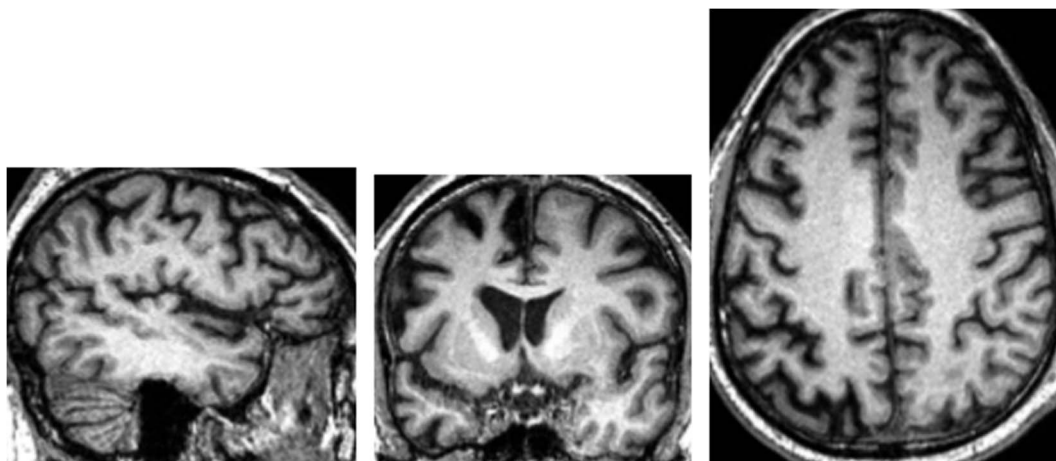
The subject produced spontaneous speech in response to questions on general topics of interest. He was then prompted to repeat a series of short phrases (see Appendix A) and single words, presented orally by the experimenter, 10 times in random order. The words were two-syllable and four-syllable words, which were chosen to contain a wide variety of consonants and vowels (see Appendix B). Moreover, many of the words were selected because they have properties known to be challenging for individuals with AOS—that is, consonant clusters and travel between different places of articulation.

In the final task in this study, the subject was asked to repeat the sequence *cop top* at his highest rate possible and for the longest duration possible without the aid of a metronome over two separate trials. This sequence has been used in several studies (Goldstein et al., 2007; Pouplier & Goldstein, 2010) to elicit speech errors in typical speakers, and an expected rate of gestural intrusion error is known.

Data Acquisition

Image data were acquired on a 1.5T GE Signa scanner, using a 13-interleaf spiral gradient echo pulse sequence (TR = 6.5 ms, FOV = 200 × 200 mm, flip angle = 15°) and a head and neck receiver coil. A midsagittal scan plane (3 mm slice thickness) was used; image resolution in the

Figure 1. Structural brain magnetic resonance imaging showing left-lateralized atrophy in study participant. Left panel: midsagittal plane. Center: coronal plane (left hemisphere to left of image). Right: axial plane (left hemisphere to left of image).



sagittal plane was 68×68 pixels (2.9×2.9 mm). New image data were acquired at a rate of 11.2 frames per second and reconstructed at 22.41 frames per second using a sliding window technique. Audio was recorded inside the scanner at 20 kHz simultaneously with the MRI acquisition and subsequently noise reduced (Bresch, Nielsen, Nayak, & Narayanan, 2006). The resulting video and audio recordings allow for dynamic visualization of the entire midsagittal plane of the subject's vocal tract during speech.

Articulatory Analysis

The audio and MRI video recordings along with the MRI frame sequences corresponding to each task in the experimental corpus were carefully examined. Recordings were inspected and audited to locate all instances of dysfluencies in production, prosodic abnormalities, and speech errors. For each utterance in each task, articulatory coordination among gestures of the lips, tongue tip, and tongue body was examined in the temporal vicinity of the target item. Articulatory activity in each region of interest was tracked over the duration of the utterance and compared to fluent speaker productions of the same utterance when possible.

Articulatory analysis was conducted using a temporal analysis method specifically designed to track the formation and release of constrictions in targeted regions of the vocal tract (labial, alveolar, velar; see Figure 2). Pixel intensity time functions estimating constriction formation and release were automatically generated by calculating the mean intensity of pixels within each region. A pixel within each region was manually chosen, and pixels falling within a radius of three pixels of the chosen pixels were included. This method of estimating articulatory activity in a specified region of interest has been found to provide a robust estimate of constriction degree in noisy data (Lammert, Proctor, & Narayanan, 2010). This technique requires less manual correction and is more computationally efficient than other techniques relying on segmentation of articulators along air-tissue boundaries. Furthermore, this method has been

Figure 2. Mean midsagittal magnetic resonance imaging slice showing vocal tract regions (labial, alveolar, velar) within which articulatory activity is estimated from correlated pixel intensity (details in Lammert et al., 2010).



shown to generate articulatory traces that are directly comparable to those obtained in electromagnetic articulograph studies (Proctor et al., 2011) and therefore facilitates comparison of patterns of apraxic gestural coordination with those previously reported for fluent speech.

Error Analysis and Quantification

In addition to constriction tracking, all acoustic and articulatory data from the self-paced word pair repetition task were analyzed quantitatively (see below) and compared with data from typical speakers performing the same tasks reported in past studies (Goldstein et al., 2007). Incidence of speech errors made by the apraxic patient was calculated as the ratio of onsets containing a gesture deletion or intrusion error (error identification method below) to the total number of onsets.

Nontarget gestures in the word pair repetition task were classified as speech errors if the error threshold in the pixel intensity function was exceeded. The error threshold for a given intensity function (i.e., region of interest) was defined as the average of the interquartile means during (a) target gestures with a constriction in that region of interest (e.g., alveolar region intensity during /t/) and (b) target gestures with a constriction distal to that region of interest (e.g., alveolar region intensity during /k/). Interquartile means were used in order to lessen the influence of potential speech errors on the error threshold itself. See Pouplier (2008) for further description of the method.

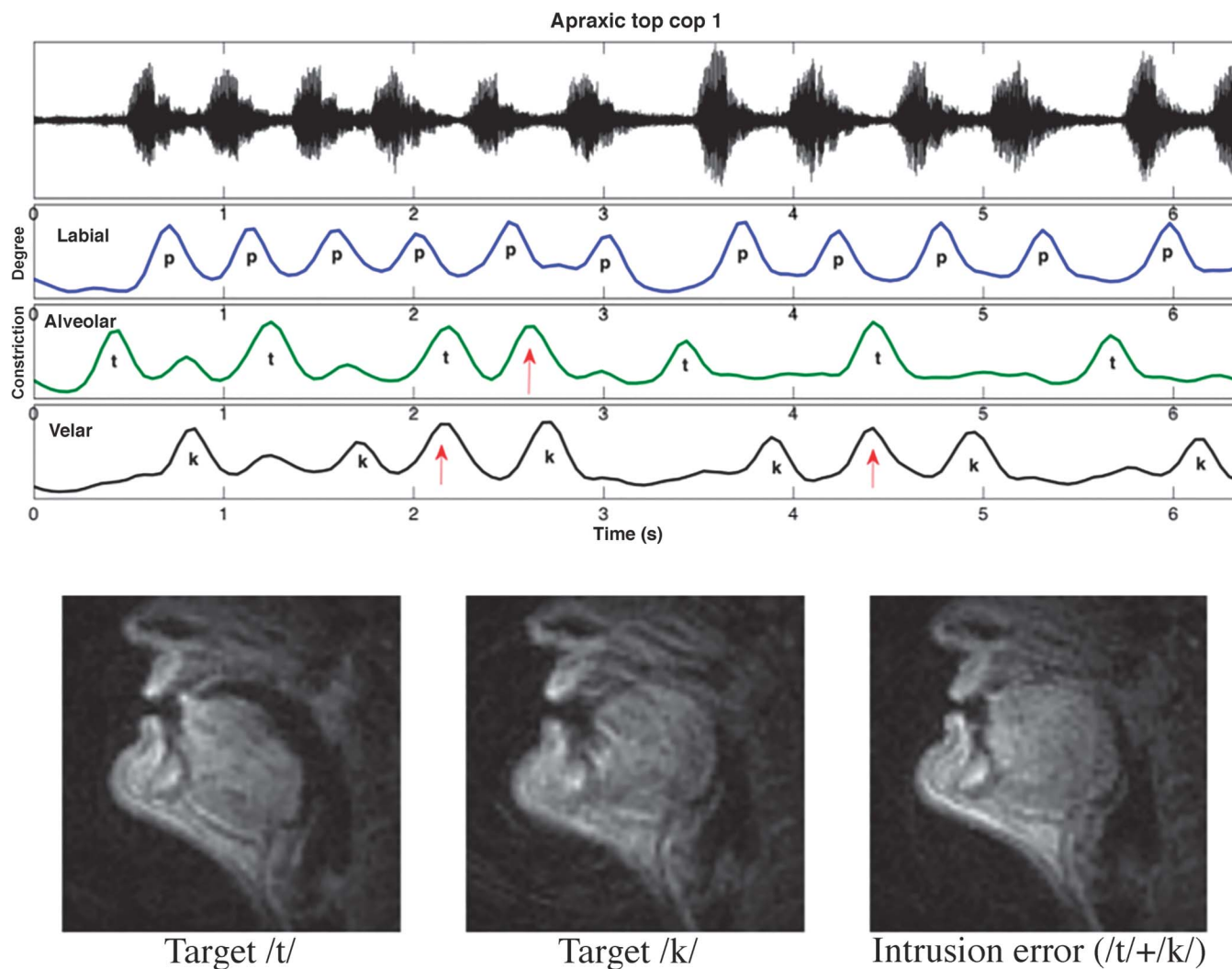
Results

Gestural Intrusions in Repeated Word Pairs

Auditory observation of the acoustic records of the two repeated word pair trials reveals a subtle, although distinctive, difference between the two; the second contains substantially more auditorily perceptible errors than the first. No difference in speech error rate between the apraxic patient and a typical speaker is perceptible for the first repeated word pair trial. However, upon examining the articulatory data, it is clear that this is not the case. Despite the lack of perceptible speech errors in the acoustic signal, constriction activity was observed in regions of the vocal tract where they are not expected given the phonological structures of the target word pairs. Time functions estimated using mean pixel intensity in the labial, alveolar, and velar regions of interest reveal that gestural intrusions were made frequently by the apraxic patient. (Figure 3 shows three such intrusions in red arrows in the first fifth of the first trial; see online Supplemental Material S6.)

When comparing the apraxic patient's gestural intrusions in the word pair repetition task to those reported in published speech error studies on typical subjects performing the same linguistic task (Goldstein et al., 2007), we find that even though the apraxic patient performed the task at a much slower rate than the typical population (76 bpm increasing to 120 bpm), he produced intrusions far more frequently than typical subjects (approximately 61% of all

Figure 3. Top: Acoustic waveform and time-aligned estimated constriction functions (labial, alveolar, velar) in /top–kop/ repetition task. Bottom: magnetic resonance imaging frames showing articulatory postures for target /t/, target /k/, and first intrusion error (coproduced /t/ + /k/).



onsets for the apraxic patient vs. approximately 15%–20% for typical subjects; $n = 7$). The apraxic patient produced more intrusions than typical subjects even in the trial in which the apraxic speaker made the fewest number of errors, and the typical speakers, even at the highest rate (120 bpm), made fewer errors than the apraxic speaker, whose repetitions were self-paced.

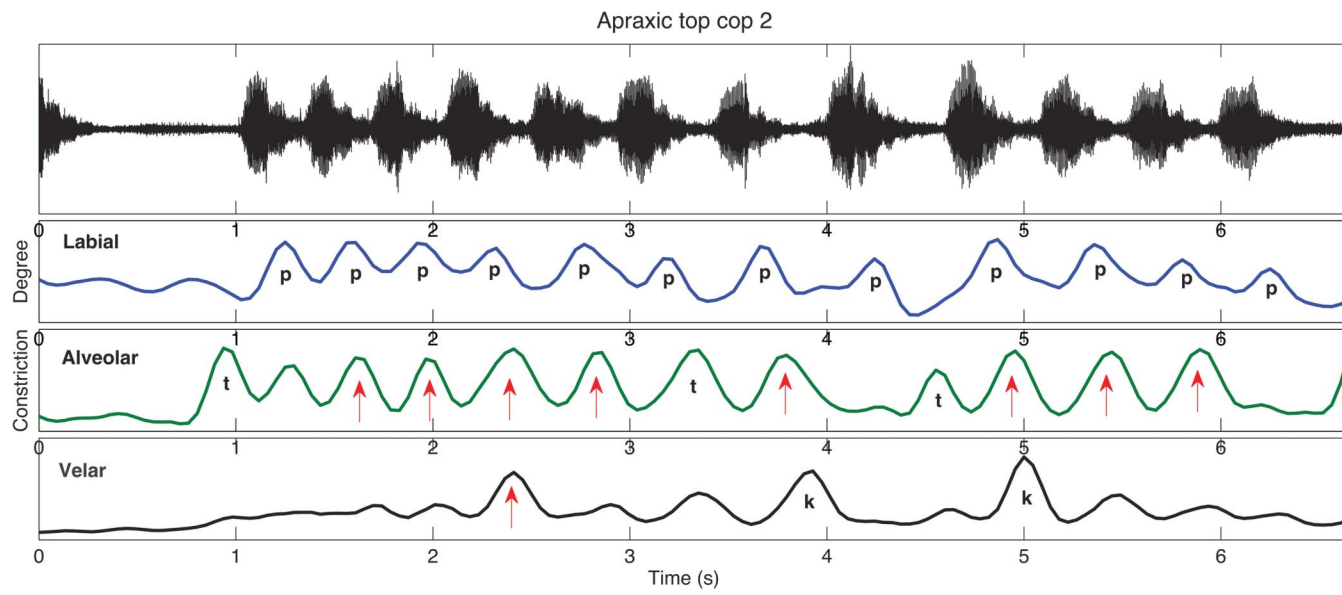
Visual inspection of the MRI frames in the vicinity of the errors, identified by the pixel intensity time functions, confirmed what is shown in the time functions, namely that they are not simply substitution errors whereby an erroneous gesture replaced the target gesture. Instead, the errors at hand are gestural intrusion errors whereby the erroneous gesture (red arrow) was produced synchronously with the target gesture. Further, we observe that in most cases of gestural intrusion errors, neither the target gesture nor the intrusive gesture appears to be reduced in amplitude.

Articulatory Coordination in Repeated Word Pair Trials

Articulatory data in Figure 3 reveal that during the first word pair repetition trial, the speaker produced alternating tongue tip (*/t/*) and tongue dorsum (*/k/*) gestures before coda */p/*. In some instances, gestural intrusion errors occurred whereby the onset gestures for */t/* and */k/* were coarticulated (red arrows).

In the first half of the second word pair repetition trial, the speaker's deviation from the target sequence was clearly audible; no regular alternations between the target onset gestures for */t/* and */k/* were produced (see Figure 4). Instead, tongue tip (*/t/*) and labial (*/p/*) gestures were generally produced synchronously, and the dorsal gesture for */k/* was rarely produced when expected. In this trial, the type of errors made differed from those made by typical speakers; the target onset */t/* and target coda */p/* were produced synchronously. In sum, on the first repetition of the word

Figure 4. Acoustic waveform (top) and time-aligned estimated constriction functions (labial, alveolar, velar) in second /top–kop/ repetition trial. Labial and tongue tip gestures coordinated in phase (synchronously; arrows). Dorsal gestures are missing at expected times.



pair sequence (Figure 3), the apraxic speaker made the same types of errors that typical speakers make, although at a higher rate. In the second repetition (Figure 4), the apraxic speaker made errors not made by typical speakers.

Hidden Intrusion Errors in Nonrepeated Speech

Articulatory data from nonrepetitive apraxic speech reveal that the occurrence of silent gestural intrusion errors is not limited to the context of repetitive speech tasks; these errors were frequently present in speech produced during a repetition task. Figure 5 (bottom panel) illustrates articulatory activity during production of the phrase “I can type ‘bow know’ five times.” Articulation of the same utterance by a typical 25-year-old male speaker of American English is provided for comparison (top panel). Constriction time course functions for both speakers were generated using the same method from regions of interest centered at equivalent locations in each speaker’s vocal tract.

As expected, the apraxic utterance is slower (5.45 s) than the fluent equivalent (1.63 s) in this case by a factor of three. Comparison of typical and apraxic productions of this utterance reveal that the fluent speech is characterized by smoother formation and release of target gestures, and the constriction functions for the apraxic speech reveal partial gestural intrusions and separation of adjacent labial gestures that would typically be coarticulated. A full tongue tip intrusion error can be observed during the labial closure for the initial /b/ of *bow*.

In addition, we find that the apraxic patient produced gestural intrusion errors while attempting to articulate individual words during a repetition task. Acoustic analysis of one of the subject’s responses to the stimulus item *federation* suggests a form that might be represented in close

transcription as [ʌdəʊɛɪfən] (see online Supplemental Material S4). Analysis of the MRI frames acquired during the production of this utterance shows that the initial acoustic segment (transcribed as /ʌ/) is not caused by simple “anticipatory substitution” whereby the gestures for the word-medial /ʌ/ replaced those required for the target /f/. Instead, the initial labial gesture of target /f/ and an (erroneous) anticipatory lingual intrusion gesture for /ʌ/ are observed to be synchronously produced. It is important to note that contrast against the wider and less protruded labial posture during the production of intervocalic /ʌ/ in *ceremony* (Figure 6, far-right frame) suggests that the labial constriction in Token 2 pertains solely to the /f/ target constriction. MRI frames acquired during target and errorful productions of the initial portion of the word *federation* (see online Supplemental Material S5) as well as during intervocalic /ʌ/ are contrasted in Figure 6.

As illustrated in Figures 5 (see online Supplemental Material S2) and 6, rtMRI is particularly useful in capturing dynamic images of gestural intrusion errors or “double articulations” that skilled clinicians have long observed in apraxic speech and that have been previously evidenced in flesh point and EPG data (Bartle-Meyer, Goozée, et al., 2009; Hardcastle, 1985, 1987).

Multiple Initiation Gestures in Imitated Speech

The pixel intensity time functions (see Figure 7) reveal that a single, silent tongue tip gesture (arrow) surfaces preceding full (and audible) production of the coronal-initial word *know* [no] in the utterance “I can type ‘bone know’ five times.” Evidence for unphonated gestural attempts or rehearsals is also illustrated in Figure 8, in which three silent tongue tip gestures are observed in the interval before

Figure 5. Acoustic waveforms and time-aligned labial, alveolar, velar constriction functions for fluent (top) and apraxic (bottom) utterances, “(I) can type ‘bow know’ five times.” Top: Fluent utterance (25-year-old typical American male) showing high degree of gestural overlap, smooth gestural transitions, and shorter overall duration. Bottom: Constriction functions reveal unphonated tongue tip intrusion (arrow) during labial closure for /b/ in apraxic utterance.

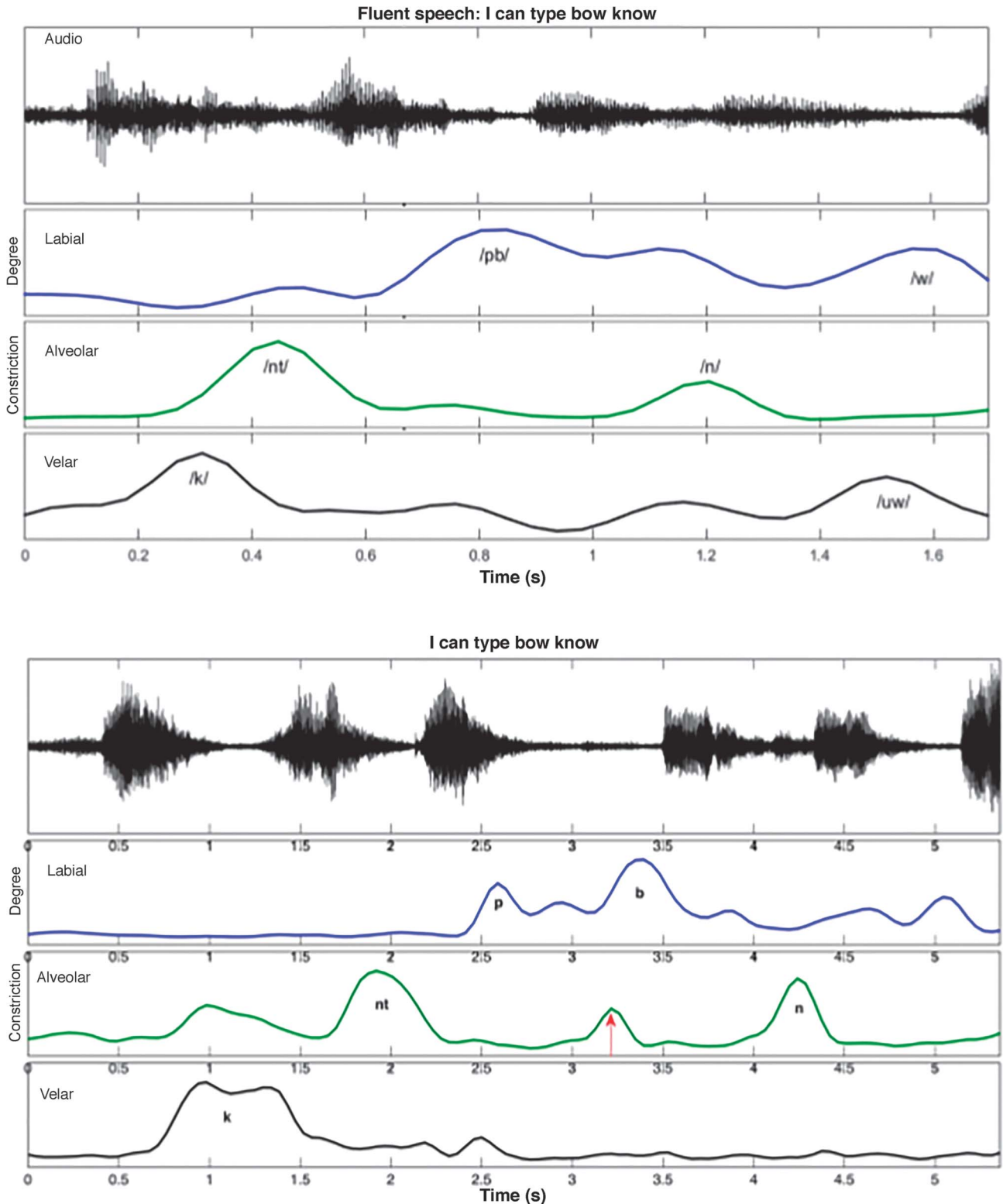


Figure 6. Two productions by apraxic subject of initial fricative in *federation*. Left: Target production. Right: Labial constriction coproduced with intrusive lingual gesture corresponding to tongue posture observed during /ɪ/ production later in same word. Labiodental frication is not discernible in acoustic signal of Token 2.



the vocalized production of the coronal stop /t/ that initiates successful production of the complete word *temperatures*.

Multiple Initiation Gestures in Spontaneous Speech

Consistent with previous observations of “silent groping” and “false starts” (Duffy, 2013; Ogar et al., 2005), the data at hand (see Figures 7 [see online Supplemental

Material S1] and 8) reveal that multiple initiation gestures are frequently made in the apraxic patient’s speech. Further, we find that multiple initiation gestures were produced in both spontaneous speech and in speech produced during the repetition task (imitated speech). The segments produced word-initially by the apraxic patient in spontaneous speech that exhibited multiple initiations included /t/, /d/,

Figure 7. Covert tongue tip gesture during first (silent) attempt at producing coronal-initial word *know* in the utterance, “I can type ‘bone know’ five times.”

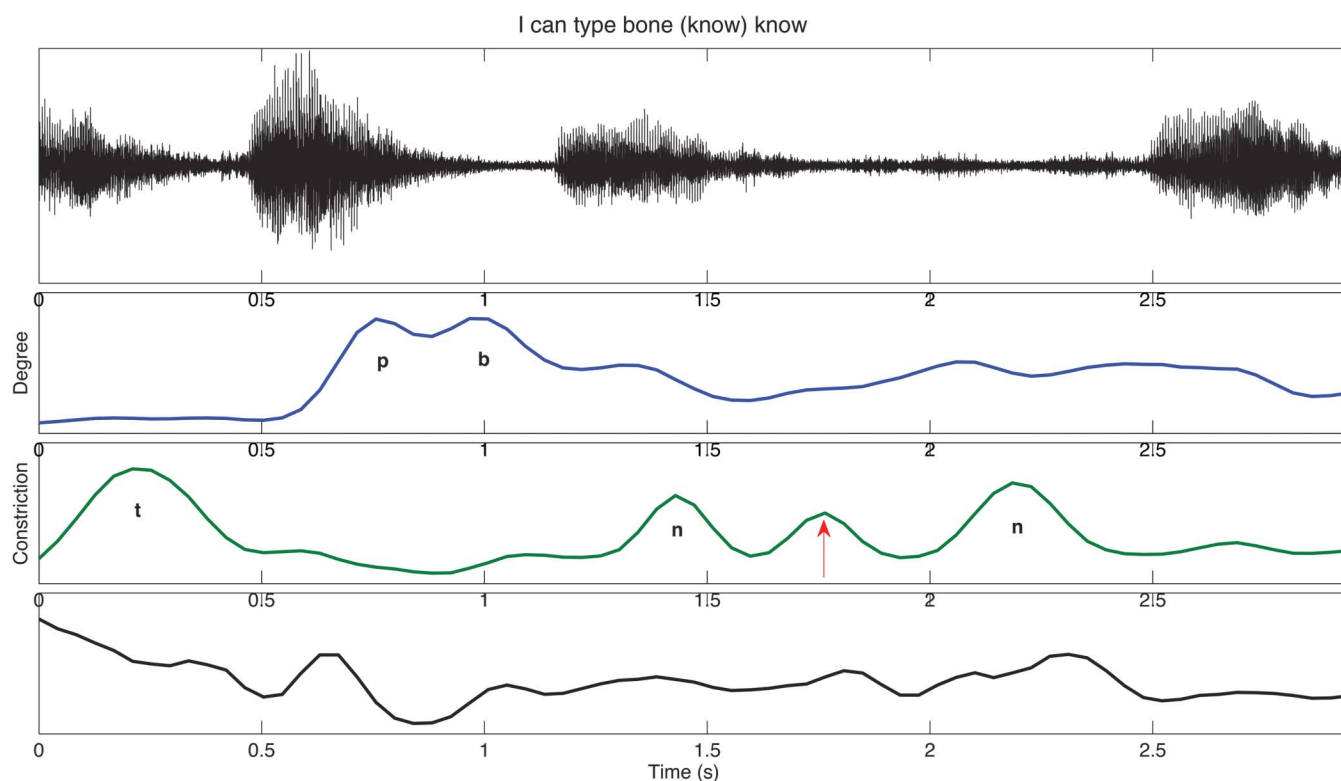
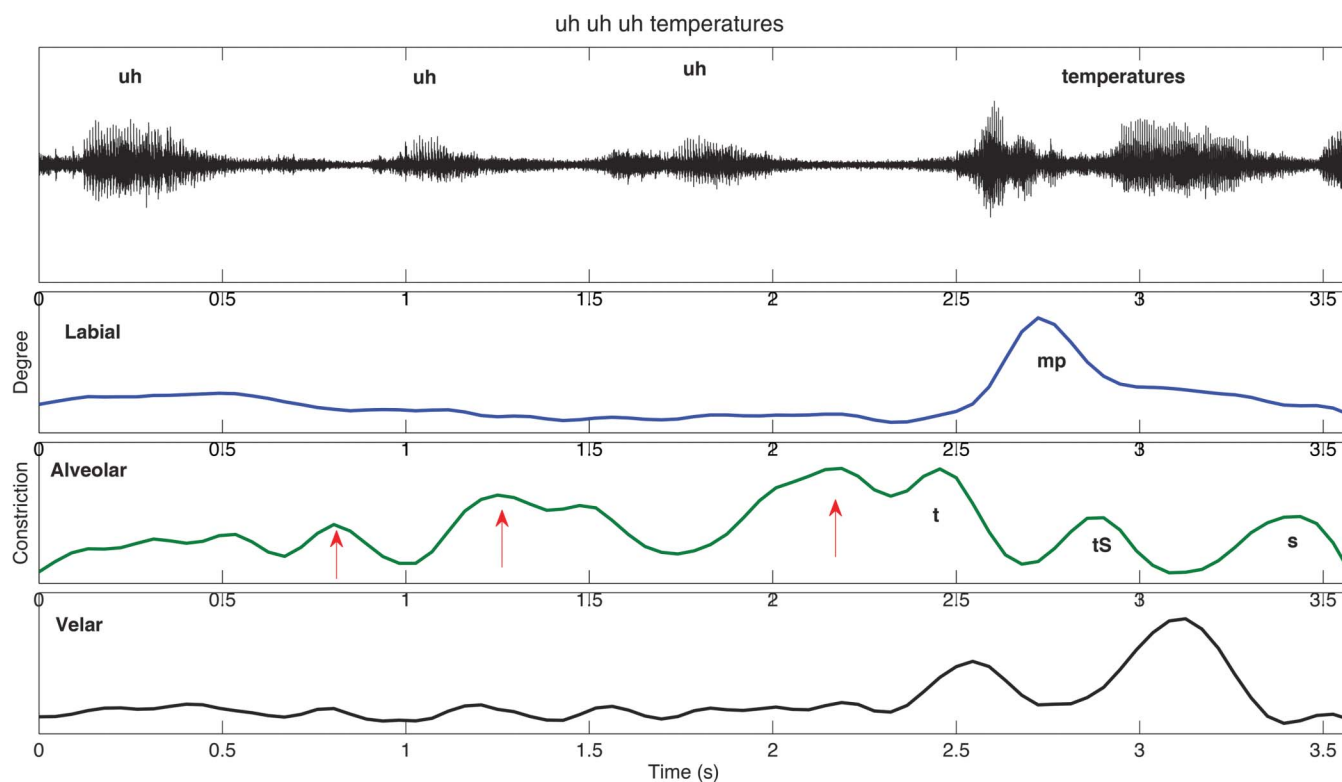


Figure 8. Three silent tongue tip gestures preceding successful (vocalized) production of coronal-initial word *temperatures*.



/g/, /dʒ/, /w/, /f/, /s/, /m/, and /l/. Tokens produced with multiple initiation gestures are defined as those in which a visible articulatory gesture occurs at least once before complete production of the word.

Table 2 illustrates the segments for which multiple initiation gestures were produced, along with their prevalence across all tokens beginning with the given segment (e.g., the ratio of /w/ productions involving multiple initiations to the total number of /w/ productions). It is noteworthy that all of the segments except /d/ and /g/ that caused difficulty for the apraxic patient require coordination of more than one vocal tract gesture (e.g., /w/ requires lingual and labial gestures; /t/ requires glottal and tongue tip gestures), and in many cases, more than one supralaryngeal gesture is required.

Table 2. Prevalence of segments requiring multiple initiation gestures.

Segment	Prevalence	<i>n</i>
/w/	100%	3
/dʒ/	50%	2
/d/	50%	2
/f/	40%	5
/s/	37.5%	8
/l/	33%	6
/g/	25%	4
/t/	25%	4
/m/	16.6%	6

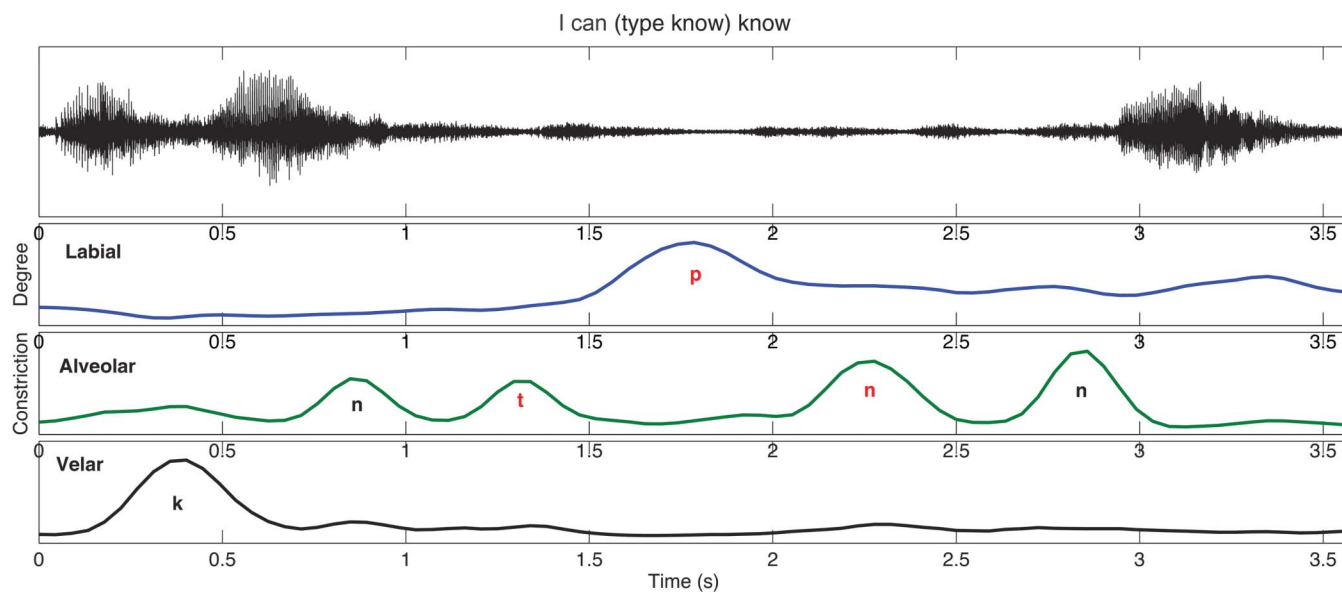
Hidden Articulation in Imitated Speech

It is interesting that we observe, in multiple instances, that the patient fully produced all appropriate supralaryngeal consonantal gestures for monosyllabic words but failed to produce them with phonation. Articulatory organization during the execution of the target sentence “I can type ‘bow know’ five times” (produced “I can type know ‘know [stumbles] bow know’ five times”) is illustrated in Figure 9 (see online Supplemental Material S3), in which the constituent lingual and labial gestures corresponding to the consonants in *type* and *know* can be clearly identified although, as evidenced by the acoustic waveform, produced silently.

Discussion

A major contribution of the present work is to demonstrate that rtMRI is capable of detecting and quantifying characteristics of apraxic speech that have been previously observed clinically and through analysis of both acoustic and kinematic data. Using the rich, dynamic data provided by rtMRI, we come to a fuller understanding of the pathomechanisms underlying AOS. A key contribution of this study is to demonstrate that although intrusive articulations appear to occur at a much higher rate in apraxic speech throughout a larger range of speech conditions, the intrusive gestures produced by the apraxic speaker demonstrate variable coordination patterns, some of which are timed in the same ways that have previously been reported for

Figure 9. Silent production (see attenuated acoustic signal [top], red text [bottom]) of entire sequence *type know* (/taɪp nou/) in utterance “I can (type know) know...”



typical populations producing some of the same tasks under duress (see Figure 3), and others are timed in ways that are not observed in typical speakers (see Figure 4). It is most important to note that the speech errors detected in the apraxic speech overwhelmingly demonstrate an in-phase (synchronous) relationship with the target gestures with which they are coproduced.

The data presented here reveal that the subject’s speech was produced with many unexpected partial and complete gestural intrusions. The details of much of this atypical articulatory activity would not necessarily always be apparent to a listener or clinician because the impact on the acoustic speech signal falls within the range of allophonic variation for the same segments produced in other phonological environments even in typical speakers’ English. Because phonological perception is categorical by nature, auditors (except those carefully trained) are predisposed to ignore many of the underlying articulatory subsegmental variations that appear to be pervasive in apraxic speech, and even if they are noticeable, their frequency and magnitude may be difficult to quantify. A major contribution of the approach to studying disordered speech that we present here is that it provides a method of overcoming this limitation on quantification.

Despite auditorily perceptible differences in the acoustics of typical speech and apraxic spontaneous speech and repeated speech, results of the self-paced word pair repetition task suggest that the auditorily perceptible error rate in the apraxic speaker’s first word pair repetition trial (see Figure 3) might be similar to that produced by a typical speaker. Upon investigation of the rtMRI data, however, we find that there are considerable differences in the articulation patterns of each and also considerable variability in the

patterns produced by the apraxic speaker. The results of the repeated word pair task illustrate two hallmarks of AOS: token-to-token variability and the tendency to produce articulatory gestures synchronously. In the first trial, the apraxic speaker erred at a high rate and produced both targeted and erroneous patterns of gestural coordination that typical speakers do. In the second trial, erroneous patterns of gestural coordination were made that are certainly deemed (audibly) atypical and that are not observed in typical speech. Despite the two error patterns observed differing in their articulatory components (i.e., one consisting of synchronous production of target/nontarget onsets and the other consisting of synchronous production of target onsets and codas), they both illustrate the apraxic speaker’s tendency to synchronize articulatory movements as described below.

When considering the tendency toward entrainment that is revealed in the data described in this study, it is important to make the distinction between phase locking and frequency locking. Frequency locking determines the ratio of two oscillators’ respective frequencies and specifies the relative number of times that each is executed per unit time. One-to-one frequency locking occurs when the oscillators corresponding to two articulatory gestures oscillate at the same frequency, causing each gesture to be executed the same number of times per unit time (and possibly, although not necessarily, synchronously). Phase locking, on the other hand, dictates whether the coupled gestures are executed synchronously or sequentially. In-phase (0°) coupling will cause the gestures to be executed simultaneously, and anti-phase (180°) coupling will cause the gestures to be executed sequentially (Nam, Saltzman, & Goldstein, 2009). Although these two higher level modes of coordination (in phase and antiphase) may account for all patterns of articulatory

coordination, cross-linguistically, there are unquestionably subtle differences between languages with respect to gestural timing within these two categories (e.g., fine differences in the timing of glottal and supraglottal gestures that are nonetheless coupled in phase, giving rise to differences in voice onset time in English and Spanish /t/). Thus, it is the case that “synchrony” and “sequentiality,” in this framework, do not imply precise absolute timing but rather are high-level categorizations of in-phase and antiphase coordination patterns. It is possible that in apraxic speech, disruption of both higher level and lower level gestural coordination occurs: at the higher level resulting in full intrusion errors and at the lower level resulting in artifacts, such as atypical voice onset times.

The data provided by this study reveal that the apraxic speaker erroneously slipped into more intrinsically simple, stable modes of articulatory coupling—namely, 1:1 frequency locking and in-phase (0°) coupling—when attempting to produce the target forms that require more complex modes of coupling (1:2 frequency locking and antiphase [180°] coupling), causing an increased number of covert (silent) gestural intrusion errors to surface compared with typical speakers. The presence of frequent, full gestural intrusion errors during the first trial of the word pair repetition task demonstrates that the apraxic speaker erroneously used the more simple, stable 1:1 frequency-locking mode of coordination whereby one coronal gesture /t/ was produced for each labial gesture /p/ in place of the target, which requires two repetitions of the labial gesture /p/ for every one repetition of coronal /t/ or dorsal /k/ (1:2 frequency-locking mode).

The tendency toward in-phase articulatory coupling, on the other hand, is illustrated during the second word pair repetition trial, during which no alternation occurs. Instead, coronal /t/ and labial /p/ gestures exhibited phase locking, being produced in phase, and dorsal /k/ sometimes exhibited phase locking with those labial and coronal gestures, being produced synchronously with them as well. In a similar manner, phase locking was illustrated in the apraxic speaker’s production of nonrepeating single words. For example, for the target word *federation*, the apraxic speaker produced a gestural intrusion error during the initial /f/, such that both target /f/ and /t/, targeted to come later in the word, were produced simultaneously. The coproduction of /f/ and /t/ involves executing the gestures in phase, one being produced at 0° with respect to the other. In sum, the apraxic speaker displayed a tendency for multiple gestures to be produced synchronously through phase locking and/or frequency locking.

The silent intrusion errors that surface in both repeated and nonrepeated apraxic speech are particularly interesting when considered along with results of past EPG studies that seem to suggest that apraxic speakers experience difficulty suppressing lingual activity, thus giving rise to errors involving substitution of /t/ and /tʃ/ for other sounds (Sugishita et al., 1987). These data, combined with our findings using rtMRI, raise the question of whether the observed “substitution” errors may, in fact, be gestural intrusion errors involving coproduction of a tongue tip gesture and another

gesture not able to be captured using EPG (e.g., a labial or dorsal gesture). If this is the case, it is likely that the patients’ deficit lies in correctly selecting and suppressing speech gestures and controlling the modes of coordination utilized. That is, a key part of the patient’s struggle lies in resisting the tendency for all articulators to entrain and move in synchronous coordination (Poupier & Hardcastle, 2005). If it is true that AOS involves a stronger tendency for all articulators to entrain, it would be expected that (a) at any given time during speech production, more moving parts of the vocal tract are measurably observed than necessary for the articulatory task at hand, and as a result, (b) tangential velocities of diverse articulator flesh points or velocities of constriction degree change in diverse regions (as estimated by rtMRI region of interest intensities) would correlate more strongly in apraxic speech than in typical speech. By determining how constriction changes in multiple regions of the vocal tract (including the velic and pharyngeal regions) covaried in apraxic and typical speech, rtMRI can be used to substantiate the notion that a key pathomechanism of AOS involves decreased functional independence of articulators, which has been proposed for adults on the basis of kinematic data (Bartle-Meyer, Goozée, et al., 2009) and for children (Cheng, Murdoch, Goozée, & Scott, 2007; Gibbon, 1999; Green, Moore, Higashikawa, & Steeve, 2000). This method may be used to differentiate or draw similarities among acquired AOS, childhood AOS, and typical speech.

In both repetitive and nonrepetitive speech, the apraxic speaker produced multiple repetitions of the same token with a great degree of variability. In nonrepetitive speech, the apraxic patient produced the word *federation* according to target on one repetition and erroneously (with a word-initial intrusion error) on the next. In the word pair repetition task, the specific errors made by the apraxic speaker differ from trial to trial. This is consistent with past work observing that speech errors of apraxic patients are inconsistently distributed, with a speaker making errors on one instantiation of the utterance, then successfully producing it according to target on the next (Staiger, Finger-Berg, Aichert, & Ziegler, 2012; Wertz et al., 1984). What is consistent, however, is that regardless of how the tokens deviating from the target are produced, the preference for the stable modes of in-phase and 1:1 coordination among vocal tract gestures is exhibited. The token-to-token variability exhibited by our patient is consistent with results of past studies describing variability in articulator kinematics (Itoh et al., 1979) and timing relationships among articulators (Itoh, Sasanuma, Hirose, Yoshioka, & Ushijima, 1980) as well as variability in linguo-palatal contact patterns (Hardcastle, 1987) and in abductory and adductory laryngeal gestures (Hoole, Schroter-Morasch, & Ziegler, 1997). As to the mechanism underlying AOS, our results are consistent with the hypothesis that the syndrome affects the formation of “molecular units” consisting of temporally and spatially coordinated articulatory gestures (Browman & Goldstein, 1992; Staiger et al., 2012). Impeded coordination of discrete articulatory gestures in a target sequence can account for the speech irregularities typically exhibited in

apraxic speech, including perceived deletion errors, distortion errors, and insertion errors. Further, although we did not attempt to classify gestural intrusion errors in terms of their magnitude, informally they appear to be gradient in nature—not always produced with the magnitude of intended vocal tract gestures (see Figure 5; Frisch & Wright, 2002; Goldrick & Blumstein, 2006; Goldstein et al., 2007; Laver, 1979). Due to the gradient nature of these errors, their aerodynamic consequences may or not be acoustically perceptible, thus further giving rise to the percept of token-to-token variability.

Consistent with previous observations of “silent groping” and “false starts” (Duffy, 2013; Ogar et al., 2005), our study reveals that multiple initiation gestures are sometimes present during periods of acoustic silence. That these initiation gestures are produced without phonation and that covert articulation of entire words is also found can be explained in two ways. It is possible that these covert (silent) gestures are true initiation attempts for which the patient is unable to initiate laryngeal activity when appropriate. As an alternative, it may be that these gestures serve as articulatory rehearsals of the target words. It is possible that the motor programs or executions for these segments take longer to access or plan, and explicit articulatory rehearsal is beneficial. It is, perhaps, less likely that the multiple preutterance initiation gestures exhibited are true articulation attempts (i.e., that the heavy accessing/planning load would lead to misselection of the required gestures causing supralaryngeal gestures to be produced in the absence of laryngeal activity) given that synchronous laryngeal–articulatory (supralaryngeal) coupling is formed with particularly strong intergestural cohesion. In visually inspecting all frames in the vicinity of multiple initiation gestures, we observe that velum behavior during multiple initiation gestures is generally congruent with velum posture in the target segment—that is, multiple initiation gestures of the tongue tip preceding /n/ are produced with a lowered velum, whereas multiple initiation gestures of the tongue tip preceding oral stops are produced with a raised velum. This serves as preliminary evidence that multiple initiation gestures do not involve miscoordination of supralaryngeal gestures. Nonetheless, in the absence of oral airflow data or data on laryngeal activity, it is possible only to speculate as to what the speaker’s intent for these productions might be.

It is interesting to note, however, that most segments for which multiple initiation gestures are produced in spontaneous speech are those requiring coordination of more than one vocal tract gesture. As stated above, it is possible that the added task of appropriately coordinating the gestures required for the production of multigestural segments presents additional challenges in planning, causing the apraxic patient to exhibit false starts or perhaps (rehearsal-like) explicit articulation as part of the planning process. However, a hypothesis that the production of multigestural segments, no matter their particular constituency, would be equally more motorically difficult for apraxic speakers than single-gesture segments would predict voiceless segments to be more error prone than voiced segments, contrary to evidence suggesting the markedness of voiced

/d/ over voiceless /t/ (de Lacy, 2002; Hamilton, 1996) and evidence from studies investigating voice onset time in apraxic speech (Itoh et al., 1982). Especially in light of the existing research on voice onset time in AOS (Itoh et al., 1982) and apraxic error modeling simulations in German (Ziegler, 2009), a more tenable explanation is that the vulnerability of segments depends on the degree of cohesion that exists between their component gestures and that this degree of gestural cohesion is itself dependent on motor learning of these particular segments and thus may be related to their relative frequencies in a given language. As Ziegler (2009, p. 655) suggests, “Due to the high frequency of occurrence of such co-ordinated patterns [e.g. *glottal abduction and an oral gesture*], gestural synchrony may in these particular instances be a highly overlearned routine, which remains stable in apraxia of speech.”

Our rtMRI data reveal that this apraxic speaker frequently produced covert (silent) gestural intrusion errors in both repetitive and nonrepetitive speech. In addition, it reveals that the patient produced multiple, silent articulatory gestures at the initiation of speech as well as unphonated supralaryngeal consonantal gestures corresponding to monosyllabic words. These findings suggest that, consistent with previous descriptions of AOS, the AOS of the patient at hand is best characterized by disordered selection and temporal organization of articulatory gestures rather than failure to reach gestural targets or a failure of sequential ordering. Because the underlying deficit seems to be one of coordination, a tactile–kinesthetic treatment, such as Prompts for Restructuring Oral Muscular Phonetic Targets, would not likely be most effective in managing the symptoms of AOS. Treatments focused on articulatory tasks that our data suggest are most problematic for the speaker—namely, the production of complex segments or segment sequences requiring less cohesive gestures to be coordinated—would likely be far more effective.

As the results of this case study illustrate, rtMRI is useful in (a) identifying and further characterizing several aspects of apraxic speech that have been previously observed both clinically and using other imaging modalities and (b) providing rich, dynamic articulatory data that inform our understanding of pathomechanisms underlying AOS. Our single-subject pilot study, however, is not without limitations. The lack of large populations of apraxic speakers willing and able to participate in rtMRI studies involving multiple elicitations of an extended test corpus limits us to a quantitative analysis that if made quantitative, would severely lack statistical power. This type of qualitative analysis is consistent with the norms of the field in articulatory phonetics (e.g., Goldstein et al., 2007; Proctor, Lo, & Narayanan, 2015; Ramanarayanan, Lammert, Goldstein, & Narayanan, 2014), in which participant populations are typically small ($n < 10$), and differences between individual speakers and vocal tract morphologies (Lammert, Proctor, & Narayanan, 2013a, 2013b) preclude application of many statistical tests involving pooled or averaged data. Nonetheless, we must be cautious when applying these norms to the investigation of clinical populations in which there is much

more within- and cross-subject variability than in typical populations. This is particularly true in the case of AOS, a disorder of which a hallmark is token-to-token variability.

A further limiting factor in our study is the relatively low number of utterances and repetitions that can be elicited from a speaker struggling with effortful and disfluent speech during a scan session of reasonable duration. Due to these limitations, our participant did not produce all consonant sounds of English in word-initial position during spontaneous speech. Future studies should be designed such that the data acquisition will include a spontaneous speech sample extensive enough to include all consonantal segments of English in word-initial position at least three times.

In addition, our single-subject pilot study was part of a larger study of which speech articulation was only a subpart. For this reason, the single word stimuli have various shortcomings from an experimental phonetic point of view; the gestural content of the onset and coda consonants elicited were not systematically controlled. Future work would benefit from inclusion of stimuli items in which gestural organization is more systematically controlled; it would be possible to translate the *top cop* paradigm into real speech by using words such as *backpack* and phrases such as “The pod cod saw a cod pod” or, placed in an even more natural context, “At my college reunion, we figured out that there were quite a few people named Ken and even more named George. By the end of the weekend, we had counted 10 Kens and 13 Georges.” With recent developments in imaging technology (Lingala et al., 2016), it is possible to detect vocal fold abduction, even in the sagittal slice, of images acquired at an extremely high frame rate, allowing for use of segments that differ only in voicing.

In our illustration of the production of consonant gestures corresponding to monosyllabic words, an extremely low amplitude signal is visible in the waveform. It is likely that this signal, time locked with the consonantal gestures, arises due to rtMRI scanner noise resonating in the speaker’s vocal tract. In order to ensure that these articulations are, indeed, produced without any phonation and are not whispered, utilization of the imaging technique described above (Lingala et al., 2016) to determine if sustained adduction or abduction of the vocal folds are observed at the appropriate points in time or collection of oral airflow data would be required.

The participant’s language and cognitive deficits were relatively mild and did not affect our ability to investigate speech motor control. No studies have systematically compared AOS in neurodegenerative cohorts (primary progressive aphasia and primary progressive apraxia of speech) to AOS with the more common etiology of stroke (Duffy & Josephs, 2012). However, several studies of AOS in progressive patients suggest that many of the same features that define AOS in stroke are also present in degenerative AOS (Duffy, 2006; Josephs et al., 2006; Ogar, Dronkers, Brambati, Miller, & Gorno-Tempini, 2007). It will be important to determine if the results of our study generalize to other patients with progressive AOS and to patients with AOS due to other etiologies.

Conclusion

Using rtMRI, a noninvasive method of observing and quantifying articulatory movement in the entire vocal tract, we demonstrate that an apraxic patient produced covert, intrusive speech gestures in repetitive and non-repetitive speech. Further, we observe that the patient produced silent articulation of consonants corresponding to monosyllabic words and multiple, hidden initiation gestures when attempting to produce segments or segment sequences requiring the coordination of more than one vocal tract gesture. These data suggest that rtMRI is indeed capable of capturing many characteristics of apraxic speech that have previously been described in the literature and can help enrich our understanding of coordination patterns in apraxic speech to provide new insights into the nature of this disorder.

Acknowledgments

Research supported by NIH Grants R01 DC007124-01, awarded to Shrikanth S. Narayanan, and DC008780-05, awarded to Louis M. Goldstein.

References

- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, 88, 148–159.
- Bartle-Meyer, C. J., Goozée, J. V., Murdoch, B. E., & Green, J. R. (2009). Kinematic analysis of articulatory coupling in acquired apraxia of speech post-stroke. *Brain Injury*, 23, 133–145.
- Bartle-Meyer, C. J., Murdoch, B. E., & Goozée, J. V. (2009). An electropalatographic investigation of linguopalatal contact in participants with acquired apraxia of speech: A quantitative and qualitative analysis. *Clinical Linguistics & Phonetics*, 23, 688–716.
- Bresch, E., Kim, Y. C., Nayak, K., Byrd, D., & Narayanan, S. (2008). Seeing speech: Capturing vocal tract shaping using real-time MRI. *IEEE Signal Processing Magazine*, 25, 123–132.
- Bresch, E., Nielsen, J., Nayak, K., & Narayanan, S. (2006). Synchronized and noise-robust audio recordings during realtime MRI scans. *The Journal of the Acoustical Society of America*, 120, 1791–1794.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155–180.
- Buckingham, H., & Yule, G. (1987). Phonemic false evaluation: Theoretical and clinical aspects. *Clinical Linguistics and Phonetics*, 1, 113–125.
- Byrd, D., & Harris, K. (2007). Identifying and evaluating apraxic speech deficits using magnetometry. In *Proceedings of the XVIIth International Congress of Phonetic Sciences* (pp. 2029–2032). Paris, France: International Phonetic Association.
- Cheng, H. Y., Murdoch, B. E., Goozée, J. V., & Scott, D. (2007). Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research*, 50, 375–392.
- de Lacy, P. (2002). *The formal expression of markedness* (Unpublished doctoral dissertation). University of Massachusetts, Amherst.
- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92, 145–177.
- Duffy, J. R. (2006). Apraxia of speech in degenerative neurologic disease. *Aphasiology*, 20, 511–527.
- Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Mosby.

- Duffy, J. R., & Josephs, K. A. (2012). The diagnosis and understanding of apraxia of speech: Why including neurodegenerative etiologies may be important. *Journal of Speech, Language, and Hearing Research, 55*(Suppl.), S1518–S1522.
- Edwards, S., & Miller, N. (1989). Using EPG to investigate speech errors and motor agility in a dyspraxic patient. *Clinical Linguistics and Phonetics, 3*, 111–126.
- Frisch, S. A., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics, 30*, 139–162.
- Fromkin, V. (1973). The non-anomalous nature of anomalous utterances. *Language, 47*, 27–52.
- Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*, 382–397.
- Goldrick, M., & Blumstein, S. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*, 649–683.
- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition, 103*, 386–412.
- Goodglass, H., Gleason, J. B., Bernholtz, N. A., & Hyde, M. R. (1972). Some linguistic structures in the speech of a Broca's aphasic. *Cortex, 8*, 191–212.
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., . . . Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology, 76*, 1006–1014.
- Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research, 43*, 239–255.
- Hamilton, P. J. (1996). *Phonetic constraints and markedness in the phonotactics of Australian languages* (Unpublished doctoral dissertation). University of Toronto, Canada.
- Hardcastle, W. (1985). Profiling lingual-palatal contact patterns in normal and dyspraxic speech. In J. E. Clark (Ed.), *The cultivated Australian: Festschrift in honour of Arthur Delbridge* (pp. 417–427). Hamburg, Germany: H. Buske.
- Hardcastle, W. (1987). Electropalatographic study of articulation disorders in verbal dyspraxia. In J. H. Ryalls (Ed.), *Phonetic approaches to speech production in aphasia and related disorders* (pp. 113–136). Boston, MA: College-Hill Press.
- Hardcastle, W., Gibbon, F., & Jones, W. (1991). Visual display of tongue-palate contact: Electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication, 26*, 41–74.
- Hoole, P., Schroter-Morasch, H., & Ziegler, W. (1997). Patterns of laryngeal apraxia in two patients with Broca's aphasia. *Clinical Linguistics and Phonetics, 11*, 429–442.
- Howard, S., & Varley, R. (1995). III: EPG in therapy using electropalatography to treat severe acquired apraxia of speech. *International Journal of Language & Communication Disorders, 30*, 246–255.
- Itoh, M., Sasanuma, S., Hirose, H., Yoshioka, H., & Ushijima, T. (1980). Abnormal articulatory dynamics in a patient with apraxia of speech: X-ray microbeam observation. *Brain and Language, 11*, 66–75.
- Itoh, M., Sasanuma, S., Tatsumi, I., Murakami, S., Fukusako, Y., & Suzuki, T. (1982). Voice onset time characteristics in apraxia of speech. *Brain and Language, 17*, 193–210.
- Itoh, M., Sasanuma, S., & Ushijima, T. (1979). Velar movements during speech in a patient with apraxia of speech. *Brain and Language, 7*, 227–239.
- Josephs, K. A., Duffy, J. R., Strand, E. A., Whitwell, J. L., Layton, K. F., Parisi, J. E., . . . Petersen, R. C. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain, 129*(Pt. 6), 1385–1398.
- Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test—Second Edition*. Philadelphia, PA: Lea and Febiger.
- Katz, W., Bharadwaj, S., & Stettler, M. (2006). Influences of electromagnetic articulography sensors on speech produced by healthy adults and individuals with aphasia and apraxia. *Journal of Speech, Language, and Hearing Research, 49*, 645–659.
- Kertesz, A. (1982). *Western Aphasia Battery*. New York, NY: Grune and Stratton.
- Lammert, A., Proctor, M., & Narayanan, S. (2010). Data-driven analysis of realtime vocal tract MRI using correlated image regions. In *Proceedings of InterSpeech2010* (pp. 1572–1575). Red Hook, NY: International Speech Communication Association.
- Lammert, A., Proctor, M., & Narayanan, S. (2013a). Interspeaker variability in hard palate morphology and vowel production. *Journal of Speech, Language, and Hearing Research, 56*, S1924–S1933.
- Lammert, A., Proctor, M., & Narayanan, S. (2013b). Morphological variation in the adult palate and pharyngeal wall. *Journal of Speech, Language, and Hearing Research, 56*, 521–530.
- Laver, J. (1979). Slips of the tongue as neuromuscular evidence for a model of speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 21–26). The Hague, the Netherlands: Mouton.
- Lingala, S. G., Zhu, Y., Kim, Y. C., Toutios, A., Narayanan, S. S., & Nayak, K. S. (2016). A fast and flexible MRI system for the study of dynamic vocal tract shaping [Epub ahead of print]. *Magnetic Resonance in Medicine*. <https://doi.org/10.1002/mrm.26090>
- Morris, J. C. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology, 43*, 2412–2414.
- Nam, H., Saltzman, E., & Goldstein, L. (2009). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marisco, & I. Chitoran (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin, Germany: Mouton de Gruyter.
- Narayanan, S., Nayak, K., Lee, S., Sethy, A., & Byrd, D. (2004). An approach to real-time magnetic resonance imaging for speech production. *The Journal of the Acoustical Society of America, 115*, 1771–1776.
- Ogar, J. M., Dronkers, N. F., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2007). Progressive nonfluent aphasia and its characteristic motor speech deficits. *Alzheimer Disease & Associated Disorders, 21*, S23–S30.
- Ogar, J. M., Slama, H., Dronkers, N., Amici, S., & Gorno-Tempini, M. (2005). Apraxia of speech: An overview. *Neurocase, 11*, 427–431.
- Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., & Jackson, M. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America, 92*, 3078–3096.
- Pouplier, M. (2008). The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics, 36*, 114–140.
- Pouplier, M., & Goldstein, L. (2010). Intention in articulation: Articulatory timing of coproduced gestures and its implications for models of speech production. *Language and Cognitive Processes, 25*, 616–649.
- Pouplier, M., & Hardcastle, W. (2005). A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica, 62*, 227–243.

- Proctor, M., Lammert, A., Katsamanis, A., Goldstein, L., Hagedorn, C., & Narayanan, S.** (2011). Direct estimation of articulatory kinematics from real-time magnetic resonance image sequences. In *Proceedings of InterSpeech2011* (pp. 281–284). Red Hook, NY: International Speech Communication Association.
- Proctor, M., Lo, C., & Narayanan, S.** (2015). Articulation of English vowels in running speech: A real-time MRI study. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proc. 18th Intl. Congress of Phonetic Sciences (ICPhS)*. Glasgow, 10–14 Aug. 2015: Paper# 220, 1–4.
- Ramanarayanan, V., Lammert, A., Goldstein, L., & Narayanan, S.** (2014). Are articulatory settings mechanically advantageous for speech motor control? *PLoS ONE*, *9*, e104168.
- Saltzman, E. L., & Munhall, K. G.** (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*, 333–382.
- Staiger, A., Finger-Berg, W., Aichert, I., & Ziegler, W.** (2012). Error variability in apraxia of speech: A matter of controversy. *Journal of Speech, Language, and Hearing Research*, *55*, 1544–1561.
- Sugishita, M., Konno, K., Kabe, S., Yunoki, K., Togashi, O., & Kawamura, M.** (1987). Electropalatographic analysis of apraxia of speech in a left hander and in a right hander. *Brain*, *110*, 1393–1417.
- Washino, K., Kasai, Y., Uchida, Y., & Takeda, K.** (1981). Tongue movements during speech production in a patient with apraxia of speech: A case study. In F. C. C. Peng (Ed.), *Current Issues in Neurolinguistics, a Japanese Contribution: Language Function and Its Neural Mechanisms, Advances in Neurolinguistics: Proceedings of the 2nd ICU Conference of Neurolinguistics*, (pp. 125–159). Tokyo, Japan: International Christian University, Language Sciences Summer Institute; Beaverton, OR: ISBS.
- Wertz, R. T., La Pointe, L. L., & Rosenbek, J. C.** (1984). *Apraxia of speech in adults: The disorder and its management*. Orlando, FL: Grune & Stratton.
- Westbury, J., Turner, G., & Dembowski, J.** (1994). *X-ray micro-beam speech production database user's handbook*. Madison: University of Wisconsin.
- Ziegler, W.** (2008). Apraxia of speech. In G. Goldenberg & M. Miller (Eds.), *Handbook of clinical neurology* (Vol. 88, 3rd series, pp. 269–285). London, England: Elsevier.
- Ziegler, W.** (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, *24*, 631–661.
- Ziegler, W.** (2010). Apraxic failure and the hierarchical structure of speech motor plans: A nonlinear probabilistic model. In R. D. Kent & A. Lowit (Eds.), *Assessment of motor speech disorders* (pp. 305–232). San Diego, CA: Plural.
- Ziegler, W., Aichert, I., & Staiger, A.** (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research*, *55*, 1485–1501.

Appendix A

Utterances Used in Short Phrase Elicitation Task

I can type BONE KNOW five times
 I can type BONE OH five times
 I can type BOW KNOW five times

Appendix B

Items Used in Word Elicitation Task

balloon
 catastrophe
 ceremony
 circumstances
 cumulative
 debate
 delight
 delivery
 double
 federation
 motivation
 motive
 negligible
 repetition
 solitary
 speculative
 statistical
 temple