Testing hypotheses about the underlying deficit of apraxia of speech through computational neural modelling with the DIVA model

Hayo Terband, Joe Rodd & Edwin Maas

Published online: 20 Oct 2019.

Submit your article to this journal ⏎

Article views: 6

View related articles ⏎

View Crossmark data ⏎

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Check for updates

Testing hypotheses about the underlying deficit of apraxia of speech through computational neural modelling with the DIVA model

HAYO TERBAND[1] , JOE RODD[2] & EDWIN MAAS[3]

[1]*Utrecht Institute of Linguistics-OTS, Utrecht University, Utrecht, The Netherlands,* [2]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands,* [3]*Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, USA*

Abstract

Purpose: A recent behavioural experiment featuring a noise masking paradigm suggests that Apraxia of Speech (AOS) reflects a disruption of feedforward control, whereas feedback control is spared and plays a more prominent role in achieving and maintaining segmental contrasts. The present study set out to validate the interpretation of AOS as a possible feedforward impairment using computational neural modelling with the DIVA (Directions Into Velocities of Articulators) model.
Method: In a series of computational simulations with the DIVA model featuring a noise-masking paradigm mimicking the behavioural experiment, we investigated the effect of a feedforward, feedback, feedforward + feedback, and an upper motor neuron dysarthria impairment on average vowel spacing and dispersion in the production of six/bVt/speech targets.
Result: The simulation results indicate that the output of the model with the simulated feedforward deficit resembled the group findings for the human speakers with AOS best.
Conclusion: These results provide support to the interpretation of the human observations, corroborating the notion that AOS can be conceptualised as a deficit in feedforward control.

Keywords: *apraxia of speech; computational modelling; vowel acoustics; feedback masking*

Introduction

Apraxia of Speech (AOS) is a neurogenic motor speech disorder that is defined as an impairment in the planning and/or programming of speech movements (Deger & Ziegler, 2002; Duffy, 2005; Van der Merwe, 1997). The speech of people with AOS is characterised by slow speech rate, abnormal prosody, abnormal speech sound and syllable segmentation, speech sound distortions, and speech errors that are inconsistently present but relatively consistent in type and location (Duffy, 2005; Maas et al., 2008). AOS typically results from brain lesions to the left cerebral hemisphere, but more specific lesion locations reported in the literature diverge. Most reports indicate lesions in left inferior frontal regions (e.g. Dronkers, 1996; Graff-Radford et al., 2014; Hickok et al., 2014; Hillis, et al., 2004; Itabashi et al., 2016; Richardson, Fillmore, Rorden, LaPointe, & Fridriksson, 2012), however, other regions have also been reported, including the parietal cortex (e.g.

Hickok et al., 2014; McNeil, Weismer, Adams, & Mulligan, 1990), basal ganglia (Seddoh et al., 1996), and right frontal cortex and basal ganglia structures (Balasubramanian & Max, 2004). The precise location of the lesion responsible for AOS thus remains subject of debate. Likewise, the precise nature of the disorder remains poorly understood.

One of the main difficulties in isolating the underlying deficit(s) is diagnostic circularity. The ability to investigate the characteristics underlying AOS requires *pure* cases of AOS selected on the basis of clear-cut criteria, which are only available as a result of research. As lesion inducing medical accidents such as strokes, brain injuries, or tumours rarely produce isolated and one-dimensional deficits, pure cases are rare and symptom profiles show considerable variation between individuals as well as a large overlap in symptomatology with other speech disorders. Additionally, when confronted with a partial breakdown, the speech system itself is likely to adapt to the deviant

circumstances and/or compensate for the impediments. Individuals may vary widely in these adaptive and compensatory mechanisms.

This problem of practical-diagnostic circularity results from the behavioural, symptom-oriented approach that is employed (McNeil, Pratt, & Fosset, 2004; Terband, Maassen, & Maas, 2017, 2019). Although the symptomatology might be aspecific, it is possible to describe a specific speech-motor core deficit from the perspective of the underlying cognitive and neurological processes. As such, we argue that to identify underlying deficits, one must begin by deriving detailed, specific hypotheses within the context of a detailed model of the behavioural and cognitive operations involved. These hypotheses should then be tested empirically, and ideally contrasted with alternative hypotheses for underlying deficits (e.g. those presumed to underlie other impairments such as dysarthria; Terband et al., 2017).

One promising, and relatively recent approach to understanding AOS in this respect, relates to the development of the DIVA (Directions Into Velocities of Articulators) model, a computationally implemented neural network model of speech acquisition and speech motor control (Guenther, 1994; Guenther, Ghosh, & Tourville, 2006). The main function of computational modelling for our purpose is to understand the effects of a particular underlying deficit. Computer simulations with computational models allow for controlling more tightly for deficit modality, manipulating parameters independently and systematically, and examining the complex response of the system to deficits. Currently, clinicians tend to interpret symptoms at face validity (e.g. errors in place of articulation as resulting from motor programming errors). Although models are simplifications of reality, the deductive nature of detailed models allows us to test such interpretations directly in a more controlled and specific manner, thereby giving us a powerful tool for validating inductive reasoning (from symptom to deficit; Terband & Maassen, 2010; Terband et al., 2017). In the current study, we utilised this modelling approach and set out to investigate the potential role of two deficits that have been hypothesised to underlie AOS (Maas, Mailend, & Guenther, 2015) in Simulink DIVA (Nieto-Castanon, 2011), a computational implementation of the DIVA model.

Overview of the DIVA model

The DIVA model consists of a neural network controller detailing feedforward and feedback control loops that are assumed to be involved in early speech development and mature speech production, focussing on the sensorimotor transformations underlying the control of articulator movements (Guenther, 1994; Guenther et al., 2006). The model strives to be neurobiologically plausible and its components have been associated with regions of the cerebral cortex

and cerebellum (Guenther et al., 2006). In order to produce an acoustic signal, DIVA controls the movements of an articulatory synthesiser (Maeda, 1990).

In the DIVA model, the production of a speech sound begins with activation of a speech sound map (SSM) cell in left inferior frontal cortex. SSM cells represent speech sounds (the size of phonemes, syllables, or frequent words and phrases) and are activated by higher-level input from the phonological encoding stage (Bohland, Bullock, & Guenther, 2010; Guenther et al., 2006). The activated SSM cell activates a feedforward control system and a feedback control system, whose motor commands are combined in primary motor cortex. Feedback control involves comparing actual auditory and somatosensory feedback signals to expected auditory and somatosensory consequences, and generating corrective motor commands to motor cortex when a mismatch (error) is detected. Expected sensory consequences are encoded as regions in auditory space (superior temporal gyrus) and somatosensory space (postcentral and supramarginal gyri). Feedforward control involves predictive motor commands from the SSM to motor cortex. Feedforward commands are learned by incorporating the feedback system's corrective commands from previous productions. With sufficient practice, the feedforward commands generate little to no errors, so that contributions of the feedback control system are minimal during normal speech, although feedback may be continuously monitored for deviations from expectations, even in adult speakers (Tourville, Reilly, & Guenther, 2008).

Aim of the present study

As noted above, the current consensus is that AOS is a speech motor planning and/or programming disorder, or, more specifically, an inability to transform an abstract linguistic code involving intact phonological representations into spatially and temporally coordinated patterns of muscle contractions that produce speech movements (e.g. Duffy, 2005; Maas et al., 2008). Within this accepted consensus, Maas et al. (2015) proposed two alternative hypotheses with respect to the underlying mechanisms. One suggestion was that the underlying (core) deficit in AOS may be viewed as one of impaired feedforward control (Feedforward System Deficit Hypothesis; FF hypothesis, see also Jacks, 2008; Rogers, Eyraud, Strand, & Storkel, 1996). The disruption in feedforward processing would cause the motor commands to be inappropriate or underspecified, thereby introducing errors. The mismatch between produced and target signal would evoke the feedback control subsystems to generate a corrective command, increasing the contribution of feedback-based corrective commands to the overall motor command, causing the system to rely more heavily on sensory feedback control subsystems. Thus, according to the

FF hypothesis, the role of feedback control is facilitatory in achieving and maintaining segmental contrast in speakers with AOS. A greater reliance on feedback control could account for slower speech rate, due to the need to process and incorporate the feedback signals (e.g. Rogers et al., 1996) or because slowing down speech rate is known to facilitate the use of feedback (e.g. Adams, Weismer, & Kent, 1993; Civier, Tasko, & Guenther, 2010; Terband & Maassen, 2010). Furthermore, corrections needed to counter incorrect feedforward commands would lead to articulatory adjustments (both online and in repeated productions) and could thus account for increased spatial and/or temporal variability (e.g. Jacks, 2008; Terband, Maassen, Guenther, & Brumberg, 2009).

The other suggestion proposed by Maas et al. (2015) was that AOS may involve impaired feedback control (Feedback System Deficit Hypothesis; FB hypothesis, see also e.g. Kent & Rosenbek, 1983; Rogers et al., 1996). Note that AOS does not involve an impairment of auditory perceptual processing (e.g. Deal & Darley, 1972; Kent & Rosenbek, 1983; Square, Darley, & Sommers, 1981), rather this hypothesis comprises an impairment in processing feedback and transforming feedback information into motor commands. This disruption of feedback processing would cause inappropriate or underspecified corrective commands, thereby introducing errors and rendering the system unable to correct for errors. As such, difficulties with using feedback could account for several features of AOS, including articulatory groping, speech sound distortions, and increased variability (Maas et al., 2015).

Maas and colleagues (2015) investigated these two hypotheses in a behavioural experiment featuring an auditory feedback masking paradigm. The rationale was that masking noise effectively prevents auditory feedback control, forcing reliance on feedforward control (and somatosensory feedback control). If the feedforward system is impaired, and people with AOS rely primarily on auditory feedback control to maintain segmental contrast, then removal of auditory feedback would reveal the – impaired – feedforward system. On the other hand, if symptoms of AOS reflect interference from the auditory feedback signal (e.g. due to generating unnecessary or inadequate corrective commands), then removing the auditory feedback should improve speech performance in terms of segmental contrast and stability.

Findings from vowels produced by six speakers with AOS revealed that at the group level, vowel spacing (acoustic contrast) was more reduced under masking noise conditions than in control speakers, consistent with the FF hypothesis. Further, a marginal interaction between group and condition emerged for vowel dispersion (the token-to-token variability of a vowel around its mean location in F1

× F2 space; Perkell et al., 2007), hinting at greater dispersion for the AOS group than the controls in the clear (no-masking) condition but comparable dispersion in the masking noise condition. This pattern would be expected if speakers with AOS rely to a greater extent than controls on auditory feedback control, which tends to be more variable due to on-line corrections to motor commands. Although not all individuals with AOS showed this pattern, these group-level findings support the notion of impaired feedforward control in AOS (or at least in a subgroup of people with AOS; Maas et al., 2015).

However, as discussed above, lesion-inducing medical accidents rarely produce isolated and one-dimensional deficits. Two out of three patients in the Maas et al. study were also diagnosed with mild dysarthria, leaving open the possibility that the findings are due to the comorbid mild dysarthria. In the present study, we set out to further validate the interpretation of AOS as a feedforward impairment by means of a series of computational simulations mimicking the experiment of auditory feedback masking in human speakers with AOS by Maas and colleagues (2015). This modelling paradigm allows us to control for deficit modality and examine the response of the system to different isolated deficits.

Method

DIVA simulations

Experimental paradigm
The modelling experiment was designed to mimic the behavioural experiment (Maas et al., 2015) as closely as possible. A total of four impairment conditions were implemented in the DIVA model; the Feedforward System Deficit (FF) and the Feedback System Deficit (FB) conform the two hypotheses of Maas and colleagues. To test the specificity of these hypotheses, two further impairment conditions were implemented; a model in which both the feedforward and feedback systems are affected (FF + FB), and a model where the integrated motor commands are disrupted, resembling upper motor neuron dysarthria (UMN-DYS). Subsequently, the behaviour of these impaired models was tested in a series of computational simulations. The same/bVt/tokens were used and the vowel productions of the AOS-models in the condition of normal auditory feedback were compared to productions in a masking condition in which auditory feedback from the model's articulatory synthesiser was blocked. In addition, an unimpaired, healthy model served as a control condition. We acknowledge that eliminating auditory feedback altogether is a simplification of the model simulation compared to a human study. However, this method of implementing auditory feedback masking captures the relevant aspects of noise masking (eliminating auditory feedback). Where human studies have to go

to great lengths to control or correct for behaviour associated with speaking in noise (e.g. increases in loudness), these confounds can be fully blocked in computer simulations. As such, the simulations in effect represent a clear test of the hypotheses (e.g. whether impaired feedforward control, without the benefit of auditory feedback, results in disproportionate changes in vowel articulation).

Impairment conditions
Modified versions of the Simulink DIVA model (Nieto-Castanon, 2011) were derived from a pre-trained model that in its original, healthy state produces stable, mature output. The DIVA model features a noise generator, by which Gaussian random noise (uncorrelated, signal-independent, zero-mean noise) can be added to the cell activations of specific parts of the motor-, auditory-, and somatosensory cortices as a means of simulating impairments.[1] By specifying the standard deviation of the Gaussian distribution, the level of the noise (simulated severity) can be manipulated. The Feedforward System Deficit and the Feedback System Deficit were implemented by adding random signal-independent noise to respectively the feedforward (FF) and the feedback (FB) command before their integration in the model's motor cortex. The combined FF + FB deficit was implemented by introducing random signal-independent noise to both the FF and FB commands simultaneously. The upper motor neuron dysarthria impairment was implemented by adding random signal-independent noise to the motor command after integration of the feedforward and the feedback commands (but before integration with the current articulatory position/state of the vocal tract). Following previous simulation studies (Terband, Maassen, Guenther, & Brumberg, 2014; Terband & Maassen, 2010), we implemented two levels of severity, 5% and 10% signal degradation respectively. In these prior

studies, noise levels up to 25% were investigated. Noise levels exceeding 10% lead to very severely impaired output and with noise levels exceeding 15%, the produced speech features so much distortion and irregularities that it no longer resembles speech (Terband et al., 2014). Against this background, and given that the speakers in the Maas et al. (2015) study had AOS severities in the mild to moderate range, we believe that the 5% and 10% degrees of noise used in the present study are reasonable.

Speech targets
The target items that the DIVA model seeks to produce are specified by time-aligned minimum and maximum limits for relevant acoustic and articulatory parameters. Acoustic dimensions consist of pitch, F1, F2, and F3 while the articulatory dimensions of pressure, voicing, and closure at pharyngeal, uvular, palatal, alveolar-dental and labial places of articulation are expressed on an arbitrary −1 to 1 scale. In the simulations, we used the same /bVt/ speech targets as in Maas et al. (2015), including the six vowels /ae/, /i/, /ɛ/, /ɪ/, /u/, and /ʌ/ (*bat, beat, bet, bit, boot, butt*; Figure 1). Vowel formant targets were derived from 95th and 5th percentile linear predictive coding peaks from twenty natural productions of each vowel produced by the second author. Articulatory targets for the consonants were based on sample target items distributed with the computational DIVA model. All items were time-normalised and set to 500 ms; thus, the total duration and vowel and consonant onset and offset times were the same in all items.

Simulation procedure
Prior to the simulations, a base model was trained with 40 initial training trials comprising the production of each of the six words in the vocabulary (corresponding to asymptotic learning for the current stimuli). This initial training stage was the same for
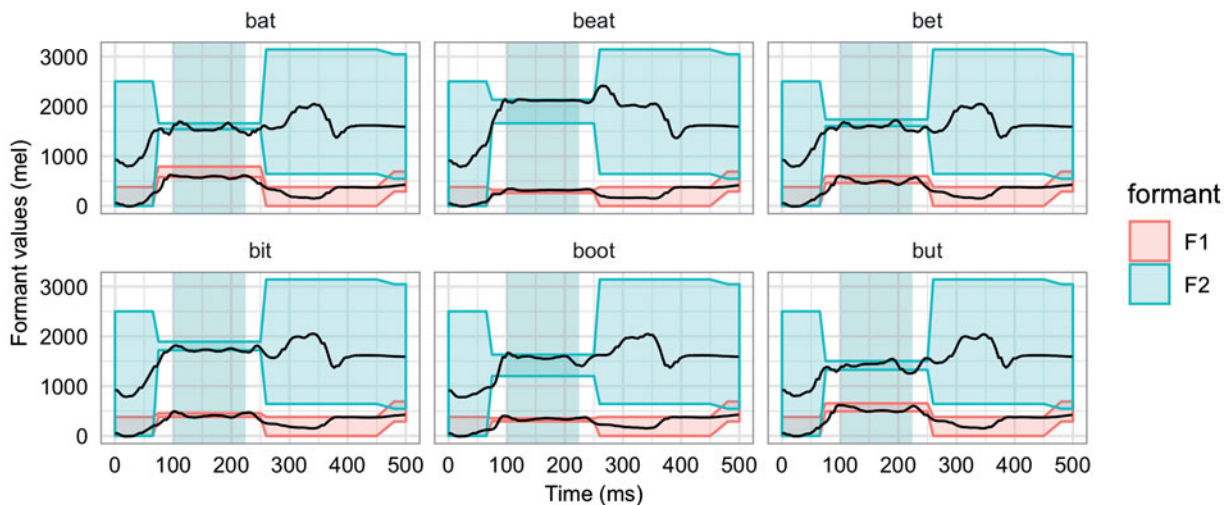

Figure 1. A schematic representation of the F1 and F2 dimensions of the /bVt/ speech sound targets used in the simulations. The acoustic targets consist of time varying upper and lower bounds for each dimension, and have a fixed duration of 500 ms. The target region frequencies (Hz, *y*-axis) for F1 and F2 over time (*x*-axis, ms) are depicted (shading). A vertical band indicates the vowel analysis window.

all conditions and did not involve any impairment (as the speakers in Maas et al., 2015 were all adult speakers prior to AOS onset). The resulting model formed the base model for our simulations.[2] For each combination of impairment condition and severity level, as well as the healthy control condition, a three-stage simulation procedure was then executed. First, 20 warm-up trials (simulations of single word productions) with auditory feedback were run, to allow the performance of the model to stabilise in its new, impaired condition (simulating the fact that the speakers in Maas et al., 2015, all had chronic, not acute, AOS). The 20 stabilisation trials were followed by 10 experimental trials with auditory feedback. Finally, auditory feedback was masked and 10 further experimental trials were run. This procedure was repeated 20 times. Values for the formants calculated by the articulatory synthesiser were analysed.

Acoustic analysis
Acoustic analysis followed the procedure used in Maas et al. (2015). For each trial, the mean Mel values of the first and second formants (F1 and F2) were calculated over all samples in a 50 ms window around vowel midpoint. Average vowel spacing (AVS) was calculated as the mean Euclidean distance between the means of each of the 15 possible pairs of vowels. AVS was calculated for each trial in the simulation procedure separately for each of the 20 repetitions. Average vowel dispersion (AVD) was calculated as the average of the Euclidian distances between each vowel token and that vowel's mean. AVD was calculated across the 20 repetitions of each vowel for each trial number.

Statistical analysis
An all-subsets approach was used to fit mixed effects models predicting AVS and AVD respectively, using the lme4 R package (Bates, Maechler, Bolker, &

Walker, 2014). Models with only random predictors were fitted first. Subsequently, all the possible combinations of fixed predictors (including interactions) were added to the models. The resulting models were compared with each other (AVS and AVD models separately), and the models with the best fit (lowest AIC) which was also a significant improvement over the previous model (χ^2 test) were selected. In both cases, these were the models including all fixed factors and interactions (masking, impairment, and the interaction between masking and impairment).

All categorical predictors were dummy coded, with the healthy, non-masking conditions on the intercept. Satterthwaite approximations for degrees of freedom were used to make it possible to calculate *p*-values (Kuznetsova, Brockhoff, & Christensen, 2017). Finally, we also ran models fitting standardised dependent variables to calculate Cohen's *d* effect sizes.

Result

Average vowel spacing (AVS)

Figure 2 presents the AVS results over the course of the simulations for the four impairment models compared to the healthy model. The results show considerable variability across trials for the UMN-DYS, FF and FF + FB models in both the no masking- and the masking condition. The trial-to-trial variability is particularly large in the FF and FF + FB models, and appears larger in the masking condition. The healthy and FB models show very little variability across trials.

Mean AVS in masking and no masking conditions for the healthy control and four impairment models are presented in Figure 3. The statistical model of AVS included predictors of masking and impairment, and the interaction masking × impairment. Wald χ^2
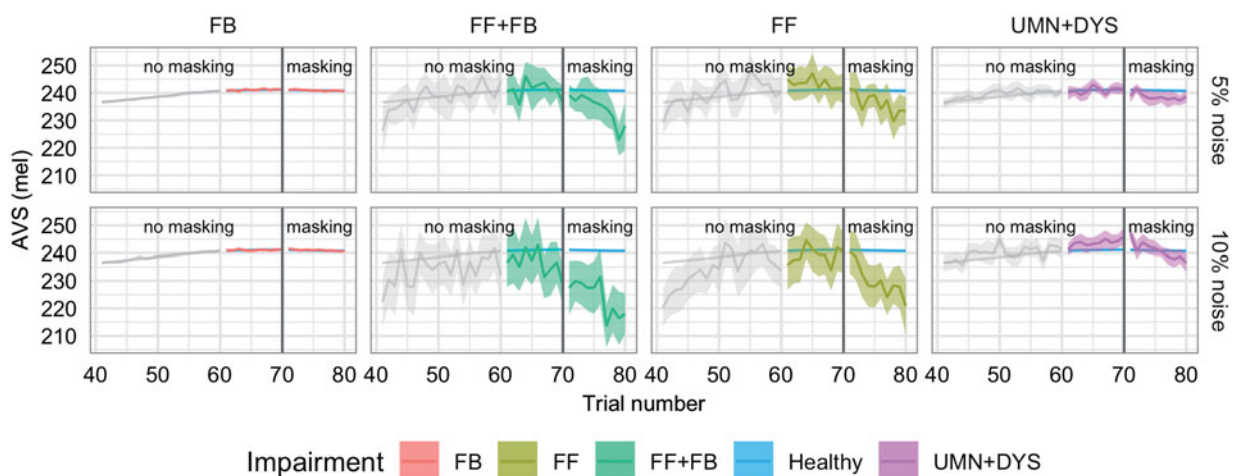

Figure 2. Mean average vowel spacing (AVS; mel, *y*-axis) over the course of the simulations (*x*-axis), for the three impairment conditions (colors, columns), at 5% and 10% impairment severity (rows). The healthy reference (purple) is repeated in each cell for comparison. Shading around the lines indicates 95% confidence intervals around the mean. Grey lines and shading indicate the stabilisation trials. A vertical line indicates the onset of auditory feedback masking.
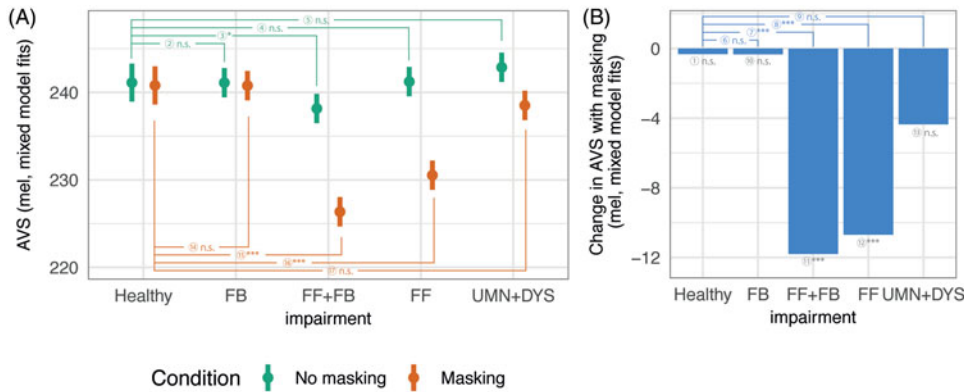
Figure 3. Panel A: Mean average vowel spacing (AVS) estimated by the mixed model (points) in mel (*y*-axis) for the healthy control model and the four impairments (*x*-axis), in masking (green, left) and no masking (orange) conditions. Bars indicate the 95% confidence intervals around the estimated means. Brackets indicate reported statistical comparisons on simple effects and pairwise least-square mean comparisons. Circled numbers identify the effects referred to in the text and in model tables in the Supplementary materials. Panel B: Bars indicate the difference (mel) between the mean AVS estimated by the mixed model in masking and no-masking conditions, in the healthy model and impaired models (*x*-axis). Brackets indicate interaction effects, beneath each bar, the relevant pairwise least-square mean comparison is indicated.

tests revealed that all fixed predictors contributed significantly to the predictiveness of the model (masking: $\chi^2 = 30.13$***; impairment: $\chi^2 = 251.95$***; masking × impairment interaction: $\chi^2 = 118.75$***).

In the healthy model, masking resulted no change in vowel separation (① in Appendix model summaries and Figure 3, $\beta = -0.14$, SE $= 1.3$, $t = -0.11$, $d = -0.012$). In the no masking condition, the combined FF + FB model showed significantly lower AVS (reduced vowel space) compared to the healthy model (③, $\beta = -2.3$, SE $= 0.95$, $t = -2.4$*, $d = -0.19$). The FB, FF and UMN-DYS models showed no difference from the healthy model (FB: ②, $\beta = 0.027$, SE $= 0.95$, $t = 0.029$, $d = 0.002$; FF: ④, $\beta = 0.24$, SE $= 0.95$, $t = 0.25$, $d = 0.02$; UMN-DYS: ⑤, $\beta = 1.3$, SE $= 0.95$, $t = 1.3$, $d = 0.11$).

The interaction between masking and impairment was significant for the FF and the FF + FB models. This means that there was significantly more difference in AVS as a result of masking in those models than in the healthy model (FF + FB: ⑦, $\beta = -9.4$, SE $= 1.3$, $t = -7$***, $d = -0.79$; FF: ⑧, $\beta = -8.1$, SE $= 1.3$, $t = -6$***, $d = -0.68$). The effect of masking on AVS in the FB and UMN-DYS models did not differ from the healthy model (FB: ⑥, $\beta = 0.012$, SE $= 1.3$, $t = 0.009$, $d = 0.001$; UMN-DYS: ⑨, $\beta = -2.5$, SE $= 1.3$, $t = -1.8$, $d = -0.21$).

Post hoc comparison of least-square means (with Bonferroni correction for multiple comparisons) revealed that the vowel space was reduced in the masking condition compared to the non-masking condition in the FF and FF + FB models (FF + FB: ⑪, $\Delta\beta = 9.5$, SE $= 1.1$, $t = 8.9$***, $d = 0.81$; FF: ⑫, $\Delta\beta = 8.2$, SE $= 1.1$, $t = 7.6$***, $d = 0.69$). No significant vowel space reduction was in evidence in the FB and UMN-DYS models (FB: ⑩, $\Delta\beta = 0.13$, SE $= 1.1$, $t = 0.12$, $d = 0.011$; UMN-DYS: ⑬, $\Delta\beta = 2.6$, SE $= 1.1$, $t = 2.4$, $d = 0.22$).

In the masking condition, relative to the healthy model, the FF and FF + FB models showed a significant difference in AVS (FF + FB: ⑮, $\Delta\beta = 12$, SE $= 0.95$, $t = 12$***, $d = 0.99$; FF: ⑯, $\Delta\beta = 7.8$, SE $= 0.95$, $t = 8.3$***, $d = 0.66$). The FB and UMN-DYS models showed no significant difference (FB: ⑭, $\Delta\beta = -0.04$, SE $= 0.95$, $t = -0.042$, $d = -0.003$; UMN-DYS: ⑰, $\Delta\beta = 1.2$, SE $= 0.95$, $t = 1.3$, $d = 0.1$).

Average vowel dispersion

Figures 4 and 5 present the AVD results. Similar to the AVS results, the results for AVD show considerable variability across trials in both the no masking- and the masking condition for the UMN-DYS, FF and FF + FB models but not for the healthy and FB models. The statistical model of AVD included predictors of masking and impairment, and the interaction of masking × impairment. Wald χ^2 tests revealed that all variables contributed significantly to the predictiveness of the model (masking: $\chi^2 = 3.25$; impairment: $\chi^2 = 18842.94$***; masking × impairment interaction: $\chi^2 = 1.01$). The model was fitted on 21600 observations.

In the healthy model, masking had no effect on AVD (㉑ in Appendix model summaries and Figure 4, $\beta = -0.5$, SE $= 0.5$, $t = -1$, $d = -0.03$).

In the no masking condition, the combined FF + FB, FF and UMN-DYS models showed significantly higher AVD (more dispersion) than the healthy model (FF + FB: ㉓, $\beta = 26$, SE $= 0.43$, $t = 60$***, $d = 1.5$; FF: ㉔, $\beta = 25$, SE $= 0.43$, $t = 57$***, $d = 1.5$; UMN-DYS: ㉕, $\beta = 6.7$, SE $= 0.43$, $t = 16$***, $d = 0.4$). The FB model did not differ from the healthy model (㉒, $\beta = 0.72$, SE $= 0.43$, $t = 1.7$, $d = 0.043$).

The interaction between masking and impairment was not significant for any of the impaired models. This means that the (lack of) difference in AVD as a result of masking was comparable to the difference in the healthy model (FB: ㉖, $\beta = 0.16$, SE $= 0.6$, $t = 0.26$, $d = 0.009$; FF: ㉘, $\beta = 0.33$, SE $= 0.6$, $t = 0.55$, $d = 0.02$; FF + FB: ㉗, $\beta = 0.36$, SE $= 0.6$,
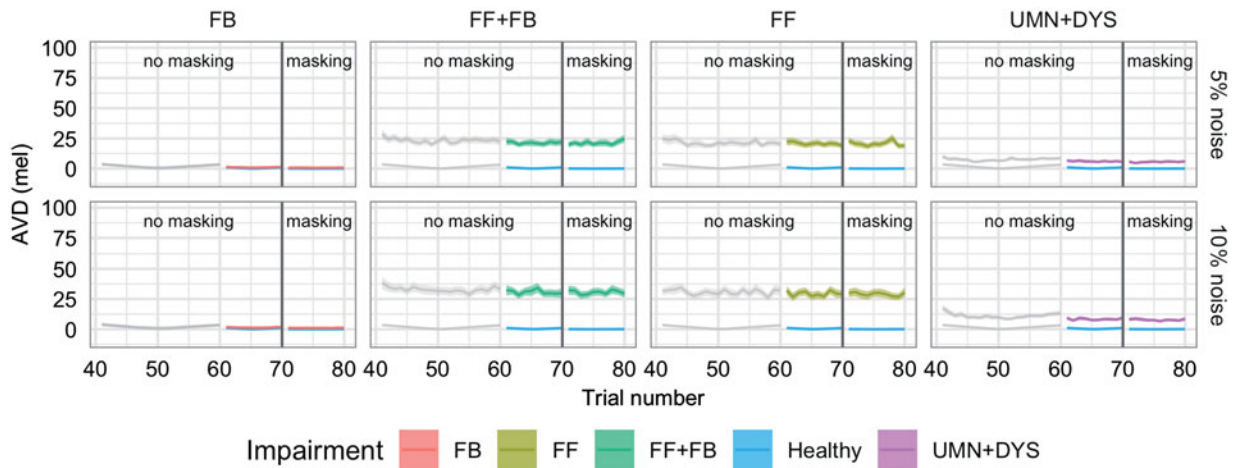
Figure 4. Mean vowel dispersion (AVD; mel, *y*-axis) over the course of the simulations (*x*-axis), for the three impairment conditions (colors, columns), at 5% and 10% impairment severity (rows). The healthy reference (purple) is repeated in each cell for comparison. Shading around the lines indicates 95% confidence intervals around the mean. Grey lines and shading indicate the stabilisation trials. A vertical line indicates the onset of auditory feedback masking.
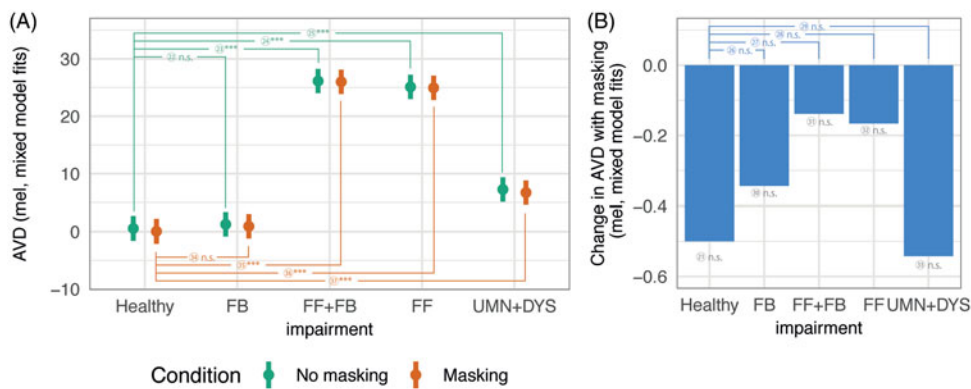

Figure 5. Panel A: Mean average vowel dispersion (AVD) estimated by the mixed model (points) in mel (*y*-axis) for the healthy control model and the four impairments (*x*-axis), in masking (green, left) and no masking (orange) conditions. Bars indicate the 95% confidence intervals around the estimated means. Brackets indicate reported statistical comparisons on simple effects and pairwise least-square mean comparisons. Circled numbers identify the effects referred to in the text and in model tables in the Supplementary materials. Panel B: Bars indicate the difference (mel) between the mean AVD estimated by the mixed model in masking and no-masking conditions, in the healthy model and impaired models (*x*-axis). Brackets indicate interaction effects, beneath each bar, the relevant pairwise least-square mean comparison is indicated.

$t = 0.6$, $d = 0.022$; UMN-DYS: ㉙, $\beta = -0.042$, SE $= 0.6$, $t = -0.069$, $d = -0.003$).

Post hoc comparison of least-square means (with Bonferroni correction for multiple comparisons) confirmed that there was no effect of masking on AVD in any of the impairment models (FB: ㉚, $\Delta\beta = 0.34$, SE $= 0.35$, $t = 0.97$, $d = 0.021$; FF + FB: ㉛, $\Delta\beta = 0.14$, SE $= 0.35$, $t = 0.39$, $d = 0.008$; FF: ㉜, $\Delta\beta = 0.17$, SE $= 0.35$, $t = 0.47$, $d = 0.01$; UMN-DYS: ㉝, $\Delta\beta = 0.54$, SE $= 0.35$, $t = 1.5$, $d = 0.032$).

In the masking condition, relative to the healthy model, the combined FF + FB, FF and UMN-DYS models showed a significant difference in AVD (FF + FB: ㉟, $\Delta\beta = -26$, SE $= 0.43$, $t = -61$***, $d = -1.6$; FF: ㊱, $\Delta\beta = -25$, SE $= 0.43$, $t = -58$***, $d = -1.5$; UMN-DYS: ㊲, $\Delta\beta = -6.6$, SE $= 0.43$, $t = -16$***, $d = -0.4$). The FB model showed no difference from the healthy model (㉞, $\Delta\beta = -0.87$, SE $= 0.43$, $t = -2$, $d = -0.052$).

Discussion

Validation of the healthy model

The present study set out to investigate the effect of different impairments that could be involved in AOS in a series of computational simulations with the DIVA model. The modelling experiment was designed to mimic a previous behavioural experiment of auditory feedback masking in speakers with AOS (Maas et al., 2015). Before evaluating the DIVA model's behaviour under the different implemented impairments, we first need to validate the healthy model against the healthy human speakers. In this respect, Maas and colleagues (2015) found no effects of masking on either AVS or AVD in the group of healthy older adults. The results of the current simulations with the healthy model also did not show significant effects of masking for AVS or AVD, for this part corroborating the healthy model as a reference

Table I. Schematic overview of main results.

	Maas et al. (2015) – human AOS speakers	This study –FB	This study –FF + FB	This study – FF	This study – UMN-DYS
AVS					
No masking condition	***Similar AVS to healthy control speakers***	**Similar AVS to healthy model** ②	*Lower AVS than healthy model* ③	**Similar AVS to healthy model** ④	**Similar AVS to healthy model** ⑤
Masking condition	***Lower AVS than healthy control speakers***	*Similar AVS to healthy model* ⑭	**Lower AVS than healthy model** ⑮	**Lower AVS than healthy model** ⑯	*Similar AVS to healthy model* ⑰
Masking effect	***Larger decrease compared to healthy control speakers***	*Similar decrease compared to healthy model* ⑥	**Larger decrease compared to healthy model** ⑦	**Larger decrease compared to healthy model** ⑧	*Similar decrease compared to healthy model* ⑨
AVD					
No masking condition	***Higher AVD than healthy control speakers***	*Similar AVD to healthy model* ㉒	**Higher AVD than healthy model** ㉓	**Higher AVD than healthy model** ㉔	**Higher AVD than healthy model** ㉕
Masking condition	***Similar AVD to healthy control speakers***	**Similar AVD to healthy model** ㉞	*Higher AVD than healthy model* ㉟	*Higher AVD than healthy model* ㊱	*Higher AVD than healthy model* ㊲
Masking effect	***Larger decrease compared to healthy control speakers***	*Similar decrease compared to healthy model* ㉖	*Similar decrease compared to healthy model* ㉗	*Similar decrease compared to healthy model* ㉘	*Similar decrease compared to healthy model* ㉙

The left column contains a description of the reference findings, the three other columns describe the present modeling results. Bold text corresponds to the reference, italic does not and bold-italic is neutral.
FF: Feedforward System Deficit; FB: Feedback System Deficit; UMN-DYS: upper motor neuron dysarthria; AVS: Average vowel spacing; AVD: Average vowel dispersion.

for comparing the model's output in the different impairments conditions.

In terms of absolute numbers, it should be noted that in particular AVD in the healthy model is extremely low relative to what was found in the human study (Maas et al., 2015). The DIVA model is a simplified version of reality, and therefore does not include all sources of variation that a real human (with or without brain damage) experiences in an experimental session. Admittedly, this is a limitation of computational modelling studies in general. However, while the absolute degree of variation is lower in the simplified model, it is the data *pattern* that is most relevant. Although it is possible that a model that includes additional (or all possible) sources of variability and shows greater variability in output would behave differently in response to the impairments and conditions, we believe that, as a first step, investigating how the data patterns for different simulated impairments compare and how they resemble the patterns that are observed in human data can provide meaningful insights into the possible mechanisms at work in speech disorders such as AOS.

Evaluation of the impairment conditions

In Table I, we have summarised a comparison of the main findings for the different impairment models with the main findings for human speakers with AOS reported by Maas and colleagues (2015). A first thing that stands out is that the results did not show any differences in AVS and AVD between the output of the FB model and the healthy model. What this shows is that with feedforward processing intact, the feedforward motor command remains accurate enough for the FB model to (Terband et al., 2014; Terband & Maassen, 2010) produce the sound without

generating auditory errors and thus without invoking the auditory feedback control subsystem. From this non-distinctiveness of the FB results compared to the healthy model results, we conclude that FB impairment alone should be rejected as a plausible option for a deficit underlying AOS. However, a feedback processing deficit might still be involved in combination with a deficit in feedforward processing. The further discussion will therefore focus on the differential results of the FF, FF + FB and UMN-DYS models.

The key findings of the human study (Maas et al., 2015) were a reduction in AVS in the masking condition for the speakers with AOS. The group of human AOS speakers showed a smaller AVS compared to the healthy control speakers in the masking condition while they did not show differences in AVS compared to healthy speakers in the no masking condition. In the current simulations, only the FF impairment model exhibited this exact pattern of results. The UMN-DYS model showed a similar AVS in the no masking condition, but did not show a significant masking effect. Whereas the FF + FB model did show a masking effect and a smaller AVS in the masking condition compared to the healthy model, AVS was also smaller in the no masking condition. The FF model thus clearly provides the best correspondence to the key findings in the human speakers with AOS.

The behaviour of the FF + FB model shows that with impaired feedback processing, the model is not able to correct for the inaccurate feedforward motor commands caused by the feedforward processing deficit, and thus not able to retain vowel quality as the FF model and the human AOS speakers did. Consequently, the involvement of a feedback processing deficit should be rejected as a plausible option for a deficit underlying AOS.

As mentioned in the introduction, several of the human speakers with AOS who showed a disproportionate masking effect in the Maas et al. (2015) study also had mild dysarthria, which raised the question whether the observed pattern in human speakers might be (partially) be ascribed to dysarthria rather than AOS. In the present simulations, the UMN-DYS model results did not show a masking effect. As such, the simulations suggest that the masking effect exhibited by the human speakers is attributable to AOS and is cannot be explained by concomitant upper motor neuron dysarthria.

Regarding AVD, the findings for the human speakers in the Maas et al. (2015) study were a larger AVD in the no masking condition for the AOS group compared to the healthy control speakers, accompanied by a reduction in AVD in the masking condition. None of the impairment conditions in the current simulations showed this exact pattern of results. The FF, FF + FB, and UMN-DYS models all three demonstrated a larger AVD compared to the healthy model in the no masking condition with the effect being larger for the FF and FF + FB models than for the UMN-DYS model. However, none showed a significant reduction in AVD in the masking condition, contrary to the findings in the human study.

Average vowel spacing (AVS) was the primary dependent variable of interest in the human study (Maas et al., 2015). As for the present simulations, a comparison of the patterns of results indicates that only the results of the FF impairment model correspond to these AVS results of the human speakers with AOS. In the context of a feedforward deficit, AOS has been associated with a greater reliance on feedback control (Jacks, 2008; Maas et al., 2015; Rogers et al., 1996). This has two aspects. First, impaired feedforward control would affect the integrity of the issued motor commands, increasing the contribution of feedback-based commands to the overall motor command. Second, impaired feedforward control combined with intact feedback control would predict a shift in relative weighting toward the intact control subsystem. In other words, the speech production system's control strategy would be biased toward sensory feedback control (Terband et al., 2009; Terband & Maassen, 2010).

Computer simulations in which the feedforward/feedback control weighting ratio was varied during production in the DIVA model showed an increase in token-to-token variability as the reliance on feedback control increased (Terband et al., 2009). In the present simulations, the weighting parameters of feedforward and feedback control were kept constant at DIVA's standard values. The lack of a difference in AVD between the no masking and the masking conditions thus suggests that the results found in the present simulations for the impairment models compared with the healthy model reflects the implemented impairments directly rather than an overreliance on feedback control (increased relative weighting) due to the impairments. Further research into the interaction between reliance on feedback control and vowel spacing and dispersion is warranted, both in human speakers with AOS and in modelling experiments. Based on the present findings, we hypothesise that in AOS the masking effect on AVS result from the feedforward impairment directly, while the masking effects on AVD stem from an overreliance on feedback control consequent to the impairment.

Limitations and further research

This study represents the first systematic computer simulation directly tied to empirical data from human speakers with AOS. As such, we did not attempt to create individualised model parameters to simulate individual participant findings. In Maas et al. (2015), individual speaker analyses revealed clear effects for some, but not all, speakers with AOS. In the present study, we focussed on simulating the group patterns. Depending on further development of the DIVA model to allow detailed specification of model parameters to match individual speakers, future research may be able to test additional specific hypotheses about potential different underlying profiles of speech motor impairment, as was suggested by Maas et al. (2015). Furthermore, future research may investigate alternative impairments as well as alternative ways to simulate deficits. In the present study, we added uncorrelated, signal-independent, Gaussian zero-mean noise in specific parts of the system directly to the generated signals, as a means of simulating impaired processing in these specific parts and allowing us to tease out different conditions (i.e. impairment models in which feedforward and feedback systems are affected in isolation or in combination, before or after integration). Although such focussed impairments are arguably different from the often complex lesions in human individuals with AOS, we believe it to be appropriate for the purpose of the present study, i.e. to investigate the conceptualisation of AOS as a deficit in feedforward or feedback control. Further research into the conceptualisation of neurological impairments in computational models is warranted.

It should also be noted that the present model simulations do not capture all features often observed in speakers with AOS, nor was this study intended to provide such a comprehensive account. For example, features such as abnormal prosody or lengthened vowel durations were not addressed here. Vowel duration was kept constant here, because segment duration is incorporated in the DIVA's sound targets and not a parameter controlled by the neural network model. Because sound targets are predefined in the present computational implementation, deviant characteristics such as abnormal prosody, prevocalic groping and speech sound prolongations cannot currently

be investigated in DIVA. Nevertheless, although the DIVA model is not comprehensive and thus cannot model all aspects of AOS, the present study represents an important initial step in modelling aspects of AOS, to examine the degree to which the model aligns with human data for those aspects that are within the current scope and capability of the model. Our findings provide the first cross-validation of human observations against a detailed computational model of speech motor control, and provide support for the hypothesis that AOS can be viewed as a deficit in feedforward control. Future studies could investigate the effects of noise to either of the feedforward and feedback control models with other stimuli (e.g. featuring fricatives), speech measures (e.g. coarticulation; searching articulatory behaviour) or experimental paradigms (e.g. varying articulatory complexity), which could subsequently be investigated in human behavioural experiments. Another exciting direction would be to derive specific predictions about the effects of various therapeutic manipulations that can be tested in human learning and treatment studies.

With respect to clinical implications, future development of customised models to match individual speakers will enable more accurate and precise diagnosis. By deriving individualised predictions from such custom model specifications and validating such predictions against behavioural data from individual speakers, we will move closer toward an individualised, process-oriented approach to diagnosis, which in turn will facilitate development and refinement of more targeted treatment approaches (Terband et al., 2017, 2019). In fact, because the DIVA model is at its core also a model of speech motor learning, it lends itself well to deriving specific predictions about the effects of various therapeutic manipulations that can be tested in human learning and treatment studies. For example, some have argued that updating of impaired feedforward commands, as hypothesised for both AOS and for childhood apraxia of speech (CAS; Terband et al., 2009) may be enhanced through the use of visual biofeedback (see e.g. Preston, Leece, & Maas, 2016; Preston, Maas, Whittle, Leece, & McCabe, 2016 in CAS; see e.g. Katz, McNeil, & Garst, 2010; McNeil et al., 2010, for similar benefits of visual biofeedback in AOS). Models like DIVA may inform the specific mechanism of such effects, and help predict the likely benefit of such manipulations for individual speakers.

Overall, this study represents a first step in a longer-term research programme. Ultimately, as computational models improve in scope and detail, we may be able to develop individually-tailored simulations for specific speakers, in order to capture variability across speakers and to inform individualised clinical practice (precision-medicine). For example, by allowing estimates of response to different treatment options, or providing individual prognoses. This is admittedly a long-term goal, but a first step in this process is to demonstrate that the model in its current stage of development can account for some specific behavioural patterns. The present study provides promising support in this regard.

Conclusion

In the present study, we set out to validate the interpretation of AOS as a feedforward impairment as posed by Maas and colleagues (2015) by means of simulating their auditory feedback masking experiment in the DIVA model under different impairment conditions, comprising a feedforward, a feedback, a combination of feedforward and feedback, and an upper motor neuron dysarthria impairment. The primary dependent variable of interest in the human study (Maas et al., 2015) was average vowel spacing. With respect to this outcome measure, the key findings for the human speakers with AOS were a similar AVS compared to healthy control speakers in the no masking condition and a reduction in AVS results in a smaller AVS in the masking condition. In the present simulations, only the two impairment conditions that included the feedforward system deficit demonstrated a reduction in vowel spacing in the masking condition. Additionally, only in the isolated simulated feedforward deficit this effect was combined with an average vowel spacing similar to the healthy model in the no masking condition. Regarding the feedback control and upper motor neuron dysarthria impairments, the simulation results did not show differences in AVS compared to the healthy model results. The findings from the present simulation study thus corroborate the notion that AOS can be conceptualised as a deficit in feedforward control.

Notes

1. Note that this noise refers to a model-inherent method of simulating impaired processing within the system, and should not be confused with the auditory *masking noise* used in the Maas et al. (2015) study. To avoid confusion, we refer to the latter as *masking noise* throughout.
2. It should be noted that the DIVA model, like any model, is a simplified version of reality, and therefore does not include all sources of variation that a real human (with or without brain damage) experiences in an experimental session. As such, DIVA model productions reflect less variability in acoustic output than human data. This is admittedly a limitation of the DIVA model (and arguably of any computational model). However, while the absolute degree of variation is lower in the model, the critical issue here is whether differences between impairments and conditions emerge.

Acknowledgements

Disclosure statement

Supplemental Material

Funding

ORCID

Hayo Terband http://orcid.org/0000-0001-7265-3711

References

Adams, S.G., Weismer, G., & Kent, R.D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, *36*, 41–54. doi:10.1044/jshr.3601.41

Balasubramanian, V., & Max, L. (2004). Crossed apraxia of speech: A case report. *Brain and Cognition*, *55*, 240–246. doi:10.1016/j.bandc.2004.02.005

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1.7*, 1–23.

Bohland, J.W., Bullock, D., & Guenther, F.H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*, *22*, 1504–1529. doi:10.1162/jocn.2009.21306

Civier, O., Tasko, S.M., & Guenther, F.H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders*, *35*, 246–279. doi:10.1016/j.jfludis.2010.05.002

Deal, J.L., & Darley, F.L. (1972). The influence of linguistic and situational variables on phonemic accuracy in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *15*, 639–653. doi:10.1044/jshr.1503.639

Deger, K., & Ziegler, W. (2002). Speech motor programming in apraxia of speech. *Journal of Phonetics*, *30*, 321–335. doi:10.1006/jpho.2001.0163

Dronkers, N.F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*, 159–161. doi:10.1038/384159a0

Duffy, J.R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management.* St Louis, MO: Mosby.

Graff-Radford, J., Jones, D.T., Strand, E.A., Rabinstein, A.A., Duffy, J.R., & Josephs, K.A. (2014). The neuroanatomy of pure apraxia of speech in stroke. *Brain and Language*, *129*, 43–46. doi:10.1016/j.bandl.2014.01.004

Guenther, F.H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*, *72*, 43–53. doi:10.1007/BF00206237

Guenther, F.H., Ghosh, S.S., & Tourville, J.A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, *96*, 280–301. doi:10.1016/j.bandl.2005.06.001

Hickok, G., Rogalsky, C., Chen, R., Herskovits, E.H., Townsley, S., & Hillis, A.E. (2014). Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: Evidence from apraxia of speech following acute stroke. *Frontiers in Human Neuroscience*, *8*, 649. doi:10.3389/fnhum.2014.00649

Hillis, A.E., Work, M., Barker, P.B., Jacobs, M.A., Breese, E.L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*, *127*, 1479–1487. doi:10.1093/brain/awh172

Itabashi, R., Nishio, Y., Kataoka, Y., Yazawa, Y., Furui, E., Matsuda, M., & Mori, E. (2016). Damage to the left precentral gyrus is associated with apraxia of speech in acute stroke. *Stroke*, *47*, 31–36. doi:10.1161/STROKEAHA.115.010402

Jacks, A. (2008). Bite block vowel production in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *51*, 898–913. doi:10.1044/1092-4388(2008/066)

Katz, W.F., McNeil, M.R., & Garst, D.M. (2010). Treating apraxia of speech (AOS) with EMA-supplied visual augmented feedback. *Aphasiology*, *24*, 826–837. doi:10.1080/02687030903518176

Kent, R.D., & Rosenbek, J.C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research*, *26*, 231–249. doi:10.1044/jshr.2602.231

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. doi:10.18637/jss.v082.i13

Maas, E., Mailend, M.L., & Guenther, F. (2015). Feedforward and feedback control in apraxia of speech (AOS): Effects of noise masking on vowel production. *Journal of Speech, Language and Hearing Research*, *58*, 185–200. doi:10.1044/2014_JSLHR-S-13-0300

Maas, E., Robin, D.A., Austermann Hula, S.N., 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*, 277–298. doi:10.1044/1058-0360(2008/025)

Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling*. Boston, MA: Kluwer Academic Publishers.

McNeil, M.R., Katz, W., Fossett, T., Garst, D., Szuminsky, N., Carter, G., & Lim, K. (2010). Effects of online augmented kinematic and perceptual feedback on treatment of speech movements in apraxia of speech. *Folia Phoniatrica et Logopaedica*, *62*, 127–133. doi:10.1159/000287211

McNeil, M.R., Pratt, S.R., & Fosset, T.R.D. (2004). The differential diagnosis of apraxia of speech. In B. Maassen, R. Kent, H.F.M. Peters, P.H.H.M. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 389–414). Oxford, UK: Oxford University Press.

McNeil, M.R., Weismer, G., Adams, S., & Mulligan, M. (1990). Oral structure nonspeech motor control in normal, dysarthric, aphasic and apraxic speakers: Isometric force and static position control. *Journal of Speech, Language, and Hearing Research*, *33*, 255–268. doi:10.1044/jshr.3302.255

Nieto-Castanon, A. (2011). Simulink DIVA model (Version v.2011): Speech Lab, Department of Speech, Language & Hearing Sciences, Sargent College, Boston University.

Perkell, J.S., Denny, M., Lane, H., Guenther, F., Matthies, M.L., Tiede, M., … Burton, E. (2007). Effects of masking noise on vowel and sibilant contrasts in normal-hearing

speakers and postlingually deafened cochlear implant users. *The Journal of the Acoustical Society of America*, *121*, 505–518. doi:10.1121/1.2384848

Preston, J.L., Leece, M.C., & Maas, E. (2016). Intensive treatment with ultrasound visual feedback for speech sound errors in childhood apraxia. *Frontiers in Human Neuroscience*, *10*, 440. doi:10.3389/fnhum.2016.00440

Preston, J.L., Maas, E., Whittle, J., Leece, M.C., & McCabe, P. (2016). Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. *Clinical Linguistics & Phonetics*, *30*, 363–381. doi:10.3109/02699206.2015.1052563

Richardson, J.D., Fillmore, P., Rorden, C., LaPointe, L.L., & Fridriksson, J. (2012). Re-establishing Broca's initial findings. *Brain and Language*, *123*, 125–130. doi:10.1016/j.bandl.2012.08.007

Rogers, M.A., Eyraud, R., Strand, E.A., & Storkel, H. (1996). The effect of noise masking on vowel duration in three patients with apraxia of speech and a concomitant aphasia. *Clinical Aphasiology*, *24*, 83–96.

Seddoh, S.A., Robin, D.A., Sim, H.-S., Hageman, C., Moon, J.B., & Folkins, J.W. (1996). Speech timing in apraxia of speech versus conduction aphasia. *Journal of Speech, Language, and Hearing Research*, *39*, 590–603. doi:10.1044/jshr.3903.590

Square, P.A., Darley, F.L., & Sommers, R.K. (1981). Speech perception among patients demonstrating apraxia of speech, aphasia, and both disorders. In R.H. Brookshire (Ed.), *Clinical aphasiology: Conference proceedings* (pp. 83–88). Minneapolis: BRK Publishers.

Terband, H., & Maassen, B. (2010). Speech motor development in Childhood Apraxia of Speech (CAS): generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica et Logopaedica*, *62*, 134–142. doi:10.1159/000287212

Terband, H., Maassen, B., & Maas, E. (2017). Toward a model of pediatric speech sound disorders (SSD) for differential diagnosis and therapy planning. In P. Van Lieshout, B. Maassen, & H. Terband (Eds.), *Speech motor control in normal and disordered speech: Future developments in theory and methodology*. Rockville, MD: American Speech-Language-Hearing Association.

Terband, H., Maassen, B., & Maas, E. (2019). A framework of speech development and disorders for diagnosis and treatment planning. *Folia Phoniatrica et Logopaedica*. Advance online publication. doi:10.1159/000499426

Terband, H., Maassen, B., Guenther, F.H., & Brumberg, J. (2009). Computational neural modeling of Childhood Apraxia of Speech (CAS). *Journal of Speech, Language and Hearing Research*, *52*, 1595–1609. doi:10.1044/1092-4388(2009/07-0283)

Terband, H., Maassen, B., Guenther, F.H., & Brumberg, J. (2014). Auditory-motor interactions in pediatric motor speech disorders: Neurocomputational modeling of disordered development. *Journal of Communication Disorders*, *47*, 17–33. doi:10.1016/j.jcomdis.2014.01.001

Tourville, J.A., Reilly, K.J., & Guenther, F.H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, *39*, 1429–1443. doi:10.1016/j.neuroimage.2007.09.054

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1–25). New York, NY: Thieme Medical Publishers Inc.