

New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders

Anita Van Der Merwe

To cite this article: Anita Van Der Merwe (2020): New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders, *Aphasiology*, DOI: 10.1080/02687038.2020.1765306

Published online: 21 May 2020

To link to this article: <https://doi.org/10.1080/02687038.2020.1765306>

New perspectives on speech motor planning and programming in the context of the four-level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders

Anita van der Merwe

Professor Emeritus.

Department of Speech-Language Pathology and Audiology,

University of Pretoria,

South Africa.

E-mail address: anita.vandermerwe05@gmail.com

Abstract

Background: The complexity of speech motor control, and the incomplete conceptualisation of phases in the transformation of the speech code from linguistic symbols to a code amenable to a motor system, tend to obscure the understanding of acquired apraxia of speech (AOS). The four-level framework (FLF) of speech sensorimotor control (Van der Merwe, 1997; 2009) suggests the differentiation between speech motor planning, programming and execution and locate the locus of disruption in AOS in the motor planning phase. Currently, terminological confusion and uncertainty regarding phases in speech motor control still complicate the characterisation of AOS. This neuromotor disorder is inconsistently described in the literature as a “planning *or* programming”, “planning *and* programming”, or as a “planning and/or programming” disorder.

Purpose: To describe a new version of the FLF, the FL (four-level) model, which further explicates and differentiates between speech motor planning, programming, and execution levels or phases of processing; to integrate concepts from computational modelling into the FL model and propose distinct control architectures for both the planning and programming levels; and to identify the loci and nature of disruption in the motor planning phase which could explain the pathophysiology and core features of AOS.

Discussion and Conclusions: A four-level model is presented that differentiates two pre-execution phases and an execution phase. The first pre-execution phase is controlled by a motor planner and involves an inverse model, an efference copy, and a forward model for each sound or over-learned utterance. This phase also involves a forward predictive planner which enables the system to handle the planning of several sounds and to plan coarticulation of sounds. The motor planner is operated according to an auxiliary forward model architecture. AOS is depicted as a breakdown at several possible points in the motor planning phase. The second pre-execution phase is driven by a motor program generator and predictive controller that is governed by an integral forward model architecture. The final execution phase is portrayed as being driven by closed loop control. The conceptualization of the programmer challenges the traditional view of execution and not only that of planning as is generally accepted. The implications for the classification of motor speech disorders are discussed. Future research should address the exact nature of articulatory movements and other features of speech across the range of planning, pure programming, programming-execution and pure execution disorders.

Key words: apraxia of speech, speech motor control, motor planning, motor programming, dysarthria.

Running head: New perspectives on planning and programming

Darley, Aronson and Brown (1975) are recognized as early pioneers in the field of speech motor disorders who identified the acquired neurogenic speech disorder which they called apraxia of speech. They described the disorder as an “impairment of speech motor programming” (p. 250) At the time, the idea of a speech programmer was quite a novel and ground-breaking concept. Subsequent researchers used the terms “programming” and “planning” interchangeably in reference to this pre-execution process. The terms were used in this way with regard to apraxia of speech and also in speech motor control theory. A theoretical framework for the characterization of pathological speech sensorimotor control, first published in 1997, proposed speech motor planning and programming as distinct pre-execution phases in motor control and posited apraxia of speech as primarily a motor planning disorder (Van der Merwe, 1997; 2009). To capture the essence of this model, it has been referred to as the four-level framework (FLF). It distinguishes linguistic-symbolic planning, which necessarily precedes speech production, and three levels of motor control namely motor planning, motor programming, and execution. Differentiation of four phases or levels in speech production as proclaimed in the FLF (Van der Merwe, 1997; 2009) and the updated FL model which is presented in this paper, is contrary to the traditional model in speech pathology which proposes only three phases – a linguistic phase, a motor programming/planning phase and the execution phase. Traditionally impairment of each is envisaged to respectively cause aphasia, apraxia, and dysarthria (or a combination of these with multiple impairment). The traditional three-level model still prevails as dominant framework for the classification of neurogenic speech disorders.

The FLF (Van der Merwe, 1997) created an awareness of the possible distinction between motor planning and programming phases and between programming and execution, but this theory has not yet significantly advanced the conceptualization of the

underlying disorder in AOS. The state of the art is reflected in the inconsistent use of terminology in summative descriptions of the nature of AOS. Some researchers still use the terms “planning” and “programming” interchangeably while others display awareness of a second hypothetical pre-execution phase (programming), which could be compromised in AOS, but few are decisive in their description. Apraxia of speech is inconsistently described as either a disorder of “speech motor planning/programming” (Ballard et al., 2016; Bislick, McNeil, Spencer, Yorkston, & Kendall, 2017), “planning *and/or* programming” (ASHA, 2007), “planning *or* programming” (Duffy, 2013), motor planning *and* programming (McNeil, Doyle, & Wambaugh, 2000) or as a motor planning disorder due to impaired motor programs (Mailand, Maas, Beeson, Story, & Forster, 2019; Brendel, et al., 2011).

Today, most of the significant segmental and suprasegmental features of acquired apraxia of speech (AOS) have been well described and are accepted widely (Ballard et al., 2016; Bislick & Hula, 2019; Duffy, 2007, 2013; McNeil, Robin, Schmidt, 1997, 2009; McNeil, Ballard, Duffy, & Wambaugh, 2016), but the exposition of the underlying pathophysiology of apraxia remains incomplete and contentious. The matter in dispute is whether AOS should be characterized as a planning, a programming, or a motor planning and programming disorder. Underlying this deliberation is the question whether these phases or levels of processing are indeed separable and have distinct roles in speech motor control. This Viewpoint paper presents updated theory substantiating the differentiation between these two pre-execution phases and elucidating the potential implications for our understanding of the pathophysiology underlying AOS and other motor speech disorders.

The purpose of this paper is: (a) to propose a new refined version of the FLF, the four-level (FL) speech production model, which further clarifies and differentiates between speech motor planning, programming, and execution levels of processing; (b) to show how

the FL model addresses control of segmental and also suprasegmental features of speech; (c) to integrate computational concepts into the explication of motor planning in the FL model and indicate how such concepts could explain the pathophysiology and core features of acquired apraxia of speech; (d) to position the phases identified in the FL model in the theoretical context of modern-day computational models of speech motor control; and (e) to posit distinct motor control architectures to account for control of motor planning and programming.

CONTEMPORARY SPEECH MOTOR CONTROL MODELS

Theorists have long implemented models of normal speech motor control to explain the concept of apraxia of speech. Neurolinguistic models and psycholinguistic models (for example, Darley et al., 1975; Guenther & Perkell, 2004; Hickok 2014; Hickok, Houde, & Rong, 2011; Van der Merwe, 1997; 2009; Ziegler, 2009; Ziegler & Aichert, 2015) were developed for this purpose. Currently computational modelling and neural engineering are greatly influencing the design of speech production models. One such model, the Directions Into Velocities of Articulators (DIVA) model (Guenther, 2016) is implemented to explain AOS. Parrell, Lammert, Ciccarelli and Quatieri (2019) describe computational models as “formal, mathematical models”. They regard the various components of speech motor control as layered modules or levels and conceptualise these as: a higher-level *linguistic processor* which also controls prosody, a *planner* (responsible for motor program generation and sequencing movements), a *controller* (that takes a speech plan and issue motor commands) and the *plant* (the vocal tract and articulators). The focus of Parrell, and colleagues (2019) is on the lower-level control layer which they regard as the “bridge” between high-level linguistic planning and the biomechanical movements of the

articulators. They proclaim that models which provide a formal description of the control layer include, for example, the DIVA model (Guenther 2016), Task Dynamics (TD) (Saltzman & Munhall, 1989), State Feedback Control (SFC) (Houde & Nagarajan, 2011) and Feedback Aware Control of Tasks in Speech (FACTS) model (Parrell, Ramanarayanan, Nagarajan, & Houde, 2019; Parrell, Lammert, et al., 2019).

As background to an account of computational models it is necessary to provide a theoretical orientation and introduce fundamental terminology. Key concepts in the field of computational motor control include: feedback control, feedforward control (open-loop control), forward predictive control (also named model-predictive control, the predictor mechanism, or internal prediction), inverse internal models (appropriate motor commands based on intended consequences), forward internal models (models that predicts the future state of the system), corollary discharge (output of the forward model), inverse-forward scheme, motor commands and efference copy (copy of the current motor command fed back to the system). The role assigned to feedback signals (proprioceptive, tactile, and auditory in the case of speech production), feedforward control, forward predictive control, and the efference copy determines the control architecture of a theoretical or cognitive model (Andersen & Cui, 2009; Cui, 2016; Davidson & Wolpert, 2005; Franklin & Wolpert, 2011; Parrell, Lammert, et al., 2019; Parrell & Houde, 2019; Pickering & Clark, 2014).

Feedback control of speech production was first proposed by Fairbanks (1954). According to this theory accurate articulation depends on auditory, proprioceptive and tactile response produced afferent feedback. However, feedback signals occur late and outdated state information is relayed to the controller. The alternative theory to feedback control is open-loop or feedforward control. Previously acquired motor commands are feedforward to the motor apparatus or articulators in the case of speech production

(Guenther, 2016; Parell & Houde, 2019; Parrell, Lammert, et al., 2019). However, pure feedforward control cannot explain how the system deals with unexpected perturbations. Apart from modelling the control architecture, models should also propose solutions for several problems inherent in sensorimotor control. These include, for example, nonlinearity, delays, uncertainty, noise (Franklin & Wolpert, 2011), interference and unexpected perturbations (Parrell & Houde, 2019). The liabilities of feedback and feedforward control are overcome by the concept of forward models (Pickering & Clark, 2014, Wolpert, Diedrichsen, Flanagan, 2011; Wolpert, Ghahramani, & Flanagan, 2001; Wolpert & Kawato, 1998) and forward predictive motor control (Cui, 2016; Davidson & Wolpert, 2005; Parrell, Lammert, et al., 2019; Wolpert & Kawato, 1998). A forward internal model is a model within the brain that can predict the likely sensory consequences of an action (Andersen & Cui, 2009; Wolpert et al., 2001; Itoh, 2008). Model predictive control does not make use of outputs from the plant to maintain control. An error is detected internally and is not derived from feedback signals (Parrell, Lammert, et al., 2019).

Divergent accounts of the role of forward models in motor control have been proposed. For example, in an auxiliary forward model (AFM) architecture, the output of an inverse model (the motor commands) is copied to the forward model in an efference copy and is then used by the forward model to estimate feedback and finesse the outcome. In an integral forward model (IFM) architecture, the predictions act as action commands and there is no need for an efference copy (Pickering & Clark, 2014). These examples illustrate that models differ in their conceptualization of the role of feedback and feedforward processing.

Building on a traditional feedback control system, the State Feedback Control (SFC) model was developed to overcome the noisy and delayed nature of reafferent sensory

feedback and the limitation this place on pure feedback control (Parrell, Lammert, et al., 2019; Houde & Nagarajan, 2011). The SFC integrates sensory feedback with internal model predictions. Predicted feedback is compared with actual feedback to determine a sensory error. The final state estimate relayed back to the model controller is based on a combination of state prediction (internal feedback) and sensory processes. The SFC is characterised as “an integrated model predictive feedback control architecture” (Parrell & Houde, 2019; Parrell, Lammert, et al., 2019, p. 1468). Hickok et al. (2011) takes this further to develop an Integrated SFC model which also involve psycholinguistic and neurolinguistic systems like the “motor phonological system”, “auditory-motor translation” and the “auditory phonological system” (p. 413). Hickok (2014) also proposes a hierarchical state feedback control model, the HSFC, which is best classified as a psycholinguistic model.

The original SFC model was recently expanded to formulate the Feedback-aware Control of Tasks in Speech or FACTS model (Parrell, Ramanarayanan, et al., 2019). This model postulates that the speech system is not organized to control individual articulator movements, but higher order tasks. The tasks they propose to be constrictions in the vocal tract. Examples are lip protrusion, location and degree of constriction between the tongue tip and palate and degree of velum opening. The tasks are controlled according to standard SFC principles. The FACTS model operates hierarchically. A “high-level task state feedback controller” operates on a “low-level articulatory state feedback controller” which drives the speech production mechanism itself (Parrell & Houde, 2019, p. 2973).

The DIVA model (Guenther, 2016) proposes an extensive computational account of speech sensorimotor control and also endeavours to explain the underlying pathophysiology of AOS. It utilizes a hybrid control system which means that it combines a feedforward controller with separate somatosensory feedback and auditory feedback controllers that

process feedback. Together, the output of the three controllers generate the motor commands to the speech system. The system can operate in a purely feedforward style, but feedback is available if unexpected circumstances arise or during speech acquisition. The highest processing level in the DIVA model is the activation of the appropriate nodes of a speech sound map. The sound map is a neural representation of a sound or a well-known syllable or word. The hypothesized location of the speech sound map is the left ventral premotor cortex (vPMC). This area is proposed to be the source of direct projections to articulator maps in primary motor cortex and also via the cortico-cerebellar loop (Guenther, 2016; Guenther & Vladusich, 2012). Damage to these projections disable feedforward commands, resulting in impaired articulation and groping movements characteristic of AOS. According to Guenther (2016) the speaker with AOS cannot access the appropriate feedforward commands or motor programs. Damage to the speech sound map will also impact feedback control. The proposed reason is that the projections from vPMC to higher-order auditory and somatosensory areas that carry sensory targets for speech sounds are damaged. Increased duration of pauses between syllables and abnormal prosody in AOS could best be explained by the GODIVA model according to Guenther (2016) and relates to slow retrieval from a phonological content buffer that temporarily stores the phonological content for upcoming utterances. Terband, Rodd and Maas (2019) tested hypotheses about the underlying deficit of AOS through computational modelling with the DIVA model. The effect of a feedforward, feedback, and combined feedback and feedforward impairment during a noise-masking condition was investigated. Results indicated that a feedforward deficit best resembled the group findings of human speakers with AOS.

PHASES IN THE PREPARATION OF SPEECH MOVEMENTS

The fundamental premise of the FL model is that speech motor planning and programming are differentiable pre-execution phases. Motor planning takes place in cortical motor areas in the dominant hemisphere while motor programming is mediated via bilateral subcortical areas and cortical-subcortical circuits in the brain. Justification for this distinction can be found in the motor control literature.

Theoretical accounts of the motor control of human kinetics focus mainly on the role of specific areas or structures in the brain, such as the basal ganglia (for example, Groenewegen, 2003; Leisman, Baraun-Benjamin, & Melillo, 2014), the cerebellum (for example: Callan, Kawato, Parsons, & Turner, 2007; Itoh, 2008; Habas, 2010; Xu, Liu, Ashe, Bushara, 2006), the supplementary motor area, the premotor area, and the primary motor area (for example, Fetz, 1993; Murata, Wen, Asama, 2016; Reis et al., 2008). Phases or levels of control in the preparation of movement during the pre-execution stage are not always identified in this body of literature. However, many researchers who adopt a holistic approach do present an overview of phases in motor control and all posit the existence of two pre-execution phases and then a final execution phase. To lend credence to the claims and proposals made in this article, some verbatim quotes are included in the following sections. The quotes are from the work of authors who are prominent researchers in their respective fields.

To explain function at the *highest level*, reference is made to central programs (Evarts, 1982), preprogramming (Allen & Tsukahara, 1974), the *idea* of a movement that is expressed in patterns of excitation in the association cortex (Allen & Tsukahara, 1974; Eccles, 1977), intention (Cui, 2014; Frith & Haggard, 2018), or movement planning at a cognitive level (Andersen & Cui, 2009; Cui, 2016). Motor planning is proposed to occur at the highest level of the motor control hierarchy and is mediated by the association cortex

and cortical motor areas. These include the prefrontal cortex, Area 6, the supplementary motor area (SMA), areas 5 and 7 (posterior parietal areas) and also Broca's and Wernicke's areas (Andersen & Cui, 2009; Brooks, 1986; Burbaud, Doegle, Cross, Bioulac, 1991; Di Pellegrino & Wise, 1991; Frith & Haggard, 2018; Itoh, 2008; Jeannerod, 1994, 1995; Magill, 2007; New et al., 2015; Widmaier, Raff, & Strang, 2006).

Mental representations in the cerebral cortex (Itoh, 2008) or motor representations which are conceived as "internal models of the goal of the action" (Jeannerod, 1995, p. 1427) denote motor plans and planning at the highest level of motor control. Jeannerod (1995) found evidence of this level of control when studying motor imagery. During covert speech production, which does not require control of physical movements, activity was observed in higher cortical areas. Patterns of activation that resemble that of action (overt) execution were observed. The aim of his study was to examine the implications of this phenomenon for a model of action generation. His consequent model differentiates between intention, planning at a cortical level, programming and then execution that is mediated by lower levels in the motor system. Using electrocorticography, Brumberg et al. (2016) studied cortical activity during continuous overt and covert speech production and they too found a common neural substrate in both conditions as well as cortical involvement that starts in the frontal-motor areas.

Parrell, Lammert, et al. (2019) also propose the existence of a *planner* which provides "reference vectors" to achieve "some higher-level sensorimotor or cognitive goal" or task such as producing a specific word. They differentiate between task and mobility space with the task space conceived as a "higher" level space. Reference vectors are proclaimed to be insufficient for use as motor commands and require transformation into motor commands in mobility space by the controller (p. 1459).

The *middle level* of control “converts plans received from the highest level to a number of smaller programs which determine the pattern of neural activation” (Magill, 2007, p. 75; Widmaier, et al., 2006; Cui, 2016). Motor programs are prepared at the middle level of the motor hierarchy, presumably by the basal ganglia (Evarts & Wise, 1984; Goldberg, Farries & Fee, 2013; Groenewegen, 2003; Johnson, Vernon, Almeida, Grantier, & Jog, 2003) and the lateral cerebellum (Ackerman, 2008; Allen & Tsukahara, 1974; Dreher & Grafman, 2002; Habas, 2010; Ishikawa, Tomatsu, Izawa, Kakei, 2016). In the words of Jeannerod (1995, p. 1427) the “global internal model of the action activates an appropriate plan, which in turn activates motor programs”. Brooks (1986) explains the planning-programming relationship by using the terms *strategy* and *tactics*. Strategies determine the “general nature of plans”, while tactics provide “particular specifications in space and time” (p. 26). In the FL model, the middle level motor phase is referred to as the *programming* level as this is where muscle-specific programs are created and specified.

Finally, at the third and *lowest (local) level*, the programs and subprograms transmitted from the middle control levels are fed forward to the brainstem and spinal cord to be executed (Brooks, 1986; Groenewegen, 2003; Jeannerod, 1995; Magill, 2007; Widmaier, et al., 2006). The motor cortex, lower motor neurons, peripheral nerves, and motor units in the muscles are the final structures that handle efferent motor signals. Efferent fibres from the cerebellum and basal ganglia reach the motor cortex via the thalamus but some fibres pass directly to the motor centres of the brainstem (Allen & Tsukahara, 1974; Ishikawa, et al., 2016; Nolte 1999). Appropriate muscle tone, strength, and coordination (the latter term intends to describe the absence of involuntary or uncoordinated movements) need to be controlled by the relevant areas in the central nervous system during execution.

The controller of speech production as conceived by Parrell and colleagues (Houde & Nagarajan, 2011; Parrell & Houde, 2019; Parrell, Lammert, et al. 2019; Parrell, Ramanarayanan, et al., 2019) and presented in the SFC and FACTS models, most likely correspond to the *programmer* as proposed in the FL model. In their conception of the controller, motor programs (note the use of the word *programs*) are represented indirectly as desired target end states that the articulators should achieve and not desired articulatory positions (trajectories) (Parrell & Houde, 2019, p. 2971). They also proclaim that the purpose of the controller is issuing motor commands that will lead to movements of the plant (Parrell, Lammert, et al., 2019, p. 1459). In the SFC and FACTS models, the operation of the controller is not distinguished from an execution phase. Their conception of the functions of the controller in these models appears too complex to be assigned to activity at the execution level. Therefore, the inference may be made that they are actually describing the activity of a programmer as portrayed in the FL model. However, the act of execution does give rise to reafferent auditory and somatosensory feedback which potentially could be implemented for online motor control by comparing actual feedback to predicted feedback (Parrell & Houde, 2019, p. 2978), but this process likewise appears to be too complex for an execution system.

The conception of the motor control hierarchy as found in the limb control literature appears to provide adequate justification for the differentiation made in the FLF and updated FL model between motor planning, programming, and execution (Van der Merwe, 1997, 2009). Motor planning takes place at the highest level of control while motor programming constitutes a transitional phase between planning and execution. This viewpoint is the point of departure for further exploration of the intricate process of speech motor control and for unravelling the nature of AOS and other motor speech disorders.

SPEECH SENSORIMOTOR CONTROL WITH PARTICULAR REFERENCE TO MOTOR PLANNING AND PROGRAMMING

The ensuing section presents a discussion of speech motor planning and programming. The content is intended to substantiate hypotheses regarding the nature of these processes as presented in the FL model.

Speech motor planning

Speech motor planning is an interface stage between phonological planning, which entails the selection and sequencing of phonemes that occur in an utterance, and the preparation of impulses to be conveyed to a motor system. During motor planning abstract phonological symbols are assigned properties amenable to a motor code. Phonemes are changed into sounds which have discrete place and manner of articulation features. During speech acquisition, a core motor plan (CMP) for the production of each sound, needs to be developed (Van der Merwe, 1997, 2009). The need for CMPs in speech production becomes obvious when a speaker attempts to produce a word containing a sound which has not previously been acquired. An example would be the voiceless, lingual, lateral, apico-alveolar click /ǀ/ in the Zulu word *ǀxǀxo* (Niesler, Louw, & Roux, 2005; Van der Merwe & Steyn, In Press). The unilateral release of air, which can be on either side, would probably need careful motor planning by the non-native speaker, even during imitated production. The primary role of single sounds is also acknowledged in the DIVA and other models (Guenther, 2016; Guenther & Perkell, 2004; Hickok, 2014). During motor learning of speech production skills, acquisition of the CMPs of the sounds of the language plays a key role and these elements act as the “building blocks” of speech production (Van der Merwe & Steyn, 2018).

In contemporary neurophysiology, motor skill learning and performance are approached from a computational perspective (Wolpert, Ghahramani, & Flanagan, 2001). From this viewpoint the brain is a processor that converts (transforms) sensory inputs to motor outputs (Andersen & Cui, 2009; Cooper, 2010; Crochet, Seung-Hee, & Petersen, 2019; Wolpert et al., 2001). Sensory input constitutes the aggregate of sensory feedback provided by our sense organs (reafferent feedback) and also internal feedback derived internally from an efference copy of the descending motor command (Wolpert et al., 2001, p. 488). Each transformation is bidirectional and contains an inverse and forward internal model. An inverse model converts intentions into motor commands and the forward model predicts the future state of the system and estimate sensory feedback (Franklin & Wolpert, 2011; Kawato, 1999; Pickering & Clark, 2014; Wolpert & Kawato, 1998). Transformations are classified as kinematic or dynamic. Kinematic transformations operate between coordinate systems such as the joint angles of the arm and position of the hand. Dynamic transformations are, as the name implies, dynamic and they correlate motor commands to the movement of the system. An example would be learning the force that should be applied to achieve a specific outcome (Wolpert, et al., 2001). The acquired representations of these transformations within the central nervous system are referred to as dynamic internal models (Cui, 2016; Itoh, 2008; Kawato, 1999; Kawato & Gomi, 1992; Wolpert et al., 2001). Internal models enable the central nervous system to determine the motor commands required to perform a task and to predict the consequences of motor commands (Kawato & Wolpert, 2007). Conceptually, internal models are regarded as *motor primitives* which are used to structure intricate motor behaviours with an extensive range. By modulating the contribution of a set of internal models, a great repertoire of actions can be generated (Wolpert et al., 2001, p. 492).

Skilled motor behaviour requires both inverse and forward models. An *inverse* model transforms a desired sensory outcome into the motor commands that could realize it (Wolpert et al., 2001). The definition of an inverse model by Kawato and Gomi (1992) refers to a neural representation of the transformation, from the desired movements to the motor commands that are required to achieve these movement goals. An inverse model provides a controller that does not need feedback and allows feedforward control of movements in which response feedback is available too late to guide the movement (Itoh, 2008). Neurophysiological evidence indicates that the cerebellum plays a role in the long-term storage of internal models for limb and trunk control, but the role of a higher order controller is also acknowledged (Ackermann, 2008; Andersen & Cui, 2009; Itoh, 2008; Ishikawa, et al., 2016; Wolpert et al., 2001).

The concept of *inverse* internal models of movement corresponds to and augments the notion of a CMP for the production of a sound as depicted in the FL model. These inverse internal models or CMPs are proposed to be encoded during the acquisition of speech and contain spatial (place and manner of articulation) and temporal (relating to inter-articulatory synchronization) specifications for movement of the speech structures involved in the production of a sound. The size of the motor plan or inverse model in speech production is not always predictable. The FL model proposes a CMP for each sound, but it is possible that there are stored syllable-sized and even word- or phrase-sized internal models for highly automated and over-learned utterances (Van der Merwe, 2007). In speech production, inverse models also need to be flexible in the sense that certain non-critical specifications can be adapted to the phonetic context in which a sound occurs and which allows for the coarticulation of sounds. This conception is viable as frontal cortical areas have been shown to encode information including subsequent movement parts, directions,

and temporal organization (Andersen & Cui, 2009). It is proposed that inverse models for speech production are centred in the dominant hemisphere of the brain and in the higher cortical and cognitive-motor planning areas of the brain.

Forward models which predict the future state of a system and the sensory consequences of a set of motor commands (Andersen & Cui, 2009; Franklin & Wolpert, 2011; Itoh, 2008; Pickering & Clark, 2014; Wolpert et al., 2001) could be implemented for monitoring movements. The concept of forward models advanced the understanding of how efference copies of motor commands are routed back to sensory areas in the brain to internally monitor movement and execute predictive sensorimotor control (Cui, 2016). Recent conceptions of motor control have extended the notion of forward models to also include active inference or action-oriented predictive processing. The generation of a motor response is accomplished by predicting the flow of sensation that would occur during performance of a target action. In this type of account there is no need for an efference copy and the predictions act as action commands (Friston, Daunizeau, Kilner, Kiebel, 2010; Pickering & Clark, 2014). The exact location of the predictor is not known, but the assumption is that forward prediction of somatosensory consequences most probably takes place in the posterior parietal cortex (Anderson & Cui, 2009; Cui, 2014; 2016; Frith & Haggard, 2018). The sensory consequences of speech production include movement related somatosensory and also auditory information. Because motor acts aim to reach sensory targets, the neural network underlying the planning of speech movements requires sensorimotor integration and translation (Hickok, 2012). According to Hickok (2012) the Spt (Sylvian parietal-temporal) area in the left planum temporale region, acts as a sensorimotor integrator and generates forward predictions in the auditory cortex during speech production (Hickok, 2012; Hickok et al., 2014).

Speech motor programming

At the middle level of the motor control hierarchy, mediated by the sensorimotor cortex, lateral cerebellum, and the basal ganglia (Groenewegen, 2003; Magill, 2007), motor plans received from the highest level are augmented with muscle-specific motor programs that “shape the final descending signal” (Groenewegen, 2003, p. 108). *Tactics* or specifications regarding muscle tone, movement direction, velocity, force, range and mechanical stiffness of joints are added to sets of motor plans (Brooks, 1986; Miall, Weir, & Stein, 1987; Rose 1997; Schultz & Romo, 1992). In other words, motor programs are superimposed on a set of CMPs, and the parameters that are added prepare the speech code for externalization and finally execution.

The new version of the FLF, the FL model, proposes that suprasegmental features are controlled at the level of motor programming. Spatiotemporal and force parameters are specified for movements of the articulatory, phonatory, and respiratory muscles during programming. Characteristics of such motor programs could be influenced by circumstantial factors (for example, the need to talk louder or faster or with increased force) and by linguistic factors. The latter would include suprasegmental features of an utterance such as intonation, stress, duration and juncture. The principle determinants of these perceptual qualities are fundamental frequency, amplitude and duration (Borden, Harris, & Raphael, 2003). Sentence intonation would require a change in the pitch of the fundamental frequency of voice at a certain point in a sentence. To achieve this, an adaptation of tension (in the cricothyroid muscles bilaterally) and length of the vocal folds, which would result in a change in fundamental frequency of vibration (Borden, et al., 2003), are necessary. The same applies to changes in pitch across a word for lexical (syllabic) tone production in tone

languages. Syllabic stress is dependent on muscle-specific increased tension (force). Stressed syllables are produced with greater loudness, higher pitch and longer duration. The added expiratory force that creates the extra air pressure required for the stressed syllables is attained by contraction of the internal intercostal muscles (Borden et al., 2003). To realize these changes, muscle-specific programs need to be specified for a set of motor plans during the motor programming phase.

In addition to the tactical specification of the descending signal (Groenewegen, 2003), repeated initiation and feed-forward of co-occurring and successive motor programs for speech production are controlled at the programming level. Prompt initiation and a smooth flow of motor programs determines rate, rhythm and fluency of speech. The selection and activation of motor programs that are appropriate for a particular context might be one of the primary functions of the basal ganglia (Groenewegen, 2003; Mink, 1996; Redgrave, Prescott, & Gurney, 1999). Well-practiced components of a particular task become automatic habitual responses which are fed forward. A basal ganglia disorder leads to a breakdown in the mechanisms responsible for automatic habitual performance (Hernandez, Obeso, Costa, Redgrave, & Obeso, 2019). This underlying problem could by inference lead to delayed initiation and temporal inaccuracies in speech.

Current thinking has contributed to a reappraisal of function concerning the basal ganglia and cerebellar loops (Leisman, et al., 2014). Traditionally, the primary signs of hypokinetic disorders (for example, Parkinson's disease) and hyperkinetic disorders (for example, chorea) were assumed to provide a model of the role of the basal ganglia in motor control. However, contemporary researchers oppose this single-faceted approach and emphasise also the cognitive, executive and emotional-motivational functions of the basal ganglia. Engagement at this level is subserved by the involvement of the basal ganglia in a

number of cortical-subcortical circuits, including primary motor, premotor, and associative (cognitive) and limbic prefrontal cortical areas (see for example Doya, 2000; Groenewegen, 2003; Leisman, et al., 2014; Marsden, 1984). Recent studies have also substantially extended the traditional view on the cerebellum “from a mere coordinator of automatic and somatic motor functions to a topographically organized and highly specialized neural mechanism” (Ackermann, 2008; Ishikawa, et al., 2016; Mariën, 2012, p. 470; Leisman et al., 2014, Xu, et al., 2006). The more precise determination of the relative contribution of the basal ganglia and cerebellum to the “learning and programming of skilled movements” is an important issue for future research (Groenewegen, 2003, p.119). The development in this particular field points to involvement of these structures/areas on a programming and execution level of motor control and challenges the assumption in speech pathology that all types of dysarthria are mere execution disorders.

Impaired *speech motor programming* as conceptualized in the FL model could lead to defective tactical specification of motor programs and to disorders of repeated initiation and feed-forward. The first-mentioned problem would cause inaccurate articulation and compromised laryngeal and respiratory function. This could be due to disrupted assignment of muscle-specific spatiotemporal and force parameters to the motor commands. The range, direction, velocity, muscle tension adaptation, and force of movements would be incorrectly programmed. The result is likely to be distortion of speech and disrupted prosodic features such as monotone speech and incorrect lexical stress assignment, all of which would probably be consistently present during production of an utterance. The second-mentioned problem at the programming level could cause slow initiation and disrupted feedforward of motor programs resulting in rate, rhythm and fluency disorders in speech.

In neuropathology the symptoms of cerebellar and basal ganglia disorders are lending support to the view that these brain areas are involved in both programming and the execution of movement. The notion of dual-symptomatology, in other words the presence of both execution and programming deficiency, in basal ganglia and cerebellar disorders, is not a novel concept (for example, Sheridan, Flowers & Hurrell, 1987). Dysarthria due to cerebellar and basal ganglia disorders are traditionally considered as *pure* execution disorders, but lesions in these areas which cause hypokinetic, hyperkinetic and ataxic dysarthria may lead to dual symptomatology (Diehl et al., 2019; Skodda, 2011; Spencer & Rogers, 2005; Spencer & Slocomb, 2007; Stipinovich & Van der Merwe, 2007). Evidence of symptoms which cannot be explained reliably from the perspective of an execution level disorder was found in ataxic, hypokinetic and hyperkinetic dysarthria. Examples of such symptoms or signs are short rushes of speech, variable rate, prolonged intervals, abnormally fast rate, repeated sounds and inappropriate silences which occur in hypokinetic and hyperkinetic dysarthria (Diehl et al., 2019; Duffy, 2007). These signs could point to disrupted feedforward of motor programs which occurs together with execution level problems such as muscle tone disorders and the occurrence of involuntary movements. In the theoretical context of the FL model, ataxic, hypokinetic and hyperkinetic dysarthria are best classified as programming-execution speech disorders (Stipinovich & Van der Merwe, 2007). Research evidence appears to support this viewpoint.

Spencer and Rogers (2005) investigated the hypotheses that people with ataxic dysarthria due to cerebellar dysfunction have reduced ability to program movement sequences before the onset of movement and that people with hypokinetic dysarthria due to Parkinson's disease are unable to maintain a programmed reaction or to rapidly switch between responses. The results provided evidence of deficits in these speakers that are

separable from execution impairments. Reilly and Spencer (2013) did a study on the effects of sequence complexity (defined in terms of phonemic similarity and phonotactic probability by Reilly and Spencer) on speech production in speakers with either hypokinetic or ataxic dysarthria. The analyses revealed significantly higher error rates and longer within-syllable vowel and pause durations in more complex utterances in both groups. Task effects are not to be expected in the instance of a pure execution disorder where flaccidity, spasticity, or involuntary movements are the cause of the speech disorder. Other studies also report task effects on the speech of individuals with Parkinson's disease (PD) and hypokinetic dysarthria (Van Lancker Sidtis, Cameron, & Sidtis, 2012; Van Lancker Sidtis, Pachana, JeVrey, Cummings, & Sidtis, 2006; Lowit, Marchetti, Corson, Kuschmann, 2018). A reaction time study also found disruption of speech initiation and programming in individuals with PD and hypokinetic dysarthria (McAllen, Spencer, France, & Shulein, 2010). Many individuals with movement disorders in due course demonstrate a coexisting cognitive deficit and the role of cognitive load in task effects needs to be considered during the interpretation of results.

Following this line of argumentation, the only *pure* execution dysarthria would be flaccid dysarthria as the lesions are in the lower motor neurons which are pathways for the descending signals. Spastic dysarthria also presents predominantly with an execution disorder. Spasticity in articulatory, phonatory and respiratory muscles impact speech production comprehensively and possibly mask any programming disruption which could theoretically be present. The primary motor cortex (including the upper motor neurons) which is implicated in spastic dysarthria is hypothetically also involved at higher levels of motor control than execution (Guenther, 2016). However, the exact role of the primary motor cortex is still being debated (Tanaka, 2016). Two other speech disorders, acquired foreign accent syndrome (Duffy, 2013; Schmullian, Van der Merwe, & Groenewald, 1997)

and stuttering, not yet attributed to a disruption of speech motor programming, could also potentially be explained more comprehensively if viewed within a four-level theoretical framework (Van der Merwe, 2009). Acquired foreign accent could be the result of compromised specification of muscle-specific spatiotemporal and force parameters of motor commands. For example, the range of movements could be affected, leading to a consistently present change in vowel formants, lending a foreign accent to speech. Stuttering, by nature, reflects a breakdown in repeated initiation and feedforward of motor programs.

THE FOUR-LEVEL MODEL OF SPEECH SENSORIMOTOR CONTROL

The new version of the four-level (FL) model is portrayed graphically in Figure 1. This depiction contains a focus on speech acquisition together with a delineation of the functions of the planning, programming, and execution levels. The underlying control architecture of the FL model is subsequently described and presented in Figure 2.

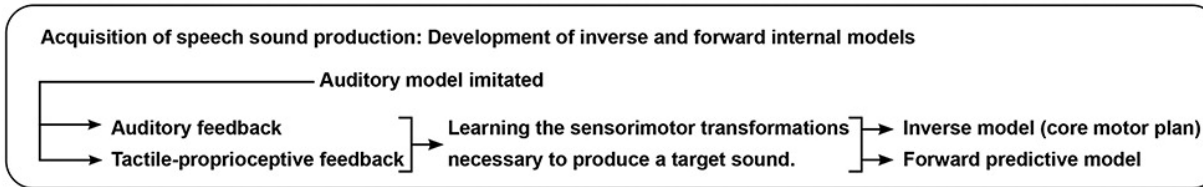
A four-level model for the characterization of pathological speech sensorimotor control

In response to the intention to communicate verbally, linguistic-symbolic planning is initiated. This process is portrayed as *symbolic* since words and phonemes are symbols intended to convey meaning. Syntactic, morphological, and phonological planning takes place at the highest level of the FL model (see Figure 1) and are mediated by the temporal-parietal and Broca's and adjacent areas. Phonemes that were selected and sequenced are relayed to the motor planning areas of the brain.

LINGUISTIC-SYMBOLIC PLANNING: Syntactic, morphological and phonological planning. The latter entails selection and sequencing of phonemes.

↓ ↓ ↓ Phonemes relayed to motor planning areas

MOTOR PLANNING: Place and manner of articulation and inter- and co-articulatory control of segmental features of speech



Inverse model → Motor commands → Internal feedback of efference copy → Forward model

Monitoring by the forward model: Comparison of efference copy to inverse model and prediction of outcome

Adjusted (if necessary) and coarticulated motor commands relayed to programming areas

↓ ↓ ↓

Linguistic input → **MOTOR PROGRAMMING:** Circumstance (context) sensitive and muscle-specific specification of spatiotemporal, force and metrical structure parameters for segmental and suprasegmental control
 Fronto-limbic system input →

Articulatory muscles Laryngeal muscles Respiratory muscles

↓ ↓ ↓

↓ ↓ ↓

↓ ↓ ↓

Repeated initiation, switching and feedforward of motor programs to the muscles

EXECUTION: Muscle tone, strength, and coordination control

↓ ↓ ↓

Manifested speech movements → Response produced feedback

Linguistic-symbolic planning
 Temporal-parietal area and Broca's and adjacent areas

Motor planning
 Prefrontal cortex, Area 6, Supplementary motor area (SMA), Areas 5 & 7 (parietal), Broca's and Wernicke's areas

Motor programming
 Cortico-basal ganglia and Cortico-cerebellar circuits, SMA, Motor cortex, Fronto-limbic system

Execution
 Cerebellum, basal ganglia, SMA, Motor cortex, Thalamus, Brainstem, Motor units

Figure 1. The four-level (FL) model of speech sensorimotor control (after the original four-level framework, FLF, in Van der Merwe, 1997, 2009). A model for the characterization of pathological speech sensorimotor control.

Speech motor planning: Motor planning of place and manner of articulation of each sound and inter- and co-articulatory control of the segmental features of speech take place during this phase of motor control. From a motor perspective the core motor plan and the different motor goals (for example, tongue and lip position) within a plan are to be recalled from sensorimotor memory. At this point the sequential organization of movements for each sound and the different sounds in the planned unit has to take place. The potential for coarticulation of the different sounds within a unit is also created and needs to be handled by the planner.

In Figure 1, an insert which is dedicated to speech acquisition is included. This description is from the perspective of feedback, feedforward, and forward predictive control. The primary purpose for its inclusion is to indicate the role of response produced auditory and tactile-proprioceptive feedback during speech acquisition and to state that sensory information is available and accessible during processing at the planning level. This will serve as background information in the discussion of the control architecture of speech motor planning in the next section.

During speech acquisition an auditory model of a sound or word is imitated (see insert in Figure 1). Auditory and tactile-proprioceptive feedback are relayed back to the system. During repeated production, sensorimotor transformations are learnt. An inverse model (a core motor plan) and forward model for each sound and over-learned word or syllable are acquired. Imbedded in an inverse model are speech structure-specific motor plan subroutines (motor goals such as lip rounding, velar lifting, glottal closure) which, when coarticulated, would guide the system to produce a target sound and combination of sounds that forms a word.

During mature speech production, inverse models that are key to motor planning generate motor commands. The motor commands from the inverse model reach the forward model by way of an efference copy during internal feedback. The forward predictive model supplies an internal prediction or an estimation of the expected feedback (Miall & Wolpert, 1996). In speech production, monitoring of the efference copy of motor commands by the forward predictive model is proposed to be accomplished by comparing the copy to the inverse model to ensure that the critical acoustical configuration of each sound will be reached. Motor commands are adjusted if necessary and coarticulated and then relayed to motor programming areas.

The FI model proposes that speech motor planning be seen as context sensitive for motor-related factors such as motor complexity, length, familiarity (high versus low frequency use), and initiation mode (imitated or self-initiated) of the target utterance (see Van der Merwe, 1997; 2009 for a discussion). These factors determine task complexity. This viewpoint is in accord with the hypothesis underlying the nonlinear gestural model proposed by Ziegler and colleagues (Ziegler 2009; Ziegler & Aichert; Ziegler, Lehner, Pfab & Aichert, 2020). According to that model the probability of AOS errors can be predicted based on familiarity and some other phonetic features.

Speech motor programming: To further unpack the motor plan, muscle tone, velocity, direction, force, and range of movements of the articulatory, laryngeal and respiratory muscles are specified in motor programs. Motor programming is not reliant on inverse models. Specification of spatiotemporal, force, and metrical structure parameters is circumstance sensitive and could be adapted, for example, to convey emotional intent, to talk louder or faster, or to assign syllabic stress. Specification is also muscle-specific. The linguistic and fronto-limbic systems provide input to the programmer (Brooks, 1986;

Leisman et al., 2014; Nisticò, Cerasa, Olivadese, Volta, Crasà, et al., 2019; Nolte, 1999; Ploog, 1981; Riva, Taddei, Ghielmetti, Erbetta & Bulgheroni, 2019; Zappa et al., 2019) to augment motor programs with specifications which contain suprasegmental and affective information. Repeated initiation, switching and feedforward of motor programs determine rate and metrical structure including the rhythm of speech. Once the motor programs are specified, these are relayed to the muscles for execution. Manifested speech will cause reafferent feedback signals which could potentially be implemented for control.

The control architecture of the four-level model

Architecture of the Speech Motor Planner: The control architecture proposed to underly speech motor planning differs from the control architecture underlying programming. The architecture of the motor planning phase in the FL model (see Figure 2) resembles the Auxiliary Forward Model (AFM) account as described by Pickering and Clark (2014). The AFM makes use of an inverse-forward scheme. This scheme posits two distinct models: an inverse and forward model. In the context of speech motor planning in the FL model, the inverse model is conceptualised as a memory pattern which contains the motor commands to speech structures. These commands are necessary to achieve a sound-specific auditory and tactile-proprioceptive outcome. The motor planner recalls the memory patterns or inverse models. An efference copy of the motor commands for the production of each sound is subsequently communicated to a forward model of that sound. The forward model can predict the sensory consequences of the motor commands. A series of forward models (of the different sounds in a planned unit) are relayed to what is here proposed to be a *forward predictive planner*.

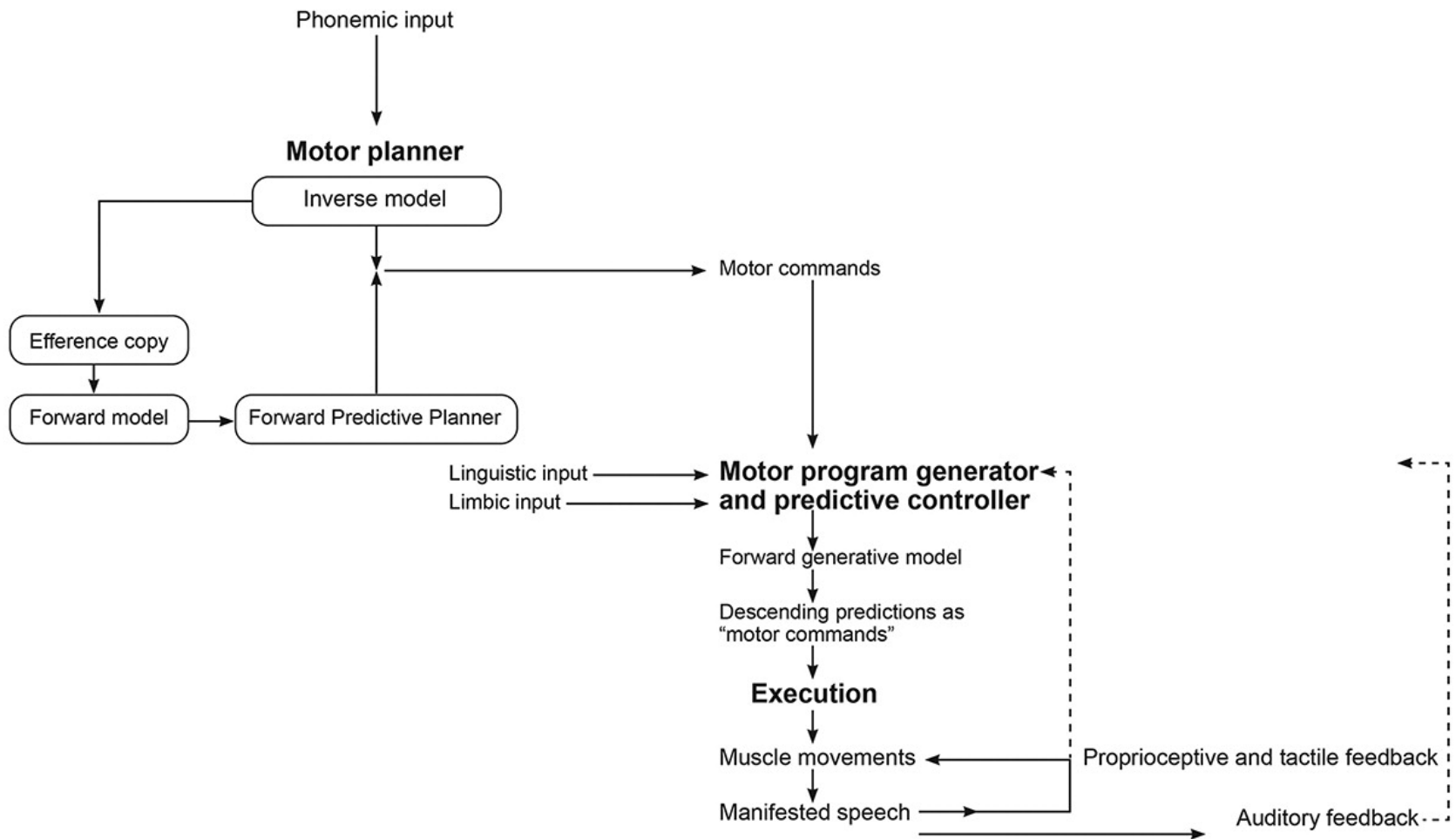


Figure 2. Control architecture of the four-level (FL) model of speech sensorimotor control. The FL model posits two different control architectures for motor planning and programming. The conception of the motor planner utilizes concepts from the auxiliary forward model (AFM) architecture, while the conception of the motor program generator and predictive controller utilizes concepts from the integral forward model (IFM) architecture after Pickering and Clark (2014).

A forward predictive planner would be able to handle planning a series of sounds. Speech production does not imply sound-by-sound production. Speech “violates what can be called the linearity and invariance conditions” (Wanner, Teyler & Thompson, 1977, p. 6). Variance in the articulatory parameters of a sound occurs due to coarticulation of sounds and adaptation of spatial specifications to the phonetic context in which a sound is produced. Coarticulation could also take place across word boundaries (Kent & Minifie, 1977). Only a mechanism such as the *forward predictive planner* conceptualised in the FL model would be able to handle such input. This highly specialized mechanism is necessary to monitor the sensory consequences of each forward model and also a series of models taken up in a planned unit *of sounds*. The output of the forward predictive planner then “work in concert” (Cui, 2016, p. 3) with inverse models to finesse and finalize the motor commands. The motor commands are subsequently relayed to the programming areas of the central nervous system.

The motor planner can operate without the implementation of reafferent feedback and is portrayed as such in Figure 2. Motor planning takes place in a feedforward mode. Predicted feedback is created by a forward model, but this is an internal feedback process. Reafferent feedback is naturally accessible for attentional control (Cooper, Ruh, & Marechal, 2014) during the acquisition of a new motor plan. This viewpoint concurs with the statement by Parrell and Houde (2019) that the speech motor system is sensitive to feedback and perturbations, but that feedback is not critical to speech production.

Architecture of the Speech Motor Program Generator and Predictive Controller: The control architecture of the speech motor program generator and predictive controller differs from that of the motor planner in the sense that inverse models and efference copies are not

utilized. The integral forward model (IFM) architecture described by Pickering and Clark (2014) is a probable control mode for the motor programmer. The general principles of this control architecture are employed with adaptations tailored to the FL model and to the requisites of the speech process. The IFM account stems from work on the role of prediction in perception (Bastos, Usrey, Adams, Mangun, Fries, & Friston, 2012; Clark, 2016). In this account the descending predictions that emanate from an integral generative forward model act as action ('motor') commands to the plant. Active inference or action-oriented processing generate predictions of sensory outcomes that would ensue. Prediction errors are generated, fed back to the forward model, and quashed. Such a system is able to handle delays, perturbations, and sensory noise and suppress errors before they occur (Pickering & Clark, 2014). Bite-block (Folkins & Zimmerman, 1981) and unexpected weight-perturbation studies (Gracco & Abbs, 1986) demonstrated instant compensatory speech movements. In the latter group compensation occurred within milliseconds (Parrell & Houde, 2019). The ensuing inference is that the speech motor program generator and predictive controller (see Figure 2) is circumstance sensitive and able to handle noise in the central nervous system. The programmer is also aware of the desired outcome of the motor plan or task at hand and could utilize internal feedback to generate compensatory motor programs, for example during perturbation. The broken lines in Figure 2 indicate that tactile-proprioceptive and auditory feedback are available and could potentially be utilized by the programmer.

Speech Execution: Motor programs are implemented to drive muscle movements during execution of speech movements. Descending signals provide instructions to the lowest levels of control where they are translated into commands to the muscles. Closed-loop tactile-proprioceptive feedback as a means of control is available during the final phase of

execution (Eccles, 1977). Proprioceptive signals from mechanoreceptors of the joints, muscles, tendons, and skin are implemented for the neural control of movement. Impairment of proprioceptive afferent feedback may impact the control of muscle tone, as well as spatial-temporal parameters of voluntary movement (*Aman, Elangovan, Yeh, & Konczak, 2015*).

THE PATHOPHYSIOLOGY OF APRAXIA OF SPEECH

In view of the preceding account, it seems evident that AOS is a speech motor planning disorder due to a disruption at the highest motor processing level. AOS is the result of a left hemisphere cortical lesion (Duffy, 2013) in the cortical motor areas responsible for the planning of movements. Lesions that lead to lower level deficits than AOS could be either unilateral (left or right sided) or bilateral. The DIVA model locates the lesion that causes AOS in the left ventral premotor cortex (Guenther, 2016). However, the exact localization of the lesion in the dominant hemisphere is still unconfirmed (McNeil, et al., 2016; Terband et al., 2019).

The portrayal of speech motor planning in the FL model provides rich theoretical scope for loci and nature of pathophysiology which could present as AOS signs. In Figure 3 the core components of motor planning are presented together with possible deficiencies that could occur in each. The nature of impairment and underlying pathophysiology could be located in the inverse models of some or all sounds, in the relay of efference copies, or in the mechanism of forward prediction. Examples of the nature of deficiency are listed in Figure 3. These are for example: damage to the internal model that results in a total loss of the inverse model of the CMP; deficient retrieval of motor commands embedded in the inverse model; and a reduced range of forward prediction and planning.

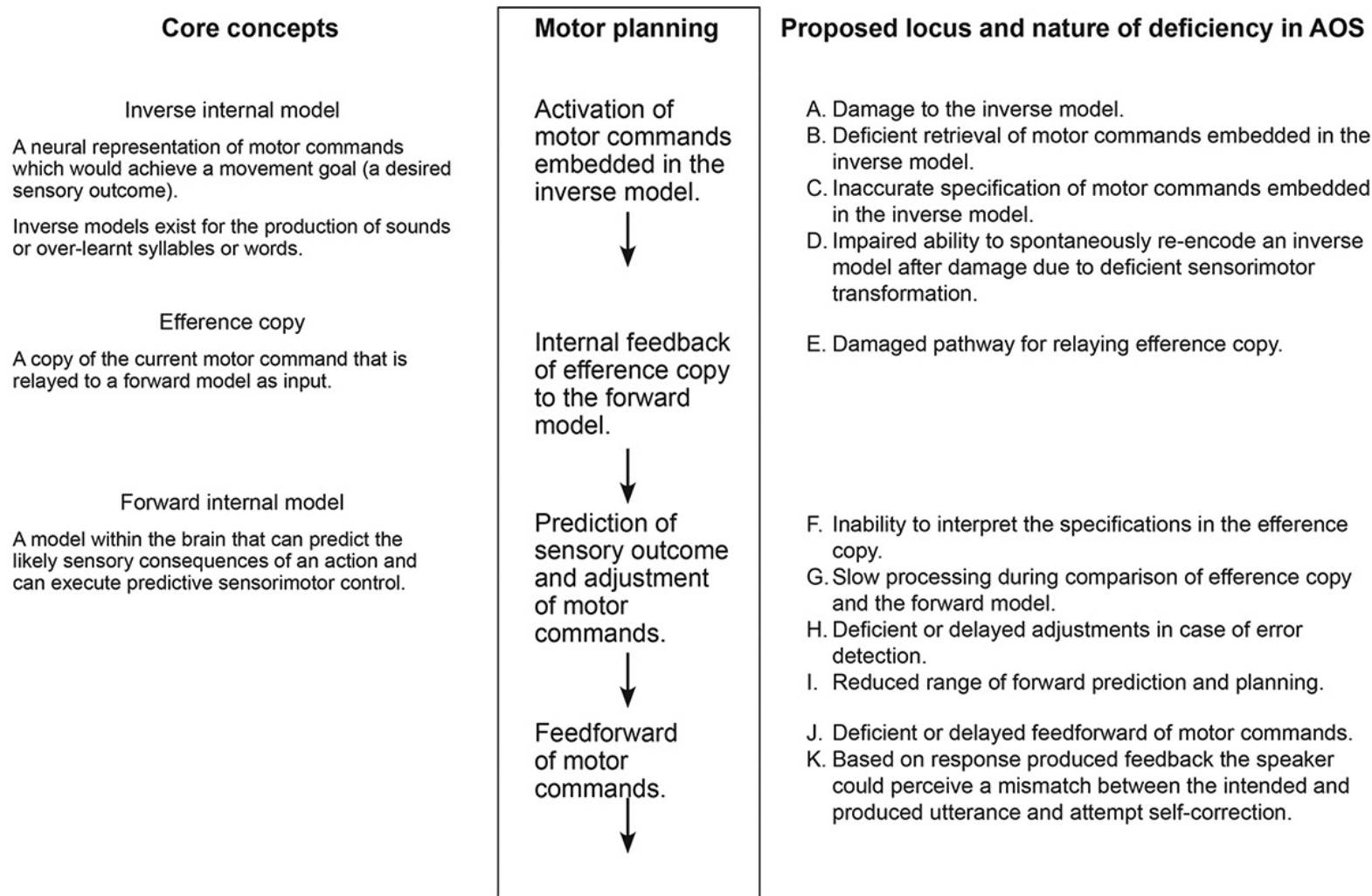


Figure 3. Pathophysiology underlying AOS portrayed as impaired inverse and forward internal models at the motor planning level of speech sensorimotor control.

The underlying pathophysiology could lead to the characteristic symptoms of AOS. Damage to the *inverse* models of sounds and an inability to reconstruct such models could predictably lead to the complete or partial inability to produce these (specific) sounds and therefore also a string of sounds. This may lead to apparent articulatory groping and start-restart behaviour. Incorrect specification of the spatial and temporal motor commands for the production of sounds (including coarticulation of sounds) could impact accurate production leading to distortion of one or more sounds in an utterance. This may also give rise to apparent substitutions and distorted substitutions. Slow retrieval and the effect of increased planning load of longer, unfamiliar, or motorically complex utterances may lead to slow speech, extended segmental duration, extended intersegment duration and syllable-by-syllable speech production. The severity or frequency of speech errors may increase or decrease due to these motor-related contextual factors. Damage to the *forward* internal models of sounds could negatively impact monitoring of the motor commands during internal feedback and lead to distorted speech sounds, slow speech, and apparent groping and start-restart behaviour. These predicted speech errors that could, by inference, occur due to impaired inverse and forward predictive models for speech production concur with the kernel characteristics of AOS as proposed in the literature (Ballard et al., 2016; Duffy, 2007, 2013; McNeil, et al., 2016; McNeil et al., 1997; 2009; McNeil, Pratt & Fossett, 2004).

The relevant question is: is there also a dysfunction at the middle programming level in AOS? Features of speech that could potentially signify disruption in both motor planning and programming in AOS are deficient prosody, speech rate, and accuracy of articulation. Prosodic errors, which are mentioned as salient features of AOS, include slow rate, equal stress across adjacent syllables in syllable segregated speech, and extended segment and intersegment durations (McNeil, et al., 2009; McNeil, et al., 2016; Vergis, et al., 2014).

Prosodic and rate disturbances are perceptual surface features of AOS (Duffy, 2013; McNeil et al., 2009), but are not necessarily primary deficiencies in motor planning. In view of the assignment of prosodic control to the programming level in the updated FL model, these errors could be interpreted as a disruption of the initiation, specification and feedforward of motor programs. However, these signs could also reflect the secondary side effects of a planning disorder. Impaired planning could potentially impact subsequent programming of a unit of sounds. Another possible explanation for signs such as slow rate and syllable segregation is that the speaker is displaying compensatory behaviour. A speech motor planning problem will inevitably cause associated or secondary disruption of the rate, rhythm, and fluency of speech (Duffy, 2013). While the speaker is trying to recall and specify the *mental representation* (Itoh, 2008; Jeannerod, 1995) or motor plan, a delay in processing could surface as, for example, slow rate, disrupted fluency, and syllabic speech.

Studies that sought to find evidence of a primary rate disruption in AOS speech present contradictory results. Spectrographic analyses of durational aspects (total click, burst and release duration) of Zulu click production in words produced by a Zulu speaker with AOS revealed variability in duration. Only nine of 30 opportunities showed duration outside the normal range. Five instances of longer duration and four of shorter duration were noted (Van der Merwe & Steyn, In Press). This study suggests that slow rate at *segmental* level is not a characteristic feature of AOS. Results of another acoustic study, involving five individuals with AOS, showed normal or shorter than normal vowel and total duration of consonant-vowel-consonant nonwords, but longer than normal duration of longer nonwords (Van der Merwe & Grimbeek, 2006). Findings of a study by Ballard and colleagues (2014) suggest that a measure of relative vowel duration from a polysyllabic word repetition task is sufficient to detect the presence of AOS in cases with progressive

aphasia. In primary progressive AOS, speech rate was found to become slower and utterance duration more extended as the disorder progresses. This may point to slow rate as a primary deficit in AOS (Duffy et al. 2015). In programming disorders, prosodic and rate disturbances are primary signs (discussed earlier).

Distortion of speech occurs across all neuromotor speech disorders. The context sensitivity and consistency of occurrence may differentiate the underlying cause of distortion. In the case of AOS, speech errors - including distortion - may be sensitive to motor-related contextual factors such as familiarity, length, and motor complexity (Van der Merwe, 2011; Ziegler et al., 2020). Conversely, distortion of speech in programming-execution and execution disorders appears to be a more consistently present characteristic (Duffy, 2013).

In summary, superficially judged, there are overlapping speech signs which can be explained at both a planning and a programming level. To differentiate these options, a detailed model of speech motor planning, programming and execution is necessary to drive research questions. The possibility that a disruption of motor programming is also present in AOS cannot conclusively be discounted. It may also be present in certain individuals due to the localization of the lesion. McNeil and co-authors (2016, p. 211) state: "Assuming that speech planning and programming are separable and multistage processes, ... researchers may begin to generate testable hypotheses about clinical features and localisation correlates of what might become recognizable subtypes of AOS". Underlying this statement is the assumption that programming is more closely connected to and associated with the planning phase than to the execution phase. The updated FL model and the recent conceptualization of computational models of the Controller (Houde & Nagarajan, 2011; Parrell, Lammert, et al., 2019; Parrell, Ramanarayanan, Nagarajan, & Houde, 2019) appear

to oppose such an assumption. The conceptualization of a programmer challenges the traditional (as in the three-level model) simplistic view of execution and not primarily the understanding of planning, as is generally accepted. The search for subtypes of AOS should display an awareness of these facts. Programming is a complex process that precedes execution, and involves muscle-specific programs and the use of deafferent feedback if necessary. These programs drive motor execution, but execution also depends on closed-loop control. Programming is therefore proposed in the FL model to be closely connected to execution. Conversely, planning of movement does not share these properties (muscle-specific commands and implementation of feedback), but depends on sensorimotor memory patterns and inverse models. Also, planning takes place at the highest levels of motor control in the dominant hemisphere while programming and execution are mediated by cortical-subcortical circuits and sub-cortical structures. At this point in time, and against the theoretical account provided in this manuscript, AOS can most logically be characterized as a disorder in speech motor planning. It would be inaccurate to portray AOS as a “planning or programming”, or as a “planning and/or programming” disorder. Also, a depiction of AOS as a “planning and programming” disorder should include a warning regarding the tentative nature of this label. To describe AOS as a “planning/programming” disorder (in other words the interchangeable use of these terms) is to disregard the advances that were made in theoretical accounts of speech sensorimotor control.

CONCLUSIONS AND FUTURE RESEARCH

Contemporary accounts of the motor control hierarchy for skilled movements substantiate the proposed three levels or phases in the motor control of speech production (Van der Merwe, 1997; 2009; Van der Merwe & Steyn, 2018). A motor planning level and a

programming level are acknowledged in many published works on the topic of motor control, though the descriptive terms employed sometimes differ across publications. In the domain of speech motor control, contemporary computational theoretical models acknowledge the existence of a planning level involved in the creation of motor plans or high-level tasks. This level is differentiated from a lower-level controller. Though the latter is not explicitly acknowledged as a programming level and not differentiated from an execution phase, future developments in this field could support the distinction between planning, programming, and execution phases.

The new version of the FLF, the FL model, which is presented in this manuscript, differentiates these three motor phases or levels. Planning of the segmental features of speech (place and manner of articulation and inter- and co-articulatory control) is handled by a motor planner. The author proposes that the motor planner functions according to an auxiliary forward model architecture and is driven by an inverse model for every speech sound, relay of an efference copy of the motor commands, and a forward model of each plan. Also posited in the FL model is the existence of a forward predictive planner that can handle planning of coarticulation of more than one sound. Speech motor planning is suggested to function without any feedback in a typical mature speaker. The proposed control mode of the planning phase is conceptually based on the auxiliary forward model (AFM) architecture as described by Pickering and Clark (2014).

During the following motor programming phase, a motor program generator and predictive controller is proposed to be responsible for the circumstance sensitive and muscle-specific specification of spatiotemporal, force, and metrical structure parameters for segmental and suprasegmental control. An important implication is that the programmer is responsible for controlling prosodic features and the metrical structure of speech output.

Perturbations and noise in the nervous system can be handled by the programmer. The control architecture is based on the integral forward model (IFM) architecture described by Pickering and Clark (2014). During execution, closed loop online control based on proprioceptive and tactile feedback takes place.

The conception of speech motor planning as proposed in the FL model provides a theoretical basis for making inferences about the locus and nature of deficiency in AOS. A breakdown may occur at multiple points in this process. The nature of pathophysiology may include, for example, actual damage to an inverse model of a sound, deficient retrieval of motor commands, inaccurate specification, slow processing of an efference copy and reduced range of forward prediction and planning. These difficulties could conceivably lead to the characteristic features of speech in AOS.

In view of the developments in the neurosciences, the traditional conceptualization of a two-phase process (planning/programming and execution) in speech motor control is too limited to characterise motor speech disorders and drive future research. Speech motor planning and programming has distinguishable roles in speech motor control. AOS is most accurately typified as a motor planning disorder. Future research could address the exact nature of articulatory movements and other features of speech in planning versus programming, programming-execution and execution disorders, using for example, real-time magnetic resonance imaging (see Hagedorn, 2017). Response to changes in contextual factors (for example, motor complexity, syllable structure, and length of utterance) could differentiate between the signs of disorders at these levels. Research with a focus on levels of severity could also reveal much about the underlying nature of impairment of the different motor speech disorders. In conclusion, future research would benefit from

consistent terminology and the consistent depiction of AOS as a motor *planning* disorder with a disorder in motor *plans*.

List of Figures

Figure 1: The four-level (FL) model of speech sensorimotor control (after the original four-level framework, FLF, in Van der Merwe, 1997; 2009). A model for the characterization of pathological speech sensorimotor control.

Figure 2: Control architecture of the four-level (FL) model of speech sensorimotor control. The FL model posits two different control architectures for motor planning and programming. The conception of the motor planner utilizes concepts from the auxiliary forward model (AFM) architecture, while the conception of the motor program generator and predictive controller utilizes concepts from the integral forward model (IFM) architecture after Pickering and Clark, 2014.

Figure 3: Pathophysiology underlying AOS portrayed as impaired inverse and forward internal models at the motor planning level of speech sensorimotor control.

REFERENCES

- Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. *Trends in Neurosciences*, 31(6), 265-272.
- Aman, J. E., Elangovan, N., Yeh, I-L, & Konczak, J. (2015). The effectiveness of proprioceptive training for improving motor function: a systematic review. *Frontiers in Human Neuroscience*, 8, Article 1075, 1-18.
- American Speech-Language-Hearing Association (ASHA). (2007). *Childhood apraxia of speech [technical report]*. Available from: www.asha.org/policy.
- Allen, G. I., & Tsukahara, N. (1974). Cerebrocerebellar communication systems. *Physiology Review*, 54, 957-997.
- Andersen, R.A., & Cui, H. (2009). Intention, action planning and decision making in parietal-

frontal circuits. *Neuron*, 63, 568–583.

Ballard, K. J., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., ..., & Robin, D. A.

(2016). A predictive model for diagnosing stroke-related apraxia of speech.

Neuropsychologia, 81, 129-139.

Ballard, K. J., Savage, S., Leyton, C. E., Vogel, A. P., Hornberger, M., Hodges, J. R. (2014).

Logopenic and Nonfluent Variants of Primary Progressive Aphasia Are Differentiated by Acoustic Measures of Speech Production. *PLoS ONE* 9(2): e89864.

<https://doi.org/10.1371/journal.pone.0089864>

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012).

Canonical microcircuits for predictive coding. *Neuron*, 76, 695-711.

Bislick, L. & Hula, W. D. (2019). Perceptual characteristics of consonant production in apraxia

of speech and aphasia. *American Journal of Speech-Language Pathology*, 28(4),

1411-1431.

Bislick, L., McNeil, M. R., Spencer, K. A., Yorkston, K., & Kendall, D. L. (2017). The nature of

error consistency in individuals with acquired apraxia of speech and aphasia.

American Journal of Speech-Language Pathology, 26, 611-630.

Brumberg J. S., Krusienski, D. J., Chakrabarti, S., Gunduz, A., Brunner, P., Ritaccio, A. L., et al.

(2016). Spatio-Temporal Progression of Cortical Activity Related to Continuous Overt and Covert Speech Production in a Reading Task. *PLoS ONE* 11(11): e0166872.

<https://doi.org/10.1371/journal.pone.0166872>.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2003). *Speech Science Primer: Physiology,*

Acoustics, and Perception of Speech. Fourth edition. Philadelphia: Lippincott Williams & Wilkins.

Brendel, B., Erb, M., Riecker, A., Grodd, W., Ackermann, H., Ziegler, Z. (2011). Do We Have a

- “Mental Syllabary” in the Brain? An fMRI Study. *Motor Control*, 15, 34-51.
- Brooks, V. B. (1986). *The Neural Basis of Motor Control*. New York: Oxford University Press.
- Burbaud, P., Doegle, C., Gross, C., & Bioulac, B. (1991). A quantitative study of neural discharge in areas 5, 2 and 4 of the monkey during fast arm movements. *Journal of Neurophysiology*, 66, 429-443.
- Callan, D. E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: The role of the cerebellum. *The Cerebellum*, 6, 321–327.
- Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. New York, NY: Oxford University Press.
- Cooper, R. P. (2010). Forward and inverse models in motor control and cognitive control. *Symposium on AI-inspired biology*. De Montfort University, Leicester, UK.
- Cooper, R. P., Ruh, N., & Mareschal, D. (2014). The goal circuit model: A hierarchical multi-route model of the acquisition and control of routine sequential action in humans. *Cognitive Science*, 38, 244–274.
- Crochet, S., Seung-Hee, L., & Petersen, C. C. H. (2019). Neural Circuits for Goal-Directed Sensorimotor Transformations. *Trends in Neurosciences*, 42(1), 66-77.
- Cui, H. (2014). From intention to action: Hierarchical sensorimotor transformation in the posterior parietal cortex. *eNeuro*, 1(1), November-December, e0017-14.2014. 1-6.
- Cui, H. (2016). Forward Prediction in the Posterior Parietal Cortex and Dynamic Brain-Machine Interface. *Frontiers in Integrative Neurosciences*, 10, Article 35, 1-6.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor Speech Disorders* (1975). Philadelphia: W. B. Saunders.
- Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. *Journal of Neural Engineering*, 2, S313–S319.

- Diehl, S. K., Mefferd, A. S., Lin, Y-C., Sellers, J., McDonell, K. E., De Riesthal, M. & Claassen, D. O. (2019). Motor speech patterns in Huntington disease. *Neurology*, *93*(22), e2042-e2052.
- Di Pelligrino, G. & Wise, S. P. (1992). A neurophysiological comparison of three distinct regions of the primate frontal lobe. *Brain*, *114*, 951-978.
- Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, *10*, 732–739.
- Dreher, J., & Grafman, J. (2002). The roles of the cerebellum and basal ganglia in timing and error prediction. *European Journal of Neuroscience*, *16*, 1609-1620.
- Duffy, J. R. (2007). Motor speech disorders: History, current practice, future trends and goals. In Weismer G. (Ed.). *Motor Speech Disorders: Essays for Ray Kent*. San Diego, CA: Plural Publishing. p. 7-56.
- Duffy, J. R. (2013). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. Third edition. St. Louis, Missouri: Elsevier Mosby.
- Duffy, J. R., Strand, E. A., Clark, H., Machulda, M., Whitwell, J. L., Josephs, K. A. (2015). Primary progressive apraxia of speech: Clinical features and acoustic and neurological correlates. *American Journal of Speech-Language Pathology*, *24*, 88-100.
- Eccles, J. C. (1977). *The Understanding of the Brain*. Second edition. New York: McGraw-Hill Book Company.
- Evarts, E. V. (1982). Analogies between central programs for speech and for limb movements. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson. *Speech Motor Control*, vol. 36. Oxford: Pergamon Press.
- Evarts, E. V., & Wise, S. P. (1984). Basal ganglia outputs and motor control. In *Ciba Foundation Symposium 107: Functions of the Basal Ganglia*. London: Pitman.

- Fairbanks, G. (1954). Systematic research in experimental phonetics. 1. A theory of the speech mechanism as a servosystem. *Journal of Speech and Hearing Disorders*, 19(2), 133–139.
- Fetz, E. E. (1993). Cortical mechanisms controlling limb movement. *Current Opinion in Neurobiology*, 3, 932-939.
- Folkins, J. W., Zimmerman, G. N. (1981). Jaw-muscle activity during speech with the mandible fixed. *Journal of the Acoustical Society of America*, 69, 1441-445.
- Franklin, D. W., & Wolpert, D. M. (2011). Computational Mechanisms of Sensorimotor Control. *Neuron* 72, 425-442.
- Friston, K. J., Daunizeau, D., Kilner, J., Kiebel, S. J. (2010) Action and behavior: a free-energy formulation. *Biological Cybernetics*, 102, 227–260.
- Frith, C. D., & Haggard, P. (2018). Volition and the Brain – Revisiting a Classic Experimental Study. *Trends in Neurosciences*, 41(7), 405-407.
- Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research*, 65, 156-166.
- Goldberg, J. H., Farries, M. A., & Fee, M. S. (2013). Basal ganglia output to the thalamus: still a paradox. *Trends in Neurosciences*, 36(12), 695-705.
- Groenewegen, H. J. (2003). The Basal Ganglia and Motor Control. *Neural Plasticity*, 10(1-2), 107-120.
- Guenther, F. H. (2016). *Neural Control of Speech*. Cambridge, MA: MIT Press.
- Guenther, F. H., & Perkell, J. S. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maassen, R. Kent, H. Peters, P. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (29-49). Oxford: Oxford University Press.

- Guenther, F. H. & Vladusich, T. (2012). A Neural Theory of Speech Acquisition and Production. *Journal of Neurolinguistics*, 25(5), 408–422.
- Habas, C. (2010). Functional imaging of the deep cerebellar nuclei: a review. *Cerebellum*, 9(1), 22-28.
- Hagedorn, C., Proctor, M., Goldstein, L., Wilson, S. M., Bruce Miller, B., Gorno-Tempini, M. L., & Narayanan, S. S. (2017). Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging. *Journal of Speech, Language, and Hearing Research*, 60, 877–891.
- Hernandez, L. F., Obeso, I., Costa, R. M., Redgrave, P., & Obeso, J. A. (2019). Dopaminergic vulnerability in Parkinson disease: The cost of humans' habitual performance. *Trends in Neurosciences*, 24(6), 375-383.
- Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding in the context of a dual-stream model. *Journal of Communication Disorders* 45(6), 393-402.
- Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience*, 29(1), 2-20.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69, 407-422.
- 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, Article 649, 1-8.
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, Article 82.

- Itoh, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews: Neuroscience*, 9, 304-313.
- Ishikawa, T., Tomatsu, S., Izawa, J., Kakei, S. (2016). The cerebro-cerebellum: Could it be loci of forward models? *Neuroscience Research* 104, 72–79.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17(2), 187-245.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33(11), 1419-1432.
- Johnson, A. M., Vernon, P. A., Almeida, O. J., Grantier, L. L. & Jog, M. S. (2003). The role of the basal ganglia in movement: effect of pre-cues on discrete bi-directional movements in Parkinson's disease. *Motor Control*, 7, 71-81.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718-727.
- Kawato, M. & Gomi, H. (1992). The cerebellum and VOR/OKR learning models. *Trends in Neuroscience*, 15, 445-453.
- Kawato, M. & Wolpert, D. (2007). Internal models for motor control. In G. R. Bock & J. A. Goode (Eds.). *Sensory Guidance of Movement*. London: Novartis Foundation Symposia, 218.
- Kent, R. D., & Minifie, F. D. (1977). Coarticulation in recent speech production models. *Journal of Phonetics*, 5, 115-133.
- Leisman, G., Baraun-Benjamin, O., & Melillo, R. (2014). Cognitive-motor interactions of the basal ganglia in development. *Frontiers in Systems Neurosciences*, 8, 1-18.
- Lowit, A., Marchetti, A., Corson, S., Kuschmann, A. (2018). Rhythmic performance in hypokinetic dysarthria: Relationship between reading, spontaneous speech and

- diadochokinetic tasks. *Journal of Communication Disorders*, 72, 26-39.
- Mailend, M-L., Maas, E., Beeson, P. M., Story, B. H., & Forster, K. I., (2019). Speech Motor Planning in The Context of Phonetically Similar Words: Evidence from Apraxia of Speech and Aphasia, *Neuropsychologia*, 127, 171-184.
- Magill, R. A. (2007). *Motor learning and control: Concepts and applications* (8th ed.). Boston, MA: McGraw-Hill.
- Mariën, P. (2012). Cerebellar control of motor speech. In Consensus paper: Roles of the cerebellum in motor control – The diversity of ideas on cerebellar involvement in movement. *The Cerebellum*, 11(2), 457-487.
- Marsden, C. D. (1984). Which motor disorder in Parkinson's disease indicates the true motor function of the basal ganglia? In: *Ciba Foundation Symposium 107: Functions of the Basal Ganglia*. London: Pitman.
- McAllen, A. W., Spencer, K. A., France, K. N., & Shulein, O. M. (2010). Speech and Manual Reaction Time as a Function of Dopaminergic Medication in Parkinson's Disease. *Journal of Medical Speech-Language Pathology*, 18(3), 59–74.
- McNeil, M. R., Ballard, K. J., Duffy, J. R., & Wambaugh, J. (2016). Apraxia of speech theory, assessment, differential diagnosis, and treatment: Past, present and future. 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: ASHA Press.
- McNeil, M. R., Doyle, P. J., & Wambaugh, J. (2000). Apraxia of speech: A treatable disorder of motor planning and programming. In S. E. Nadeau, L. J. Gonzalez Rothi, B. Crosson. (Eds). *Aphasia and Language: Theory to Practice*. Ne York, NY: The Guilford Press.
- McNeil, M. R., Pratt, S. R., & Fossett, T. R. D. (2004). The differential diagnosis of apraxia of

- speech. In B. Maassen, R. D. Kent, H. Peters, P. van Lieshout, & W. Hulstijn. (Eds.), *Speech Motor Control in Normal and Disordered Speech*. Oxford: Oxford University Press.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: Definition and differential diagnosis. In M. R. McNeil. (Ed.). *Clinical management of sensorimotor speech disorders*. New York, NY: Thieme.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition and differential diagnosis. In M. R. McNeil. (Ed.). *Clinical management of sensorimotor speech disorders* (pp. 1-25). New York, NY: Thieme.
- Miall, R. C., Weir, D. J., & Stein, J. F. (1987). Visuo-motor tracking during reversible inactivation of the cerebellum. *Experimental Brain Research*, *65*, 455-464.
- Miall, R. C., & Wolpert, D., M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265-1279.
- Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*, 381-425.
- Murata, A., Wen, W., & Asama, H. (2016). The body and objects represented in the ventral stream of the parieto-premotor network. *Neuroscience Research*, *104*, 4-15.
- New, A. B., Robin, D. A., Parkinson, A. L., Duffy, J. R., McNeil, M. R., Piguet, O., ..., & Ballard, K. J. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage: Clinical*, *8*, 429-439.
- Niesler, T., Louw, P., & Roux, J. (2005). Phonetic analysis of Afrikaans, English, Xhosa and Zulu using South African speech databases. *Southern African Linguistics and Applied Language Studies*, *23(4)*: 459-474.

- Nisticò, R., Cerasa, A., Olivadese, G., Volta, R. D., Crasà, M., et al. (2019). The embodiment of language in tremor-dominant Parkinson's disease patients. *Brain and Cognition*, *135*, 103586. 1-7.
- Nolte, J. (1999). *The Human Brain: An Introduction to its Functional Anatomy*. 4th Edition. St Louis: Mosby.
- Parrell, B., & Houde, J. (2019). Modeling the role of sensory feedback in speech motor control and learning. *Journal of Speech, Language, and Hearing Research*, *62*, 2963-2985.
- Parrell, B., Lammert, A. C., Ciccarelli, G., & Quatieri, T. F. (2019). Current models of speech motor control: A control-theoretic overview of architectures and properties. *Journal of the Acoustical Society of America*, *145(3)*, 1456-1481.
- Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLOS Computational Biology*, *15(9)*, e1007321. 1-26.
- Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18(9)*, 451-456.
- Ploog, D. (1981). Neurobiology of primate audio-vocal behaviour. *Brain Research Review*, *3*, 35-61.
- Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*, *89*, 1009-1023.
- Reilly, K. J., & Spencer, K. A. (2013). Sequence Complexity Effects on Speech Production in Healthy Speakers and Speakers with Hypokinetic or Ataxic Dysarthria. *PLoS ONE* *8(10)*: e77450.

- Reis, J., Swayne, O.B., Vandermeeren, Y., Camus, M., Dimyan, M. A., Harris-Love, M., ..., & Cohen, L. G. (2008). Contribution of transcranial magnetic stimulation to understanding of cortical mechanisms in motor control. *The Journal of Physiology*, *586*(2), 325-351.
- Riva, D., Taddei, M., Ghielmetti, F., Erbetta, A., & Bulgheroni, S. (2019). Language Cerebro-cerebellar Reorganization in Children After Surgery of Right Cerebellar Astrocytoma: a fMRI Study. *The Cerebellum*, *18*, 791–806.
- Rose, D. J. (1997). *A Multilevel Approach to the Study of Motor Control and Learning*. Boston: Allyn & Bacon.
- Saltzman, E., & Munhall, K. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.
- Sheridan, M. R., Flowers, K. A., & Hurrell, J. (1987). Programming and execution of movement in Parkinson's disease. *Brain*, *110*, 1247–1271.
- Schmullian D., Van der Merwe A. & Groenewald E. (1997). An exploratory study of an undefined acquired neuromotor speech disorder within the context of the Four Level Framework for Speech Sensorimotor Control. *South African Journal of Communication Disorders*. *44*, 87-97.
- Schultz, W., & Romo, R. (1992). Role of primate basal ganglia and frontal cortex in the internal generation of movements. I. Preparatory activity in the anterior striatum. *Experimental Brain Research*, *91*, 363-384.
- Skodda, S. (2011). Aspects of speech rate and regularity in Parkinson's disease. *Journal of the Neurological Sciences*, *310*, 231–236.
- Spencer, K. A., & Rogers, M. A. (2005). Speech motor programming in hypokinetic and ataxic dysarthria. *Brain and Language*, *94*, 347–366

- Spencer, K. A., & Slocumb, D. L. (2007). The neural basis of ataxic dysarthria. *The Cerebellum*, 6, 58-65.
- Stipinovich, A., & Van der Merwe, A. (2007). Acquired Dysarthria within the Context of the Four-Level Framework of Speech Sensorimotor Control. *South African Journal of Communication Disorders*, 54, 67-76.
- Tanaka, H. (2016). Modeling the motor cortex: Optimality, recurrent neural networks, and spatial dynamics. *Neuroscience Research*, 104, 64-71.
- Terband, H., Rodd, J., & Maas, E. (2019). Testing hypotheses about the underlying deficit of apraxia of speech through computational neural modelling with the DIVA model. *International Journal of Speech-Language Pathology*, 35, 246–279.
- 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.
- Van der Merwe, A. (2007). Self-correction in apraxia of speech: The effect of treatment. *Aphasiology*, 21(6-8), 658-669.
- Van der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 3-18). New York, NY: Thieme.
- Van der Merwe, A., & Grimbeek, J. (2006). Variability of voice onset time, vowel duration and utterance duration in apraxia of speech. *Stem-, Spraak- en Taalpathologie*, 14 Supplement, Juni 2006. 5th International Conference on Speech Motor Control Nijmegen: Abstracts.

- Van der Merwe, A., & Steyn, M. (2018). Model-driven treatment of childhood apraxia of speech: Positive effects of the speech motor learning approach. *American Journal of Speech-Language Pathology, 27*, 37-51.
- Van der Merwe, A. & Steyn, M. (In Press). Production of click sounds in acquired apraxia of speech: A view to the motoric nature of the disorder. In B. Sands (Ed.). *The Click Book*. Leiden, The Netherlands: Brill Publishers.
- Van Lancker Sidtis, D., Cameron, K., & Sidtis, J. J. (2012). Dramatic effects of speech task on motor and linguistic planning in severely dysfluent parkinsonian speech. *Clinical Linguistics & Phonetics, 26*, 695–711.
- Van Lancker Sidtis D., Pachana, N., JeVrey, C., Cummings, L., & Sidtis, J. J. (2006). Dysprosodic speech following basal ganglia insult: Toward a conceptual framework for the study of the cerebral representation of prosody. *Brain and Language 97*, 135–153
- Vergis, M. K., Ballard, K. J., Duffy, J. R., McNeil, M. R., Scholl, D., & Layfield, C. (2014). An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology, 28(5)*, 554-575.
- Wanner, E., Teyler, T. J., & Thompson, R. F. (1977). The psychobiology of speech and language – an overview. In J. E. Desmedt (Ed.), *Language and Hemispheric Specialization in Man: Cerebral Event-Related Potentials, Progress in Clinical Neurophysiology*, Volume 3. Basel: S. Karger.
- Widmaier, E. P., Raff, H., & Strang, K. T. (2006). *Vander's Human Physiology: The Mechanism of Body Function* (10th edition). New York: McGraw-Hill.
- Wolpert, D. M., Diedrichsen J., & J. Randall Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*, 739-751.

- Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, 5(11), 487-494.
- Wolpert, D. M. & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11, 1317-1329.
- Xu, D., Liu, T., Ashe, J., & Bushara, K. O. (2006). Role of the Olivo-Cerebellar System in Timing. *The Journal of Neuroscience*, 26(22), 5990–5995.
- Zappa A., Bolger, D., Pergandia, J-M., Mallet, P., Dubarry, A-S. Mestre, D., Frenck-Mestre, C. (2019). Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment. *Brain and Cognition* 134, 44–57.
- Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, 24, 631-661.
- Ziegler, W. & Aichert, I. (2015). How much is in a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex*, 69, 24-39.
- Ziegler, W., Lehner, K., Pfab, J., Aichert, I. (2020) The nonlinear gestural model of speech apraxia: Clinical implications and applications. *Aphasiology*,
DOI: [10.1080/02687038.2020.1727839](https://doi.org/10.1080/02687038.2020.1727839)