


Article

In Time with the Beat: Entrainment in Patients with Phonological Impairment, Apraxia of Speech, and Parkinson's Disease

Ingrid Aichert ^{1,*}, Katharina Lehner ¹, Simone Falk ², Mona Späth ³, Mona Franke ¹ and Wolfram Ziegler ¹ 

¹ Clinical Neuropsychology Research Group, Institute of Phonetics and Speech Processing, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; katharina.lehner@phonetik.uni-muenchen.de (K.L.); Mona.Franke@phonetik.uni-muenchen.de (M.F.); wolfram.ziegler@ekn-muenchen.de (W.Z.)

² International Laboratory for Brain, Music and Sound Research (BRAMS), Département de Linguistique et de Traduction, Université de Montréal, Montréal, QC H3C 3J7, Canada; simone.falk@umontreal.ca

³ Neolexon, Limedix GmbH, 80538 Munich, Germany; mona.spaeth@neolexon.de

* Correspondence: ingrid.aichert@ekn-muenchen.de

Abstract: In the present study, we investigated if individuals with neurogenic speech sound impairments of three types, Parkinson's dysarthria, apraxia of speech, and aphasic phonological impairment, accommodate their speech to the natural speech rhythm of an auditory model, and if so, whether the effect is more significant after hearing metrically regular sentences as compared to those with an irregular pattern. This question builds on theories of rhythmic entrainment, assuming that sensorimotor predictions of upcoming events allow humans to synchronize their actions with an external rhythm. To investigate entrainment effects, we conducted a sentence completion task relating participants' response latencies to the spoken rhythm of the prime heard immediately before. A further research question was if the perceived rhythm interacts with the rhythm of the participants' own productions, i.e., the trochaic or iambic stress pattern of disyllabic target words. For a control group of healthy speakers, our study revealed evidence for entrainment when trochaic target words were preceded by regularly stressed prime sentences. Persons with Parkinson's dysarthria showed a pattern similar to that of the healthy individuals. For the patient groups with apraxia of speech and with phonological impairment, considerably longer response latencies with differing patterns were observed. Trochaic target words were initiated with significantly shorter latencies, whereas the metrical regularity of prime sentences had no consistent impact on response latencies and did not interact with the stress pattern of the target words to be produced. The absence of an entrainment in these patients may be explained by the more severe difficulties in initiating speech at all. We discuss the results in terms of clinical implications for diagnostics and therapy in neurogenic speech disorders.

Keywords: apraxia of speech; dysarthria; phonological impairment; perception of rhythm; production of rhythm; auditory cuing; acoustic analysis



Citation: Aichert, I.; Lehner, K.; Falk, S.; Späth, M.; Franke, M.; Ziegler, W. In Time with the Beat: Entrainment in Patients with Phonological Impairment, Apraxia of Speech, and Parkinson's Disease. *Brain Sci.* **2021**, *11*, 1524. <https://doi.org/10.3390/brainsci11111524>

Academic Editor: Anja Lowit

Received: 2 August 2021

Accepted: 12 November 2021

Published: 18 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rhythm characterizes many kinds of human activities. For example, the movement sequences of many physical activities such as swimming, hurdling, or fencing have typical rhythmic structures. In music, rhythm plays a particularly important role as well. In traditional compositions, rhythm is conveyed by a beat structure characterized by a more or less regular sequence of different tone durations and pauses [1]. If a piece of music is based on a regular rhythm, people can dance or clap to the rhythm while listening. Underlying this observation is the ability to entrain body movements to music, i.e., the ability to synchronize with the beat [2–7]. It is assumed that children are sensitive to rhythm from an early age [8–10].

In all languages, rhythmic structure is a fundamental prosodic feature of spoken utterances. Rhythm in language refers to the alternation of strong and weak elements in spoken words and phrases, which structures the production and perception of utterances in time. The phonetic features that convey rhythm and prosody are duration, loudness, and pitch [11]. In phonology, the rhythmic grouping of a stressed with one or more unstressed syllables is framed in the concept of the metrical foot. Although word stress in English and German can vary across syllable positions, there is a preference for trochaic stress in disyllabic words, i.e., for a metrical foot composed of a strong syllable followed by a weak syllable [12–14].

1.1. The Role of Rhythm in Speech Production

Models of spoken language production differ in their assumptions regarding the extent to which rhythmic parameters are intertwined with the segmental structure of utterances. Two stages in the speech production process are relevant here: on the one hand, the phonological processing of word forms, and on the other hand, the subsequent phonetic encoding stage where motor programs for speech gestures are retrieved. Access to the phonological form of a word involves the encoding of metrical and segmental information, which is generally viewed as two separate processes. In the speech production model proposed by Levelt and colleagues [15,16], the retrieval of the phonological word form from the mental lexicon involves information about the segmental and metrical structure, which is accessed in parallel, mutually independent stages. The metrical frame of an intended word is assumed to specify the number of syllables and, in cases of irregular stress, the main stress position. For words with regular stress, i.e., when stress assignment is predictable (e.g., the trochaic stress pattern in English or German), the metrical structure is generated by default [17,18].

At the phonetic encoding stage, articulatory-motor programs are retrieved, which then guide speech articulation [19]. The linguistic unit at the center of the phonetic encoding stage in Levelt's model is the syllable. It is postulated that speakers access a long-term store of gestural scores for the frequently occurring syllables of their language; i.e., they download the motor information for each syllable as a pre-compiled, holistic package [20]. For the production of multisyllabic words and words in phrases, the retrieved syllables are assembled in a linear fashion. In contrast to this account, we have proposed a nonlinear gestural (NLG) model of speech motor planning along lines suggested by articulatory phonology [21]. The NLG model explains the phonetic planning demands for spoken words by a hierarchical architecture, expanding from articulatory gestures over syllable constituents to metrical structures [22–24]. The gestural arrangement is nonlinear in the sense that gestures (and the segments they compose) are not planned in a strictly left-to-right sequence; rather, the probability of correctly producing each gesture is a function of its position in the metrical-gestural model of the word as whole. This idea is also in line with the "prosody-first" account, where word form encoding is integrated into an emerging, hierarchically organized prosodic structure [25].

1.2. The Influence of Rhythm on Speech Production in Patients with Neurogenic Speech Sound Impairments

In speakers with aphasia, speech production can be affected by a phonological encoding impairment. Usually, patients with aphasic phonological impairments show failures concerning the segmental structure of words. In contrast, the metrical form of words is considered much less vulnerable in this patient group. However, a limited number of aphasic patients were reported who showed metrical errors in naming, repetition, or reading [26–29].

Apraxia of speech (AOS) is considered an acquired disorder of speech motor programming or phonetic planning [30,31]. The speech production of patients with AOS is characterized by frequent phonetic sound distortions and phonemic errors, as well as by dysfluent and dysprosodic speech [32–34]. Until recently, metric influences on the error pattern in patients with AOS were widely neglected. In a study of German speakers with

AOS, we directly compared the production of disyllabic trochaic words (e.g., 'Puma, puma) with disyllabic iambic words (e.g., Ko'pie, copy) [35]. The regular metrical pattern of German (i.e., the trochee) had a facilitating effect on word production in apraxic speakers, with trochees being produced with greater articulatory accuracy and fluency than iambs. This effect was replicated in a more recent study of English-speaking patients with AOS [36]. Therefore, the vulnerability of speech segments to apraxic impairment strongly interacts with word-level prosody—an observation that is compatible with the NLG model described above [37].

Dysarthria refers to a heterogeneous group of neurogenic speech disorders resulting from paresis of the muscles needed to produce speech, akinesia, ataxia, or dyskinesia. Over and above articulation impairment, patients with dysarthria may also exhibit impairments of speech breathing, voice, and prosody [38]. Rhythmic abnormalities in dysarthric speech may be reflected in altered speech rate, dysfluencies, and reduced stress. In particular, patients with Parkinson's disease and hypokinetic dysarthria are described with disturbances in prosody [39,40].

1.3. The Role of Rhythm in Speech Perception: Theories of Rhythmic Entrainment

The "quasi-rhythmic" nature of spoken language has also been studied in terms of its influences on speech perception. There has long been ample evidence for the relevance of metrical structures to auditory segmentation and parsing processes in language acquisition and adult language processing [41–43]. More recently, this evidence has been substantially extended by electrophysiological studies, which hint to the importance of metrical structures for phonological, syntactic, and semantic processing, as well as for monitoring functions and for auditory attentional control [44–48]. For example, rhythmically regular sentence contexts have been found to facilitate the resolution of syntactic ambiguities in auditory sentence processing [49] and the detection of syntactic violations [45,50].

A physiological basis for these effects is assumed to reside in neural entrainment, that is, the phase-locking of oscillatory activity of cortical neuron associations in the auditory cortex to periodicities of an external auditory stimulus [51]. In the case of speech, neural entrainment is hierarchically modulated at different frequency bands reflecting hierarchical temporal modulations in the speech envelope. For example, temporal envelope information pertaining to the syllable level has been found to primarily modulate neural activity in the theta range (4–8 Hz), while delta and beta/gamma bands respond to higher prosodic and segmental information, respectively [51,52]. Neural entrainment is a mechanism that supports the dynamic selection of information in the brain towards a specific goal [53], and thereby underpins speech intelligibility and comprehension during speech perception [51,54]. It is also a fundamental process when aligning motor rhythms to auditory rhythms ("rhythmic entrainment") [53].

Generally, theories of rhythmic entrainment assume that sensorimotor predictions about upcoming rhythmic structure allow humans to synchronize their actions with the rhythm of external stimuli [5]. For language processing, it is assumed that listening to metrically regular speech establishes rhythmical expectations, which, in turn, might also facilitate speech production processes [3,48]. In other words, regular rhythm in speech provides listeners with a temporal expectancy structure, i.e., the brain extracts rhythmic regularities and sets up expectancies for subsequent production. In everyday conversation, rhythmic entrainment may also play a role, because predictions about another speaker's utterances for rhythm and rate enable a listener to anticipate the end of a speaker's turn and thereby facilitate smooth turn taking [55,56].

1.4. Rhythmic Auditory Cueing in Patients with Speech Sound Impairments

Speaking to the rhythm of an external timekeeper has a long and successful tradition in the treatment of various speech disorders. For example, the metronome effect was already described in 1969 for the treatment of stuttering. In persons who stutter, coupling speech movements to metronomic beats led to a considerable reduction in stuttering symp-

toms [57,58]. In the treatment of patients with AOS, different treatment techniques focus on the rhythmic–melodic aspects of speaking [59]. Some of these techniques use internal “pacemakers” (e.g., tapping or counting) [60] or a pacing board [61], while others rely on external cues (e.g., tactile) [62] or metronome pacing [63,64]. Whereas the metronome generates a uniform pace, a metrical pacing therapy developed by Brendel and Ziegler [65] respected the rhythmic properties of natural speech. This study used acoustic pacing signals that were rhythmically adapted to the metrical properties of natural speech. Metrical pacing therapy was found to improve segmental accuracy and fluency of sentence production. Rhythmic stimulation, along with the melodic structuring of speech, is also a central component of melodic intonation therapy (MIT), which has a long tradition in the treatment of patients with aphasia and AOS [66–68]. It is assumed that the rhythmic aspect of MIT leads to major improvements in apraxic speakers [69].

Although rhythmic–auditory cueing is also a well-established method in the treatment of dysarthria, the influence of rhythmic cues on speech has rarely been investigated in treatment trials. In persons with dysarthria, rhythmic cues are commonly used as pace setters, with the aim of reducing speech rate and increasing speech intelligibility [70–72]. For patients with Parkinson’s disease, it is assumed that dysfunctions at the level of the basal ganglia may impair the internal generation of rhythmic movements, leading to problems with the initiation, generation, and maintenance of speech rhythm [73]. There are numerous studies of these patients that demonstrate the effectiveness of rhythmic auditory stimulation on the (gait) motor function [74]. However, few studies in this patient group addressed the influence of rhythmic stimulation on speech motor function [72]. An auditory priming experiment [75] investigated if the metrical regularity of sentences influences speech motor control in individuals with Parkinson’s disease. In this experiment, participants heard a metrically regular or irregular sentence spoken by a model speaker (prime) and were subsequently requested to read aloud a rhythmically regular or irregular sentence as a response (target). The study showed that speech initiation in individuals with Parkinson’s disease was facilitated by a perceived regular speech rhythm. The authors concluded that regular stress distributions in the speech of an interlocutor allow for enhanced temporal predictions and help persons with dysarthria better accommodate to the speech of their conversational partner.

Following the above mentioned auditory priming experiment with dysarthric speakers [75], we developed a similar paradigm to investigate if articulation of patients with AOS and patients with phonological impairment (PI) also benefit from auditory priming by speech with a regular rhythm [76]. In this study, we conducted a sequential synchronization paradigm based on a sentence completion task. The task required participants to complete a phrase spoken by a model speaker by a disyllabic target word while maintaining the rhythm and fluency of the auditory model. Both patient groups produced fewer segmental errors on target words preceded by regularly stressed prime sentences as compared to prime sentences with a more irregular metrical pattern. They therefore seemed to exploit natural rhythmic cues in heard speech for the segmental realization of the target words. Furthermore, our study confirmed earlier results showing that the symptoms of AOS can be modulated positively by a regular (trochaic) stress of words to be produced [35,36]. The benefit from the regularity of target words could, for the first time, also be demonstrated in individuals with PI. Therefore, there seems to be a robust metrical influence on speech at both the phonetic and the phonological planning stages of speech production.

1.5. Aims

In the present study, we asked if individuals with different sound production impairments accommodate their speech to the natural speech rhythm of an auditory model similar to neurotypical speakers (rhythmic entrainment). In addition to patients with AOS and PI, which have already been described in [76], we also studied patients with Parkinson’s disease using the same paradigm to see if the effects spread to lower levels of speech production. To investigate entrainment effects, we related the response latencies

of the participants to the spoken rhythm of the prime. The main question was whether accommodation effects, if they occur at all, are more significant after hearing sentences with a metrically regular as compared to an irregular pattern (“Prime effect”). Furthermore, we asked if there is a facilitation effect of regular (trochaic) word stress on speech production in these patients (“Target effect”).

2. Materials and Methods

2.1. Participants

Thirty-six persons with three types of neurogenic speech sound impairment participated (16 female; mean age 62 years, range 30–83, s.d. = 13.7): twelve patients were diagnosed with apraxia of speech (AOS), all of whom had varying degrees of coexisting aphasia; twelve participants presented aphasic phonological impairments (PI) without AOS; and twelve individuals had hypokinetic dysarthria associated with Parkinson’s disease (PD). All patients were recruited in cooperation with several clinical institutions (see acknowledgements). Furthermore, 24 neurologically healthy speakers (CON) were examined (18 female, mean age 50 years, range 30–70, s.d. = 13.3). All patients and healthy controls were native German speakers.

Patients with PD had a confirmed diagnosis of Parkinson’s disease by a neurologist and a diagnosis of hypokinetic dysarthria determined by a speech therapist. All patients showed typical symptoms of hypokinetic dysarthria, such as reduced loudness, breathy voice, imprecise articulation, reduced stress, or increased speech rate. Motor impairment in PD was described by the Hoehn and Yahr [77] scale (stage I: $n = 1$; stage II: $n = 3$; stage III: $n = 7$; stage IV: $n = 1$). Though we excluded patients with a diagnosis of dementia, eight of twelve patients with PD showed a mild cognitive impairment according to the Montreal Cognitive Assessment [78].

In the aphasic patients, a clear diagnosis of AOS or aphasic phonological output impairment constituted a prerequisite for participation in the study. In these patients, no or only very mild dysarthria was allowed. The diagnosis of AOS was made on the basis of the following criteria: (a) inconsistent occurrence of phonetic distortions; (b) presence of perceived phonemic errors (i.e., well-articulated phoneme substitutions, additions, and deletions); (c) prosodic disturbances such as syllable segregation, phoneme lengthenings, or inadequate pauses; (d) articulatory groping, self-corrections, and effortful speech. Patients with a diagnosis of an aphasic PI primarily exhibited phonological errors, whereas word-level prosody and speech rate were largely intact. All aphasic patients suffered from a left-hemisphere cerebral lesion (19 ischemic, 5 hemorrhagic). According to the profile height of the Aachen Aphasia Test (AAT) [79], all patients had mild to moderate aphasia.

In order to assess functional deficits underlying the repetition and naming abilities, all patients were administered three subtests of the model-based assessment battery LEMO 2.0 (German acronym for Lexikon modellorientiert) [80]: auditory nonword discrimination (36 of 72 items), oral naming (20 items), and auditory word-to-picture matching (20 items; same set of items as in oral naming). In all patients, auditory function assessed by auditory discrimination abilities was considered sufficient to accomplish the sentence completion task administered in this study. Most of the aphasic patients showed some lexical access deficit in spoken naming. However, largely intact word comprehension for the same items, as revealed by the auditory word-to-picture matching task, implied that semantic processing abilities were preserved. Patients with PD showed very good naming abilities; only two patients exhibited performance at threshold. Moreover, auditory word comprehension was intact in all these patients. For a summary of the LEMO 2.0 test results, see Table 1.

Table 1. Demographic details and clinical characteristics of the patient sample (12 patients with apraxia of speech, 12 patients with aphasia, 12 dysarthric speakers with Parkinson's).

Demographics and Test Scores	Apraxia of Speech (AOS; <i>n</i> = 12)	Phonological Impairment (PI; <i>n</i> = 12)	Hypokinetic Dysarthria (PD; <i>n</i> = 12)
Age in years: median (range)	55 (30–78)	60 (41–73)	75 (54–83)
Months post onset: median (range)	63 (2–191)	6 (1–97)	63 (9–213)
Word repetition accuracy ¹			
Phonetic	44 (2–75)	93 (73–100)	97 (89–100)
Phonemic	64 (19–94)	78 (29–96)	100 (99–100)
Fluency	59 (6–90)	78 (42–98)	99 (94–100)
Auditory word discrimination (LEMO, test V1 ²)	10 unimpaired, 1 at threshold, 1 impaired	6 unimpaired, 6 impaired	9 unimpaired, 3 impaired
Auditory word-picture matching (LEMO, test 11 ²)	8 unimpaired, 1 at threshold, 3 impaired	7 unimpaired, 3 at threshold, 2 impaired	12 unimpaired
(LEMO, test 11 ²)			
(LEMO, test 11 ²)			
Verbal naming (LEMO, test 13 ^{2,3})	3 unimpaired, 9 impaired	1 unimpaired, 2 at threshold, 9 impaired	10 unimpaired, 2 at threshold

¹ Percentage of correct items in the “Hierarchical Word Lists—compact” [81]: Mdn (minimum–maximum), ² Subtests of LEMO 2.0 [80], ³ Responses with errors on only one phoneme (e.g., [klait] (engl. dress) → [kait]) were scored as correct.

To describe the severity and quality of the speech sound impairment at word level, we applied the modified version of the hierarchical word lists, the HWL-compact [81], to all patients (see also Table 1). HWL-compact contains 32 one- to four-syllable picturable words with simple and complex syllable structures. To capture error variability, four selected items are tested multiple times (five times). Word production accuracy was assessed for the presence or absence of phonetic distortions, perceived phonemic errors, and dysfluencies (initiation problems, phoneme lengthenings, pauses, and self-corrections). Regarding word repetition accuracy in the HWL-compact testing, patients with AOS revealed the most severe impairments with the lowest number of correct items for all three error categories. In addition to a high percentage of phonological errors, patients with AOS also showed a considerable number of phonetic errors and word fluency errors in the test. In contrast, patients with PI primarily produced phonemic errors. The occurrence of fluency errors in some PI patients was mainly due to *conduite d’approche* and self-corrections. The word repetition performance in patients with PD was largely unimpaired in the HWL-compact testing. The presence of consistently hypernasal speech in one patient with PD was not considered in the evaluation.

Though all patients with Parkinson's had good segmental abilities in word repetition, most of them showed clear impact of the dysarthric impairment on spontaneous speech, where all patients exhibited reduced articulatory precision. Some patients showed additional changes to voice quality, reduced overall loudness, decreased prosodic modulation, and/or an increased speech tempo. Severity of dysarthria was judged by the clinic's speech language therapists on site (3 mild, 6 moderate, 3 severe).

2.2. Materials and Procedure

A sequential synchronization paradigm based on a sentence completion task was conducted. A detailed description of the procedure was made in [76]. The stimulus set

contained 96 sentences consisting of five two-syllable words each. The first four words of each sentence served as an auditory prime, while the last word was the target word to be produced by the participant. The prime sentence as well as the target word were controlled for the regularity of the underlying speech rhythm. Half of the prime sentences were metrically regular, comprising a series of four trochaic words (i.e., an alternating strong–weak pattern; Xx Xx Xx Xx). The other half of the prime sentences were metrically irregular, consisting of an alternation of trochaic and iambic words (trochee–iamb–iamb–trochee; Xx xX xX Xx). Target words were either trochees (regular; Xx) or iambs (irregular; xX). There were 48 target words in total. Each target word occurred twice in two different prime conditions, once after a regular prime sentence and once after an irregular prime sentence. All target words were concrete nouns and were low in frequency with regard to spoken/written word frequency (CELEX) [82]. Apart from syllable complexity (i.e., CV and CVC syllables), target words were also controlled for the sound class of the first phoneme.

To ensure that the target words were semantically unpredictable by the preceding prime sentence, a written sentence completion task including 43 healthy participants (28 women, 15 men; mean age: 41.8 years) was conducted. Participants were instructed to complete each prime sentence with the first word they thought of. For only two items, the original wording had to be changed to make the target words less predictable.

Due to the systematic variation of prime sentences and target words, there were four different rhythmic conditions:

- (1) Regular prime sentence—trochaic target word (Xx Xx Xx Xx-Xx) Example: 'Lena 'pflanzte 'damals 'diese-'Tulpe (engl. Lena planted then this-tulip)
- (2) Regular prime sentence—iambic target word (Xx Xx Xx Xx-xX) Example: 'Friedrich 'neckte 'häufig 'diesen-'Te'nor (engl. Friedrich teased often this-tenor)
- (3) Irregular prime sentence—trochaic target word (Xx xX xX Xx-Xx) Example: 'Jule ver'schenkt je'doch 'diese-'Tulpe (engl. Jule gives away though this-tulip)
- (4) Irregular prime sentence—iambic target word (Xx xX xX Xx-xX) Example: 'Mira be'trog zu'nächst 'diesen-'Te'nor (engl. Mira cheated at first this-tenor)

The experiment was implemented on a computer where the sentences were presented in a pseudorandomized order. The sequential synchronization paradigm requires participants to complete a phrase spoken by a model speaker by maintaining the rhythm and fluency of the auditory model.

The participants were first familiarized with the target word by hearing a male model speaker producing the whole sentence, including the target word (full sentence model). Immediately after this model, the participants only heard the prime sentence spoken by a female model speaker and were required to complete the sentence fragment with the previously heard target word. Participants were instructed to join in by continuing the speech rhythm of the auditory model. Before starting with the experimental items, the patient was familiarized with the procedure through four practice items, which could be repeated several times.

Prior to the experiment, model stimuli were recorded by two experienced speakers (male and female) using an audio interface (Focusrite Scarlett 2i2) and a Rode NTG-2 microphone. Some of the stimuli of both model speakers were adjusted by minimally shortening or lengthening pauses or segments to ensure that there were no differences in mean speech rate between metrically regular and irregular prime sentences (one-way analysis of variance, $F(3, 95) = 0.561, p > 0.1$).

The experiment was administered using the Software PsychoPy (Psychophysics software v1.90.1) [83]. Stimuli were presented over mobile loudspeakers (Philips multimedia speakers 2.0). New trials were always initiated manually by the experimenter. The participants' responses were recorded digitally at a sampling rate of 44.1 kHz, using the same technique as for the recording of the model speakers. All participants signed an informed consent form but were naive as to the specific purpose of the study.

2.3. Data Analyses

As a primary outcome parameter, response latencies—defined as the time interval between the onset of the last word of the prime sentence and the onset of the spoken reaction—were measured. Latencies were manually marked using Praat [84]. Once the segmentation was completed, a Praat script was used to export the data automatically. A total of 5760 responses were evaluated (96 per participant).

For the acoustic analyses, we excluded responses following articulatory groping (Example: Tenor → p.. t.. t.. tenor) or phonemic searching behavior (Example: Tenor → pen.. ton.. ten.. tenor; $n = 339$) as well as words with phonological errors on the first syllable of the target word ($n = 329$). Furthermore, to a lesser extent, we excluded null reactions ($n = 70$), perseverative errors (i.e., repetition of the preceding target word; $n = 27$), semantic paraphasias (Examples: package → gift, tenor → opera; $n = 38$), and responses where an indefinite or definite article was inserted before the target word ($n = 17$). In 15 trials, the audio recording was missing, or the quality of the recording was too poor for an acoustic analysis due to noise. Latencies of more than 2000 ms ($n = 32$) were also excluded. As a result, a total of 4.893 reactions (84.9%) could be considered in the statistical analyses (i.e., exclusion of 867 reactions).

While a similarly high proportion of reactions had to be excluded in the aphasic subgroups (AOS: $n = 398$, PI: $n = 410$), a significantly smaller proportion of responses were excluded in the patients with PD ($n = 53$). In the healthy controls, only six responses were discarded. There were similar numbers of targets that could not be acoustically analyzed following regular ($n = 439$) and irregular ($n = 428$) prime sentences. However, with regard to the regularity of the target words, more iambic than trochaic words had to be excluded (trochaic: $n = 331$, iambic: $n = 536$).

In order to assess the reliability of the acoustic analyses, 5% of the data were re-analyzed (total: $n = 288$). Equal proportions were selected from the samples produced by the patient groups and the controls (2 patients per group, 36 items per participant). A high correlation between the values measured by the two examiners (first and fifth author) was found (intraclass correlation coefficient, $r = 0.997$, $p < 0.001$).

To relate the response latencies to the spoken rhythm of the individual prime sentences, we further determined the average foot duration (i.e., duration of iambs/trochees, in ms) of the model speaker's prime sentences. For this, the duration of the complete prime sentence was measured and divided by the number of words (i.e., four words per prime). If the participants are sensitive to the temporal predictability of the prime sentences, we expect that they delay their response by the average foot duration of the preceding prime. For an illustration of the procedure, see Figure 1.

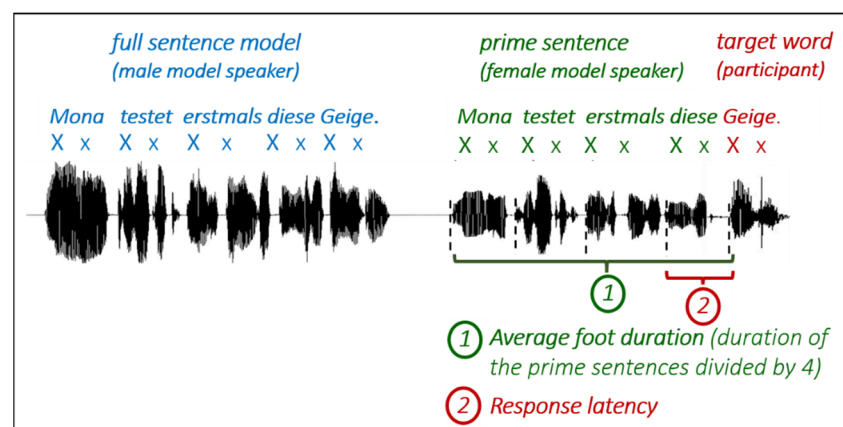


Figure 1. Overview of the procedure and data analysis.

2.4. Rhythm Discrimination Task

In addition to the sequential synchronization task, all participants performed a rhythm discrimination task (see also [76] for a detailed description of the procedure). The purpose of the task was to ensure that participants were able to perceive the metrical differences between the two priming conditions. Two harmonic tones that differed in duration, volume, and pitch were combined to form a trochaic and iambic tone pair, respectively. Stressed tones were longer in time, higher in pitch, and louder compared to unstressed tones. The trochaic and iambic tone pairs had the same overall duration of 655 ms. Trochaic and iambic tone pairs were then combined to match the rhythmically regular and irregular patterns of the prime sentences in the sentence completion task (regular: Xx Xx Xx Xx, irregular: Xx xX xX Xx). In the experiment, participants heard two tone sequences in succession, separated only by a short pause of 75 ms. They had to decide if the second tone sequence was the same or different from the first tone sequence. Four combinations of regular and irregular tone sequences were created: regular–regular, regular–irregular, irregular–regular, and irregular–irregular. In total, the subjects listened to 32 items (8 items per condition), which were presented in a randomized order. For the analysis, the error numbers were evaluated across the four conditions.

2.5. Statistical Analyses

All statistics were performed using R [85]. Linear mixed effects models were calculated to estimate the effects of prime regularity (fixed effects factor PRIME) and target regularity (factor TARGET) on response latencies in the four groups (fixed effects factor GROUP), with ITEMS and SUBJECTS as random slope effects. In several cases, model complexity was reduced after likelihood-ratio testing of a nested vs. the full model using the R-function ANOVA.

3. Results

3.1. Rhythm Discrimination Task

Across all conditions, neurologically healthy individuals made only 6% errors, whereas the patient groups made 20% (PD) and 25% errors (both AOS and PI), respectively. All groups produced considerably fewer errors when the first tone sequence was regular. In these conditions (regular–regular, regular–irregular), none of the healthy speakers made more than one error (out of a total of 16 responses). Comparable performance with no or only one error was also shown by a considerable proportion of patients in all three patient groups (AOS and PI: 8/12 participants, PD: 9/12 participants). Table 2 shows mean error rates for the four rhythmic conditions (regular–regular, regular–irregular, irregular–regular, and irregular–irregular).

Table 2. Error rates in % (mean, s.d.) in the rhythm discrimination task across the four rhythmic conditions. Figures represent grand averages across patients and items. AOS: apraxia of speech ($n = 12$); PI: phonological impairment ($n = 12$); PD: Parkinson’s dysarthria ($n = 12$); CON: healthy controls ($n = 24$). Eight items per condition.

Stimulus 1	Regular			Irregular		
	Stimulus 2	Regular	Irregular	Total	Regular	Irregular
AOS	8 (27.8)	3 (17.5)	6 (9.7)	59 (49.4)	29 (45.7)	44 (31.5)
PI	5 (22.3)	10 (30.7)	8 (13.2)	43 (49.7)	42 (49.6)	42 (25.5)
PD	11 (31.9)	6 (23.3)	9 (16.5)	32 (46.8)	33 (47.3)	32 (28.0)
CON	1 (11.7)	1 (8.3)	1 (3.5)	14 (34.3)	9 (28.7)	11 (18.4)
Total	5 (21.3)	4 (18.4)	4 (10.1)	29 (45.4)	22 (41.2)	25 (27.8)

The most remarkable result was that error rates were substantially higher when the first stimulus had an irregular rhythm compared to a regular first stimulus, across all groups and regardless of the regularity of stimulus 2.

A generalized linear mixed-effects model was calculated using the function *glmer* of the R-package “lme4” [86]. STIMULUS 1 (regular, irregular), STIMULUS 2 (regular, irregular), and GROUP (CON, PD, AOS, PI) were modeled as fixed effects, and SUBJECT as random intercept effects. The complexity of the full model could be reduced by ignoring the three-way interaction and all two-way interactions with STIMULUS 1, as the variance explained by the simplified model did not differ from that of the full model (likelihood-ratio test using R-function *ANOVA*, $\chi^2(7) = 9.0$, $p = 0.25$). The simplified model revealed that each of the three patient groups made significantly more errors than the neurotypical controls (AOS: $\beta = 2.07$, $p < 0.001$; PI: $\beta = 2.21$, $p < 0.001$; PD: $\beta = 1.76$, $p < 0.001$), but there were no significant pairwise differences between the three patient samples ($p > 0.05$). A significant increase in the likelihood of a discrimination error was observed across all groups when Stimulus 1 was irregular ($\beta = 1.21$, $p < 0.001$). Since there was no interaction of GROUP with STIMULUS 1, the regularity effect of Stimulus 1 was statistically the same for all groups, demonstrating that the participants of all groups were sensitive to the regularity of the first stimulus. Regarding STIMULUS 2, the likelihood of discrimination errors decreased when the stimulus was irregular ($\beta = -0.26$, $p < 0.05$), a result that was mainly driven by the large number of errors of the AOS patients when an irregular first stimulus was followed by a regular second stimulus (STIMULUS 2 interactions with AOS: CON: $\beta = 0.39$, $p < 0.05$; PI: $\beta = 0.73$, $p < 0.001$; PD: $\beta = 0.56$, $p < 0.01$), while there were no pairwise differences between CON, PI, and PD regarding the regularity effect of the second stimulus ($p > 0.05$).

3.2. Sequential Synchronization Paradigm

For the sequential synchronization paradigm, we first present the data for only the healthy controls. Figure 2 plots the absolute onset latencies of the target words, i.e., the distance of the target word onset from the onset of the last word of the prime sentence, as a function of the regularity of the prime and the target for the control group (CON). The horizontal line represents the average foot duration of the prime sentences, which was almost identical for the regular (654 ms) and the irregular (655 ms) primes.

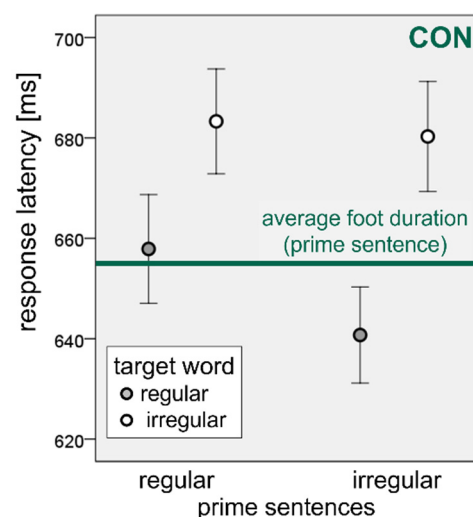


Figure 2. Response latencies (in milliseconds) of the neurotypical participants/CON (mean values, 95% CI of the mean), as a function of the regularity of the prime sentences (abscissa) and the target words (filled and unfilled symbols). The horizontal line represents the grand average of the mean foot durations across all prime sentences.

The figure shows that, in the neurotypical controls, responses following regular prime sentences were almost precisely in time with the average prime foot durations when the target word had a trochaic stress pattern. After irregular prime sentences, the trochaic

target words had shorter onset latencies. Iambic target words were initiated consistently later, with a delay of ca. 30 ms relative to the “metrical beat”.

A linear mixed effects regression model was calculated for response onset latencies using the function ‘lmer’ of the R-package “lme4” [86]. PRIME (regular, irregular) and TARGET (regular, irregular) were modeled as fixed effects, and SUBJECT as random slope effects. A type III analysis of variance using the function ‘anova’ of the R-package *stats* [85] revealed a significant effect of TARGET ($F(1, 23) = 6.8, p < 0.05$) but not of PRIME ($F(1, 23) = 0.3, p = 0.59$) and a significant PRIME \times TARGET interaction ($F(1, 2169) = 25.0, p < 0.001$). The regression coefficients for irregular primes ($\beta = -28.8$) and irregular targets ($\beta = 22.6$) were both non-significant, while the effect size of the TARGET irregularity in the irregular PRIME condition was highly significant ($\beta = 37.6, p < 0.001$).

For the analysis of the entire data set, all response latencies were corrected by the average foot duration of the corresponding prime sentence, so that a response that is exactly in time with the metric beat of the model speaker received a corrected value of 0 (horizontal line in Figure 3). Negative and positive values represent speech initiation latencies shorter or longer than this unit, respectively.

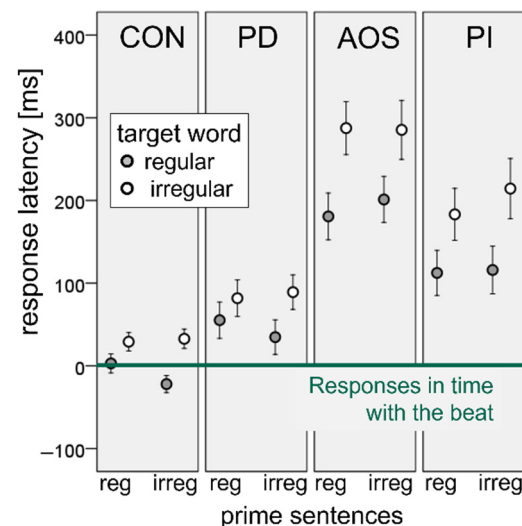


Figure 3. Normalized response latencies (in milliseconds) as a function of prime and target regularity for the healthy control speakers (CON) and the patients with Parkinson’s disease (PD), apraxia of speech (AOS), and phonological impairment (PI). Normalization was performed relative to a predicted latency of exactly the mean duration of the four metrical feet of each individual prime sentence (=0). Error bars are 95% confidence intervals, as in Figure 2.

Figure 3 plots these values in milliseconds for each of the four experimental groups, again as a function of the regularity of primes and targets.

The figure illustrates that the CON and the PD group differed only slightly in their response latencies and had similar patterns regarding their sensitivity to the rhythm of the prime sentences and the stress pattern of the target words, whereas the AOS and PI patients had considerably longer onset latencies, which also differed in their sensitivity to the two regularity factors. As with the data of Figure 2, a linear mixed effects model was calculated, this time with GROUP (CON, PD, AOS, PI) as a third fixed effects factor and participants as random slope effects. The complexity of the full model could be reduced by ignoring the three-way interaction, as a likelihood-ratio χ^2 -test revealed that the variances explained by the full and the simplified models were not different ($\chi^2(3) = 4.8, p = 0.19$).

In the simplified model, a type III analysis of variance using the R-function ‘anova’ revealed significant main effects of GROUP ($F(3, 59) = 12.2, p < 0.001$) and TARGET ($F(1, 30) = 17.4, p < 0.001$), but not of PRIME ($F(1, 25) = 0.1, p = 0.80$) and significant interactions of GROUP \times PRIME ($F(3, 1058) = 2.7, p < 0.05$), GROUP \times TARGET ($F(3, 47) = 6.3, p < 0.01$),

as well as PRIME \times TARGET ($F(1, 4665) = 8.1, p < 0.01$). The AOS and the PI groups had significantly longer response delays than the healthy speakers (AOS: $\beta = 211.7, p < 0.001$; PI: $\beta = 161.0, p < 0.01$) and the Parkinson's patients (AOS: $\beta = 165.5, p < 0.01$; PI: $\beta = 105.8, p < 0.05$), whereas no differences were found between CON and PD ($\beta = 55.7, p = 0.25$) and between AOS and PI ($\beta = -50.7, p = 0.36$). Response onset latencies were significantly longer for iambic target words than for trochees ($\beta = 59.9, p < 0.05$), whereas there was no overall effect of the regularity of PRIME ($\beta = 13.6, p = 0.52$).

Regarding the two-way effects, the GROUP \times PRIME interaction was exclusively driven by a larger effect of irregular vs. regular primes in AOS as compared to CON ($\beta = 34.4, p < 0.01$), whereas this effect was comparable between all other groups ($|\beta| < 30, p > 0.05$ in all pairs). The GROUP \times TARGET interaction was mainly driven by longer response latencies of iambic relative to trochaic target words in PI compared to CON ($\beta = 71.6, p < 0.001$) and to PD ($\beta = 71.3, p < 0.001$), while all other between-group comparisons of the TARGET effect had non-significant regression coefficients ($|\beta| < 40.0, p > 0.05$).

Since there was an obvious divergence between the healthy speakers and the patients with Parkinson's dysarthrias, on the one hand, and the two left hemisphere stroke groups on the other hand, further linear mixed effects models were computed for each of these pairs. In the PD vs. CON comparison, the only significant effect was the two-way interaction between PRIME and TARGET that was already seen in the analysis of the CON-data of Figure 2 ($\beta = 37.1, p < 0.001$). The fact that there was no main effect of and no interaction with GROUP supports that the Parkinson's patients did not differ from the healthy speakers in their response patterns. In the AOS vs. PI comparison, the only significant effect was an overall increase in onset latencies of iambic vs. trochaic target words across both groups and both priming conditions ($\beta = 79.6, p < 0.05$), an effect that was also present as a tendency in the CON and PD groups, but did not reach significance there (see Figure 3). All other effects and all interactions were non-significant. In particular, the finding that there was no main effect of and no interaction with GROUP supports that the AOS patients did not differ from the PI patients in their response patterns. Moreover, the absence of a significant PRIME effect demonstrates that the patients in these two groups were not sensitive to the rhythm of the prime sentence.

4. Discussion

The aim of this study was to investigate if individuals with acquired speech sound disorders entrain their speech to the natural speech rhythm of an auditory model similar to neurologically healthy speakers. While interventions based on auditory cueing with rhythmic tone sequences are well-established in the treatment of patients with sound production impairments [59,70], there are only few exploratory studies indicating that these patients might benefit from the rhythm of natural speech input [75,76]. This is the first study to examine rhythmical entrainment in three patient groups with different underlying pathomechanism using the rhythmical priming paradigm described in [76]. In addition to aphasic patients with phonological impairment (PI) or apraxia of speech (AOS), we also studied a group of individuals with hypokinetic dysarthria associated with Parkinson's disease (PD). Participants were presented with prime sentences with a regular or an irregular rhythmical pattern, respectively, and were instructed to complete them by producing single target words (sequential synchronization task). The target words were also controlled for the regularity of word stress (regular trochaic vs. irregular iambic).

4.1. Rhythm Discrimination Task: Were Participants Able to Perceive the Metrical Differences between the Two Priming Conditions?

First of all, we wanted to ensure that participants perceived the rhythmic difference between the regular vs. irregular prime sentences in the sequential synchronization task. For this purpose, we conducted a rhythm discrimination task in which participants had to decide whether there was a difference between two subsequent rhythmic tone sequences. An influence of rhythmic regularity of the first tone sequence was observed in all participants. All groups had much less difficulties in the conditions in which the first stimulus

was a regular tone sequence, irrespective of the rhythm of stimulus 2. The healthy speakers and a considerable proportion of the subjects in the patient groups showed no or almost no errors in these conditions. In contrast, all three patient groups had considerable problems when the sentences started with an irregular tone sequence. Therefore, our results show that both the unimpaired subjects and the patients were sensitive to the rhythmical properties of the auditory stimuli.

The advantage of a regular first stimulus in this task is not surprising, considering that participants were able to use the regular rhythm to make perceptual predictions for the processing of incoming inputs. This explanation is in accordance with principles of rhythmic entrainment, where rhythmical expectancies are assumed to be induced by the regular metrical structure of an auditory input [3,48]. Within the framework of the dynamic attending theory [87,88] it is postulated that rhythmically regular tone sequences entrain attentional oscillations and thereby can enhance stimulus processing. Moreover, our results are also consistent with the assumption of an influence of rhythmical stimuli on the capacity of verbal short-term memory. For example, a study from the 1960s already described an advantage of rhythmically grouped items for short-term memory performance [89]. Within the framework of the BUMP model (bottom-up multi-scale population oscillator), Hartley and colleagues explained the relationship between the mechanisms of rhythm perception and the capacity of verbal short-term memory [90]. In their model, sensory processes that are driven by input set up neural oscillations based on the rhythm of speech. In the case of a regular grouping, more oscillators are activated, which then provide a richer and more distinctive representation of the sensory input. Though we used nonverbal tone stimuli in our experiment, it could also be assumed that the rhythmic regular grouping of the first stimulus made the stimulus easier to remember than the irregular grouped stimulus, which, in turn, facilitated the comparison with the rhythmical pattern of the second stimulus.

Especially for dysarthric patients with PD, some authors assume deficits in the auditory processing of temporal cues [91]. For example, in a study by Grahn and Brett, patients with PD were less able to discriminate non-speech rhythms than healthy subjects [92]. This result was explained by deficits in the internal generation of rhythms in this patient group. However, in the experiment performed here, the participants with PD could profit from the regularity of the first stimulus, suggesting that auditory processing can be supported by rhythm (for a discussion of this effect, see Section 4.2). Our results also confirm a former study that revealed that individuals with PD were also able to detect rhythmic differences between tone sequences [75]. The discrepancy between the studies might be explained by the fact that Grahn and Brett [92] used longer and metrically more complex tone sequences in their study.

4.2. Sequential Synchronization Paradigm-“Prime Effect”: Do Participants with Neurogenic Speech Sound Impairments Accommodate Their Speech to the Natural Speech Rhythm of an Auditory Model?

The main question was whether accommodation effects, if they occur at all, are larger after hearing sentences with a metrically regular as compared to an irregular pattern. For healthy speakers, our study revealed an entrainment of trochaic target words preceded by regularly stressed prime sentences: participants delayed their response almost exactly by the average foot duration of the preceding prime. In contrast, trochaic target words following irregular prime sentences were produced faster than expected. We assume that, in this condition, a temporal reference frame could not be set up to establish a rhythmic scaffold for the subsequent word production. While hearing a rhythmically irregular prime sentence, the participants may have been unable to create a rhythmical framework for their response and therefore tried to produce the target word as quickly as possible. The results of our study are in line with the assumption that expectancy-induced rhythmic priming can also be used to entrain the speech production of a speaker with the previously heard natural speech signal [3,48]. Although the subjects in our experiment were not involved in a natural conversation, the results also fit into the idea that rhythmic and neural entrainment underlies smooth turn taking processes [55,56]. Importantly, to stay

in the rhythm, for the control participants, it was necessary that the target word is also rhythmically regular. Irregular target words were produced slower in both conditions, i.e., an entrainment effect was not evident for these words. In a more general sense, our results confirm a close intertwining of language perception and production processes in neurotypical speakers [93].

Even though somewhat slower across all conditions, individuals with PD showed a pattern similar to that of the healthy speakers. Therefore, dysarthric patients with PD were also able to temporally anticipate and accommodate to the rhythmic regular pattern of an input. This finding confirms our earlier investigation [75], in which patients with PD entrained their speech rhythm especially with another person's metrically regular speech. Individuals with PD suffer from a progressive loss of dopamine in the basal ganglia, which results in the characteristic motor symptoms of these patients. The basal ganglia are considered to be involved, among other things, in the integration of auditory input into speech motor movements [94]. In particular, the involvement of the basal ganglia in rhythm processing and temporal prediction is widely discussed [95–97]. For patients with PD, it was shown that they might benefit from exposure to a temporally predictable, regular auditory cue [98]. Furthermore, numerous studies also revealed that patients with damage to the basal ganglia show improvements in gait function after perceiving rhythmical auditory beats [74,99], suggesting that the patients are still able to access rhythmic entrainment mechanisms to improve their gait patterns. Our results provide further evidence for this assumption within the speech domain and confirm that these patients are still able to use regular temporal cues to improve their motor impairment, i.e., in our case, their impaired speech production abilities.

In contrast to the dysarthric speakers with PD and the healthy controls, no entrainment could be observed for the aphasic patients with apraxia of speech (AOS) or phonological impairment (PI), respectively. Both patient groups showed considerably longer response latencies and different patterns regarding their sensitivity to the regularity conditions of the experiment. In particular, the regularity of the prime sentences had no influence on their response latencies. A linear model including only these two groups showed no significant main effect of or interaction with the factor PRIME.

The absence of an entrainment in these groups contrasts with the results of our former study, in which we evaluated error production in the same samples of patients with AOS and PI based on the same paradigm and materials [76]. In this former study, both patient groups profited from the metrical regularity of speech input, in that they produced fewer segmental errors on target words preceded by regularly stressed prime sentences as compared to prime sentences with an irregular metrical pattern. Therefore, patients with AOS and PI obviously seemed to exploit rhythmic cues in speech for the segmental realization of the target words. We already discussed this effect within the framework of rhythmic entrainment [3–5] and saw the results as further evidence for the assumption that listening to metrically regular speech establish rhythmical expectations, which, in turn, might facilitate speech production processes [48]. The absence of an entrainment effect regarding speech initiation latencies in these patients may be explained by the more severe difficulties in initiating speech at all. If patients take too long to even start producing a word, it is no longer possible for them to stay in the rhythm of the prime sentence.

4.3. Sequential Synchronization Paradigm-“Target Effect”: Is There a Facilitation Effect of Regular (Trochaic) Word Stress on Speech Production in Patients with Neurogenic Speech Sound Impairments?

Both the healthy speakers and all three patient groups showed a strong effect of target regularity. Regardless of the regularity of the first sentence, participants initiated irregular (iambic) words with considerably longer latencies compared to regular (trochaic) words. The effect was small in the healthy participants and the Parkinson's patients and much larger in the two patient groups with AOS or phonological impairment after left hemisphere lesions, and it was independent of the regularity of the prime sentences. For patients with AOS, this result confirms earlier investigations based on repetition tasks,

which revealed an advantage of trochaic over iambic patterns [35–37]. Furthermore, the results are in agreement with the error data reported from the same experiment for patients with AOS and PI [76]. In both patient groups, the regular metrical pattern of German (i.e., the trochee) had a facilitating effect on word production, with trochees being produced with greater articulatory accuracy and fluency than iambs.

For patients with AOS, the faster initiation of words with trochaic as compared to iambic stress patterns suggests that the encoding of speech segments to apraxic impairment strongly interacts with word-level prosody. Therefore, our study once again supports the assumption of the NLG model that intergestural cohesion expands across the whole metrical tree of a word [22–24]. The NLG model assumes a stronger bonding between encoding units within the domain of a trochaic foot (i.e., a stressed unstressed rhythm) than across metrical boundaries, which may explain the facilitating effect of regular words in our patients with AOS. The facilitating effect of trochees in patients with PI could be explained by an entrenchment of segmental articulation in metrical rhythm, as, for example, predicted by the “Prosody First” framework outlined by [25]. It might be that, in patients with PI, the access to nondefault, irregular metrical word forms is more demanding than the spell-out of default, metrically regular forms [27,100].

5. Conclusions

Persons with Parkinson’s dysarthria accommodated their speech to the natural speech rhythm of an auditory model similar to neurotypical speakers. Both groups accommodated the initiation of their responses to rhythmically regular speech input (rhythmic entrainment). The sensitivity to the temporal predictability in the individuals with Parkinson’s disease also supports therapeutic approaches based on rhythmic cueing to improve intelligibility in these patients. The absence of an entrainment effect in patients with apraxia of speech and phonological impairment was explained by the more severe difficulties in initiating speech at all, which could be seen in considerably longer response latencies of the target word production. As an additional result, we observed considerably longer latencies for irregular (iambic) words compared to regular (trochaic) words. These acoustic data complemented our previous finding that the stressed–unstressed rhythm promotes accuracy in these patient groups.

Author Contributions: Conceptualization, I.A., S.F. and W.Z.; methodology, I.A., S.F., M.S. and W.Z.; validation I.A. and W.Z.; formal analysis, I.A. and W.Z.; investigation, I.A., K.L. and M.F.; resources, I.A.; data curation, I.A., K.L. and M.F.; writing—original draft preparation, I.A. and W.Z.; writing—review and editing, I.A., S.F., W.Z. and K.L.; visualization, I.A.; supervision, W.Z.; project administration, I.A.; funding acquisition, I.A. and W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG), grant number Zi 469/14-2.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the Faculty of Medicine, Ludwig-Maximilians-Universität Munich, Germany (protocol code 129-14; date of approval: 22 April 2014).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

Acknowledgments: We are grateful to the patients and to the healthy control speakers for their participation. We also thank the speech-language teams of 14 collaborating centers where patient examinations took place: Asklepios Klinik Schaufling (Holger Grötzbach), Asklepios Klinik Bad Tölz (Doris Klintwort), Benedictus Krankenhaus Feldafing (Sophie Held), Bezirksklinikum Mainkofen, Deggendorf (Andrea Hys), Berufsfachschule für Logopädie Augsburg (Ralf Peters), Döpfer Berufsfachschule für Logopädie München (Karen Fasch), Fachklinik Bad Heilbrunn (Wolf Finger), Therapie Sendling (Brigitta Kraus), Mutabor e.V. (Steffy Marinho), Praxis Wort für Wort (Anja Wunderlich), Praxis für Logopädie und klinische Lerntherapie (Charlotte Averdick), Staatliche Berufsfachschule

für Logopädie München (Pia Goldstein), Schön Klinik Bad Aibling-Harthausen (Gudrun Klingenberg), and Schön Klinik Schwabing (Edith Wagner-Sonntag). We thank Ramona Schreier for assistance in data collection of the healthy speakers.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Cason, N.; Astésano, C.; Schön, D. Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta Psychol.* **2015**, *155*, 43–50. [[CrossRef](#)]
2. Thaut, M.H.; McIntosh, G.C.; Hoemberg, V. Neurobiological foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Front. Psychol.* **2015**, *5*, 1185. [[CrossRef](#)]
3. Falk, S.; Dalla Bella, S. It is better when expected: Aligning speech and motor rhythms enhances verbal processing. *Lang. Cogn. Neurosci.* **2016**, *31*, 699–708. [[CrossRef](#)]
4. Lagrois, M.; Palmer, C.; Peretz, I. Poor Synchronization to Musical Beat Generalizes to Speech. *Brain Sci.* **2019**, *9*, 157. [[CrossRef](#)] [[PubMed](#)]
5. Repp, B.H. Sensorimotor synchronization: A review of the tapping literature. *Psychon. Bull. Rev.* **2005**, *12*, 969–992. [[CrossRef](#)] [[PubMed](#)]
6. Dalla Bella, S.; Białuńska, A.; Sowiński, J. Why movement is captured by music, but less by speech: Role of temporal regularity. *PLoS ONE* **2013**, *8*, e71945. [[CrossRef](#)] [[PubMed](#)]
7. Patel, A.D. *Music, Language, and the Brain*; Oxford University Press: New York, NY, USA, 2008.
8. Drake, C.; Jones, M.R.; Baruch, C. The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition* **2000**, *77*, 251–288. [[CrossRef](#)]
9. Post, B.; Payne, E. Speech rhythm in development: What is the child acquiring? In *Prosodic Development in First Language Acquisition*; Esteve-Gibert, N., Prieto, P., Eds.; John Benjamins: Amsterdam, The Netherlands, 2018; pp. 125–144.
10. Chen-Hafteck, L. Music and Language Development in Early Childhood: Integrating Past Research in the Two Domains. *Early Child Dev. Care* **1997**, *130*, 85–97. [[CrossRef](#)]
11. Fletcher, J. The prosody of speech: Timing and rhythm. In *The Handbook of Phonetic Sciences*, 2nd ed.; Wiley-Blackwell: Hoboken, NJ, USA, 2010; pp. 521–602.
12. Kager, R. Feet and metrical stress. In *The Cambridge Handbook of Phonology*; Lacy, P., Ed.; Cambridge University Press: Cambridge, UK, 2007; pp. 195–228.
13. Domahs, U.; Wiese, R.; Bornkessel-Schlesewsky, I.; Schlesewsky, M. The processing of German word stress: Evidence for the prosodic hierarchy. *Phonology* **2008**, *25*, 1–36. [[CrossRef](#)]
14. Féry, C. German word stress in optimality theory. *J. Comp. Ger. Linguist.* **1998**, *2*, 101–142. [[CrossRef](#)]
15. Levelt, W.J.M.; Roelofs, A.; Meyer, A.S. A theory of lexical access in speech production. *Behav. Brain Sci.* **1999**, *22*, 1–38. [[CrossRef](#)] [[PubMed](#)]
16. Levelt, W.J.M. *Speaking. From Intention to Articulation*; MIT Press: Cambridge, MA, USA, 1989.
17. Roelofs, A.; Meyer, A.S. Metrical structure in planning the production of spoken words. *J. Exp. Psychol. Learn. Mem. Cogn.* **1998**, *24*, 922–939. [[CrossRef](#)]
18. Schiller, N.O.; Fikkert, P.; Levelt, C.C. Stress priming in picture naming: An SOA study. *Brain Lang.* **2004**, *90*, 231–240. [[CrossRef](#)]
19. Laganaro, M. Phonetic encoding in utterance production: A review of open issues from 1989 to 2018. *Lang. Cogn. Neurosci.* **2019**, *34*, 1193–1201. [[CrossRef](#)]
20. Levelt, W.J.M.; Wheeldon, L.R. Do speakers have access to a mental syllabary? *Cognition* **1994**, *50*, 239–269. [[CrossRef](#)]
21. Goldstein, L.; Fowler, C.A. Articulatory Phonology: A phonology for public language use. In *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*; Schiller, N.O., Meyer, A.S., Eds.; Mouton de Gruyter: Berlin, Germany, 2003; pp. 159–207.
22. Ziegler, W. A nonlinear model of word length effects in apraxia of speech. *Cogn. Neuropsychol.* **2005**, *22*, 603–623. [[CrossRef](#)]
23. Ziegler, W.; Lehner, K.; Pfab, J.; Aichert, I. The nonlinear gestural model of speech apraxia: Clinical implications and applications. *Aphasiology* **2021**, *35*, 462–484. [[CrossRef](#)]
24. Ziegler, W. Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Lang. Cogn. Process.* **2009**, *24*, 631–661. [[CrossRef](#)]
25. Keating, P.A.; Shattuck-Hufnagel, S. *A Prosodic View of Word Form Encoding for Speech Production*; UCLA: Los Angeles, CA, USA, 2002; pp. 112–156.
26. Aichert, I.; Ziegler, W. Segmental and metrical encoding in aphasia: Two case reports. *Aphasiology* **2004**, *18*, 1201–1211. [[CrossRef](#)]
27. Laganaro, M.; Vacheresse, F.; Frauenfelder, U.H. Selective impairment of lexical stress assignment in an Italian-speaking aphasic patient. *Brain Lang.* **2002**, *81*, 601–609. [[CrossRef](#)] [[PubMed](#)]
28. Nickels, L.; Howard, D. Effects of lexical stress on aphasic word production. *Clin. Linguist. Phon.* **1999**, *13*, 269–294. [[CrossRef](#)]
29. Janßen, U.; Domahs, F. Going on with optimised feet: Evidence for the interaction between segmental and metrical structure in phonological encoding from a case of primary progressive aphasia. *Aphasiology* **2008**, *22*, 1157–1175. [[CrossRef](#)]

30. Ballard, K.J.; Granier, J.P.; Robin, D.A. Understanding the nature of apraxia of speech: Theory, analysis, and treatment. *Aphasiology* **2000**, *14*, 969–995. [[CrossRef](#)]
31. Ziegler, W. Phonology vs. phonetics in speech sound disorders. In *Speech Motor Control in Normal and Disordered Speech. Future Developments in Theory and Methodology*; van Lieshout, P., Maassen, B., Terband, H., Eds.; ASHA Press: Rockville, MD, USA, 2016; pp. 223–255.
32. McNeil, M.R.; Ballard, K.J.; Duffy, J.R.; Wambaugh, J.L. Apraxia of speech. Theory, assessment, differential diagnosis and treatment: Past, present and future. In *Speech Motor Control in Normal and Disordered Speech: Future Developments in Theory and Methodology*; Van Lieshout, P., Maassen, B., Terband, H., Eds.; ASHA Press: Rockville, MD, USA, 2016; pp. 195–221.
33. Croot, K. Diagnosis of AOS: Definition and criteria. *Semin. Speech Lang.* **2002**, *23*, 267–280. [[CrossRef](#)] [[PubMed](#)]
34. Wambaugh, J.L.; Duffy, J.R.; McNeil, M.R.; Robin, D.A.; Rogers, M.A. Treatment Guidelines for Acquired Apraxia of Speech: A Synthesis and Evaluation of the Evidence. *J. Med. Speech Lang. Pathol.* **2006**, *14*, xv.
35. Aichert, I.; Späth, M.; Ziegler, W. The role of metrical information in apraxia of speech. Perceptual and acoustic analysis of word stress. *Neuropsychologia* **2016**, *82*, 171–178. [[CrossRef](#)] [[PubMed](#)]
36. Bailey, D.J.; Bunker, L.; Mauszycki, S.; Wambaugh, J.L. Reliability and stability of the metrical stress effect on segmental production accuracy in persons with apraxia of speech. *Int. J. Lang. Commun. Disord.* **2019**, *54*, 902–913. [[CrossRef](#)] [[PubMed](#)]
37. Ziegler, W.; Aichert, I. How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. *Cortex* **2015**, *69*, 24–39. [[CrossRef](#)]
38. Duffy, J.R. *Motor Speech Disorders. Substrates, Differential Diagnosis, and Management*, 4th ed.; Elsevier: St. Louis, MO, USA, 2020.
39. Skodda, S.; Schlegel, U. Speech rate and rhythm in Parkinson’s disease. *Mov. Disord.* **2008**, *23*, 985–992. [[CrossRef](#)]
40. Lowit, A.; Marchetti, A.; Corson, S.; Kuschmann, A. Rhythmic performance in hypokinetic dysarthria: Relationship between reading, spontaneous speech and diadochokinetic tasks. *J. Commun. Disord.* **2018**, *72*, 26–39. [[CrossRef](#)] [[PubMed](#)]
41. Shields, J.L.; McHugh, A.; Martin, J.G. Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *J. Exp. Psychol.* **1974**, *102*, 250–255. [[CrossRef](#)]
42. Meltzer, R.H.; Martin, J.G.; Mills, C.B.; Imhoff, D.L.; Zohar, D. Reaction time to temporally-displaced phoneme targets in continuous speech. *J. Exp. Psychol. Hum. Percept. Perform.* **1976**, *2*, 277–290. [[CrossRef](#)]
43. Pitt, M.A.; Samuel, A.G. The use of rhythm in attending to speech. *J. Exp. Psychol. Hum. Percept. Perform.* **1990**, *16*, 564–573. [[CrossRef](#)] [[PubMed](#)]
44. Cason, N.; Schon, D. Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia* **2012**, *50*, 2652–2658. [[CrossRef](#)]
45. Schmidt-Kassow, M.; Kotz, S.A. Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *J. Cogn. Neurosci.* **2009**, *21*, 1693–1708. [[CrossRef](#)] [[PubMed](#)]
46. Rothermich, K.; Schmidt-Kassow, M.; Schwartze, M.; Kotz, S.A. Event-related potential responses to metric violations: Rules versus meaning. *NeuroReport* **2010**, *21*, 580–584. [[CrossRef](#)]
47. Rothermich, K.; Schmidt-Kassow, M.; Kotz, S.A. Rhythm’s gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia* **2012**, *50*, 232–244. [[CrossRef](#)]
48. Falk, S.; Lanzilotti, C.; Schön, D. Tuning Neural Phase Entrainment to Speech. *J. Cogn. Neurosci.* **2017**, *29*, 1378–1389. [[CrossRef](#)]
49. Roncaglia-Denissen, M.P.; Schmidt-Kassow, M.; Kotz, S.A. Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *PLoS ONE* **2013**, *8*, e56000. [[CrossRef](#)] [[PubMed](#)]
50. Schmidt-Kassow, M.; Kotz, S.A. Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis. *Brain Res.* **2008**, *1226*, 144–155. [[CrossRef](#)] [[PubMed](#)]
51. Giraud, A.; Poeppel, D. Cortical oscillations and speech processing: Emerging computational principles and operations. *Nat. Neurosci.* **2012**, *15*, 511–517. [[CrossRef](#)] [[PubMed](#)]
52. Donhauser, P.W.; Baillet, S. Two Distinct Neural Timescales for Predictive Speech Processing. *Neuron* **2020**, *105*, 385–393. [[CrossRef](#)] [[PubMed](#)]
53. Lakatos, P.; Gross, J.; Thut, G. A New Unifying Account of the Roles of Neuronal Entrainment. *Curr. Biol.* **2019**, *29*, R890–R905. [[CrossRef](#)] [[PubMed](#)]
54. Etard, O.; Reichenbach, T. Neural Speech Tracking in the Theta and in the Delta Frequency Band Differentially Encode Clarity and Comprehension of Speech in Noise. *J. Neurosci.* **2019**, *39*, 5750–5759. [[CrossRef](#)]
55. Scott, S.K.; McGettigan, C.; Eisner, F. A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. *Nat. Rev. Neurosci.* **2009**, *10*, 295–302. [[CrossRef](#)] [[PubMed](#)]
56. Cummins, F. Rhythm as an affordance for the entrainment of movement. *Phonetica* **2009**, *66*, 15–28. [[CrossRef](#)] [[PubMed](#)]
57. Brady, J.P. Studies on the metronome effect on stuttering. *Behav. Res. Ther.* **1969**, *7*, 197–204. [[CrossRef](#)]
58. Ingham, R.J.; Bothe, A.K.; Jang, E.; Yates, L.; Cotton, J.; Seybold, I. Measurement of speech effort during fluency-inducing conditions in adults who do and do not stutter. *J. Speech Lang. Hear. Res.* **2009**, *52*, 1286–1301. [[CrossRef](#)]
59. Zumbansen, A.; Tremblay, P. Music-based interventions for aphasia could act through a motor-speech mechanism: A systematic review and case-control analysis of published individual participant data. *Aphasiology* **2019**, *33*, 466–497. [[CrossRef](#)]
60. Square, P.; Martin, R.E.; Bose, A. Nature and treatment of neuromotor speech disorders in aphasia. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*, 4th ed.; Chapey, R., Ed.; Lippincott Williams & Williams: Philadelphia, PA, USA, 2001; pp. 847–884.

61. Wertz, R.T.; Lapointe, L.L.; Rosenbek, J.C. *Apraxia of Speech in Adults: The Disorder and Its Management*; Grune & Stratton: Orlando, FL, USA, 1984.
62. Rubow, R.T.; Rosenbek, J.; Collins, M.; Longstreth, D. Vibrotactile stimulation for intersystemic reorganization in the treatment of apraxia of speech. *Arch. Phys. Med. Rehabil.* **1982**, *63*, 150–153. [[PubMed](#)]
63. Wambaugh, J.L.; Martinez, A.L. Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech. *Aphasiology* **2000**, *14*, 851–871. [[CrossRef](#)]
64. Shane, H.C.; Darley, F.L. The effect of auditory rhythmic stimulation on articulatory accuracy in apraxia of speech. *Cortex* **1978**, *14*, 444–450. [[CrossRef](#)]
65. Brendel, B.; Ziegler, W. Effectiveness of metrical pacing in the treatment of apraxia of speech. *Aphasiology* **2008**, *22*, 77–102. [[CrossRef](#)]
66. Sparks, R.W.; Holland, A.L. Method: Melodic Intonation Therapy for aphasia. *J. Speech Hear. Disord.* **1976**, *41*, 287–297. [[CrossRef](#)] [[PubMed](#)]
67. Albert, M.L.; Sparks, R.W.; Helm, N.A. Melodic intonation therapy for aphasia. *Arch. Neurol.* **1973**, *29*, 130–131. [[CrossRef](#)] [[PubMed](#)]
68. Zumbansen, A.; Peretz, I.; Hébert, S. Melodic Intonation Therapy: Back to Basics for Future Research. *Front. Neurol.* **2014**, *5*, 7. [[CrossRef](#)] [[PubMed](#)]
69. Stahl, B.; Kotz, S.A.; Henseler, I.; Turner, R.; Geyer, S. Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain* **2011**, *134*, 3083–3093. [[CrossRef](#)]
70. Mainka, S.; Mallien, G. Rhythmic speech cueing (RSC). In *Handbook of Neurologic Music Therapy*; Oxford University Press: New York, NY, USA, 2014; pp. 150–160.
71. Pilon, M.A.; Mcintosh, K.W.; Thaut, M.H. Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic-ataxic dysarthric speakers: A pilot study. *Brain Inj.* **1998**, *12*, 793–803. [[CrossRef](#)]
72. Thaut, M.H.; Mcintosh, K.W.; McIntosh, G.C.; Hoemberg, V. Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. *Funct. Neurol.* **2001**, *16*, 163–172.
73. Skodda, S.; Flasskamp, A.; Schlegel, U. Instability of syllable repetition as a model for impaired motor processing: Is Parkinson's disease a “rhythm disorder”? *J. Neural. Transm.* **2010**, *117*, 605–612. [[CrossRef](#)]
74. Spaulding, S.J.; Barber, B.; Colby, M.; Cormack, B.; Mick, T.; Jenkins, M.E. Cueing and Gait Improvement Among People with Parkinson's Disease: A Meta-Analysis. *Arch. Phys. Med. Rehabil.* **2013**, *94*, 562–570. [[CrossRef](#)] [[PubMed](#)]
75. Späth, M.; Aichert, I.; Ceballos-Baumann, A.; Wagner-Sonntag, E.; Miller, N.; Ziegler, W. Entraining with another person's speech rhythm: Evidence from healthy speakers and individuals with Parkinson's disease. *Clin. Linguist. Phon.* **2016**, *30*, 68–85. [[CrossRef](#)]
76. Aichert, I.; Lehner, K.; Falk, S.; Späth, M.; Ziegler, W. Do Patients With Neurogenic Speech Sound Impairments Benefit From Auditory Priming With a Regular Metrical Pattern? *J. Speech Lang. Hear. Res.* **2019**, *62*, 3104–3118. [[CrossRef](#)] [[PubMed](#)]
77. Hoehn, M.M.; Yahr, M.D. Parkinsonism: Onset, progression, and mortality. *Neurology* **1967**, *17*, 427–442. [[CrossRef](#)] [[PubMed](#)]
78. Nasreddine, Z.S.; Phillips, N.A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J.L.; Chertkow, H. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **2005**, *53*, 695–699. [[CrossRef](#)] [[PubMed](#)]
79. Huber, W.; Poeck, K.; Weniger, D.; Willmes, K. *Aachener Aphasie Test (AAT)*; Hogrefe: Göttingen, Germany, 1983.
80. Stadie, N.; Cholewa, J.; De Bleser, R. *LEMO 2.0. Lexikon Modellorientiert—Diagnostik Für Aphasie, Dyslexie Und Dysgraphie*; NAT-Verlag: Hofheim, Germany, 2013.
81. Ziegler, W.; Aichert, I.; Staiger, A.; Schimeczek, M. *HWL-Kompakt*. Available online: <https://neurophonetik.de/sprechapraxie-wortlisten> (accessed on 30 July 2021).
82. Baayen, R.H.; Piepenbrock, R.; Gulikers, L. *The CELEX Lexical Database (CD-ROM)*. *Linguistic Data Consortium*; University of Pennsylvania: Philadelphia, PA, USA, 1995.
83. Peirce, J.W. PsychoPy—Psychophysics Software in Python. *J. Neurosci. Methods* **2007**, *162*, 8–13. [[CrossRef](#)] [[PubMed](#)]
84. Boersma, P.; Weenink, D. PRAAT: Doing Phonetics by Computer [Computer Program], Version 5.3.42; 2013. Available online: <https://www.fon.hum.uva.nl/praat/> (accessed on 8 February 2021).
85. Team, R.C. R: A Language and Environment for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 30 July 2021).
86. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **2015**, *67*, 48. [[CrossRef](#)]
87. Large, E.; Jones, M. The dynamics of attending: How people track time-varying events. *Psychol. Rev.* **1999**, *106*, 119–159. [[CrossRef](#)]
88. Jones, M.; Moynihan, H.; MacKenzie, N.; Puente, J. Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *Psychol. Sci.* **2002**, *13*, 313–319. [[CrossRef](#)] [[PubMed](#)]
89. Ryan, J. Grouping and short-term memory: Different means and patterns of grouping. *Q. J. Exp. Psychol.* **1969**, *21*, 137–147. [[CrossRef](#)]
90. Hartley, T.; Hurlstone, M.J.; Hitch, G.J. Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cogn. Psychol.* **2016**, *87*, 135–178. [[CrossRef](#)] [[PubMed](#)]

91. Grahn, J.A.; Brett, M. Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* **2009**, *45*, 54–61. [[CrossRef](#)] [[PubMed](#)]
92. Jones, C.R.; Jahanshahi, M. Motor and perceptual timing in Parkinson's disease. *Adv. Exp. Med. Biol.* **2014**, *829*, 265–290. [[CrossRef](#)]
93. Watkins, K.E.; Strafella, A.P.; Paus, T. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* **2003**, *41*, 989–994. [[CrossRef](#)]
94. Ziegler, W.; Ackermann, H. Subcortical Contributions to Motor Speech: Phylogenetic, Developmental, Clinical. *Trends Neurosci.* **2017**, *40*, 458–468. [[CrossRef](#)] [[PubMed](#)]
95. Grahn, J.A. The Role of the Basal Ganglia in Beat Perception. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 35–45. [[CrossRef](#)]
96. Shi, E.R.; Zhang, Q. A domain-general perspective on the role of the basal ganglia in language and music: Benefits of music therapy for the treatment of aphasia. *Brain Lang.* **2020**, *206*, 104811. [[CrossRef](#)]
97. Kotz, S.A.; Schmidt-Kassow, M. Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex* **2015**, *68*, 48–60. [[CrossRef](#)]
98. Kotz, S.A.; Gunter, T.C. Can rhythmic auditory cuing remediate language-related deficits in Parkinson's disease? *Ann. N. Y. Acad. Sci.* **2015**, *1337*, 62–68. [[CrossRef](#)] [[PubMed](#)]
99. McIntosh, G.C.; Brown, S.H.; Rice, R.R.; Thaut, M.H. Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **1997**, *62*, 22–26. [[CrossRef](#)] [[PubMed](#)]
100. Howard, D.; Smith, K. The effects of lexical stress in aphasic word production. *Aphasiology* **2002**, *16*, 198–237. [[CrossRef](#)]