



## A neurophonetic approach to articulation planning: The case of apraxia of speech

**Wolfram Ziegler**, EKN, Institute of Phonetics and Speech Processing, Ludwig-Maximilians-University, Munich, Germany, [wolfram.ziegler@ekn-muenchen.de](mailto:wolfram.ziegler@ekn-muenchen.de)

---

This article addresses the question of whether data from neurological populations can inform basic phonetic research. An approach is reviewed in which speech error data from patients with articulation disorders after a stroke were used to model articulation planning processes.

The first section addresses a controversy in cognitive neuroscience about the extent to which data from neurological patients can enrich research in cognitive science at all. It introduces post-stroke apraxia of speech (AOS) as a clinical model of a process termed speech motor planning, suggesting that speech data from patients with AOS can inform phonetic theories about articulatory control.

The second section explains how speech error data from patients with AOS were used to develop a measure of articulatory ease that integrates facilitating and complicating factors across different phonological levels, from the articulatory gesture to the phonological word.

Section 3 presents previously unpublished data to exemplify how this model can be applied to typical speech, e.g., in studies investigating interactions between lexical and articulatory parameters or in developmental studies of word learning. In a final section, the main line of reasoning is summarized and limitations of this work are discussed.

---



# 1. Introduction

## 1.1. Neurophonetics

Neurophonetics is a research field that uses the knowledge and methodological tools developed in phonetics to study the neurological aspects of speaking and speech perception. It covers investigations of the impact of brain dysfunctions on speech production and perception in patients with neurologic conditions, but also of the neural substrates of speech in typical speakers/listeners. The work reviewed here deals exclusively with speech production aspects of neurophonetics.

Two different lines of neurophonetic research can be distinguished. First and foremost, phonetic knowledge and methodologies are applied in the service of clinical neurology or neurologic rehabilitation. Investigations are devoted to describing the speech patterns of patients with neurologic disorders, such as stroke, Parkinson's disease, cerebellar ataxia, or neurologic conditions acquired at birth or during childhood. Most generally, this research aims at understanding the impact of neural dysfunctions of specific aetiologies (e.g., Parkinson's disease; cf. Ruzs et al., 2015) or localizations (e.g., cerebellar degeneration; cf. Brendel et al., 2015) on the patients' speech characteristics. Applications of this research seek to establish physiologic, acoustic, or auditory-perceptual parameters that are sensitive and specific to neurologic speech impairments, with the ultimate goal of developing phonetically based assessment tools. For a comprehensive overview of clinical issues, see Duffy (2020).

A second, much less frequented way of neurophonetic research goes in the opposite direction, aiming to uncover principles of normal speech by investigating its breakdown due to dysfunctions of relevant brain networks. An example of this approach will be reviewed in this article. Generally, the study of disordered motor or cognitive functions after lesions to the brain to learn more about the neural organization of 'normal' functions has long been a prolific principle in cognitive neuropsychology, especially in neurolinguistics. In this approach, it is assumed that brain lesions—with little empathy sometimes referred to as 'experiments of nature'—open a window on how the brain generates behaviour.

However, this account is not undisputed. It tacitly builds on the so-called 'transparency' assumption, which claims that brain lesions do not create 'new' cognitive architectures, but rather cause 'local transformations' of the pre-existing system, such that the resulting condition still allows the structure of the underlying, unimpaired cognitive or linguistic functions to be recognized (for a profound discussion of this controversy see Caramazza, 1988). As a major precondition, this approach must be based on the solid foundations of a clinical model whose relationship to the cognitive function in question is sufficiently well established. If this condition is not met, as is the case with other speech impairments such as neurogenic stuttering, for example (see Section 1.2.3.), the relationship between the impaired behaviour and the cognitive or linguistic mechanisms underlying normal behaviour remains opaque.

The argument made in this paper rests on a speech disorder called apraxia of speech (AOS) and on the claim that this disorder afflicts a specific capacity, i.e., the capacity of planning the articulator movements for the production of speech. Articulation planning is conceived as a process upstream of the elementary motor functions regulating muscular tone or the amplitudes, speed, and coordination of speech movements. It is considered highly dependent on speech motor learning during language acquisition and therefore language-specific. The speech impairment of AOS is considered to result from lesions to brain centres housing the acquired implicit knowledge about how the articulator movements must be planned to produce the syllables, words, and sentences of a speaker's native language (e.g., Ziegler, Aichert, & Staiger, 2012). In line with the transparency assumption mentioned above, I hold that the disintegration of articulation planning in speakers with AOS discloses the structure of the articulation plans of healthy adult speakers. For a better understanding of the argument, a brief overview of this clinical condition and its theoretical basis is given in the following sections.

## 1.2. Apraxia of speech (AOS)

### 1.2.1. Clinical pattern

Apraxia of speech occurs predominantly in adult patients who have suffered a left hemisphere stroke.<sup>1</sup> Initially, the patients are often completely unable to articulate, but within several hours or days their speech may gradually recover to a greater or lesser extent. Their articulation is usually slow, dysfluent, and effortful, and they produce errors such as substitutions, omissions and distortions of speech sounds, schwa intrusions in consonant clusters or cluster reductions, slowed or distorted transitions between speech sounds, and impaired coarticulation. They show laborious groping movements of the articulators, with repeated attempts, false starts, and restarts in the initiation of utterances. Such aberrant articulations have also been demonstrated by analyses of speech movements of the tongue (Hagedorn et al., 2017; Pouplier & Hardcastle, 2005) or the vocal folds (Hoole, Schröter-Morasch, & Ziegler, 1989).

Apraxic speech errors are hardly predictable, i.e., a word that causes severe articulation problems for a patient at a certain moment can be produced almost flawlessly by the same patient a few seconds later. As an example, on five separate attempts to say the word /kapite:n/ (*Kapitän*, English “captain”), a patient with AOS produced:

---

<sup>1</sup> For neuroanatomical reasons, AOS usually occurs concomitantly with aphasia (Ziegler, Aichert, & Staiger, 2012; Ziegler et al., 2022). It has also been described in patients with primary progressive aphasia, a rare condition occurring in the course of frontotemporal lobar degeneration or a language-onset variant of Alzheimer's disease, but it is still not clear whether the clinical picture is the same as in post-stroke AOS (J. R. Duffy, Strand, & Josephs, 2014; Staiger et al., 2021). AOS is distinguished from childhood apraxia of speech, which has similar symptomatology to adult AOS but affects speech motor planning at or before the nascent stage.

(1) /k'a:pi't<sup>hh</sup>e:m/ /p<sup>·</sup>ka<sup>·</sup>ti'tse:n/ /tapi'teñ/ /k<sup>·</sup>kap<sup>·</sup>pi'te:n/ /<sup>?</sup>ap<sup>·</sup>t<sup>hh</sup>e:n/.

The same patient produced several other words without making any substantial errors, e.g.:

(2) /dax/ (roof) → [dax]      /ʃa:f/ (sheep) → [ʃa:f]  
 /to'ma:tə/ (tomato) → [t..to'ma:tə]

The syndrome may manifest itself over a wide range of severity levels, from almost complete mutism to only mildly dysfluent speech with occasional articulation errors, and with different recovery dynamics across patients. For more detailed descriptions of the clinical pattern of AOS see e.g., Duffy (2020) or Ziegler (2008).

### 1.2.2. Neuro-anatomy and pathomechanism

Research interest in this condition has existed for almost 160 years. Broca's seminal case study of a man who lost "la faculté du langage articulé" [*the faculty of articulate language*] after a lesion to the posterior part of the left inferior frontal gyrus (Broca, 1861) is acknowledged by some as the birth of systematic clinical brain-behaviour research (Schiller, 1979). Broca underscored, in the terminology of his time, that the patient was unable to articulate, although he had no obvious motor restrictions ('paralysis') of the tongue and lips and no generalized language or cognitive impairment. He allocated the ability to articulate to the left frontal cortical region that bears his name until today, and characterized the speech impairment as a "loss of the memory of the procedures required for the articulation of words" (Broca, 1861; p. 333). In a modern paraphrase of Broca's terminology one would call it a loss of the acquired, implicit motor knowledge that underlies articulation in adults.

Since Broca's time, the neuro-anatomic basis and the functional characterization of this condition have been discussed extensively—probably more than any other neurobehavioural dysfunction. Over decades, evidence has accumulated that lesions to left posterior inferior frontal cortex including the opercular part of Broca's area (area 44) and the adjacent pre-motor and motor cortex, as well as subjacent anterior insular cortex, are responsible for the development of severe and often persistent articulation impairment (Hillis et al., 2004; Richardson, Fillmore, Rorden, LaPointe, & Fridriksson, 2012). Moreover, numerous imaging studies of non-brain-damaged speakers have identified this region as a higher-order speech motor centre (Riecker, Brendel, Ziegler, Erb, & Ackermann, 2008; Shuster & Lemieux, 2005). Historically, the dysfunction resulting from lesions to this cortical area has been characterized in varying terms as 'apraxia of the language muscles,' 'phonetic disintegration of speech,' a 'programming deficit,' a breakdown of the 'functional coalitions' of articulation, or, more recently, a disorder of 'phonetic planning,' or 'speech motor planning' (for a historical review, see Ziegler et al., 2012).

Implied in this thinking is that the pathomechanism of AOS disrupts the language-specific motor patterns for the production of syllables, words, and sentences that are acquired during speech development. A recently proposed model delineates how articulation learning in childhood is mediated by subcortical structures known to be implicated in motor learning in general, i.e., the basal ganglia and the cerebellum, which results in an accumulation of a ‘knowledge base for articulation’ in the left frontal cortex of the adult brain (Ziegler & Ackermann, 2017). Thus, neural plasticity mechanisms transform this cortical region into an area that is highly specialized for the motor act of speaking (as distinct from nonspeech vocal tract movements; Bonilha, Moser, Rorden, Baylis, & Fridriksson, 2006), and, in particular, for the production of the articulatory patterns that are typical of a speaker’s native language. A plausible assumption, based on knowledge about the experience-specific plasticity of the human brain (e.g., Adkins, Boychuk, Remple, & Kleim, 2006), is that speech motor patterns that are more strongly integrated through speech learning are represented more redundantly within the functional network of this brain region. By implication, they are less vulnerable to a loss of cortical tissue in this area. Conversely, articulatory patterns that are less typical of the speaker’s native language, or less cohesive, have less redundant neural representations and are therefore more vulnerable to cortical damage (Ziegler, Lehner, Pfab, & Aichert, 2021).

### 1.2.3. Taxonomy

An important issue in the classification of this disorder is to differentiate it from other types of neurogenic sound production impairment. On the one hand, the symptom pattern of AOS is not explainable by ‘elementary’ motor pathologies of the vocal tract muscles, such as paresis, ataxia, hypo- or hyperkinesia etc., which are summarized under the clinical term *dysarthria*. The speech errors illustrated in (1) and (2) above can hardly be explained by dysarthric mechanisms like muscle weakness, poverty or slowness of movements, articulatory over- or undershoot, or dyscoordination, as these pathomechanisms create more predictable and consistent patterns. Dysarthric patients would also not produce well-articulated phoneme substitutions, as, e.g., in the third rendition of the word *Kapitän* in (1), and completely symptom-free productions, as in (2), are very unlikely in dysarthria. Moreover, unlike the dysarthrias, which are typically caused by bilateral brain lesions, AOS is a syndrome of the language-dominant hemisphere. The label ‘apraxia’ historically relates to exactly these circumstances and characterizes the disorder as an impairment of ‘higher,’ left-hemisphere motor functions.

On the other hand, the syndrome is different from a condition termed *aphasic-phonological impairment*, which is characterized by phoneme errors in essentially fluent speech, without any apparent signs of articulomotor involvement (Romani, Olson, Semenza, & Granà, 2002). The examples listed in (1) above are not compatible with this condition, because they are severely

disfluent and contain numerous sound distortions, i.e., articulations not explainable by omissions, substitutions, or additions of phonemes. For example, extensive plosive aspirations (e.g., /t/ → [t<sup>hh</sup>]) or sound transition errors (e.g., /ka/ → [k'a]) are not explainable by purely phonological mechanisms. Several hypotheses have been put forward to explain phonological paraphasia. The most prominent one describes it as a disconnection of auditory-motor integration mechanisms caused by lesions to the long fibre tracts connecting superior temporal (= auditory) with inferior frontal (= higher motor) cortex, or to parieto-temporal cortex in the left hemisphere (e.g., Buchsbaum et al., 2011; Dick & Tremblay, 2012). In 'pure' cases of phonological impairment, the motor planning area of left inferior frontal cortex is spared (Schwartz, Faseyitan, Kim, & Coslett, 2012).

AOS is also different from acquired neurogenic stuttering, at both the symptom and the neuropathological level. Interestingly, speech characteristics resembling developmental stuttering—a condition with an extremely high prevalence rate in children (Craig, Hancock, Tran, Craig, & Peters, 2002)—are rare in adult neurologic patients. Furthermore, due to their low incidence it is unclear to what extent they are comparable with the symptoms of developmental stuttering and if they constitute a homogeneous syndrome at all. They may occur in different etiologies and after lesions to widely differing brain structures of both hemispheres, such as the corpus callosum, supplementary motor area, basal ganglia, thalamus, cerebellum, midbrain, and brainstem, and their underlying neuropathology is still undetermined (Lundgren, Helm-Estabrooks, & Klein, 2010). Apart from the unspecific feature of disfluency there is nothing that makes acquired neurogenic stuttering comparable with AOS.

Though these differential diagnostic considerations have been disputed over more than a century, AOS is widely accepted as a separate clinical unit, and an enormous amount of research has been devoted to its clinical pattern and its neuro-anatomic substrate. For a discussion of the characteristics of apraxic speech errors and problems of differential diagnosis see Ziegler et al. (2012), Ziegler (2016), and Duffy (2020).

#### **1.2.4. Conclusion**

As a conclusion, the speech patterns of patients with AOS provide insight into the structure of the acquired implicit knowledge about how the words and sentences of a language are articulated. The argument is similar to that brought up in studies of speech errors in healthy speakers (e.g., Pouplier & Hardcastle, 2005), but with the difference that a much larger corpus of data can be acquired from patients with AOS under relatively natural speaking conditions, and that the errors made by patients with AOS can be allocated to a rather circumscribed functional and neuro-anatomic source, i.e., a dysfunction of the acquired articulation planning processes located in the left posterior inferior frontal lobe.

## 2. The nonlinear gestural (NLG) model of word articulation in apraxia of speech

### 2.1. Accuracy of apraxic word production as a yardstick of ‘articulatory ease’

For the above reasons, analyses of the errors made by patients with AOS can reveal what is easy or difficult to articulate for adult native speakers of a given language. In phonetics, ‘ease of articulation’ is a disputed concept, at least for adult speech, because due to the highly overlearned nature of speaking everything is equally easy for adults to say in their native language (Ladefoged, 1990). Small differences that may exist between the articulatory demands of different phonological patterns of our language are masked by the ceiling effect of a highly overlearned speech motor skill. On this account, AOS can act like a magnifying glass and reveal differences in the cohesion of speech motor components that are not visible in typical speech.

In the present context, a word is considered easy to pronounce if patients with apraxia of speech have relatively few problems producing it. More specifically, patients with only mild impairment will most often produce it correctly, and only those with severe AOS will make errors on it. Conversely, difficult phonetic patterns are those that provoke errors even in patients with mild AOS. Considering the examples (1) and (2) above, the words *Dach*, *Schaf*, and *Tomate* listed in (2) appear to be easier than the word *Kapitän* in (1), with the caveat that individual observations do not necessarily reflect the general pattern. Under the transparency assumption mentioned above, the relative susceptibility of words to apraxic failure is considered to reflect the redundancy of the representation of its motor components in left inferior frontal cortex, which, in turn, is considered the neural substrate of the degree to which language-specific articulatory patterns have been settled through speech motor learning during childhood.

### 2.2. Factors that influence apraxic speech errors

In a large number of clinical studies, factors that modulate the probability of a word to elicit a speech error in AOS have been found at different phonological levels.

In most of the earlier studies, apraxic error rates were related to the phoneme level, mostly with the finding that consonants are less error prone than vowels, plosives and nasals less than fricatives or affricates, voiceless obstruents less than voiced obstruents, etc. (e.g., Bislick & Hula, 2019; Romani, Galuzzi, Guariglia, & Goslin, 2017). Other findings were related to syllable structure, e.g., that coda consonant errors are less frequent than onset errors, or simplex syllables are less vulnerable than complex syllables (e.g., Aichert & Ziegler, 2004; Staiger & Ziegler, 2008). Finally, at the supra-syllabic level a rather common finding was that the likelihood of an error increases with the number of syllables in a word (e.g., Bislick & Hula, 2019). More recently, we could also demonstrate an effect of lexical stress, showing that in German trochaic words were ‘easier’ to pronounce than iambic words (Aichert, Späth, & Ziegler, 2016). Most

notably, the iambic patterns led to greater numbers of segmental errors, which pointed at an interlacing of segmental with supra-segmental levels of speech motor planning. This effect has meanwhile been replicated for American English (Bailey, Bunker, Mauszycki, & Wambaugh, 2019). By inference, although most apraxic speech errors surface as segment-bound symptoms, as in the examples (1) above, the source of these errors does not necessarily have to be assigned to the affected phonemes themselves, but can also lie at some higher, e.g., syllabic or metrical level of articulation planning. Though experimental evidence for such links is relatively new, the interactions between the prosodic and the segmental layers of articulation have long been exploited by speech language therapists in the treatment of AOS (for an overview see Ziegler, Aichert, & Staiger, 2010; Aichert, Lehner, Falk, Späth, Franke, & Ziegler, 2021). An overview of the findings summarized above is listed in **Table 1** of Ziegler et al. (2021).

### 2.3. Design of the NLG model

Considering that the hierarchy of articulatory planning requirements sketched above extends across all levels of the phonological architecture of words, from segments to metrical structures, an approach that integrates all these levels is needed to account for the possible interactions between them. The Nonlinear Gestural (NLG) model of AOS, which is based on concepts borrowed from Articulatory Phonology (e.g., Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007), considers articulatory gestures as the basic elements of articulation planning and postulates that these units are successively integrated to form syllable structure components, syllables, metrical feet, and phonological words. The model has been described in more detail elsewhere (Ziegler, 2011; Ziegler & Aichert, 2015; Ziegler, Aichert, & Staiger, 2017; Ziegler et al., 2021).

In brief, the NLG model postulates that in a speaker with AOS of a given degree of severity, the probability that an articulatory gesture is selected and activated accurately assumes a certain value  $p \in ]0,1[$ .<sup>2</sup> Accurate gesture activation involves appropriate specification of all gestural parameters, i.e., in the NLG framework, the constriction organ, location, and degree. The more severe the disorder, the more likely is it that the proper activation of a gesture fails, hence the smaller is  $p$ .

In the building of articulatory plans, gestures are coordinated to form larger units (Goldstein & Fowler, 2003; Tilsen, 2016). By combinatorial laws, the probability that a combination of two gestures is correct drops to  $p^2$ , assuming that each one is selected accurately with the same probability  $p$ . However, this combinatorial calculus is only valid under the unlikely assumption

---

<sup>2</sup> As there is no a-priori reason to assume that gestures of the lips, tongue, velum, or glottis have different probabilities to be selected and activated correctly, the value of  $p$  is assumed to be the same for all constriction organs in a given patient. The different vulnerabilities of gestures arise exclusively from their constriction types and their locations within the gestural score of a word, both of which are specified by model coefficients (see Figure 1).

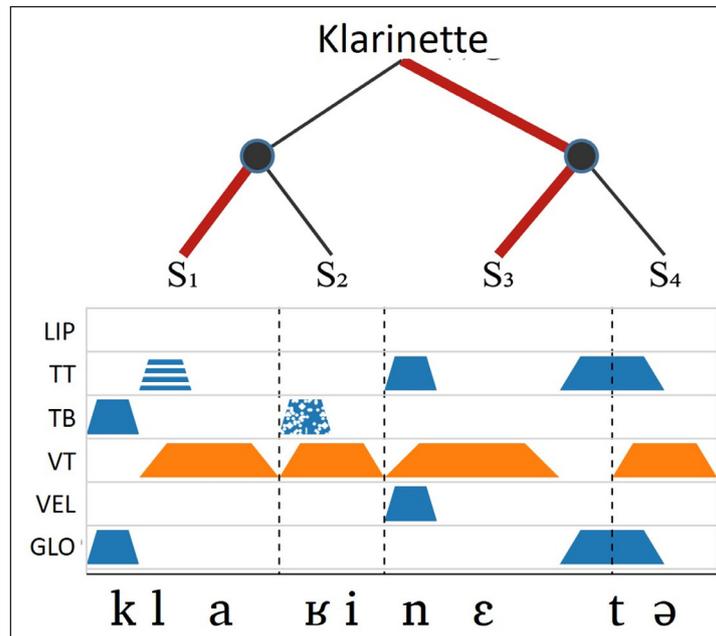
that the two gestures are activated independent of each other. More realistically, the likelihood that a combination of two gestures is accurate is expressed as  $p^2 * c$ , with an unknown coefficient  $c$  expressing whether the probability of accurate production increases ( $c > 1$ ) or decreases ( $c < 1$ ). An increase would point at some inherent bonding of the two gestures, either by biomechanical conditions or through learned gestural integration, which makes them less vulnerable because they do not count as two independent occasions to fail. This is what one would expect for healthy adult speakers, assuming, for instance, Tilsen’s hypothesis that speech acquisition converges to a ‘coordinative mode’ of gestural organization (Tilsen, 2016). Conversely, a decrease would suggest that putting the two gestures together requires additional effort, which makes them more vulnerable than predicted from purely combinatorial considerations. This ‘competitive mode’ of gestural organization (Tilsen, 2016) is considered typical for early speech acquisition, and under the assumptions of the NLG model it is a core mechanism underlying the condition of AOS. Hence, the coefficients in the model represent the strength of the glue between the gestural components of syllables and words. Computationally, different coefficients are introduced to represent the specific degree of integration of a gesture within the word’s gestural architecture, depending on its structural relationship with other gestures in the word.

**Figure 1** illustrates, as an example, the gestural score of the four-syllabic German word *Klarinette* (English “clarinet”), with primary stress on syllable 3 and secondary stress on syllable 1. The figure shows a variety of inter-gestural relationships implemented in the NLG-model, such as synchrony of a labial or lingual gesture with a glottal or velar aperture gesture (e.g., in /k/ or /n/, respectively), consonantal gestures in pre- or post-vocalic positions, gestures clustered within a pre- or post-vocalic position, gestures in the tail of a metrical foot (/bi/, /tə/) as opposed to those in the head (/kla/, /net/), and those within a left- or right-branching non-prominent metrical foot, as in (left-branching) /klaʁi/. Computationally, each word is modelled by a nonlinear, multiplicative term representing each gesture’s base probability of correct production and the coefficients coding for the gesture’s bonding strengths at different locations within the word’s gestural score (**Figure 1**). As an example, the term for the word *Klarinette* is:

$$(3) \quad p_{\text{klarinette}} = p^{12} * c_{\text{on}}^7 * c_{\text{cd}}^1 * c_{\text{glo}}^2 * c_{\text{vel}}^1 * c_{\text{clus}}^3 * c_{\text{crit}}^1 * c_{\text{nar}}^1 * c_{\text{alt}}^3 * c_{\text{tail}}^4 * c_{\text{L-branch}}^6,$$

with  $p$  denoting the base probability of correct gesture production and the  $c$ -coefficients specified in **Table 1**. The exponents indicate how often gestures of a particular type occur within the word’s gestural score.

Overall, the model is defined by twelve coefficients, i.e., the parameter  $p$  denoting the base probability of correct gesture production and eleven coefficients representing a gesture’s tract variables and position within the word’s gestural score, as explained in **Table 1**.



**Figure 1:** Gestural score representing the 4-syllabic German word *Klarinette* as generated by a web-application available at <https://neurophonetik.de/gesten-koeffizientenrechner>. Each articulatory gesture is represented by a trapezoid. Consonant gestures are arranged on different layers depending on the involved constriction organ (LIP: lips, TT: tongue tip, TB: tongue back, VEL: velar aperture, GLO: glottal aperture). Closed constrictions are represented by filled trapezoids, other constriction types are represented by different fillings (e.g., ‘narrow’ in /l/, ‘critical’ in /ʀ/). For simplification, vowel gestures (in yellow) are allocated to a separate layer representing the gross vocal tract deformations of vowel articulation, including jaw, tongue body, and, if applicable, lip rounding movements. Dashed vertical lines indicate syllable boundaries, the tree structure characterizes the word’s stress pattern (bold lines indicating prominence).

## 2.4. Model evaluation

Model coefficient estimates were calculated in two nonlinear regression analyses using corpora of 96 and 136 words of different lengths and complexities, with 120 and 66 renditions per word, respectively. The words were elicited in word repetition experiments and spoken by native German speakers diagnosed with AOS after a left-hemisphere stroke,<sup>3</sup> with or without concomitant aphasia. The materials and patient samples were described in detail in Ziegler (2005, 2009) and Ziegler and Aichert (2015). Each spoken word was scored as accurate or inaccurate by independent, clinically experienced raters. A produced word was scored inaccurate when one or several of its segments were omitted, distorted, or substituted by other phonemes, when segments were added, or when the utterance was dysfluent due to pauses or phoneme lengthening. For reliability analyses concerning these ratings see Liepold, Ziegler, and Brendel (2003); Ziegler (2005, 2009). Because all attempts to fit the data using log-linear mixed models

<sup>3</sup> Note that the presence of AOS was determined independently from the materials fed into the NLG model. The diagnosis of AOS was made on the basis of all available clinical data and the patients’ error patterns in various clinical examinations.

failed to converge, modelling the patients as random effects was not possible. Therefore, each word of the training sample was represented by its average accuracy score and modelled by the nonlinear term describing the word's gestural score, as in equation (3) above. Depending on the cohort, the twelve model coefficients were fitted using 96 or 136 equations, respectively. A cross-validation was reported in Ziegler and Aichert (2015).

## 2.5. Model shape

The sizes of the model coefficients indicate where a particularly strong or weak cohesion exists within the gestural patterns of phonological words. The shape of the NLG model is defined by the model coefficients listed in **Table 1**. It largely corroborates earlier findings about the factors influencing accuracy of apraxic speech, but unlike earlier investigations of isolated factors, the model coefficients represent an integrative view and take account of the interactions between the hierarchical levels of the model.

layer	coefficient	description	estimate
Constriction organ <sup>1</sup>	$C_{alt}$	alternation between LIP, TT, TB within phonological word	<b>.935*</b>
	$C_{vel}$	velic aperture (synchronous to LIP, TT, TB)	1.177*
	$C_{glo}$	glottal aperture (synchronous to LIP, TT, TB)	1.189*
Constriction type <sup>2</sup>	$C_{crit}$	fricative gestures of LIP, TT, TB. GLO	<b>.896*</b>
	$C_{nar}$	liquid/approximant gestures of LIP, TT, TB	1.025
Syllable level <sup>3</sup>	$C_{on}$	pre-vocalic gestures of LIP, TT, TB, VEL, GLO	<b>.899*</b>
	$C_{cd}$	post-vocalic gestures of LIP, TT, TB, VEL, GLO	.945
	$C_{clus}$	multiple gestures in pre- or post-vocalic position	1.006
Foot level <sup>4</sup>	$C_{tail}$	gestures in unstressed syllable(s) of metrical foot	1.034*
Phonological word level <sup>5</sup>	$C_{R-branch}$	gestures in non-prominent foot, right branching	1.006
	$C_{L-branch}$	gestures in upbeat or non-prominent, L-branching foot	<b>.965*</b>

**Table 1:** Empirically derived NLG coefficients. Coefficients significantly greater than 1 (marked by asterisks) indicate an inherent cohesion of articulatory gestures with their associated gestures. Coefficients that are statistically not different from 1 (not marked) indicate a dissolution of the cohesive ties between associated gestures. Coefficients significantly lower than 1 (marked by bold face and asterisks) suggest that even an extra effort is required to activate the corresponding gesture in its particular context.

<sup>1</sup>default: LIP/TT/TB <sup>2</sup>default: close, open <sup>3</sup>default: nucleus <sup>4</sup>default: head <sup>5</sup>default: prominent foot.

Only three out of 11 NLG coefficients assumed values significantly greater than 1 (**Table 1**, rightmost columns, italics), namely the two coefficients specifying the coordination of velar and glottal aperture gestures synchronous to the gestures of the lips and the tongue, and the coefficient specifying gestures within the unstressed syllable(s) of a metrical foot. These results suggest that the cohesion of gestures pertaining to the same segment (a nasal or a voiceless consonant, respectively) is relatively retained in AOS speakers, and that expansions of prominent syllables by unstressed syllables within a metrical foot are relatively inexpensive. This latter result conforms to earlier observations that in German monosyllabic words are not substantially easier than (disyllabic) trochaic words (Ziegler, 2011).

Four coefficients suggested an extra-effort of gesture selection or coordination by assuming values significantly below 1 (cf. **Table 1**, rightmost column, bold): First, in cases where alternating constriction organs have to be selected within a phonological word (as in moving from /k/ to /l/, from /l/ to /ʃ/, and from /ʃ/ to /n/ in Klarinette, **Figure 1**), the activation of an upcoming tongue or lip gesture has to overcome a tendency to re-activate the articulator that was activated shortly before, which increases the likelihood that an error occurs.

Second, the presence of gestures with a ‘critical’ constriction degree, i.e., fricative gestures, also contributes to a disproportionate decrease of the probability of correct word production. This result is consistent with the outcomes of many earlier studies, according to which fricatives and affricates are particularly sensitive to apraxic pathology (cf. Bislick & Hula, 2019, for an overview). One explanation could be that due to the highly nonlinear articulatory-to-acoustic relationship of fricative production, any motor inaccuracy, regardless of its origin, can easily lead to a perceptible distortion.

Third, the coordination of a pre-vocalic consonantal gesture with the corresponding vowel gesture turned out as a source of increased vulnerability. Notably, the onset coefficient was substantially smaller than the coda coefficient, indicating that the cohesion of consonant and vowel gestures is stronger in the post-vocalic than in the pre-vocalic position. This may be considered to support the rime cohesion hypothesis (Fudge, 1987), although only on a relative basis, because the coda coefficient was  $< 1$  as well, suggesting that AOS leads to a loss of cohesive ties also within rime structures. A remarkable finding on syllable structure effects was that two gestures involved in a cluster count as independent in AOS, i.e., they are neither integrated into a cohesive unit, nor are they penalized specifically for sharing the same onset or coda slot.

Fourth, left-branching extensions of a prominent metrical foot, either by an extra-metrical syllable (‘upbeat’) or by a non-prominent foot (as in the example of **Figure 1**), has turned out to cause extra planning effort and thereby reduce the probability of accurate word production beyond a mere combinatorial effect. A special case is the finding that German words with an iambic pattern are substantially more error prone than trochaic words (Aichert et al., 2016). In contrast,

right-branching extensions by non-prominent metrical feet, as they occur in many compounds in German (e.g., *Klapperschlange*, English “rattle snake”), simply count as an independent source of apraxic speech error, with a coefficient not differing from 1 (**Table 1**).

It is worth noting, as a caveat, that a transfer of these results to the motor organization of typical, neurologically healthy speech is only possible on a relative scale. For example, the finding that the coefficient representing the cohesive ties of two cluster gestures is close to 1 suggests an independent recruitment and vulnerability of both gestures, i.e., no integration, in patients with AOS. This does not preclude that gestures in consonant clusters are organized as cohesive units in healthy speakers (Hoole & Pouplier, 2015). However, the data suggest, for example, that the ties between a tongue back closure and glottal aperture gesture in the production of /k/, with a coefficient considerably greater than 1, are stronger than those between the same tongue back gesture and a tongue tip closure in /kn/.

### 3. Non-clinical applications

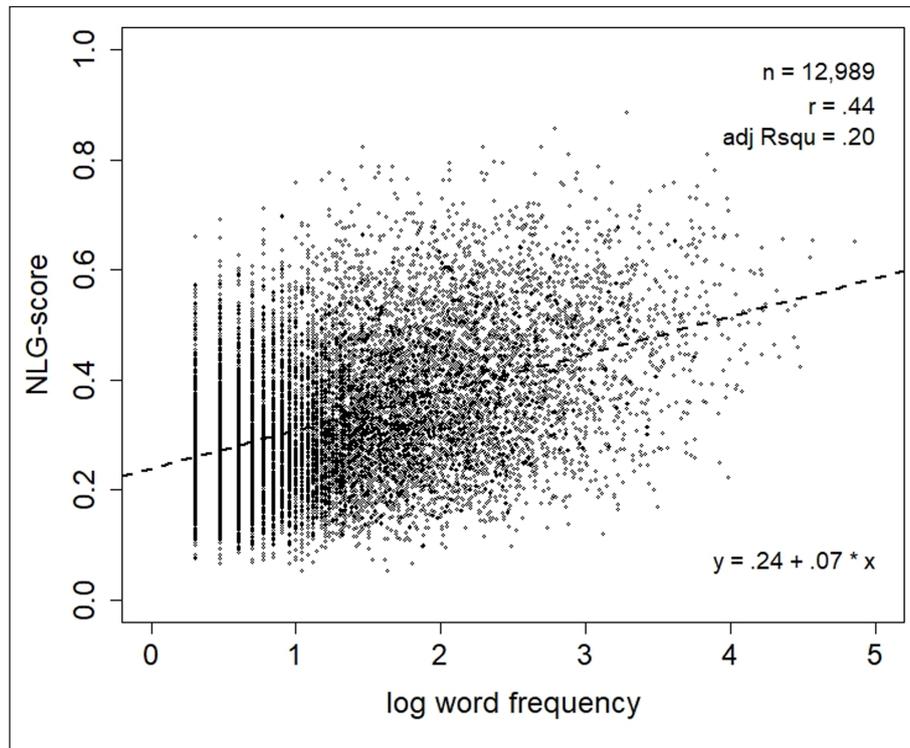
Beyond a variety of clinical applications of the NLG model (Ziegler et al., 2021), this research offers options for investigations of typical speech in adults or in children. The following two sections present some preliminary ideas on how the NLG model can stimulate investigations of non-clinical research questions in phonology/phonetics.

As a tool for such research, a web application was created that allows users to calculate an NLG score for each one- to four-syllabic word or non-word that meets German phonotactic constraints (<https://neurophonetik.de/sprechapraxie-gestenmodell>).

#### 3.1. NLG scores and lexical variables

This tool was used to compute NLG scores for large lexical databases, e.g., SUBTLEX-DE (free download at <http://crr.ugent.be/SUBTLEX-DE>), CELEX (Aichert, Marquardt, & Ziegler, 2005), or the CLEARPOND database (Marian, Bartolotti, Chabal, & Shook, 2012; free download at <https://clearpond.northwestern.edu>), with the aim of investigating the relationships between articulatory ease on the one hand and diverse phonological and lexical parameters on the other (e.g., Lehner & Ziegler, 2021).

As a first example presented here, the relationship between the frequency of words and their ease of articulation in terms of their NLG scores was examined. **Figure 2** plots the NLG score of 12,989 German words listed in *subtlex-np* (<https://neurophonetik.de/subtlex-np>) against log word frequency. Despite a huge variance, there is a clear and significant relationship between the two variables, as revealed by a linear regression analysis, with an average increase in the NLG score of .07 per log frequency unit (standard error: .001).

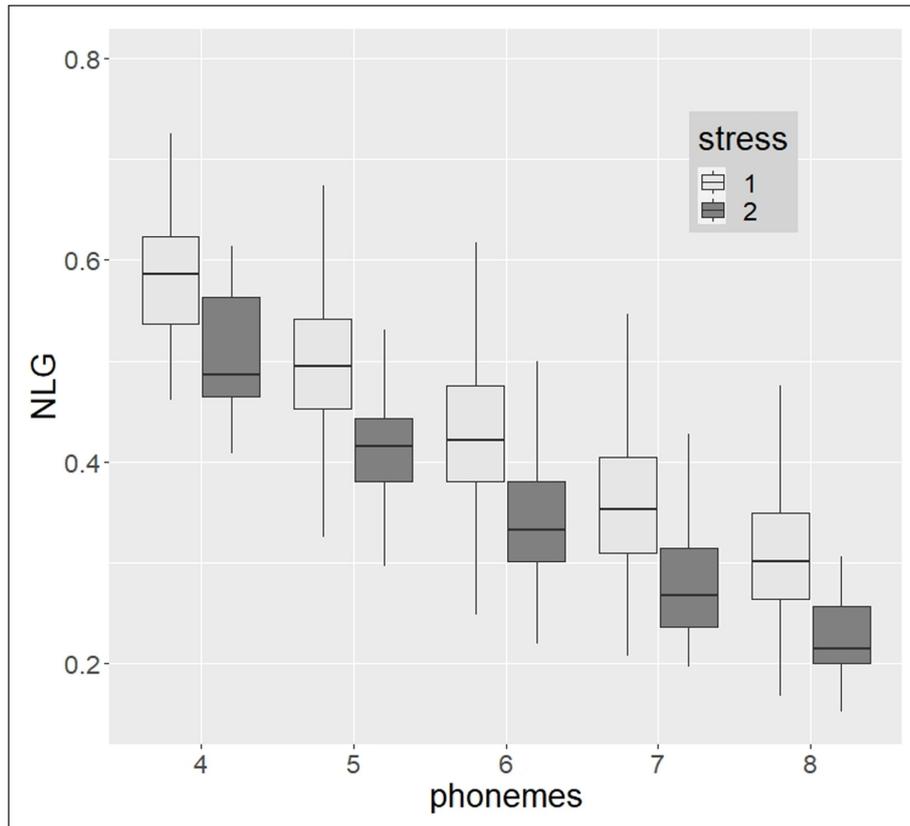


**Figure 2:** Scatterplot relating the NLG scores of 12,989 one- to three-syllabic words to their lexical frequencies. Data are provided at <https://neurophonetik.de/subtlex-np>. The striation in the low frequency range of the diagram is due to the coarse resolution of the log-scale in the lower range of the frequency-per-million data. The regression model is specified in the lower right corner.  $r$ : correlation coefficient (Pearson);  $\text{adj } R^2$ : adjusted  $R^2$ .

Considering Zipf’s observation that “the length of a word tends to bear an inverse relationship to its relative frequency” (Zipf, 1935, p. 38), a partial correlation (Spearman) was calculated between NLG- and frequency values, controlling for the confounding variable word length (in syllables). The correlation coefficient dropped from a regular  $\rho$  of .44 (95% CI: [.44, .46]) to a partial  $\rho$  of .29, but was still significantly different from 0 ( $p < .001$ ). Hence, there is more structural information than only word length that is reflected in the lexical frequencies of words.

As a second example, the hypothesis that disyllabic words with iambic stress patterns have lower NLG scores than trochees (Aichert et al., 2016) was tested on a larger scale, using again the German SUBTLX database (<https://neurophonetik.de/subtlex-np>). **Figure 3** depicts the NLG scores of 4,221 disyllabic words of 4 to 8 phonemes length in this corpus, 3,695 of them stressed on the first and 526 on the second syllable. Broken down by phoneme number, the boxplot shows that there was a clear difference between the two types of lexical stress across all phoneme numbers, with trochaic words being uniformly ‘easier’ (in terms of higher NLG scores) than iambic words. A linear regression model revealed significant main effects of both factors, with

an estimated decrease of .091 (standard error: .010) in iambic versus trochaic words ( $t = -9.6$ ,  $p < .001$ ) and of .068 (standard error: .001) for each additional phoneme ( $t = -76.0$ ,  $p < .001$ ). There was no significant interaction ( $t = .3$ ,  $p = .79$ ).



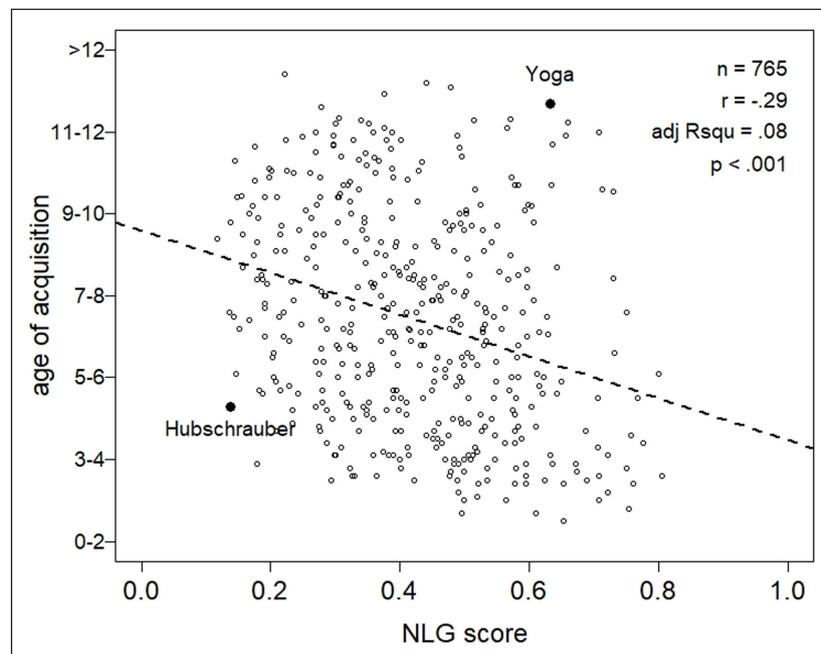
**Figure 3:** NLG scores of 4,221 disyllabic German words from the subtlex-np database (<https://neurophonetik.de/subtlex-np>) as a function of lexical stress and phoneme number. Light grey: first syllable stressed; dark grey: second syllable stressed.

This result may seem trivial, considering that in the computation of NLG scores the gestures of the unstressed syllables in the iambic words are penalized by a coefficient of .965, whereas the unstressed syllables in the trochees get a bonus of 1.34 (see **Table 1**). However, the outcome is notable nonetheless, because lexical stress obviously dominated all other influencing factors, and the difference persisted uniformly over all word lengths.

### 3.2. NLG scores and acquisition data

Since the inter-gestural bonding strengths reflected by the coefficients of the NLG model are considered to be shaped through speech acquisition, traces of the structural architecture of this model should be found in acquisition data.

To start with, **Figure 4** plots adults' retrospective estimates of the age of acquisition of 765 German words of 1 to 4 syllables length (Schröder, Gemballa, Ruppig, & Wartenburger, 2012) against their NLG scores. The ordinate specifies the response categories (age in years) in which the 60 participants indicated the age “when they thought they had learned the words.” Not surprisingly, there was only a weak (though still significant) association between the two variables. This is so because children name the objects that attract their attention, without considering the complexity of their names. As an example, the word *Hubschrauber* (English “helicopter”), which is among the most complex words in the corpus reported by Schröder et al. (2012), was judged to be acquired rather early, i.e., between 3 and 5, whereas a less complex and probably more accessible word like *Yoga* was assessed to be acquired much later (see **Figure 4**).

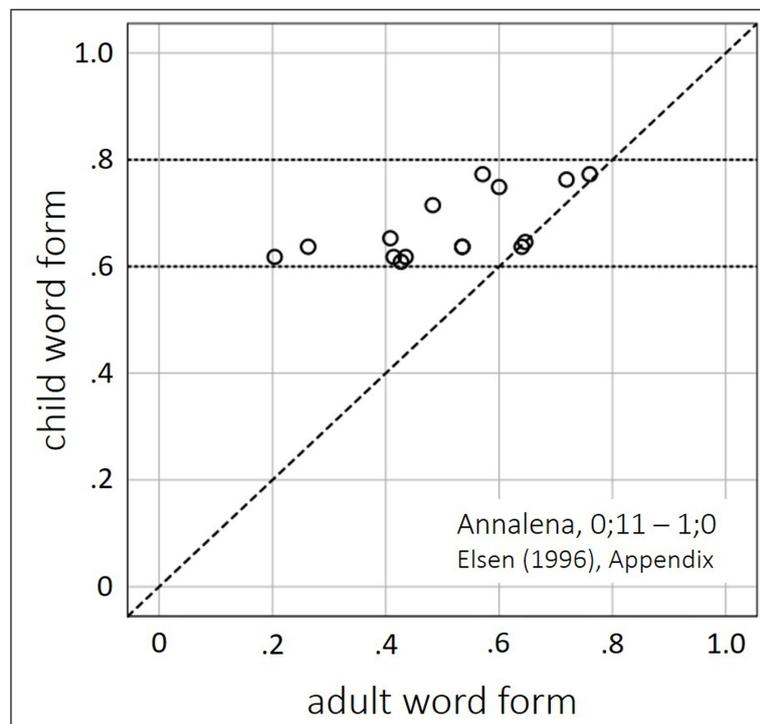


**Figure 4:** Age of acquisition (retrospective ratings by 60 participants) versus NLG scores for 765 one- to four-syllabic German words (from Schröder et al., 2012; with special thanks to Astrid Schröder for kindly providing the data). For abbreviations see Figure 3.

Most obviously, ‘acquisition’ in the usage of Schröder et al. (2012) has more to do with the acquisition of concepts and word meanings than with the ability to pronounce the words accurately. That is, when a young child actively uses the word *Hubschrauber*, she/he will probably not produce the adult form, but adjust it to a phonological pattern that is more accessible to her/his speech motor capabilities. M. Vihman hypothesized that young children draw on well-practiced motor routines when they acquire new words through an ‘articulatory filter’ (e.g., Vihman, 1993): They recognize similarities of new input with their already existing repertoire and use these similarities as phonetic templates for their own productions. This leads to a filtering

of the complexity of adult word forms down to a lower complexity level of the corresponding child forms, i.e., *Hubschrauber* may become something like [‘hʊ.haʊ.bɐ] (Simon, 16 months; personal communication by Ingrid Aichert). In this particular example, the NLG score of around .2 for the adult form rises to a score higher than .5 in Simon’s form.

As an illustration, **Figure 5** depicts the NLG scores of child versus adult forms from the vocabulary of a German girl (Annalena, 11–12 months) taken from Elsen (1996). The data show that Annalena uses words across a wide complexity range of between circa 0.2 and almost 0.8, but the forms she produces range within a small corridor of between 0.6 and 0.8. In this corridor, words like for instance *Zahnbürste* (English “toothbrush”) and *Bauch* (English “belly”) converge upon two structurally similar templates, i.e., [‘na.na] and [‘ba.ɓa]. Interestingly, the filtering is not completely uniform, but is sensitive to the complexity of the adult form, as can be recognized from the positive correlation between the adult- and child NLG scores ( $r = .64, p = .01$ ). Hence, in this example, like in several other child corpora, the NLG score provides a descriptive way of quantifying the processes hypothesized in Vihman’s articulatory filter theory.



**Figure 5:** Display of the NLG scores of 15 nouns and verbs from the vocabulary of one-year old Annalena (Elsen, 1996). Each circle represents a word. The fact that all values are on or above the diagonal (dashed line) indicates that the complexity of the adult forms is preserved at best, but is more often reduced in the child form. The child forms are within a rather narrow range of complexities, yet still with a positive correlation between the adult and the child’s word forms ( $r = .64, p = .01$ ; regression line not shown). A smaller section of the Annalena corpus was reported and discussed in Vihman, DePaolis, and Keren-Portnoy (2014).

## 4. Summary and discussion

This article exemplifies how data obtained from a clinical model of neurologic speech impairment may stimulate non-clinical experimental speech research.

The first section was devoted to the controversial question of whether at all and under which conditions it is possible to infer knowledge about normal cognitive, linguistic, or motor functions from observations of neurologically impaired functions (Caramazza, 1986). Apraxia of speech was introduced as a clinical model which is suitable for such an approach for several reasons: (i) There are almost 160 years of clinical research into this disorder, concomitant with a productive debate about its putative underlying pathomechanisms (for an overview see Ziegler et al., 2012); (ii) during this period, reliable knowledge of its neuroanatomical substrate in the left frontal lobe has been acquired (e.g., Richardson et al., 2012), and (iii) the contribution of this brain region to unimpaired speech has been verified in numerous imaging studies (Papoutsis et al., 2009).

Furthermore, the importance of AOS as a clinical model of articulation planning was substantiated by the notion that fluent articulation requires implicit knowledge of the specific patterning of articulatory gestures through which the sound patterns of a speaker's language are conveyed. This knowledge, also described as *phonetic planning* (Levelt, Roelofs, & Meyer, 1999) or *speech motor planning* competence, is necessarily language specific. It is constituted by extensive motor learning during speech acquisition in childhood and gets stored in the left inferior frontal cortex (Ziegler & Ackermann, 2017). Dysfunctions of this brain region may cause the speech error patterns typical of AOS. It therefore seems justified to regard the characteristics of speech apraxia as a window into the structure of articulation planning.

The second part of this article reviewed an approach in which accuracy data from word production samples of apraxic speech were used to build a non-linear gestural model of the architecture of articulatory plans for words. The NLG model describes phonological words as hierarchically organized patterns of articulatory gestures, similar to the Articulatory Phonology framework (e.g., Goldstein & Fowler, 2003), and postulates that apraxia of speech affects the ability of speakers to correctly select, activate, and coordinate the gestures for the production of words (cf., Tilsen, 2014). The likelihood of a word to be produced accurately by speakers with AOS is modelled by a hierarchical combination of the probabilities of accurate activation of each single gesture, corrected by coefficients representing the binding forces between gestural elements at each level of the word's gestural score. For the training of the model, probabilities of correct articulation were estimated for a large number of German one- to four-syllabic words by calculating rates of accurate word production for a large group of patients with AOS, and a nonlinear regression model was used to estimate the model coefficients. These coefficients permit calculation of an 'ease of articulation' score, the NLG score, for each word or nonword that satisfies German phonotactic rules. In short, the NLG score is a one-dimensional measure of articulatory

complexity that integrates complicating and facilitating factors of articulation planning across all levels of phonological word structure, from articulatory gestures to combinations of metrical feet, and is validated by accuracy scores of patients with AOS.

In the final section of this article I presented an exemplary selection of still unpublished applications of this parameter in non-clinical speech research. As a first field of application, expansions of lexical data bases by NLG scores were mentioned, with the objective to analyze large word corpora for relationships between the articulatory demands of words and other lexical or phonological factors, e.g., lexical familiarity, neighbourhood density, or lexical frequency. In a second example it was proposed that the NLG model can be applied to speech corpora of children, to answer questions regarding the role of articulatory demands in infant word learning. Other applications may be conceived, such as parametric analyses of the role of articulatory parameters in phonological variation and change, to mention only one example.

There are several limitations that should be addressed here. One is that the relationship of the NLG model to typical adult speech is still largely hypothetical. There is only one neuroimaging study so far that revealed that the BOLD activation in the left inferior frontal area of healthy participants during word production is modulated by the NLG scores of the produced words (Kellmeyer et al., 2010). More of such work is required to demonstrate that the NLG model outperforms parameters like word length or syllable complexity on the prediction of neural activity in the cortical regions that are known to be relevant for speech motor planning. Furthermore, in the absence of sufficiently large speech error corpora from normal speakers, other behavioural measures, e.g., vocal reaction times, should be examined to validate the NLG model for unimpaired speakers.

There are other limitations and caveats. For example, it still remains to be investigated if the NLG model is specific for motor planning impairment, or if it also predicts the speech error patterns of patients with other speech impairments. Of particular interest in this regard are patients with an impairment described as *phonological paraphasia* after more posterior perisylvian lesions. If their error patterns were to prove structurally homomorphic to those of apraxic speakers, this would shed new light on the relationship between the motor planning and the auditory-motor integration processes that are associated with different cortical regions along the pathway between auditory and motor cortical areas (Ziegler et al., 2021). A further limitation, on yet another level, lies in the fact that fine phonetic variation cannot be represented in the NLG model. Finally, as the NLG model has been developed at the level of single word utterances, expansions to higher metrical levels would be desirable.

---

## Ethics and consent

The research reviewed here was performed in accordance with the Declaration of Helsinki.

## Acknowledgements

This article was drafted from a keynote speech given at the *3<sup>rd</sup> Phonetics and Phonology in Europe (PaPE) Conference 2019*. I am grateful to Ingrid Aichert and Anja Staiger for their long-standing valuable contributions to the work reviewed here. Jakob Pfab developed the NLG web app, Katharina Lehner prepared the SUBTLEX-np data base, and Astrid Schröder thankfully supplied her age of acquisition data. I would also like to thank the many patients who have agreed to participate in our research.

## Funding information

The research reviewed here was supported by grants from the German Research Council (DFG, Zi 469 / 6-1/8-1/14-1,2).

## Competing interests

The author has no competing interests to declare.

---

## References

- Adkins, D. L., Boychuk, J., Remple, M. S., & Kleim, J. A. (2006). Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. *Journal of Applied Physiology*, *101*(6), 1776–1782. DOI: <https://doi.org/10.1152/jappphysiol.00515.2006>
- Aichert, I., Lehner, K., Falk, S., Späth, M., Franke, M., & Ziegler, W. (2021). In time with the beat: Entrainment in patients with phonological impairment, apraxia of speech, and Parkinson's disease. *Brain Sciences*, *11*, 1524. DOI: <https://doi.org/10.3390/brainsci11111524>
- Aichert, I., Marquardt, C., & Ziegler, W. (2005). Frequenzen sublexikalischer Einheiten des Deutschen: CELEX-basierte Datenbanken. *Neurolinguistik*, *19*, 55–81.
- Aichert, I., Späth, M., & Ziegler, W. (2016). The role of metrical information in apraxia of speech. Perceptual and acoustic analyses of word stress. *Neuropsychologia*, *82*, 171–178. DOI: <https://doi.org/10.1016/j.neuropsychologia.2016.01.009>
- Aichert, I., & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, *88*, 148–159. DOI: [https://doi.org/10.1016/S0093-934X\(03\)00296-7](https://doi.org/10.1016/S0093-934X(03)00296-7)
- Bailey, D. J., Bunker, L., Mauszycki, S. C., & Wambaugh, J. L. (2019). Reliability and stability of the metrical stress effect on segmental production accuracy in persons with apraxia of speech. *International journal of language & communication disorders*. DOI: <https://doi.org/10.1111/1460-6984.12493>

- 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. DOI: [https://doi.org/10.1044/2019\\_AJSLP-18-0169](https://doi.org/10.1044/2019_AJSLP-18-0169)
- Bonilha, L., Moser, D., Rorden, C., Baylis, G. C., & Fridriksson, J. (2006). Speech apraxia without oral apraxia: Can normal brain function explain the physiopathology? *NeuroReport*, 17(10), 1027–1031. DOI: <https://doi.org/10.1097/01.wnr.0000223388.28834.50>
- Brendel, B., Synofzik, M., Ackermann, H., Lindig, T., Schölderle, T., Schöls, L., & Ziegler, W. (2015). Comparing speech characteristics in spinocerebellar ataxias type 3 and type 6 with Friedreich ataxia. *Journal of neurology*, 262(1), 21–26. DOI: <https://doi.org/10.1007/s00415-014-7511-8>
- Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé; suivies d'une observation d'aphémie (perte de la parole). *Bulletins et memoires de la Société Anatomique de Paris*, XXXVI, 330–357.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregate analysis of lesion and fMRI data. *Brain and Language*, 119(3), 119–128. DOI: <https://doi.org/10.1016/j.bandl.2010.12.001>
- Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition*, 5, 41–66. DOI: [https://doi.org/10.1016/0278-2626\(86\)90061-8](https://doi.org/10.1016/0278-2626(86)90061-8)
- Caramazza, A. (1988). Some aspects of language processing revealed through the analysis of acquired aphasia: The Lexical System. *Annual Reviews of Neuroscience*, 11, 395–421. DOI: <https://doi.org/10.1146/annurev.ne.11.030188.002143>
- Craig, A., Hancock, K., Tran, Y., Craig, M., & Peters, K. (2002). Epidemiology of stuttering in the community across the entire life span. *Journal of Speech, Language, and Hearing Research*, 45, 1097–1105. DOI: [https://doi.org/10.1044/1092-4388\(2002/088\)](https://doi.org/10.1044/1092-4388(2002/088))
- Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. *Brain*, 135(12), 3529–3550. DOI: <https://doi.org/10.1093/brain/aws222>
- Duffy, J. (2020). *Motor Speech Disorders. Substrates, Differential Diagnosis, and Management*. 4<sup>th</sup> ed. St. Louis: Elsevier.
- Duffy, J. R., Strand, E. A., & Josephs, K. A. (2014). Motor speech disorders associated with primary progressive aphasia. *Aphasiology*, 28(8–9), 1004–1017. DOI: <https://doi.org/10.1080/02687038.2013.869307>
- Elsen, H. (1996). Two routes to language: Stylistic variation in one child. *First Language*, 16(47), 141–158. DOI: <https://doi.org/10.1177/014272379601604701>
- Fudge, E. (1987). Branching structure within the syllable. *Journal of linguistics*, 23(2), 359–377. DOI: <https://doi.org/10.1017/S0022226700011312>

- Goldstein, L., & Fowler, C. (2003). Articulatory Phonology: A phonology for public language use. In A. Meyer & N. O. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* (pp. 159–207). Berlin: Mouton de Gruyter.
- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition*, *103*(3), 386–412. DOI: <https://doi.org/10.1016/j.cognition.2006.05.010>
- Hagedorn, C., Proctor, M., Goldstein, L., Wilson, S. M., 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*(4), 877–891. DOI: [https://doi.org/10.1044/2016\\_JSLHR-S-15-0112](https://doi.org/10.1044/2016_JSLHR-S-15-0112)
- Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain*, *127*, 1479–1487. DOI: <https://doi.org/10.1093/brain/awh172>
- Hoole, P., & Pouplier, M. (2015). Interarticulatory Coordination. Speech Sounds. In M. A. Redford (Ed.), *The Handbook of Speech Production* (pp. 133–157). Chichester: Wiley. DOI: <https://doi.org/10.1002/9781118584156.ch7>
- Hoole, P., Schröter-Morasch, H., & Ziegler, W. (1989). Disturbed laryngeal control in apraxia of speech. *Folia Phoniatrica*, *41*, 177–177.
- Kellmeyer, P., Ewert, S., Kaller, C., Kümmerer, D., Weiller, C., Ziegler, W., & Saur, D. (2010, 6/6/2010). *Neural networks for processing articulatory difficulty – a model for apraxia of speech*. Human Brain Mapping Conference, Barcelona. <https://www.aievolution.com/hbm1001/index.cfm>
- Ladefoged, P. (1990). Some reflections on the IPA. *Journal of Phonetics*, *18*, 335–346. DOI: [https://doi.org/10.1016/S0095-4470\(19\)30378-X](https://doi.org/10.1016/S0095-4470(19)30378-X)
- Lehner, K., & Ziegler, W. (2021). The impact of lexical and articulatory factors in the automatic selection of test materials for a web-based assessment of intelligibility in dysarthria. *Journal of Speech, Language, and Hearing Research*, *64*, 2196–2212. DOI: [https://doi.org/10.1044/2020\\_JSLHR-20-00267](https://doi.org/10.1044/2020_JSLHR-20-00267)
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*, 1–38. DOI: <https://doi.org/10.1017/S0140525X99001776>
- Liepold, M., Ziegler, W., & Brendel, B. (2003). Hierarchische Wortlisten. Ein Nachsprechtest für die Sprechapraxiediagnostik. Dortmund: Borgmann.
- Lundgren, K., Helm-Estabrooks, N., & Klein, R. (2010). Stuttering following acquired brain damage: A review of the literature. *Journal of neurolinguistics*, *23*(5), 447–454. DOI: <https://doi.org/10.1016/j.jneuroling.2009.08.008>
- Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS one*, *7*(8), e43230. DOI: <https://doi.org/10.1371/journal.pone.0043230>
- Papoutsis, M., de Zwart, J. A., Jansma, J. M., Pickering, M. J., Bednar, J. A., & Horwitz, B. (2009). From phonemes to articulatory codes: An fMRI study of the role of Broca's area in speech production. *Cerebral cortex*, *19*(9), 2156–2165. DOI: <https://doi.org/10.1093/cercor/bhn239>

- Pouplier, M., & Hardcastle, W. (2005). A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica*, 62(2–4), 227–243. DOI: <https://doi.org/10.1159/000090100>
- Richardson, J. D., Fillmore, P., Rorden, C., LaPointe, L. L., & Fridriksson, J. (2012). Re-establishing Broca's initial findings. *Brain and Language*, 123(2), 125–130. DOI: <https://doi.org/10.1016/j.bandl.2012.08.007>
- Riecker, A., Brendel, B., Ziegler, W., Erb, M., & Ackermann, H. (2008). The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain and Language*, 107(2), 102–113. DOI: <https://doi.org/10.1016/j.bandl.2008.01.008>
- Romani, C., Galuzzi, C., Guariglia, C., & Goslin, J. (2017). Comparing phoneme frequency, age of acquisition, and loss in aphasia: Implications for phonological universals. *Cognitive Neuropsychology*, 34, 449–471. DOI: <https://doi.org/10.1080/02643294.2017.1369942>
- Romani, C., Olson, A., Semenza, C., & Granà, A. (2002). Patterns of phonological errors as a function of a phonological versus an articulatory locus of impairment. *Cortex*, 38(4), 541–567. DOI: [https://doi.org/10.1016/S0010-9452\(08\)70022-4](https://doi.org/10.1016/S0010-9452(08)70022-4)
- Rusz, J., Bonnet, C., Klempíř, J., Tykalová, T., Baborová, E., Novotný, M., ... Růžička, E. (2015). Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Journal of neurology*, 262(4), 992–1001. DOI: <https://doi.org/10.1007/s00415-015-7671-1>
- Schiller, F. (1979). Paul Broca, founder of French anthropology, explorer of the brain: University of California Press. DOI: <https://doi.org/10.1525/9780520315945>
- Schröder, A., Gemballa, T., Ruppín, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior research methods*, 44(2), 380–394. DOI: <https://doi.org/10.3758/s13428-011-0164-y>
- Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, 135(12), 3799–3814. DOI: <https://doi.org/10.1093/brain/aws300>
- Shuster, L. I., & Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language*, 93(1), 20–31. DOI: <https://doi.org/10.1016/j.bandl.2004.07.007>
- Staiger, A., Schroeter, M. L., Ziegler, W., Schölderle, T., Anderl-Straub, S., ... Diehl-Schmid, J. (2021). Motor speech disorders in the nonfluent, semantic and logopenic variants of primary progressive aphasia. *Cortex*, 140, 66–79. DOI: <https://doi.org/10.1016/j.cortex.2021.03.017>
- Staiger, A., & Ziegler, W. (2008). Syllable frequency and syllable structure in the spontaneous speech production of patients with apraxia of speech. *Aphasiology*, 22(11), 1201–1215. DOI: <https://doi.org/10.1080/02687030701820584>
- Tilsen, S. (2014). Selection and coordination of articulatory gestures in temporally constrained production. *Journal of Phonetics*, 44(26), 46. DOI: <https://doi.org/10.1016/j.wocn.2013.12.004>
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55, 53–77. DOI: <https://doi.org/10.1016/j.wocn.2015.11.005>

- Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, 21(1–2), 61–82. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31321-X](https://doi.org/10.1016/S0095-4470(19)31321-X)
- Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2014). The role of production in infant word learning. *Language Learning*, 64(s2), 121–140. DOI: <https://doi.org/10.1111/lang.12058>
- Ziegler, W. (2005). A nonlinear model of word length effects in apraxia of speech. *Cognitive Neuropsychology*, 22, 603–623. DOI: <https://doi.org/10.1080/02643290442000211>
- Ziegler, W. (2008). Apraxia of speech. In G. Goldenberg & M. Miller (Eds.), *Handbook of Clinical Neurology* (88 ed., pp. 269–285). London: Elsevier. DOI: [https://doi.org/10.1016/S0072-9752\(07\)88013-4](https://doi.org/10.1016/S0072-9752(07)88013-4)
- Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, 24(5), 631–661. DOI: <https://doi.org/10.1080/01690960802327989>
- Ziegler, W. (2011). Apraxic failure and the hierarchical structure of speech motor plans: A non-linear probabilistic model. In A. Lowit & R. D. Kent (Eds.), *Assessment of Motor Speech Disorders* (pp. 305–323): Plural Publishing Group.
- Ziegler, W. (2016). Phonology vs. phonetics in speech sound disorders. In P. Van Lieshout, B. Maassen, & H. Terband (Eds.), *Speech Motor Control in Normal and Disordered Speech. Future Developments in Theory and Methodology* (pp. 233–256). Rockville, MD: ASHA Press.
- Ziegler, W., & Ackermann, H. (2017). Subcortical Contributions to Motor Speech: Phylogenetic, Developmental, Clinical. *Trends in neurosciences*, 40(8), 458–468. DOI: <https://doi.org/10.1016/j.tins.2017.06.005>
- Ziegler, W., & Aichert, I. (2015). How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex*, 69, 24–39. DOI: <https://doi.org/10.1016/j.cortex.2015.04.001>
- Ziegler, W., Aichert, I., & Staiger, A. (2010). Syllable-and rhythm-based approaches in the treatment of apraxia of speech. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, 20(3), 59–66. DOI: <https://doi.org/10.1044/nnsld20.3.59>
- Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research*, 55(5), S1485–S1501. DOI: [https://doi.org/10.1044/1092-4388\(2012/12-0128\)](https://doi.org/10.1044/1092-4388(2012/12-0128))
- Ziegler, W., Aichert, I., & Staiger, A. (2017). When words don't come easily: A latent trait analysis of impaired speech motor planning in patients with apraxia of speech. *Journal of Phonetics*, 64, 145–155. DOI: <https://doi.org/10.1016/j.wocn.2016.10.002>
- Ziegler, W., Aichert, I., Staiger, A., Willmes, K., Baumgärtner, A., Grewe, T., ... Breitenstein, C. (2022). The prevalence of apraxia of speech in chronic aphasia after stroke: A Bayesian hierarchical analysis. *Cortex*, 151, 15–29. DOI: <https://doi.org/10.1016/j.cortex.2022.02.012>
- Ziegler, W., Lehner, K., Pfab, J., & Aichert, I. (2021). The nonlinear gestural model of speech apraxia: clinical implications and applications. *Aphasiology*, 35, 462–484. DOI: <https://doi.org/10.1080/02687038.2020.1727839>
- Zipf, G. K. (1935). *The psycho-biology of language*. Boston, MA: Houghton Mifflin.

