APHASIOLOGY
2022, VOL. 36, NO. 7, 854–867
https://doi.org/10.1080/02687038.2021.1914815

REPORT

Contribution of acoustic analysis to the detection of vocoid epenthesis in apraxia of speech and other motor speech disorders

Marion Bourqui[a], Michaela Pernon[b,c], Cécile Fougeron[b] and Marina Laganaro[a]

[a]Faculty of Psychology and Educational Science, University of Geneva, Geneva, Switzerland; [b]Laboratoire de Phonétique et Phonologie, UMR, France; [c]Department of Clinical Neurosciences, Geneva University Hospital, Switzerland

ABSTRACT

Background: Vocoid epenthesis within consonant clusters has been claimed to contribute to the diagnosis of apraxia of speech. In clinical practice, the clinicians often doubt about the correct production of clusters as the C-C transition may be minimally disrupted.

Aims: To demonstrate the value of acoustic analysis in clinical practice as a reliable complement to perceptive judgment.

Methods & Procedures: We compared the acoustic signature and the perceptive detection of vocoid epentheses in unvoiced consonant clusters within pseudo-words produced by 40 participants presenting different subtypes of motor speech disorders (including apraxia of speech (AoS) and dysarthria) and matched neurotypical controls.

Outcomes & Results: The results indicate that vocoid epenthesis was acoustically visible in 3 out of 10 participants with AoS, and in one out of 30 participants with dysarthria. One-quarter of these vocoid epentheses was not detected via auditory perception by expert listeners (speech and language therapists) who also made false detections.

Conclusions: The current results indicate that vocoid epenthesis is not systematic at least in mild AoS. Moreover, an important proportion is misdetected by ear, even by expert clinicians, meaning that visualisation of the acoustic signal can be of precious help.

ARTICLE HISTORY
Received 4 January 2021
Accepted 1 April 2021

KEYWORDS
Apraxia of speech; dysarthria; differential diagnosis; acoustic analysis; vocoid epenthesis

Introduction

Apraxia of speech (AoS) and dysarthria are two subtypes of motor speech disorder accompanying different underlying neuropathological and speech planning/programming processes, and they share many clinical signs. One clinical sign that has often been observed in AoS and to a lesser extent in dysarthria is the insertion of schwa/vocoid within consonant clusters. However, such an insertion is not always easy to detect perceptually. In our research, we compared acoustic and perceptual assessment of vocoid epenthesis in the production of pseudo-words containing unvoiced consonant clusters

CONTACT Marion Bourqui marion.bourqui@unige.ch Boulevard du Pont d´Arve 28,1205,Genève,Suisse

produced by four groups of participants with different types of motor speech disorders. First, we will briefly review the definition of AoS and its overlap with dysarthria before focusing on the phenomenon of schwa/vocoid epenthesis.

The definition of AoS and its underlying impairment have long been discussed in the literature. AoS has been defined in a clinical/semiological perspective, based on symptoms description, as well as in a functional perspective, based on the underlying impaired speech production processes.

From the semiological perspective, AoS is characterized by laborious speech (articulatory efforts), non-fluent speech, syllable segregation, variability over time, dysprosody, phonetic distortions and phonemic errors, and frequent false starts and restarts (Allison et al., 2020; Molloy & Jagoe, 2019; Ziegler et al., 2012).

A broad consensus has emerged in the literature about an impaired ability in AoS to retrieve and/or assemble the different elements of the phonetic plans (Blumstein, 1990; Code, 1998; Varley & Whiteside, 2001; Ziegler, 2008, 2009). Patients with AoS have a preserved knowledge/encoding of the phonological form of the utterance (as opposed to phonological impairment) and they do not have a motor disorder per se, such as paresis, ataxia or akinesia that would prevent them from achieving the required speech movements (as opposed to impaired motor speech execution in dysarthria). Thus, the impairment has been located in the transformation of a (preserved) phonological code into a motor plan, which is a form readable by the motor system. This process is called "motor speech planning" or "phonetic encoding" by different theoretical frameworks. On the other hand, dysarthria has been attributed to impaired motor execution or motor speech programming depending on the type of dysarthria (Guenther, 2016; Van Der Merwe, 2020). In such frameworks, motor speech programming refers to the enhancement and specification of the motor plans with motor programs that are muscle specific and form the final signal for motor execution. Thus, motor speech programming encodes the following parameters: muscle tone, direction of movement, speed, strength, range and mechanical stiffness (Van Der Merwe, 2020). It should be noted however that the distinction between motor speech planning and programming is not done in all theoretical framework, and AoS is often described as an impairment at the level of "motor speech programming" as opposed to impaired execution in dysarthria.

Even though AoS and dysarthria may have different underlying impairments, a large amount of speech symptoms overlap between these disorders, ranging from phonetic distortions, reduced speech rate and impaired prosody, even if the degree of overlap varies according to severity and subtypes of dysarthrias. In particular, phonetic distortions, which are our main focus here, are also observed in all types of dysarthria (Darley et al., 1975). In particular, the insertion of schwa/vocoid between syllables or within consonant clusters is frequently reported in the literature (Romani & Galluzzi, 2005; Ziegler, 2009). This phenomenon was already identified in the elaboration of the Apraxia Battery for Adults (Dabul, 1986) and it was claimed to be one of the few speech properties that appears to be unique to AoS, along with abnormal prosody (McNeil et al., 2009). Actually, lengthened intersegment durations (within or between syllables, words, or phrases, and possibly including the intrusive schwa) is still one of the criteria for the differential diagnosis of AoS (Josephs et al., 2012).

Overall, speech has been shown to be particularly vulnerable in AoS for complex syllable structures (consonant clusters) (Galluzzi et al., 2015; Romani & Galluzzi, 2005). Besides simplifications of clusters due to the omission of one (or more) consonants, misproductions of clusters also include distortions, such as the insertion of schwa/vocoid between two consecutive consonants. Aichert and Ziegler (2004) reported that word/syllable onset clusters were more error prone in AoS than word/syllable final clusters and heterosyllabic clusters.

In these studies, distortions in the production of consonant clusters were analysed based on perceptual classification. However, it is acknowledged that phoneme distortions in general and epenthesis in particular are not easily identified perceptively (Bunton et al., 2007; Fonville et al., 2008; K.L. Haley et al., 2012). Given this limitation, acoustic analyses have been applied to better characterize and quantify vocoid epenthesis in neurotypical speakers (see Bürki et al., 2011; Ridouane & Fougeron, 2011). To our knowledge, vocoid epenthesis has not been analyzed acoustically in motor speech disorders, although acoustic analyses have been conducted on other acoustic parameters such as voice (see Kent & Kim, 2003 for a review), speech rate (Kent et al., 2000; Wang, 2002), rhythm (Stuntebeck, 2002), vowel and syllable duration (Blumstein et al., 1980; Laganaro et al., 2012), and voice onset time – VOT – (Auzou et al., 2000; Blumstein et al., 1980; Marczyk et al., 2018; Melle & Gallego, 2012; Ouden et al., 2018), formant frequencies (K. L. Haley et al., 2001; Ouden et al., 2018) and in coarticulation (Tjaden, 2003).

Finally, very few studies have applied the same methods of analysis for the two populations (AoS and dysarthria). To the best of our knowledge, a direct comparison between AoS and dysarthria has only been reported on vowel formants in a small group of participants by Melle and Gallego (2012), and an indirect comparison has been made on VOT (Auzou et al., 2000).

In summary, some of the speech properties of impaired speech such as the schwa/vocoid epenthesis are claimed to be specific to AoS and may be captured more easily with an acoustic signal than by the human ear. This point needs further research. On the one hand, studies on phoneme distortions and on vocoid epenthesis in patients usually rely on perceptual judgment as it is applied in clinical practice. On the other hand, acoustic studies on impaired speech have rather focused on other speech parameters.

The purpose of this study is to demonstrate the value of acoustic analysis in clinical practice as a reliable complement to perceptive judgment. In this paper, we will examine the insertion of vocoids in consonant clusters because of their potential role in the differential diagnosis between AoS and dysarthria. Vocoid epenthesis may be missed by the human ear and may therefore be more reliably captured with acoustic analysis. More specifically, we investigate (1) whether the visualisation of the acoustic signal of the production of consonant clusters can allow us to detect the insertions of vocoids which are not perceived by the listeners and (2) whether vocoid epenthesis as detected with acoustic analysis is indeed specific to AoS.

We acoustically analysed the production of voiceless consonant clusters inserted into pseudo-words in a group of 40 participants with disordered speech, including 10 participants with AoS and 30 participants with different forms of dysarthria, and 40 neurotypical matched participants, and then submitted them to a group of speech and language therapists for perceptual classification.

Methods

Participants

The participants were 40 adults suffering from different types of motor speech disorders (MSD) (dysarthria or AoS) and 40 matched controls recruited in the framework of a larger research project on motor speech disorders (the Swiss National Science Found (SNF) Sinergia Project on Motor Speech Disorders (MoSpeeDi)). Both groups of participants were part of a larger data collection (Fougeron et al., 2018; Laganaro et al., 2020).

All participants were French native speakers, from various parts of French-speaking Switzerland and of northern France. The participants with MSD presented mild or moderate acquired or progressive speech difficulty noticed by the patient and by a speech and language pathologist. We included participants suffering from three types of dysarthria: hypokinetic dysarthria (participants with Parkinson's disease – PD -), mixed dysarthria (participants with amyotrophic lateral sclerosis -ALS-) and another type of mixed/dystonic dysarthria (participants with Wilson's disease – WD-, predominance of hypokinetic and hyperkinetic dystonic types). Participants with AoS suffered from a stroke in the left hemisphere, with no or only very mild language impairment. The demographic and clinical characteristic of the patient groups are reported in Table 1 along with the severity of their MSD. Two measures of severity are provided in Table 1: 1. individual scores from the MonPaGe-2.0.s speech assessment protocol (Laganaro et al., 2020; Pernon et al., 2020), which evaluates different speech dimensions in various speech tasks and provides an overall severity score ranging from 0 to 32 and 2. The severity score from the BECD (Auzou et al., 2000) which is a perceptive judgment of severity ranging from 0 to 4 on 5 characteristics (voice quality, phonetic realization, prosody, intelligibility and naturalness) assessed by 6 independent raters.

The 40 neurotypical participants were matched on age, sex and French regiolect. All the participants gave their informed consent for their participation in this experiment, which was approved by the local ethics committees.

Material

The stimuli are taken from the pseudo-word repetition task in the MonPaGe protocol (Fougeron et al., 2018), which is used to test all the phonemes of the French language and includes four sets of stimuli. For the purpose of the present study, only the "Cluster-pseudo-word set" in which the production of some French clusters is tested in various

Table 1. Biographical and speech characteristics of the groups of participants.

Group	Number of participants (N of females)	Mean age (years)	Mean BECD perceptual severity (range)	Mean MonPaGe severity (range)
Apraxia of speech (post stroke)	10 (N = 6)	52.5	9 (5–14)	6.4 (2–10)
Hypokinetic dysarthria (PD)	10 (N = 2)	64	5 (3–12)	2.2 (0–4)
Mixed dysarthria (ALS)	10 (N = 6)	74.3	4 (1–9)	3.8 (1–9)
Mixed-dystonic dysarthria (WD)	10 (N = 1)	35.2	10 (5–14)	5.6 (5–9)
Controls	40 (N = 16)	53.5	NA	NA

positions within the word was considered. Syllables with onset unvoiced clusters, either CC (for example/sp/and/st/) or CCC (/spl/and/stR/) are associated with CV syllables to compose bisyllabic pseudo-words (e.g.,/laspe/-/spela/). For six items, the consonant cluster syllable was in S1 position (ex./spila/) and for eight items, it was in S2 (ex./la_spe/). The complete list of stimuli is provided in the Appendix A.

Procedure and overall scoring

Participants were installed in a desk chair in a quiet room approximately 70 cm in front of a computer screen. The MonPaGe computerised tool (Fougeron et al., 2018) was used for the stimuli presentation and data collection. The recording was done by a speech therapist specialised in motor speech disorders. Professional audio material was used for the recording (external audiocard Foscurite Scarlet and headmounted Shure SM35-LXR microphone).

The stimuli were presented on the computer screen in both orthographic and pre-recorded audio form simultaneously in order to prevent possible reading difficulties. The presentation time is set by the clinician who moves to the next item by pressing a button after the patient's production. Only one production is recorded for each item.

Each of the 14 CC(C)- pseudo-words was assessed in terms of articulation accuracy via the guided computerized perceptual MonPaGe scoring tool (Fougeron et al., 2018) by two independent raters. The procedure includes: one question per item targeting accuracy in the cluster production, and further probing for potential types of errors in case of error detection of inaccurate production (distortion, substitution, omission, insertion). The kappa inter-rater agreement between the two judges was .76 which is a substantial agreement. A third rater ruled on the non-consensual ratings.

Acoustic analysis

Examination of the transition between consonants in the CC(C)-pseudo-words was done with *Praat* (Boersma & Weenink, 2001) based on the visualization of the signal and spectrogram representations. We have analyzed the signal between the/s/and the plosive unvoiced consonant which follows. Only unvoiced clusters were analysed because the presence of a transitional vocoid is easily detected and may be used in clinical practice. Criteria for determining the presence of a transitional vocoid included: the presence of periodic vibration in the signal, the presence of formants in the mid frequency region (F1/F2 region), and/or an increase in energy which could be attributed to an increase in sonority or periodic event between the voiceless consonants in the cluster (Bürki et al., 2011; Ridouane & Fougeron, 2011). An example of vocoid insertion is provided in Figure 1. When these criteria were met, the duration of the vocoid was measured and the pseudo-word was included in the perceptual analysis (see below).

Perceptive study

As the corpus used for the perceptive study is based on the results of the acoustic analyses, it will be described in the result section.

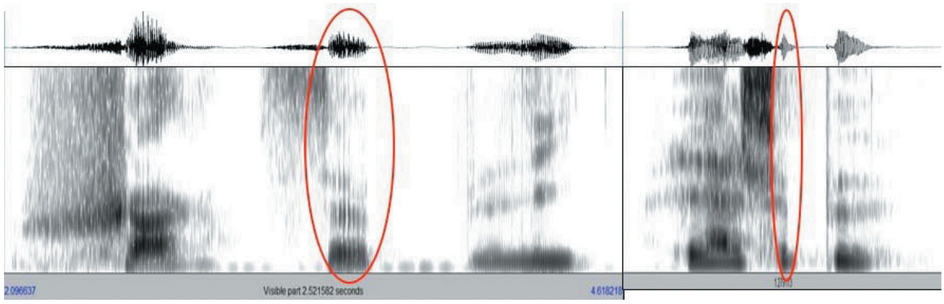

Figure 1. Example of pseudo-word with vocoid epenthesis produced by two AoS participants (/ʃa_spli/on the left,/la_spy/on the right).

The corpus of pseudo-words included exemplars with and without acoustically determined vocoid epenthesis. This was submitted for perceptive judgment to six speech and language therapists (SLTs) with at least 1 year of practice with MSD (mean 4.5 years).

The perceptual judgment was done online via the Qualtrics software, Version July 2020 of Qualtrics. Copyright© 2020 Qualtrics. Each stimulus was presented auditorily in random order and the raters had to answer three questions focusing (1) on the overall accuracy of the pseudo-word, (2) on the production accuracy of the cluster and (3) on whether they hear a voicoid insertion within the consonant cluster (see the experimental design in Appendix B). Before answering each question, they could re-listen the production as many times as necessary.

Results

The results on the overall clusters production accuracy rated with the MonPaGe assessment tool and the acoustically detected vocoids are presented in Table 2. Perceptually mispronunciations of the clusters are observed in all patient groups and not in the control group. The non-parametric Kruskal–Wallis test[1] (run on IBM SPSS Statistics for Windows, Version 25.0) show a significant difference across patient groups (H (df = 3) = 21.116, p < .001) on the number of perceptually detected mispronounciations, with only WD differing from PD in the multiple comparisons (p = .007). Other comparisons are not statistically different (PD vs AoS: p = 0.062; ALS vs WD: p = 0.75; all other p > 0.1). The types of perceptually detected errors were either distortions, substitutions, omissions, or insertions. While all patients may have occurrences of distortions, cases of insertion detected perceptually were relatively rare in the MSD population (18 out of 89 detected mispronunciations were scored as insertions).

By contrast, the vocoid epentheses detected acoustically were observed only in AoS and WD (H (df = 3) = 43.894, p < .001). The multiple comparisons across groups show that insertion of vocoids makes it possible to differentiate AoS from all other dysarthria groups, including WD group (p < .001).

Table 2. A. Mean results per group on the overall production accuracy on clusters rated perceptually and on acoustically detected vocoid epenthesis (range into brackets)B. Results by participant for participants with at least one epenthesis (range into brackets) .

A			
Groups	proportion of perceived cluster mispronunciation[a]	vocoid epenthesis acoustically detected	vocoid duration (in ms)
AoS (*N* = 10)	0.257 (0–0.86)	0.12 (0–0.64)	116 (30–220)
Parkinson's disease (*N* = 10)	0.014 (0–0.07)	0	0
ALS (*N* = 10)	0.043 (0–0.21)	0	0
Wilson's disease (*N* = 10)	0.328 (0–0.93)	0.02 (0–0.21)	26 (22–30)
Healty controls (*N* = 40)	NA	0	0
B			
Participant	MonPaGe severity	Number of epenthesis	Mean and range of vocoid duration (ms)
AoS-1	10	7	143 (120–180)
AoS-2	8.5	9	106 (30–220)
AoS-3	6.5	2	125 (80–170)
WD-1	7	3	26 (22–30)

AoS: Apraxia of speech; WD: Wilson Disease

[a]via the general MonPaGe scoring procedure

The 21 occurrences of vocoid insertion were produced by 3 out of 10 participants with AoS and by one participant with dysarthria from the WD group (see Table 2B). For the participant with WD (WD1), all the epentheses are very short (≤30 ms), suggesting a transitional element between two consonants within a cluster. In AoS participants, there are two types of epentheses: participant AoS-2 produced both very short transitional insertions and longer vowel-like insertions. Participants AoS-1 and AoS-3 only produced long vowel-like insertions (>80 ms).

On this subgroup of four patients, the proportion of epentheses was identical across S1 and S2 position (37.5% in each position). The same holds for the complexity of the cluster: 57.14% in CC and 42.86% in CCC.

Corpus and results of the perceptive study

Based on the results of the acoustic analysis, a corpus of 59 pseudo-words was created including 18 pseudo-words for which acoustic cues of a vocoid epenthesis was detected (all lasting at least 30 ms and produced by participants with AoS) and 34 pseudo-words produced with no acoustically defined epenthesis. These latter tokens included 6 pseudo-words produced without epenthesis by the same 3 participants who also produced items with epenthesis and 28 items produced by 25 other participants with AoS or dysarthria who did not produce any visible insertion.

Overall the six SLTs identified the acoustically visible vocoid insertion with 77% accuracy (range of SLT individual correct classification scores: 47.62–90.48%). They misidentified 5% (range across SLTs: 0 to 19.5%) of the epentheses on clusters where the acoustic analysis did not reveal presence of a transitional vocoid (false positives). Ten epenthesis (all lasting more than 90 ms) were detected by all experts. Non-parametric Spearman rho test was run to determine whether there is a relationship between the duration of vocoid epenthesis and the number of correct identifications by the six SLTs. The results show a positive correlation, which is not statistically significant ($r(18)_s = .438$, $p = .069$). None of the SLTs reached 100% accuracy in detecting the acoustically visible vocoid epenthesis.

Discussion

In this study, we used a visualization of the acoustic signal and spectrogram representations to detect the insertions of vocoids within voiceless consonant clusters produced by participants with different types of MSD and compared the results to the perceptual classification done by SLTs. First, we acoustically analysed all unvoiced CC clusters produced by 40 speakers with different types of MSD and 40 matched controls. Then, we asked six trained SLTs to perceptually detect the presence of vocoids in the unvoiced clusters.

Our results highlighted that vocoid epenthesis was only observed in 3 out of 10 participants with AoS and in 1 participant with mixed dysarthria, and that they are not all detected perceptually. In the following section, we will discuss the gap between acoustic and perceptive detection of vocoid epenthesis before interpreting the results in relationship with the speech features in AoS.

Acoustic versus perceptive detection of vocoid epenthesis

The results show that about one-quarter of vocoid epenthesis detected by the acoustic analysis is not perceived by ear in a group of SLTs. Moreover, none of the six SLTs identified all of the epenthesis. Some of them only detected about half of the epentheses even though they could listen to the pseudo-word as many times as necessary, and despite the fact that the classification only focused on the presence/absence of vocoid epenthesis. In clinical practice, the scoring is mostly done on-line, i.e., without re-listening the production, it is therefore very likely that a larger number of insertions will go undetected by ear. Although most vocoids with duration over 90 ms are identified by all SLTs, the detection of an epenthesis appears to be only marginally related to the acoustic duration of vocoid. Similar mismatch between perception and acoustic results have been reported in studies on covert contrasts in children, where phonemic contrasts are produced by children but are not perceived by the human ear and are therefore transcribed by the listener with the same phonetic symbols (Gibbon & Lee, 2017).

Moreover, some SLTs misidentified some insertions that were not objectively identified through the visualization of the signal and spectrogram representations (false positive). Such misidentifications can be explained by the assimilation of blurred coarticulation from the vocoid epenthesis

Hence, the present results are aligned with previous data supporting the idea that the human ear is often not sufficient to recognize and categorize small changes in fluent speech, which could help us to make an informed clinical differential diagnosis. This was the case in this experiment even when listeners were asked to focus attention to epenthesis and were given the possibility to re-listen to the recorded productions. Therefore, we can conclude that it is important to integrate acoustic analyses into clinical practice. The phenomenon analysed here, vocoid insertion in voiceless consonant cluster, is easy to detect acoustically and may therefore be used in clinical setting to complement perceptive analyses. This same phenomenon has been previously used for the differential diagnosis between AoS and dysarthria. Here for the purpose of the study, pre-established diagnostic groups were used, but in clinical practice, the problem of differentiating AoS from dysarthria may arises for comparable etiologies as for instance, in case of motor speech disorders in the framework of neurodegenerative diseases. In future, it may therefore be interesting to compare the acoustic signature of vocoid epenthesis in dysarthrias and in primary progressive apraxia of speech.

Vocoid epenthesis in AoS

In our literature review, research stated that vocoid epenthesis can be observed particularly with AoS and to a lesser extent in dysarthria. We actually found that only a minority of participants with AoS produced epentheses of vocalic elements in unvoiced clusters, and that one participant with mixed dysarthria also produced such epentheses. Hence, vocoid epenthesis does *not* characterise speech of *all* patients with acquired AoS, and it can also be observed in other pathologies. Josephs et al. (2012) already pointed out that patients with spastic dysarthria also produce insertions of vocalic elements in their CC productions. Hence, this phenomenon, although it appears to be more prevalent in participants with AoS, does not appear to be specific to this disorder as it has been reported in spastic dysarthria previously and in mixed dysarthria here. It is possible that we did not identify more vocoid insertions, especially among AoS participants, because only the transitions between two unvoiced consonants were analyzed, which represents a very small number of items per participant (14). It is likely that larger occurrences of insertion would be observed in this population if a larger corpora, including clusters containing voiced consonants were also to be analyzed.

One might wonder whether vocoid epenthesis can be related to the severity of the disorder. Indeed, among the groups of participants with MSD, AoS and Wilson's disease participants had more severe MSD than the other groups (see Table 1). Nevertheless, within these groups, not all participants with severe disorders presented vocoid epenthesis (Table 2B). For example, participant AoS-4 had a MonPaGe severity of 10; ALS1 had a MonPaGe severity of 7 and they did not produce vocoid epenthesis, while epentheses were detected in the productions of participant AoS-3 who had a MonPaGe severity of 4. Furthermore, the vocoid epentheses in the patient with a mixed dysarthria are characterised by the shortest range of acoustic durations, well beyond the duration associated to vowel-like elements closer to transitional element between two consonants within a cluster (Fougeron et al., 2007).

The observed vocoid epenthesis can have two different interpretations. First, some patients may insert vocalic elements into unvoiced clusters in order to facilitate the articulation of the two successive consonantal elements. This could be seen as a compensatory mechanism, as hypothesized by Marczyk et al. (2018). These authors reported that participants with AoS produced anticipated voicing for the consonants/b, d, g/only in one-third of the cases and that this portion was strongly related to a longer nasal murmur. This observation highlights the compensatory motor adjustment strategies used (in this case, pre-nasalization) to achieve voicing. This same compensation mechanism does not apply here as all the consonants are unvoiced, although other compensation may be at play. Also, Ziegler (2017) showed that, contrary to what is expected, monosyllabic words containing a consonant cluster are less complex than their bisyllabic equivalent containing a vocoid epenthesis due to the formation of iambic forms. By inserting a vocoid between two consonants of a cluster, speakers will violate both faithfulness and markedness constraints.

The second interpretation is that the results on vocoid epenthesis may be consistent with previous studies reporting coarticulation/coordination difficulties in participants with AoS (e.g., Whiteside, 1998; Ziegler & Von Cramon, 1986). One of the manifestations of such impaired coarticulation is represented by lengthened intersegment durations between sounds, syllables, words, or phrases, possibly including the intrusive schwa (McNeil et al., 2009). AoS speech seems to be particularly vulnerable on consonant clusters (e.g., Aichert & Ziegler, 2004; Romani & Galluzzi, 2005). Aichert and Ziegler (2004) claim that the tendency to reduce two abutting consonants depends on the location of the cluster (syllable boundary or intra-syllabic cluster location), which seems to be in line with coarticulation impairment. In our group, no position effect was observed, but other results point to a relationship with coarticulation. D'Alessandro and colleagues (2019) are also undertaking a study in this area and they are currently reporting that participants with AoS and WD have weaker vowel to vowel coarticulation, longer vowels and longer vowel intervals than other groups of MSD. It is particularly interesting to note that these authors conducted their study on the same cohort of participants as in this study. In comparing the results of the two studies, it is particularly interesting to note that the same three AoS and the WD participants who produced vocoid epentheses were also those who coarticulate the least compared to their respective groups. This observation further suggests that insertions can be related to a coarticulation deficit.

Conclusion

In this study, we have tracked the acoustic signature of vocoid epenthesis in consonant clusters in participants from different subtypes of motor speech disorders (MSD). Although vocoid epenthesis in voiceless clusters was observed mostly in AoS, this was true only for a minority of them, as not all participants with AoS presented with insertions of vocalic elements in unvoiced clusters. Only one participant with mixed dysarthria also produced such insertions. The present study shows the importance of acoustic analyses in clinical practice with patients who present with a MSD, as a non-negligible proportion of CC was misclassified by SLTs. Detecting vocoid on the spectrogram is a quick task to which all clinicians may be easily trained. However, the presence of vocoid epenthesis, although quite specific to AoS, does not seem to be systematic. Further research is needed to determine other reliable acoustic features as, such as the presence of vocoids epenthesis, weaker vowel to vowel coarticulation, longer vowels and longer vowel intervals. The computation of the sum/proportion of all of these features could give us a better identification of acoustic features of AoS.

Note

1. We were unable to perform the analyses using mixed models because some of our groups had no variance.

Acknowledgments

This research was supported by Swiss National Science Foundation grants no. CRSII5_173711.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Swiss National Science Foundation Sinergia program [CRSII5_173711, 10.2017-9.2020].

ORCID

Marina Laganaro http://orcid.org/0000-0002-4054-0939

References

Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*(1), 148–159. https://doi.org/10.1016/S0093-934X(03)00296-7

Allison, K. M., Cordella, C., Luzzini-Seigel, J., & Green, J. R. (2020). Differential diagnosis of apraxia of speech in children and adults: A scoping review. *Journal of Speech, Language and Hearing Research*, *63*(9), 2952–2994. https://doi.org/10.1044/2020_JSLHR-20-00061

Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: Review. *Clinical Linguistics & Phonetics*, *14*(2), 131–150. https://doi.org/10.1080/026992000298878

Blumstein, S. (1990). Phonological deficits in aphasia: Theoretical perspectives. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics* (pp. 33–53). Lawrence Erlbaum.

Blumstein, S. E., Cooper, W. E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. *Brain and Language*, *9*(2), 153–170. https://doi.org/10.1016/0093-934X(80)90137-6

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345. [Computer program]. http://www.praat.org

Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener agreement for auditory-perceptual ratings of dysarthria. *Journal of Speech, Language and Hearing Research*, *50*(6), 1481–1495. https://doi.org/10.1044/1092-4388(2007/102)

Bürki, A., Fougeron, C., Gendrot, C., & Frauenfelder, U. (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics*, *39*(3), 279–288. https://doi.org/10.1016/j.wocn.2010.07.003

Code, C. (1998). Major review: Models, theories and heuristics in apraxia of speech. *Clinical Linguistics & Phonetics*, *12*(1), 47–65. https://doi.org/10.3109/02699209808985212

Dabul, B. (1986). *Apraxia battery for adults*. Tigard OR: CC Publications.

D'Alessandro, D., Pernon, M., Fougeron, C. & Laganaro, M. (2019). Anticipatory VtoV coarticulation in French in several Motor Speech Disorders. In *Phonetics and Phonology in Europe*, Lecce, Italy. hal–02427864

Darley, F., Aronson, A., & Brown, J. (1975). *Motor speech disorders*. Philadelphia: W.B. Saunders.

Fonville, S., Van Der Worp, H. B., Maat, P., Aldenhoven, M., Algra, A., & Van Gijn, J. (2008). Accuracy and inter-observer variation in the classification of dysarthria from speech recordings. *Journal of Neurology*, *255*(10), 1545–1548. https://doi.org/10.1007/s00415-008-0978-4

Fougeron, C., Delvaux, L., Ménard, L., & Laganaro, M. (2018). The monPaGe_HA database for the documentation of spoken French throughout adulthood. In *Proc. 11th International Conference on Language Resources and Evaluation*, Miyazaki, Japan.
Fougeron, C., Gendrot, C., & Bürki, A. (2007). On the acoustic characteristics of French Schwa. *In Proceedings of the 16th international congress of phonetic sciences*, 941–944. Saarbrücken, Germany.
Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, *71*, 64–83. https://doi.org/10.1016/j.neuropsychologia.2015.03.007
Gibbon, F. E., & Lee, A. (2017). Preface to the special issue on covert contrasts. *Clinical Linguistics & Phonetics*, *31*(1), 1–3. https://doi.org/10.1080/02699206.2016.1254684
Guenther, F. (2016). Neurological disorders of speech production. In Guenther (Ed.), *Neural control of speech* (pp. 273–320). The MIT Press.
Haley, K. L., Jacks, A., De Riestha, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1502–1517. https://doi.org/10.1044/1092-4388(2012/11-0318)
Haley, K. L., Ohde, R. N., & Wertz, R. T. (2001). Vowel quality in aphasia and apraxia of speech : Phonetic transcription and formant analyses. *Aphasiology*, *15*(12), 1107–1123. https://doi.org/10.1080/02687040143000519
Josephs, K. A., Duffy, J. R., Strand, E. A., Machulda, M. M., Senjem, M. L., Master, A. V., Lowe, V. J., Jack, C. R., Jr, & Whitwell, J. L. (2012). Characterizing a neurodegenerative syndrome: Primary progressive apraxia of speech. *Brain*, *135*(5), 1522–1536. https://doi.org/10.1093/brain/aws032
Kent, R. D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic dysarthria. *Journal of Speech, Language, and Hearing Research*, *43*(5), 1275–1289. https://doi.org/10.1044/jslhr.4305.1275
Kent, R. D., & Kim, Y. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, *17*(6), 427–445. https://doi.org/10.1080/0269920031000086248
Laganaro, M., Croisier, M., Bagou, O., & Assal, F. (2012). Progressive apraxia of speech as a window into the study of speech planning processes. *Cortex*, *48*(8), 963–971. https://doi.org/10.1016/j.cortex.2011.03.010
Laganaro, M., Fougeron, C., Pernon, M., Levêque, N., Borel, S., Fornet, M., Catalano, S., Lopez, U., Trouville, R., Ménaard, L., Burkhard, P. R., Assal, F., & Delvaux, V. (2020). Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: The MonPaGe-screening protocol. *Clinical Linguistics & Phonetics*. https://doi.org/10.1080/02699206.2020.186546
Marczyk, A. K., Meynadier, Y., Gaydina, Y., & Solé, M.-J. (2018). Dynamic acoustic evidence of nasalization as a compensatory mechanism for voicing in Spanish apraxic speech. In Q. Fang, J. Dang, P. Perrier, J. Wei, L. Wang, & N. Yan (Eds.), *Studies on speech production* (pp. 225–236). Springer International Publishing.
McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition, differentiation, and treatment. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 249–268). Thieme.
Melle, N., & Gallego, C. (2012). Differential diagnosis between apraxia and dysarthria based on acoustic analysis. *The Spanish Journal of Psychology*, *15*(2), 495–504. https://doi.org/10.5209/rev_SJOP.2012.v15.n2.38860
Molloy, J., & Jagoe, C. (2019). Use of diverse diagnostic criteria for acquired apraxia of speech: A scoping review. *International Journal of Language & Communication Disorders*, *54*(6), 875–893. https://doi.org/10.1111/1460-6984.12494
Ouden, D. B., Galkina, E., Basilakos, A., & Fridriksson, J. (2018). Vowel formant dispersion reflects severity of apraxia of speech. *Aphasiology*, *32*(8), 902–921. https://doi.org/10.1080/02687038.2017.1385050
Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Trouville, R., Fougeron, C., & Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatisé d'évaluation perceptive et

acoustique des troubles moteurs de la parole. *Rééducation Orthophonique*, *281*, 169–197. https://halshs.archives-ouvertes.fr/halshs-02504027

Ridouane, R., & Fougeron, C. (2011). Schwa elements in tashlhiyt word-initial clusters. *Journal of Laboratory Phonology*, *2*(2), 275–300. https://doi.org/10.1515/labphon.2011.010

Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory disorders and phonological difficulties. *Cognitive Neuropsychology*, *22*(7), 817–850. https://doi.org/10.1080/02643290442000365

Stuntebeck, S. (2002). *Acoustic analysis of the prosodic properties of ataxic speech* (Unpublished master's thesis, University of Wisconsin).

Tjaden, K. (2003). Anticipatory coarticulation in multiple sclerosis and Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *46*(4), 990–1008. https://doi.org/10.1044/1092-4388(2003/077)

Van der Merwe, A. (2020). New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology*, *34*(4), 1–27. https://doi.org/10.1080/02687038.2020.1765306

Varley, R., & Whiteside, S. (2001). What is the underlying impairment in acquired apraxia of speech. *Aphasiology*, *15*(1), 39–49. https://doi.org/10.1080/02687040042000115

Wang, Y. T. (2002). *Acoustic analysis of prosodic disturbance in traumatic brain injury* (Unpublished doctoral dissertation, University of Wisconsin).

Whiteside, S. P. (1998). Coarticulation in apraxia of speech : An acoustic study of non-words. *Logopedics, Phoniatrics, Vocology*, *23*(4), 155–163. https://doi.org/10.1080/140154398434059

Whiteside, S. P., Robson, H., Windsor, F. & Varley, R. (2012). Stability in voice onset time patterns in a case of acquired apraxia of speech. *Journal of Medical Speech-Language Pathology*, *20*, 17–28.

Ziegler, W. (2008). Apraxia of speech. In G. Goldenberg & B. L. Miller (Eds.), *Neuropsychology and behavioral neurology* (Vol. 3, pp. 269–286). Elsevier.

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, *24*(5), 631–661. https://doi.org/10.1080/01690960802327989

Ziegler, W. (2017). Complexity of articulation planning in apraxia of speech: The limits of phoneme-based approaches. *Cognitive Neuropsychology*, *34*(7–8), 482–487. https://doi.org/10.1080/02643294.2017.1421148

Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language and Hearing Research*, *55*(5), 1485–1501. https://doi.org/10.1044/1092-4388(2012/12-0128)

Ziegler, W., & Von Cramon, D. (1986). Disturbed coarticulation in apraxia of speech: Acoustic evidence. *Brain and Language*, *20*(1), 34–47. https://doi.org/10.1016/0093-934X(86)90032-5

Appendix A. list of stimuli (pseudo-words containing unvoiced clusters)

	Spelling	Transcription	Structure
1	Laspa	/laspa/	CVCCV
2	Laspé	/laspe/	CVCCV
3	Laspi	/laspi/	CVCCV
4	Laspu	/laspy/	CVCCV
5	Spéla	/spela/	CCVCV
6	Spila	/spila/	CCVCV
7	Vastra	/vastra/	CVCCCV
8	Vastré	/vastre/	CVCCCV
9	Splécha	/splea/	CCCVCV
10	Splicha	/splia/	CCCVCV
11	Strava	/strava/	CCCVCV
12	Stréva	/streva/	CCCVCV
13	Chasplé	/asple/	CVCCCV
14	Chaspli	/aspli/	CVCCCV

Appendix B. perceptive analysis design