Title: Using ultrasound visual biofeedback to treat persistent primary speech sound disorders.

Abstract

Growing evidence suggests that speech intervention using visual biofeedback may benefit people for whom visual skills are stronger than auditory skills (for example, the hearing impaired population, Bacsfalvi et al., 2007), especially when the target articulation is hard to describe or see. Diagnostic ultrasound can be used to image the tongue and has recently become more compact and affordable leading to renewed interest in it as a practical, non-invasive visual biofeedback tool. In this study we evaluate its effectiveness in treating children with persistent speech sound disorders that have been unresponsive to traditional therapy approaches.

A case series of seven different children (aged 6;0 to 11;0) with persistent speech sound disorders was evaluated. For each child high-speed ultrasound (121fps), audio and lip video recordings were made whilst probing each child's specific errors at five different time points (before, during and after intervention). After intervention all of the children made significant progress on targeted segments, evidenced by both perceptual measures and changes in tongue-shape.

Keywords

Ultrasound, visual biofeedback, speech sound disorders

Introduction

Developmental speech sound disorders (SSDs) are a common communication impairment in childhood, affecting between 10% and 15% of pre-school children and 6% of school-aged children (American Speech–Language–Hearing Association (ASHA) 2000; McLeod and Harrison, 2009). They are a heterogeneous group with differential diagnoses such as (consistent/inconsistent) phonological delay/disorder, developmental verbal dyspraxia (childhood apraxia of speech) and articulation delay/disorder (see Waring and Knight, 2013 for a review of the classification of SSDs).

While there is evidence that SSDs, especially phonological delay/disorder, are amenable to remediation using auditory based methods, such as minimal pairs and core vocabulary therapy (Broomfield and Dodd, 2011; Law, Garett and Nye, 2003) there remains a proportion of children who are unresponsive to conventional intervention, defined as persistent SSD, or intractable SSD (Wood and Scobbie, 2003). For these children, visual biofeedback methods such as electropalatography (EPG) and ultrasound visual biofeedback (U-VBF) are thought to provide the missing piece of the puzzle by making internal cues such as tongue shape and position explicit (Gibbon et al., 1999) and thus allowing these children to learn new articulations or sequences of articulations. Over the last 30 years a large number of small n studies have shown that EPG is potentially effective at remediating intractable SSDs (Gibbon, 2013). However, larger studies are lacking with a Cochrane review of EPG for the treatment of cleft palate speech (Lee, Law and Gibbon, 2009) finding that only one small (n=6) study (Michi et al., 1993) met inclusion criteria. The lack of larger randomised control studies may be due to logistical problems with manufacturing individualised EPG palates. In this study we look at the potential of a different visual biofeedback technique for the treatment of persistent SSDs, ultrasound tongue imaging (ultrasound visual biofeedback, U-VBF).

Visual biofeedback intervention for intractable SSDs.

The evidence for U-VBF therapy is small but promising, with only 19 small case or group studies reported in the literature. Most studies originate from the USA and Canada, with earlier studies focussing on children and young people with hearing impairment and later studies beginning to look at persistent SSDs (including Childhood Apraxia of Speech). Therapy has mainly addressed the production of the consonant /r/, because its accurate pronunciation is communicatively and socially important in North America. However, its use with other lingual speech sounds is less reported and it cannot be assumed that approaches that are effective for /r/ misarticulations will necessarily translate to other error types since /r/ is particularly late acquired and articulatorily complex (Gick, Bernhardt, Bacsfalvi and Wilson, 2008) which in turn effects motor learning (Maas et al., 2008). It is surprising that other error types have been largely neglected given the range of lingua-palatal targets reported in the EPG literature and considering the underlying mechanisms behind the success of both techniques are likely to be the same. For example, to our knowledge no studies have reported on the use of U-VBF in the treatment of persistent velar fronting despite the fact that it has been the subject of several EPG studies (e.g. Gibbon, McNeill, Wood and Watson, 2004) and coronal/dorsal distinctions are clearly imaged with ultrasound. However, a recent study by Preston, McCabe, Rivera, Whittle, Landry and Maas (2014) shows promising results for U-VBF in treating residual SSDs. The authors employed U-VBF in the treatment of eight children (childhood apraxia of speech excluded, this group was the subject of a previous study, Preston Brick and Landi, 2013), while all of the children focused on the correct production of /r/ either mainly or to some extent, half of the participants also underwent treatment for sibilants. It is not clear from the paper what the nature of the error types were, however we presume the children presented with typical developmental /r/ errors and lateralisation of sibilants is alluded to, although not explicitly described. Nevertheless, results were generally positive, with most participants making gains (an increase in percentage target segment correct of 39% in the first treated condition and 32% in the second) in untreated words and sentences. However, the results are somewhat limited because only one word context (e.g. word initial) with two variants (i.e. if word initial /r/ was treated only /re/ and /ro/ were taught) was treated in each block of therapy. Results are based on perceptual ratings with no ultrasound data given to confirm changes in tongue-shape post-therapy, possibly because the system used to collect ultrasound data does so only at 18 frames per second, which is adequate for viewing dynamic movements during therapy but makes analysis of specific time points tied to acoustic events difficult. Ultrasound systems with frame rates between 60 and 200fps are optimal for speech analysis (Wrench and Scobbie, 2011).

Ultrasound tongue imaging.

Ultrasound Tongue Imaging (UTI) uses standard medical ultrasound to image the tongue in real-time. The probe is placed under the chin, allowing real-time visual feedback of most of the surface of the tongue in either the mid-sagittal or coronal plane. In both views, the imageable area is constrained by shadows from bone, with the tongue tip in particular being susceptible to a shadow from the mandible. Unlike EPG, the image is an anatomically correct representation of a slice of the speaker's own tongue. However, other relevant anatomical information, such as the lateral margins of the tongue and the relation of the tongue to the hard palate, are not visible in a typical mid-sagittal UTI image. Stone (2005) gives an overview of ultrasound tongue imaging, including how to interpret the image and different methods of stabilising the probe. Many clinical studies use hand-holding of the probe only, this is convenient and comfortable for the client, but leads to difficulties stabilising the images and measuring differences in tongue shape. Alternatively, the probe can be stabilised, to allow measurements to be taken and allow the clinician and speaker to be "hands free". In our study the probe was stabilised using a headset (see figure 3 below) to allow us to compare ultrasound images before and after therapy, however, it does result in some restriction of jaw movement. Figure one shows typical ultrasound images.

Insert figure one about here

The mechanisms underlying ultrasound visual biofeedback therapy.

The most commonly used interventions for SSDs (minimal pairs, auditory discrimination, phonological awareness are the most commonly used approaches, Joffe and Pring, 2008) rely on auditory skills; clients must listen to their own productions and modify them using auditory cues. Visual biofeedback provides another modality, supplementing or even potentially circumventing the auditory system. Preston et al. (2013 and 2014) argue convincingly that U-VBF is a motor based therapy. Children who make inappropriate phonetic realisations of certain speech sounds do so because they have an inappropriate motor plan for that sound (Preston et al., 2014). We expand this idea by suggesting that the inappropriate motor plan can be ascribed one of three categories. It may be identical to that of another phoneme resulting in homophony in the system (as in classic velar fronting). Alternatively, the motor plan is abnormal or underspecified resulting in something homophonous-sounding but nonetheless different in some way. In other words, a “covert contrast” (Gibbon and Scobbie, 1997) appears where a child who was previously thought to have collapsed two categories (i.e. velars and alveolars are both transcribed as [t] resulting in the clinician believing the child presents with the phonological process of velar fronting) does in fact have two phonetically distinct realisations, for example /t/→[t] and /k/→[t̠]). In this case, instrumental analysis shows that the child is producing two distinct phonemes that are imperceptible to the listener. Lastly, the motor plan is abnormal to the extent that it results in the realisation of a non-native speech sound, as is the case for lateral lisps in English speakers.

U-VBF provides the learner with an opportunity to acquire a new motor plan because it supplies a novel type of “knowledge of performance” (Preston et al., 2013 and Preston et al., 2014). Prior to U-VBF some knowledge of performance is available to speakers via both the auditory consequence of the movement and somatosensory feedback (they can feel their own tongues moving), but it seems that , even though the technique uses an unfamiliar type of image of a poorly specified aspect of the speaker’s own articulation, the real-time visual information somehow boosts feedback, hence leading to acquisition of articulations which was not previously possible (i.e. through traditional articulation therapy, which uses placement cues). Once a new general motor plan for a movement is established, transitioning from constant (e.g. producing the new articulation in only CV) to variable practice leads to better retrieval and flexibility of the motor plan, or essentially generalisation of the new articulation across speech. In Preston et al.'s (2014) study they do this by varying prosody (which was actually no more effective than not varying prosody) and teaching the new segments in limited contexts. In his study (Preston et al., 2014), /r/ is typically taught only in CV with two different vowels. In contrast, in the current study we pursue a different treatment model. We aim to move quickly from any context in which there is a correct production by varying the position within words and vowel context (all vowels of Scottish English) to promote rapid generalisation across the child's system, thus maximising the amount of therapeutic feedback on articulatory shape, location and dynamics on the one hand, and broadening experience of specific or holistic aspects of the acoustic output on the other.

Previous research (e.g. Berhardt, Bacsfalvi, Gick, Radanov and Williams, 2005; Byun, Hitchcock and Swartz, 2014) on U-VBF have presumed that it is crucially the biofeedback aspect of the technique that leads to its success, with the contribution of demonstration of tongue movements by another speaker rarely explored. Preston et al. (2013 and 2014) has focused on the theory that the success of U-VBF is attributable to the knowledge of performance supplied by the visual display of the speaker's own tongue. While we agree that U-VBF is a strongly motor-based therapy, this theoretical perspective does not account for the contribution which the demonstration of typical target tongue shapes by normal speakers plays (primarily the clinician in this setting). Research in other fields, particularly second language acquisition (see Ouni, 2011), has highlighted the potential value of demonstrating articulatory movements to learners. Moreover, Kröger, Graf-Borttscheller and Lowit (2008) showed that 25 children aged 4;6 to 10;7 with articulation disorders were able to recognise the phonetic features of a range of consonants and vowels with accuracy rates of 63% from a 3D model and 61% from a 2D-model. The children were asked to view silent movies of the phonemes and vocalise what they considered the best fit for the movie. Whilst this study did not investigate whether the articulatory animations improved progress in therapy, the authors conclude that the study does show an intuitive ability to tongue-read from a Talking Head model in school-aged children.

Another study (Cleland, McCron and Scobbie, 2013) showed that tongue-reading is possible from both EPG and ultrasound displays. In this study phonetically naïve typical adults viewed real-time and slow motion EPG and ultrasound silent movies of ten different linguo-palatal consonants and four vowels. Participants selected which segment they perceived from four forced-choice options where only one choice was possible (for example, options did not differ only in voicing since EPG/Ultrasound is not able to show voicing). Participants performed above chance for both EPG and ultrasound for consonants but at chance level for vowels with ultrasound. This study concludes that ultrasound and/or EPG might be used as models in studies which wish to investigate the use of Talking Head-type models. Indeed, Bernhardt et al. (2005) suggest that EPG and U-VBF therapy begin with demonstration of the target articulation, usually by the clinician. It is likely that articulatory models (or "Talking Heads") are the starting point of all VBF interventions, but their use is often not explicitly mentioned or explored, nor do any studies report on the use of EPG or ultrasound purely as visual articulatory models without the speaker-specific biofeedback.

Purpose and hypotheses

The purpose of the current study was to pilot the use of U-VBF with children with intractable SSDs with a variety of lingual articulatory errors. We sought to recruit children with errors other than developmental /r/ errors (gliding) to test that hypothesis that U-VBF can be used for a much wider variety of articulatory errors. In addition, we explicitly incorporated the use of ultrasound articulatory models into our motor-learning paradigm (see below) by showing all children dynamic ultrasound video of typical speakers in the same age range. Whilst previous studies allude to using clinician models, to our knowledge no study has systematically used age-matched models. Lastly, we hypothesised that the introduction of multiple articulatory contexts early-on in therapy would lead to generalisation across word position and vowel contexts of the taught segments.

Method

Participants

Eight participants aged 6;0 to 10;1 with persistent primary SSDs were initially recruited from local Speech and Language Therapists. We requested children aged over six with lingual errors which had been unresponsive to treatment. All participants were monolingual speakers of English (though participant 04M, later excluded, had lived in Greece until three years old) and all had received speech therapy in the past, but not using any visual biofeedback methods. Children were excluded if they had more than 30% segment correct at baseline probes (see below).

Design

The study is best described as a case series. Each child underwent two baseline probes (weeks 1 and 6 with no contact between) to establish stability of speech errors (to confirm reports from the referring clinicians) and to allow us to do an in-depth diagnostic analysis of their speech. Intervention occurred over 12 weekly therapy sessions, with a mid-therapy probe at sessions 6 (week 12). Finally there were a pair of post-therapy maintenance probes 6 weeks apart, first immediately post-therapy (week 19) and again, after no contact, at week 25. See Fig 2 for a timeline for each participant.

Insert figure 2 about here.

Pre-Therapy Assessments: Baseline 1 & 2

Language and Speech Assessments

The participants completed a battery of standardised speech and language assessments undertaken at either baseline 1 or 2.

Receptive Vocabulary

The British Picture Vocabulary Scales-3 (BPVS-3, Dunn, Dunn, Styles and Sewell, 2009) were used as a measure of receptive vocabulary. This assessment covers a wide age range. It is a multiple-choice test in which participants must select one of four pictures to match a single word spoken by the tester.

Receptive and Expressive Language

The Clinical Evaluation of Language Fundamentals- 4UK (CELF-4UK, Semel, Wiig and Secord, 2006) core language score was used to measure receptive and expressive language.

Oromotor Function

Oromotor function was assessed using the Robbins and Klee clinical assessment of oropharyngeal motor development in young children (RK, Robbins and Klee, 1987). In this assessment, children are required to perform speech and non-speech oral movements, which are scored as either adult-like (2 points), approaching adult-like (1 point) or absent (0).

Ultrasound Recording of Speech Measures

Unlike any previous studies of U-VBF intervention (for example Bacsfalvi and Bernhardt, 2011; Byun et al., 2014) we used a high-speed, probe-stabilised ultrasound system (Scobbie, Wrench and Van der Linden, 2006, see figure 3). This allows analysis of changes in tongue shape longitudinally, in particular pre vs. post-therapy, allowing us to supplement perceptual data with qualitative and quantitative analysis of the mid-sagittal profile of the tongue. The headset was fitted in such a way that the mandible and hyoid shadows were symmetrical on the image, thus ensuring we could see as much of the tongue as possible. It is not possible to see the hard-palate during speech due to the air boundary at the surface of the tongue, with the shape of the palate only able to be inferred when the tongue is pressed against the palate. Children were therefore asked to swallow at the beginning and end of recording sessions to allow us to record the location of the hard palate. The headset can become heavy and uncomfortable over time and for this reason sessions were restricted to a maximum of 50mins and removed sooner at participant's request. All children tolerated the headset during all recording sessions without having to remove and refit it. Whilst the headset does restrict jaw movement somewhat, we ensured that all speakers were able to articulate the Scottish English vowels /i,a,o/ comfortably before beginning as part of our protocol for fitting the headset.

Ultrasound data was acquired using an Ultrasonix SonixRP machine remotely controlled via Ethernet from a PC running Articulate Assistant Advanced software™ (Articulate Instruments Ltd, 2012) version 2.14 which internally synchronised the ultrasound and audio data. The echo return data was recorded at ~121 frames per second (fps), i.e. ~8ms per frame with a 135 degree field of view (FOV) in a mid-sagittal plane. A bespoke version of AAA was developed to allow us to use the software for therapy. The software closely emulates the AA software widely used for EPG therapy (Articulate Instruments Ltd, 2010). Static target tongue-shapes are displayed on the right of the computer screen and the child's live tongue image on the left. Tongue shapes can be saved from another person's speech (therapist's, or another child's) or be based on an individual's own productions. The live display can be frozen to allow the therapist to point out salient features. A hard-palate trace can be super-imposed on the live image and quick playback (including slow-motion playback) of participant's attempts at articulations during therapy or for analysis is easily achieved. All of the participants were recorded with ultrasound for all speech measures at all five probe time points and the same system was conveniently used to provide the real-time visual feedback therapy. By using the probe stabilisation headset during therapy we were able to avoid the situation where the ultrasound image is subject to excessive movement during the session. Only during intervention were the children able to see the ultrasound image and therefore there was no biofeedback available during probes. Recordings and therapy took place in a sound-treated studio with the SLT sitting alongside the participant. Simultaneous acoustic and lip-camera recordings (~60fps) were also made, using an audio technica 803D clip-on microphone sampling at 22050Hz and a NTSC micro-camera synchronised to the audio.

Insert Figure 3 About Here

Standardised Speech Measures

At baseline 1 the participants completed the articulation and phonology subtests of the Diagnostic Evaluation of Articulation and Phonology (DEAP, Dodd, Hua, Crosbie, & Holm, 2002). The DEAP was administered and scored by the treating clinician (the first author) only as it was used to determine therapy targets rather than as a quantitative measure of change. All speech measures were recorded with simultaneous ultrasound (see above). The articulation subtest is a picture naming task consisting of 30 words with each phoneme of English in at least one phonotactic position in a word. Any phonemes which are produced inaccurately are subjected to a stimulability test whereby participants are asked to imitate the segment in CV/VC (or vowel in isolation) or in isolation. This test allowed us to determine which, if any segments, participants were unable to articulate prior to intervention as VBF is thought to be most useful for establishing motor programmes for new articulations (Gibbon & Wood, 2010). In addition the phonology subtest was repeated at both baselines. This test is a measure of consonant production in 50 single words, covering most consonants of English in word initial and final positions. The phonology subtest allows an overall calculation of percentage consonants correct (PCC). Table one shows the group results for the pre-therapy measures, with numbers expressed as standard scores or percentages as appropriate.

Insert Table One About Here

In order to determine which consonants should be the focus of therapy, all errors produced in the phonology subtest of the DEAP at B1 were subjected to a process/pattern analysis, following instructions in the DEAP manual. This should not be taken to mean that the errors are a result of a specifically phonological impairment, rather, whilst some errors are traditionally thought to be phonological in nature, for example fronting of /k/ to [t] (Grunwell, 1985), other processes are more usually thought to be phonetic/articulatory in nature, for example, lateralisation of /s/ (Gibbon & Hardcastle, 1987). Nevertheless, the pattern analysis was undertaken as a method of describing the children's speech sound errors with no prior assumptions about the underlying cause of those errors. To determine which errors were most likely to impact on intelligibility, any error occurring fewer than three times in the DEAP was discounted (Dodd et al., 2002). Since U-VBF is only useful for remediating errors related to lingual speech-sounds, any errors which affected non-lingual consonants were also set aside. Lastly, errors involving socially acceptable variants or developmentally delayed productions of /r/ (such as [w] or [ʋ]) were discounted, as typically this segment is not a focus of treatment in the UK. This left a smaller set of errors from which therapy targets could be determined. The process with the highest number of errors, excluding gliding, was therefore treated first in therapy. For 4/7 of the participants this left only one error that could possibly be treated. For the other three participants, other errors were targeted when the primary target reached 80% accuracy during the course of therapy. Participants 03F and 08M met this criterion (see below for further details).

Probes

Following the DEAP phonology and articulation subtests at B1, a wordlist/s targeting each child's specific lingual errors was selected from a battery. Children were required to produce less than 30% probe segments accurately in order to be included in the study, participant 04M was therefore excluded at this point as his issues were more to do with sequencing of sounds at sentence level than correct production in single words. These probes or wordlists were recorded at each time-point: baseline, mid-therapy and post-therapy maintenance. Wordlists consisted of 50 untreated words on average. These words were never used in the course of therapy, allowing us to check for generalisation of targets. Each wordlist contains the "in error" lingual targets (denoted by shading in table 2) in singleton word initial, medial and final positions in a variety of vowel environments, plus clusters. Minimal pair comparison words (as appropriate, where the child's error resulted in homophony) and sentences were also collected. In this paper we present data from only one therapy target per child, however, several of the children presented with multiple errors (see Table 2). Other errors were targeted if the primary target reached 80% accuracy during the course of therapy, but here we reported only what was treated first in therapy in order to allow us to present consistent maintenance data.

A narrow phonetic transcription of the probe data (target segments only) was performed by the treating clinician using the acoustic, ultrasound and lip-camera data. This allowed us to calculate a specific percentage segment correct at each time point. Two final year Speech and Language Therapy students who had completed their phonetics training provided inter-rater reliability for the 50% of the untreated wordlists.

In addition a qualitative ultrasound analysis was undertaken at B1 and Post-therapy time point M1 of singleton consonants in the single non-treated words. Using AAA software (Articulate Instruments, 2012) fricatives were annotated at the acoustic midpoint and stops at the burst. For each segment, the nearest ultrasound frame to the midpoint/burst was selected and a spline indicating the tongue surface fitted to the image using the semi-automatic function in AAA software (Articulate Instruments, 2012). The user draws a spline in the region of the ultrasound image of the tongue and AAA's local edge-tracking function is used to search for the best edge locally, to assign a confidence value to each of the control points on the equally-spaced 42-fan measurement grid, and to smooth the spline to the requested degree. It should be noted that the accuracy of the actual location of each spline, though semi-automatically placed, is the responsibility of the trained analyst. AAA allows multiple splines from multiple tokens to be exported to a "workspace" to allow direct comparison of tongue surface shapes and locations from a single target. Tongue splines were then averaged across a session to give a holistic impression of the child's articulatory target for each segment, where the average can be compared to other segment averages, other sessions, or the articulatory landmark of the hard palate. Splines were normalised by rotation and/or transformation across sessions using hard-palate traces as a key reference, to allow for different placements of the stabilisation headset that inevitably occur when the headset is refitted anew across sessions.

Therapy

Treatment was provided by a specialist SLT experienced in the use of visual biofeedback. Each participant (except 05M, who received two blocks of U-VBF, one year apart, after re-referral) received a single block of 12 individualised therapy sessions. Sessions lasted around one hour, including family liaison, ultrasound sessions, and other activities. The amount of actual ultrasound feedback therapy varied from 10 to 40 minutes within a session. This was dependent on the child's tolerance of the stabilisation headset, their success with trying to achieve a new articulation and their own view on whether they wanted to continue within an individual session. Typically the first 30 minutes focussed on U-VBF and the second 30 minutes on traditional table-top activities focusing on the same target. This allowed us to build generalisation into sessions early and also allowed us to work on input activities (following the Stackhouse and Wells, 1997, framework) where children showed discrimination problems, for example with minimal pair discrimination. Each participant was also given individualised homework exercises and was instructed to practise for 10 to 15 minutes, five days a week. However, all parents reported practicing only once or twice a week due to other family commitments. The focus of therapy was based on the error analysis outlined above. Children did not receive any other speech and language therapy whilst taking part in the research project.

Since this was a pilot study, designed to test whether it was possible to treat a wider-range of lingual errors than previously reported, we did not operationalise the treatment protocol. Children's therapy was individualised to suit their age and types of speech errors. However, a general protocol was followed for the trajectory of sessions. The first therapy session for each child focused on learning to associate the movement of the ultrasound image on the screen with the movement of their own tongue by demonstrating control of tongue shapes already in their inventories. All children were able to do this within the first 15mins of the first session, except 06M who required further work across sessions two and three. The second phase of treatment (in either the first or second treatment session) was the use of ultrasound as a visual articulatory model, whereby the dynamic nature of speech was highlighted using pre-recorded videos of ultrasound as target movements. Target movements were taken from the Ultrax corpus of child speech (http://www.ultrax-speech.org) and hence were tongue movements of typically developing children the same age, sex and accent background of the participants. Using videos from other children, rather than live demonstration by the clinician allowed us to try different phonetic variants as targets and allowed the children to view tongues of a similar size to their own. Byun et al., (2014) found in their study of /r/ that it was necessary for children to be offered both bunched and retroflex variants of /r/ as targets, using a large corpus of tongue movements allowed us to do this easily, for both /r/ (for 02M) and other segments. It also allowed us to play the videos in real-time, slow motion (see the Seeing Speech web resource, 2014, for examples of slow motion ultrasound) and frame-by-frame (focusing on maximum point of articulation). The children were encouraged to watch these videos and their attention was drawn to the salient features of the target articulation. For example, a velar stop is characterised by a raising of the back of the tongue towards the soft palate. Production practice was individualised, but followed a basic articulation hierarchy (Van Riper and Emerick, 1984), or motor-based approach (similar to Preston et al., 2014), starting with the target phoneme in CV or VC (with a vowel likely to facilitate production, for example a back vowel for velars), progressing to CV and VC with differing vowels, through to words (building complexity) and then phrases and conversational speech. The children progressed to the next level of complexity as soon as they were able to achieve the previous level at 80% correct (i.e. If they could say /ko/ eight times out of ten they attempted all other vowels straight away and then moved to simple single words such as "came, cow, coop"). Most of the children were not initially stimulable for the target articulation (i.e. 01M, 05M, 06M, 07F were all unable to produce any velar in any context or in isolation) so therapy began with shaping the new articulation from phonetically similar segments. Velars were shaped from labial-velar /w/ or from /o/ (since the /o/ of GOAT is a mid-close back rounded monophthongal vowel [o̹], while GOOSE's /u/, is central and mid height (Scobbie, Gordeeva and Matthews, 2006). /r/ was shaped from / ʃ/ whilst encouraging retroflexion (asking the speaker to "curl his tongue backwards", as demonstrated by video). /t/ was shaped from /d/ (which was occasionally correct in participant 08M), and / ʃ/ was shaped from /tʃ/ which participant 03F was able to achieve inconsistently.

When the participant achieved an acceptable tongue shape which was also perceptually acceptable (as judged live by the treating clinician) then the participant's own best attempt could be used as a target tongue-shape. Generalisation was built into the sessions, with progression to productions without visual biofeedback as soon as the participant was able to achieve a new articulation in at least one vowel context ten times consistently. Use of the auditory feedback loop was also build in by encouraging participants to reflect on their own productions, both immediately after producing them and offline by listening to/watching their own pre-recorded attempts.

Results: Baseline

Language and Cognitive Measures

The speech, language and cognitive profile of the participants were in line with primary SSD. Participant 03F presented with significant delays in language, but normal nonverbal ability, consistent with a diagnosis of specific language impairment with SSD. These measures were not available for 07F and 08M, but were reported to be normal by the referring SLT and these children presented with normal receptive vocabulary scores.

Selecting Therapy Targets: Error Analysis

Fourteen definable processes were identified in the single word productions of the DEAP phonology subtest. After eliminating the structural processes and processes not related to changes in place of articulation (i.e. processes where the change in motor programme cannot be readily viewed in the mid-sagittal plane), only five error types remained: gliding, velar fronting, post-alveolar fronting, backing and idiosyncratic /r/ production. The process with the highest number of errors, excluding gliding, was treated first in therapy. Four children were therefore treated for velar fronting; one for post-alveolar fronting; one for idiosyncratic backing (alveolars were produced as voiceless palatal lateral affricates) and one for idiosyncratic /r/ productions (/r/ was realised as a palatal lateral approximant). Table two shows the analysis of the DEAP errors and which segment was chosen for therapy, denoted by shading.

Insert Table Two About Here

Results: Pre- and post-therapy

Probes

Figures 4a to 10a show the percentage segments correct (transcription data) at each time point with Figures 4b to 10b alongside showing pre- and post-therapy average ultrasound tongue-shapes for each participant as the basis of a qualitative analysis. Inter-rater reliability for 50% of the time-points was 95% (ranging from 83% to 100% agreement), with complete agreement that time-points rated as 100% correct were indeed 100% correct.

Insert Figures 3a to 9a and 3b to 9b about here. "a" figures should appear on the left with their "b" counterpart paired on the right.

Velar Fronting: Participants 01M, 05M, 06M and 07F.

Participant 01M received therapy targeting production of velar stops. Pre-therapy, he was unable to produce any velar stops in any context or in isolation, substituting consistently with alveolars, hence zero scores at both baselines. Mid-therapy session notes show that he was able to produce velars in CV and in treated words, but the low score in the probe demonstrates that he had not yet generalised. In contrast, his post-therapy probe scores are 100% correct, with 100% non-overlapping data suggesting highly-effective therapy. This is borne out by the ultrasound tongue curves which are clearly alveolar pre-therapy and clearly velar post-therapy, with little variation. Participants 06SSDM and 07SSDF show very similar patterns to 01M, with zero scores pre-therapy and 100% correct at post-therapy (07F) or maintenance (06M). Again, both of these participants were categorical in their errors, with incorrect productions transcribed exclusively as alveolars. At immediate post-therapy 06M produced exactly half of productions as velars and half as alveolars, the tongue-shape figure 7b therefore shows only the average of the correct productions since a mean drawing of two categorically different realisations would be meaningless.

Participant 05M was unsuccessful with his first block of U-VBF, remaining at 0% correct at post therapy and maintenance, again with classic velar fronting errors. It is possible that as one of our younger participants with diagnosed attentional problems this may have impacted on his ability to make use of the visual biofeedback. Alternatively, dosage in the first block of therapy may have been insufficient or he may simply have been unsuitable for VBF therapy. However, he was offered a further course of therapy one year later and at that point achieved velar stops at 100% accuracy. The tongue-curve figure 7b therefore compares baseline 1 tongue shapes with post 2 (i.e. after the second block of therapy) tongue shapes.

Post-Alveolar Fronting: Participant 03F

Pre-therapy participant 03F produced 30% of /ʃ/ productions correctly. The remainder were fronted to [s] (occasionally with lip-rounding), as borne out by the ultrasound-data showing a clear alveolar constriction. Post-therapy productions were consistently correct and she achieved a tip-down [ʃ]. However, given that she was stimulable for the post-alveolar sibilant before therapy, it is possible that she was in the process of acquisition naturally, making it more difficult to tell if the U-VBF was responsible for the progress.

Idiosyncratic Backing: Participant 08M

Participant 08M presented with an unusual case of backing where alveolar plosives were realised variably as velar stops, palatal stops or palatal lateral affricates. His speech was characterised overall by lateral release and lateral sibilants. Pre-therapy, he produced one production of [d] but was not stimulable for [t]. Treatment focused on establishing a consistently correct production of [t] and [d] and again was highly successful with 100% correct after therapy and 100% non-overlapping data. This is borne out in the tongue shape data with a clear tip-down, dorsum raised production before therapy and typical [t] tongue-shape after. Moreover, although the ultrasound was used only in the mid-sagittal plane he was able to eliminate the lateral release from his plosives. However, both /s/ and /ʃ/, which were realised as labials with central airflow before therapy, had improved in place of articulation but become lateral after therapy (/s/→[f] and /ʃ/ → [ɸ]).

Idiosyncratic /ɹ/: Participant 02M

Participant 02M presented with an unusual phonetic realisation of /r/ in the absence of any other errors, or a history of any other errors. He was unable to achieve a bunched or retroflex /r/ pre-therapy, instead producing a palatal lateral approximate. The pre-therapy tongue-shape shows close approximation to the hard-palate and raised tongue-tip. During therapy he was taught a retroflex tongue shape and quickly achieved this but, post-therapy, tongue shape data shows a classic bunched shape. Nevertheless, post-therapy his productions of /r/ were 100% correct.

Results Summary and Quantification

The results are overwhelmingly positive with a mean increase in percentage segments correct from baseline to six weeks post-therapy of 95 percentage points (range 70 to 100) and clear differences in the tongue shape data. All children show 100% non-overlapping data suggesting that U-VBF may be highly effective at remediating previously intractable speech errors in children with primary SSDs, however further robust clinical trials are required to extend this pilot.

It is difficult to quantify statistically the difference between tongue shapes, but table 3 presents an analysis of the pre vs. post therapy tongue shapes based on the built-in difference function in AAA. See also the overall means in Figures 4-10, which can be interpreted visually much in the same way as SS-ANOVA's confidence intervals: non-overlapping standard deviations indicate a significant difference (Davidson, 2006). Table 3 adds more detail by estimating the size of the difference. Together, they locate and quantify the significant pre/post difference in tongue shape that was found for all participants. Significance is tested along each fan-line by built-in t-test. Our threshold for reporting and estimating the size of the significant difference between means is more stringent than finding a single significant difference between curves and reporting its size. Instead, we require a minimum number of adjacent t-tests to be significant at $p<0.05$ over a longer, contiguous region of the tongue surface, reflecting the fact that the adjacent parts of the tongue, and their distance from the fan-grid's origin (i.e. the centre of the probe) are not independent. Specifically, we set a threshold at 6 adjacent fan lines (approximately 3cm of tongue surface) based on a measured lack of correlation between points that far apart. This means we do not claim any difference based on only a few significant t-tests from a small area of tongue surface. Depending on its distance from the probe, the actual length of tongue surface between each fan varies, and so the actual length of surface through which the 6 or more adjacent significant fan-lines pass is also reported (or rather the average based on the two conditions pre/post, since one is further from the probe, one closer). A pair of different tongue shapes may contain more than one of these conservatively defined regions of difference. In addition, we report a less cautious calculation in an attempt to capture the overall behaviour. We extend this region into and beyond other adjacent areas in which there is either no significant difference, or where the confidence of the location of splines as defined by AAA's edge-detection programme is low. This bolder measure requires at least one other significant comparison beyond the original conservative region of difference to act as its limit. Both the more and the less conservative results are reported for the length of tongue surface along which we found a difference. The mean difference between conditions, and the maximum difference on any single fan lines are also presented.

Insert table 3 about here

Discussion

This study assessed the effectiveness of U-VBF in seven children with primary intractable SSDs. We sought to extend previous uses of U-VBF to a wider variety of lingual articulatory errors and as such we report three previously unreported types of articulatory errors: incorrect production of velars, idiosyncratic backing, and idiosyncratic /r/ production. Each of the children had a history of SSD which had been unresponsive to conventional treatment; it is therefore likely that the large increase in percentage segments correct in the untreated probes was due the U-VBF. One child, 05M, made no progress in the first block of 12 sessions of U-VBF, but went on to make rapid progress almost a year later in a second block of the same therapy. This highlights the need to investigate more thoroughly in larger clinical studies the dosage required for U-VBF. Unlike previous studies, which relied upon perceptual judgement to assess change, we have used a high-speed, probe-stabilised ultrasound system which allows us to present pre and post-treatment tongue shape data for the first time. By averaging across multiple word positions and vowel contexts we are able to show large general changes in tongue-shape post-therapy (particularly in cases of velar fronting and alveolar backing).

Our results are very much in line with Preston et al. (2014) study of U-VBF for residual speech sound errors, despite differences in methodology, adding evidence to a growing body of research that suggests visual biofeedback might be a useful technique for providing a “break-through” for children with intractable SSDs. Both the Preston et al. study (2014) and ours capitalised on motor learning by providing learners with explicit knowledge of performance and by employing the principles of motor learning (Maas et al., 2008) including changing variables (for example vowel context) to encourage extension of a general motor plan to more complex movements, that is more complex contexts, such as use of a segment in multisyllabic words. However, it may not be the case that the success of U-VBF lies only in its provision of knowledge of performance. A distinction is made by Preston between knowledge of performance (qualitative feedback on the movement, as seen on U-VBF) and knowledge of results, which is the correctness of the attempt, essentially the acoustic consequence followed by a judgement of correct/incorrect by the clinician or the speaker themselves. However, articulatory movements (the “performance”) cannot be entirely uncoupled from the “result”, and this is especially the case when working at CV or VC level (rather than silent tongue movements), as we do even at the beginning of therapy. Even in traditional articulation therapy, speakers are aware of the acoustic consequence of their movements and therapists will encourage them to focus on articulatory placement which draws somatosensory feedback, or knowledge of performance, into sharp focus. Nevertheless, studies of VBF (particularly EPG in a large body of small n studies) seem to suggest that visual biofeedback offers children something new, or a breakthrough moment in acquisition of new articulations when other treatments have failed (Gibbon & wood, 2010). It therefore certainly offers a novel, perhaps more explicit knowledge of performance.

The contribution of articulatory demonstration to U-VBF.

One aspect which has been so far neglected is the contribution of visual articulatory models, or "Talking Heads" to visual biofeedback techniques. Berhardt et al. (2005) summarise that U-VBF and EPG begin with the clinician demonstrating the target articulation to the client. Since ultrasound is used for demonstrating and treating articulations which are hard to see, and crucially hard to describe, the use of the visual model avoids unnecessarily complex language and instructions that may be difficult for children to understand (Cleland, Timmins, Wood, Hardcastle, and Wishart, 2009). Anecdotally, parents of children with SSDs undergoing U-VBF often report that they did not fully understand the movements required for certain articulations prior to viewing ultrasound and one of the children in our study (07F) commented that a velar was "impossible" the first time she viewed an ultrasound movie of that segment, highlighting the lack of understanding she had as to the movements required to achieve a velar despite previous therapy targeting this very sound.

Further, there may be some implicit learning involved in the viewing of tongue movements. In normal audio-visual speech perception, being able to see the lips of a speaker enhances perception, for example in noise (Benoît and Le Goff, 1998). All typical speakers implicitly learn to integrate lip information into their perceptual system, a phenomenon attested to in the McGurk effect (McGurk and MacDonald, 1976). Whilst lips are easily visible during interactions the tongue is not and we might therefore predict that such an effect would be limited to visible articulators only. However, a small number of studies have sought to determine whether there is also an intuitive capacity to “tongue-read” using various Talking Heads, showing a mid-sagittal dynamic animation of the tongue (often based on Magnetic Resonance Imaging or electromagnetic articulograph). Badin, Tarabalka, Elisei, and Bailly (2010) investigated whether viewing a Talking Head enhanced perception of speech in various noise conditions. That is, a similar experiment to Benoît and Le Goff (1998), but with a condition where listener-viewers saw the tongue. Results showed that the mean phoneme identification rate increased when audio-visual information was added (including a lip-only condition), but crucially phoneme recognition was significantly greater (68.1%) when a mid-sagittal view with the tongue visible was compared to a mid-sagittal view with no visible tongue (63.7%). Badin et al. (2010) hypothesise that this is due to a natural, intuitive, capacity for listeners/viewers to tongue-read and suggest that this provides support for a perception/production link which they say could relate to the theory of mirror neurones. Further support for a capacity to tongue-read comes from a recent study by Cleland et al. (2013) looking specifically at whether naïve participants could interpret silent movies of EPG and ultrasound. Again, a capacity to do this at above chance levels was found. Whilst the study offered no firm account of the mechanism by which speakers come to be able to “tongue-read”, it does highlight the need to take the contribution of ultrasound displays as a visual articulatory model into account.

Gibbon and Wood (2010) highlight the role that observing an action plays in enhancing a motor behaviour and cite mirror neurones as the underlying mechanism by which a person learns a complex motor movement (which speech is). They suggest that studies should make use of this capacity by providing children with speech sound disorders with a visual articulatory model via EPG. No studies have yet used ultrasound or EPG in this way, but there is great potential to develop a Talking Head based on ultrasound. This has an advantage over models based MRI and EPG as data can be collected quickly and easily from multiple speakers, including children.

Whilst it would be unethical and impractical to compare U-VBF without demonstration/models to U-VBF with it, it would be entirely feasible to construct a randomised control trial where one arm of the trial involved the use of an ultrasound-based visual articulatory model, without biofeedback available. Indeed a recent pilot study (Roxburgh, Scobbie, Cleland and Wood, 2014) found that children with cleft palate did just as well with a visual articulatory model to learn new articulations as they did with U-VBF. However, this study was limited in that it only looked at two participants, both of whom had not had previous therapy to address their speech problems (i.e. they were not 'intractable').

Conclusions

There is growing evidence that Ultrasound Visual Bio-feedback is highly effective in remediating previously intractable speech sound disorders. Our study adds to a growing body of evidence by using U-VBF to treat previously unreported error types in children, particularly velar-fronting. However, the mechanisms by which U-VBF enhances learning remain unclear. Whilst it is plainly a motor-based therapy, the mechanisms by which clients learn to interpret visual information about tongue movements is not well understood, nor is the likely crucial contribution of accurate articulatory models/demonstration.

Acknowledgments

Thank you primarily to the children and their parents who attended weekly therapy sessions and follow up recordings. Thank you also to the Speech and Language Therapists who referred the children. Thank you to XXXXX XXXX for technical assistance with the ultrasound recordings and to XXX XXXXXXX for help with some of the data labelling.

Declarations of Interest

The first author was supported by an EPSRC grant (Grant no: XXXXXXXXX).

References

American Speech–Language–Hearing Association (ASHA). (2000). *Communication Facts* (Rockville, MD: American Speech–Language–Hearing Association).

Articulate Instruments Ltd 2010. *Articulate Assistant User Guide: Version 1.18*. Edinburgh, UK: Articulate Instruments Ltd.

Articulate Instruments Ltd. (2011). *Articulate Assistant Advanced User Guide: Version 2.13*. Edinburgh: Articulate Instruments Ltd.

Bacsfalvi, P., Bernhardt, B. M., & Gick, B. (2007). Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *International Journal of Speech-Language Pathology*, *9*(1), 36-45.

Bacsfalvi P. & Bernhardt BM. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: ultrasound and electropalatography. *Clinical Linguistics and Phonetics 25*, 1034-43.

Badin, P., Tarabalka, Y., Elisei, F., & Bailly, G. (2010). Can you 'read' tongue movements? Evaluation of the contribution of tongue display to speech understanding. *Speech Communication, 52,* 493–503.

Benoît, C. & Le Goff, B. (1998). Audio-visual speech synthesis from French test: eight years of models, designs and evaluation at the ICP. *Speech Communication*, *26*, 117-129.

Bernhardt, B., Bacsfalvi, P., Gick, B., Radanov, B., & Williams, R. (2005). Exploring the use of electropalatography and ultrasound in speech habilitation. *Journal of Speech-Language Pathology & Audiology, 29,* 169-182.

Broomfield, J. & Dodd, B. (2011). Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial. *International Journal of Language & Communication Disorders, 46*, 628–640

Byun, T. M., Hitchcock, E. R., & Swartz, M. T. (2014). Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *Journal of Speech, Language, and Hearing Research*,*57*(6), 2116-2130.

Cleland, J., McCron, C. & Scobbie, J.M. (2013) Tongue reading: Comparing the interpretation of visual information from inside the mouth, from electropalatographic and ultrasound displays of speech sounds*. Clinical Linguistics and Phonetics, 27 (4),* 299-311.

Cleland, J., Timmins, C., Wood, S., Hardcastle, W. & Wishart, J. (2009). Electropalatographic therapy for children and young people with Down's syndrome. *Clinical Linguistics and Phonetics, 23(12),* 926-939.

Cross, E.S., Hamilton, A.F.C., & Grafton, S.T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage, 31,* 1257–1267.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America, 120(1*): 407-15.

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale III – 3rd Edition*. London: GL Assessment.

Gick, B. Bernhardt, B.M., Bacsfalvi, P. & Wilson, I. Ultrasound imaging applications in second language acquisition. (2008). In J. Hansen & M. Zampini (eds.) *Phonology and Second Language Acquisition*. Ch. 11, pp. 309-322. Amsterdam: John Benjamins.

Gibbon, F. (2013). *Bibliography of Electropalatographic (EPG) Studies in English (1957-2013).* Available from: http://www.articulateinstruments.com/EPGrefs.pdf

Gibbon, F., & Hardcastle, W. (1987). Articulatory description and treatment of "lateral/s/" using electropalatography: A case study. *International Journal of Language & Communication Disorders*, *22*(3), 203-217.

Gibbon, F., McNeill, A., Wood, S. & Watson, J. (2004). Changes in linguapalatal contact patterns during therapy for velar fronting in a 10-year-old with Down's syndrome. *International* journal of language and communication disorders, 38, 47-64.

Gibbon, F. & Scobbie, J.M. (1997). Covert contrasts in children with phonological disorder. *Australian Communication Quarterly. Autumn,* 13-16.

Gibbon, F., Stewart, F., Hardcastle, W. J., & Crampin, L. (1999). Widening access to electropalatography for children with persistent sound system disorders. *American Journal of Speech-Language Pathology*, *8*(4), 319-334.

Gibbon, F. & Wood, S. (2010) *Visual Feedback Therapy with Electropalatography.* In: Interventions for Speech Sound Disorders in Children. Paul H Brookes Pub Co., Baltimore, pp. 509-536.

Grunwell, P. (1985). *Phonological assessment of child speech (PACS)*. San Diego, CA: College-Hill Press.

Joffe, V. and Pring, T. (2008), Children with phonological problems: a survey of clinical practice. *International Journal of Language & Communication Disorders, 43*: 154–164.

Kröger, B., Graf-Borttscheller, V. & Lowit, A. (2008). Two- and Three-Dimensional Visual Articulatory Models for Pronunciation Training and for Treatment of Speech Disorders. *Interspeech,* 2008, 22-26 September, Brisbane, Australia. P2639-2642.

Law, J., Garrett, Z. & Nye, C. (2003). Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Systematic Review, 3,* CD004110.

Lee, A. S. Y., Law, J., & Gibbon, F. E. (2009). Electropalatography for articulation disorders associated with cleft palate. *Cochrane Database Syst Rev,* 3.

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, *17*(3), 277-298.

McLeod, S. & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *Journal of Speech, Language, and Hearing Research, 52,* 1213–1229.

McGurk H. & MacDonald J. (1976). "Hearing lips and seeing voices.". *Nature,* 264 (5588): 746–8.

Michi, K. I., Yamashita, Y., Imai, S., Suzuki, N., & Yoshida, H. (1993). Role of visual feedback treatment for defective /s/sounds in patients with cleft palate. *Journal of Speech, Language, and Hearing Research, 36(2)*, 277-285.

Ouni, S. (2011). Tongue Gesture Awareness and Pronunciation Training. *Interspeech 11*, 28-31 August 2011, Florence, Italy.

Roxburgh, Z., Scobbie, J.M., Cleland, J., & Wood, S. (2014). The use of ultrasound in the treatment of speech sound disorders associated with cleft palate. *International Conference on Clinical Linguistics and Phonetics (ICPLA-15).* Karolinska Institutet, Stockholm: 11th – 13th June 2014.

Preston, J. L., Brick, N., & Landi, N. (2013). Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *American Journal of Speech-Language Pathology, 22(4),* 627-643.

Preston, J. L., McCabe, P., Rivera-Campos, A., Whittle, J. L., Landry, E., & Maas, E. (2014). Ultrasound Visual Feedback Treatment and Practice Variability for Residual Speech Sound Errors. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2102-2115.

Raven, J., Raven, J.C., & Court, J.H. (2003, updated 2004). Manual for Raven's Progressive Matrices and Vocabulary Scales. San Antonio, TX: Harcourt Assessment.

Robbins, J. & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52(3),*271-7.

The Seeing Speech Website http://www.seeingspeech.arts.gla.ac.uk/uti

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals, Fourth UK Edition.* San Antonio, TX: Harcourt Assessment.

Scobbie, J.M., Gordeeva, O.B. & Mathews, B. (2007). *Scottish English speech acquisition.* In: The International Guide to Speech Acquisition. Clifton Park, NY: Thomson Delmar Learning, (pp. 221-240). Thomson Delmar Learning, New York.

Scobbie, J.M., Wrench, A. & van der Linden, M. (2008). Head-Probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. *Proceedings of the 8th International Seminar on Speech Production*, 373-376.

Stackhouse, J. & Wells, W. (1997). *Children's Speech And Literacy Difficulties: A Psycholinguistic Framework* London: Wiley.

Ultrax. (2014). Overview of Research for Ultrax [online] Available at: http://www.ultrax-speech.org/research [Accessed September 12 2014].

Van Riper, C. & Emerick, L. (1984). *Speech Correction (7th edition).* Englewood Cliffs, NJ: Prentice Hall.

Waring, R. & Knight, R-A. (2013). 'How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems'. *International Journal of Language and Communication Disorders, 48*, 24 -40.

Wood, S. & Scobbie, J.M. (2003). *Evaluating the Clinical Effectiveness of EPG in the Assessment and Diagnosis of Children with Intractable Speech Disorders.* Proceedings of the Vth European CPLOL Congress. Evidence Based Practice: a Challenge for Speech and Language Therapists.

Wrench, A. & Scobbie, J.M. (2011). Very High Frame Rate Ultrasound Tongue Imaging. *Proceedings of the 9th International Seminar on Speech Production* (pp. 155-162), Montreal, Canada.

Table 1

CHILD	SEX	CHRON AGE	BPVS SS	CELF SS	RK%	Raven's	DEAP % B1	DEAP % B2
01M	M	6;0	109	109	91.23	>95	74.67	74.67
02M	M	10;1	95	111	99.12	95	88	87.33
03F	F	8;7	94	73	85.96	25	73.15	77.85
04M	M	8;11	<70	<70	82.46	25	69	69
05M	M	6;5	107	82	85.96	10 to 25	52	52
06M	M	5;11	80	97	85.09	>95	62	58
07F	F	7;6	109	Missing	99.12	Missing	86	86
08M	M	7;7	112	Missing	91.23	Missing	60.28	60.28
MEAN		7.63	100.86	94.40	90.02	48.33	70.64	70.64
STDEV		1.50	11.63	16.64	6.36	40.41	12.49	13.09

Table 1: Baseline Assessments. BPVS: British Picture Vocabulary Scale; CELF: Clinical Evaluation of Language Fundamentals, Core Language Subtest; RK: Robbin's Klee test of oro-motor function, % function correct; Raven's: Raven's matrices test of nonverbal ability, standard score; DEAP: phonology subtest, percentage consonants correct at baseline 1 and 2.

Table 2

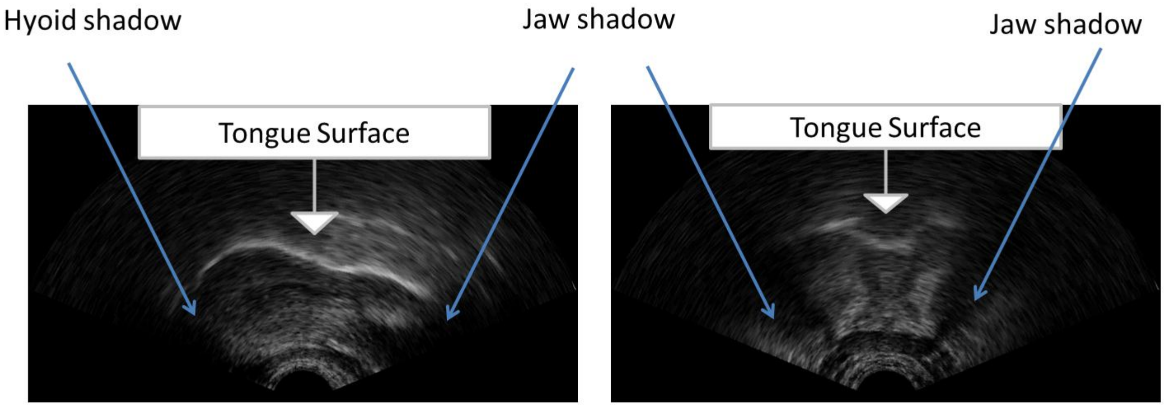
		CHILD							
	PROCESS	01M	02M	03F	05M	06M	07F	08M	Total	%Total Errors
Lingual place related errros	Velar Fronting	16		2	14	12	18		62	24.03
	Labialisation of sibilants							9	9	3.49
	Post alveolar fronting	2		7	5				14	5.43
	Backing							16*	0	0.00
	Gliding	13		13	10	12		9	57	22.09
	Idiosyncratic /r/- [ʎ]		12						12	4.65
	Stopping					1			1	0.39
	Deaffrication			2	1		1	3	7	2.71
	Voicing errors				1	4		2	7	2.71
	Weak syllable deletion				1	1			2	0.78
	FCD			3	1				4	1.55
	ICD			1					1	0.39
	MCD				1				1	0.39
	Cluster Reduction			6	7	9			22	8.53
	Other			4	7	2		1	14	5.43
	TOTAL	31	12	38	48	41	19	24	213	100
	% Total Errors	12.016	4.651	14.729	18.605	15.891	7.364	9.302	82.6	100
	* Alveolar plosives backed to palatal lateral affricates									

Table two: Phonological and phonetic process analysis of errors produced in the phonology subtest of the DEAP (Dodd et al., 2005). Shading denotes the lingual error (gliding excluded) chosen for treatment.

Table 3

	01M	02M	03F	05M	06M	07M	08M
Mean diff (mm)	6.4	2.1	2.6	3.2	5.9	4.0	8.4
sd (mm)	3.3	1.7	0.9	2.2	0.7	1.8	2.3
max diff (mm)	11.0	5.3	3.5	6.8	6.5	6.5	11.2
max length (cm)	5.2	6.0	4.3	4.3	1.7	3.0	3.5
min length (cm)	4.2	3.3	4.0	1.6	1.7	3.0	3.5

Table 3 – Average distance between tongue curves for each participant and the total length of surface over which this mean is calculated and the maximum difference found. The minimum length of tongue surface corresponds to a single shorter length of tongue in which there are at least 6 contiguous areas of significant difference, omitting cross-overs or lengths in which there is only a trend. Note 02M has two such areas (2.0cm & 1.3cm).



Mid-sagittal Coronal

Figure 1: Typical ultrasound image in mid-sagittal (left) and coronal (right) views. The surface of the tongue is visible as a white line.

	Week Number																
Week no	1	....	6	7	8	9	10	11	12	13	14	15	16	17	18	19	...	25
PROBE	B1		B2	12 Sessions of therapy												M1		M2
								Probe										

Figure 2: Probe and treatment schedule.

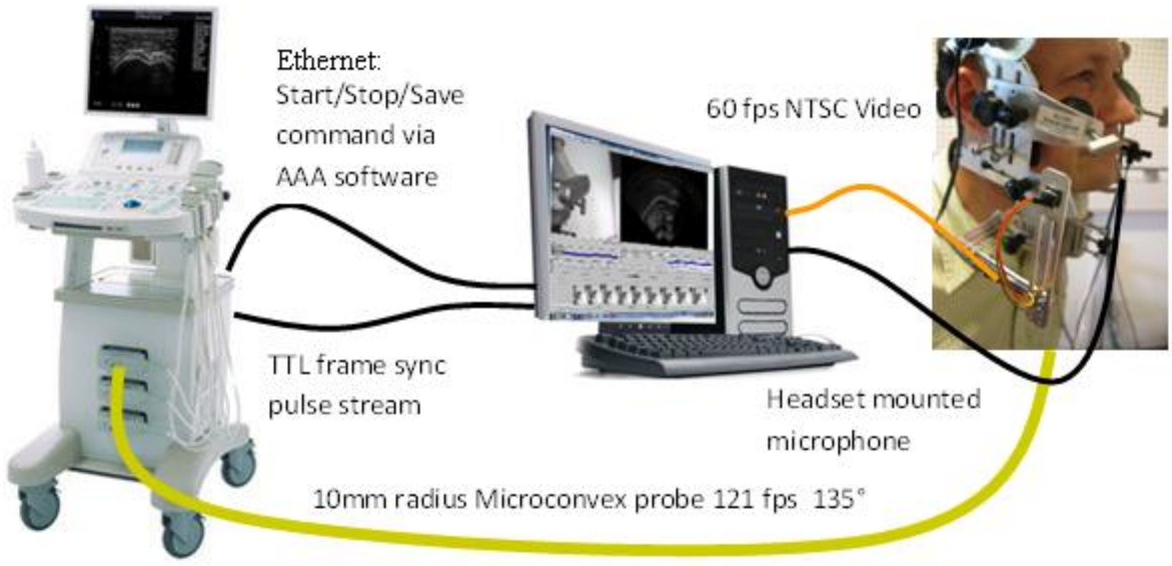
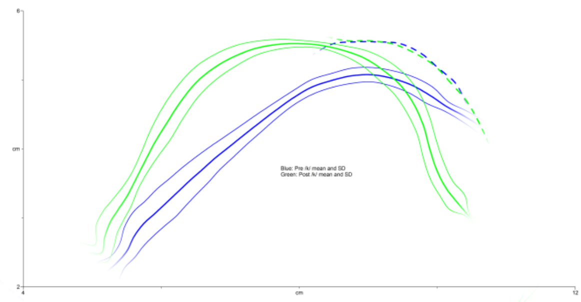


Figure 3: Ultrasound recording set-up showing probe-stabilisation headset.



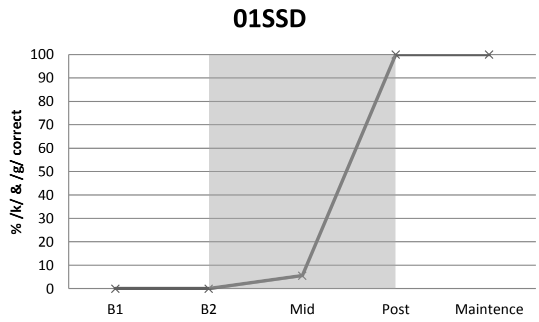
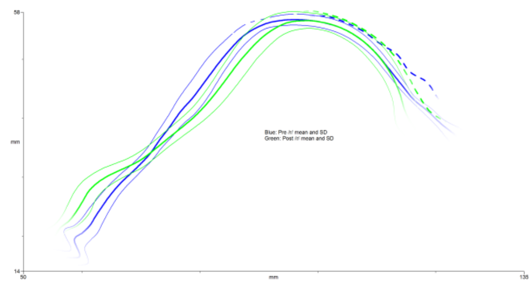


Figure 4a: **01M:** Percentage segments correct (transcription data) at each time point

Figure 4b: **01M:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes



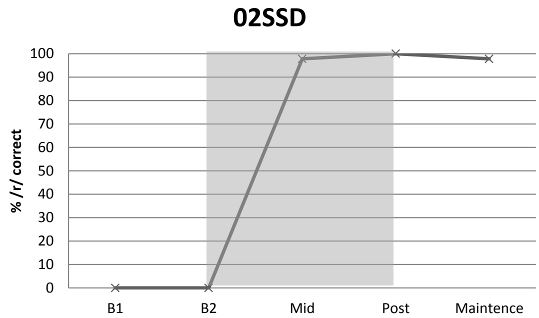
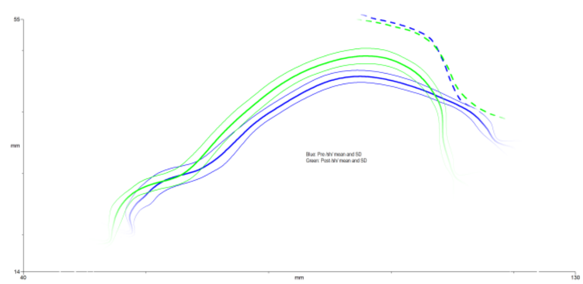


Figure 5a: **02M:** Percentage segments correct (transcription data) at each time point

Figure 5b: **02M:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes



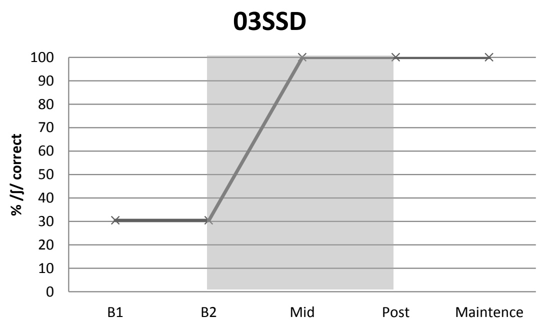
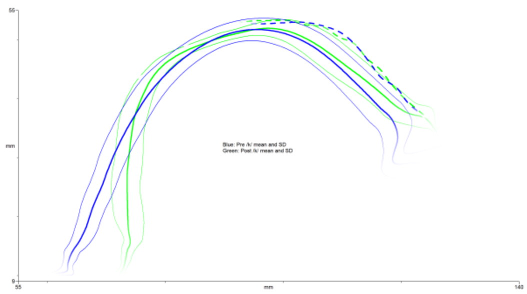


Figure 6a: **03F:** Percentage segments correct (transcription data) at each time point

Figure 6b: **03F:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes



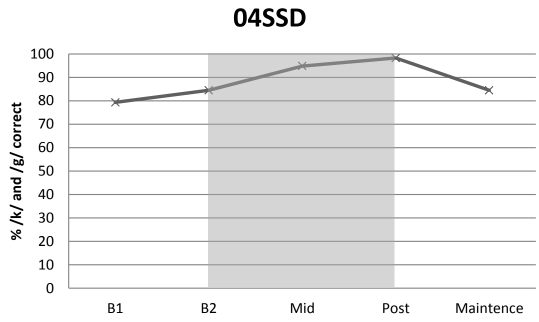
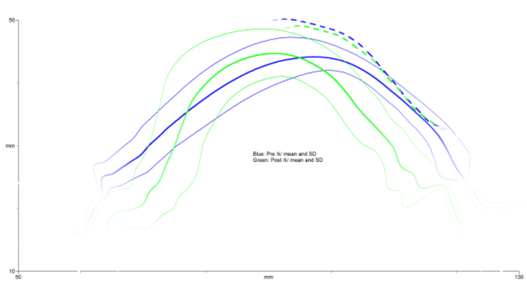


Figure 7a: **05M:** Percentage segments correct (transcription data) at each time point

Figure 7b: **05M:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes



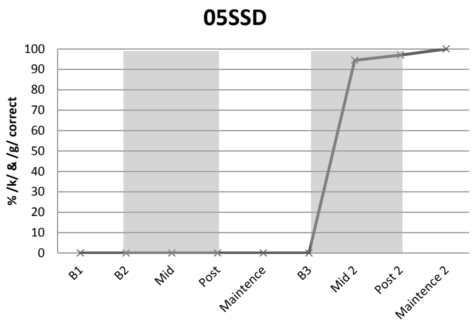
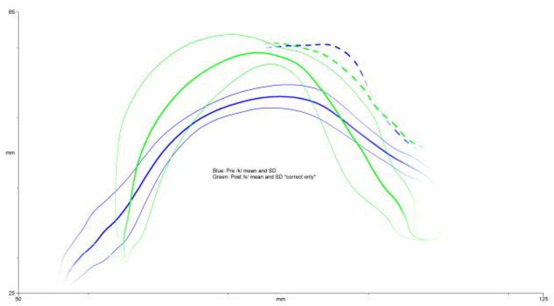
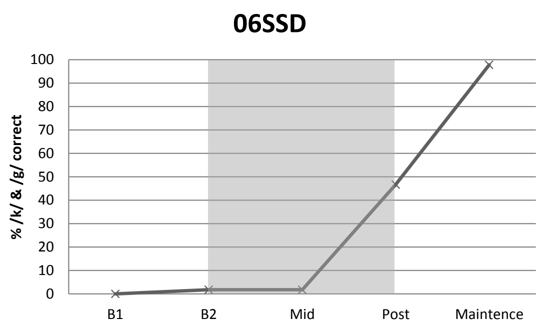
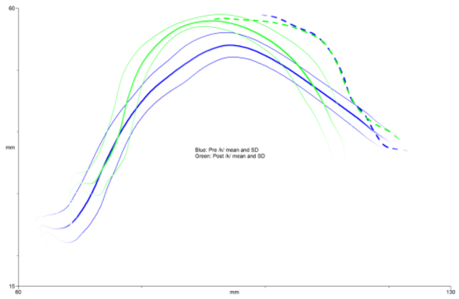


Figure 8a: **06M:** Percentage segments correct (transcription data) at each time point

Figure 8b: **06M:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes, correct segments only







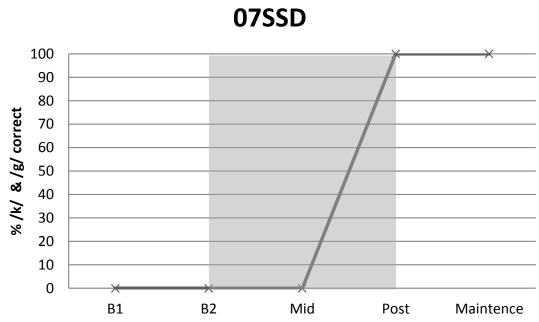
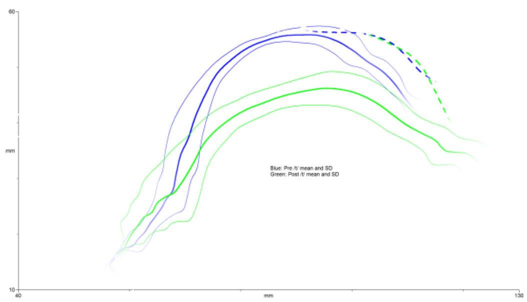


Figure 9a: **07F:** Percentage segments correct (transcription data) at each time point

Figure 9b: **07F:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes



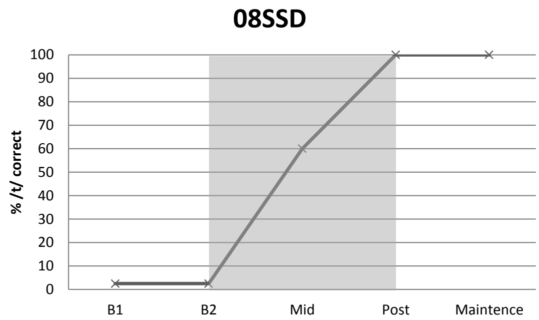


Figure 10a: **08M:** Percentage segments correct (transcription data) at each time point

Figure 10b: **08M:** Pre- (blue) and post-therapy (green) average ultrasound tongue-shapes