PHONETIC-PHONOLOGICAL FACTORS IN STUTTERING IN GERMAN

Annett B. Jorschick, Dinah Rachko, Joana Cholin

Faculty of Linguistics and Literary Studies, Bielefeld University, Germany
annett.jorschick@uni-bielefeld.de, dinah.rachko@gmail.com, joana.cholin@uni-bielefeld.de

ABSTRACT

Relating stuttering to phonetic-phonological complexity has a long tradition in clinical and psycholinguistic research. Some findings point toward a relationship between stuttering and phonetic-articulatory effort, while others emphasize lexical-representational aspects. In our study, we focus on phenomena specifically arising in the transition from phonological to phonetic-articulatory encoding. Here, abstract phonological segments are assembled to form motor plans for articulation by taking various factors into account.

We investigated stuttering in a German word-reading-task using words and nonsense-words starting with a [CV]-structure, manipulating *manner* and *place* of articulation of word-initial consonants and their distance to adjacent vowels (*transition*), *vowel length,* controlling for *word length*, *syllable frequency* and *biphone frequency*. Results show *word length* as a robust predictor of stuttering, while the factors *transition*, *manner* and *vowel length* (and their interaction) also played a role. We argue that only such systematic approaches allow insights into the complex interplay between encoding levels in stuttering.

Keywords: stuttering; phonology-phonetics interface; phonological-phonetic complexity

1. INTRODUCTION

Persistent developmental stuttering is a multifactorial speech disorder causing recurring speech disrupting symptoms such as prolongations and repetitions of speech sounds and silent blocks that severely hamper fluent speech production [1, 2]. The investigation of linguistic factors relevant for the occurrence of these symptoms aims to pinpoint deviating speech behavior to individual encoding levels in order to decipher the underlying pathomechanism of stuttering [3, 4]. Most research in this area is focused on aspects of either lexical representations and their corresponding syntactic and phonological encoding processes or phonetic-articulatory procedures. Current data do not allow to unequivocally associate the recurrent dysfluencies with deficient representations or processes.

In our study, we focus on phenomena arising in the conversion from phonological to phonetic-articulatory encoding as this interface between linguistic planning and motor programming might be particularly susceptible to stuttering. During phonological encoding, the individual phonemes of words are successively grouped together to form abstract phonological syllables, following universal and language-specific syllabification rules [5]. What constitutes an optimal syllable in a given context, crucially depends on the phonological-phonetic environment and the availability of subsequent motor programming.

Several authors [6, 7, 8] have attributed stuttering to the inability to transition between adjacent segments. According to [9], this is especially the case at word onsets and accelerated within stressed syllables. Therefore, we aim to investigate whether there are specific constellations between adjacent segments prompting stuttering. Especially relevant is the question, whether these combinations are problematic per se or whether they are problematic only if linked to specific lexical representations (i.e., words but not nonsense-words). Importantly, the many interrelated factors that are at play in the phonology-phonetics interface need to be disentangled. Previous studies mainly tested simple contrasts in corpus studies, often ignoring interdependencies between these factors and their interactions.

Thus, in our material, we systematically manipulated manner and place of articulation (hereinafter manner and place) of consonants in word onsets. The transition between initial consonants and adjacent vowels was also systematically manipulated to test whether a larger distance between consonants' place and vowels' backness accelerates stutter occurrences. Additionally, we varied vowel length across words and nonsense-words to investigate the influence of different vowel quantities on stuttering rates, which has been shown as a relevant predictor in corpus studies with children who stutter [10]. German offers the opportunity to contrast vowel length in different vowel positions. Other factors that were found to influence stuttering rates in previous studies (for an overview and discussion see [3, 4]) such as stress, word frequency, number of syllables, and word function were controlled across conditions (see below).

2. METHOD

The ethics committee of Bielefeld University approved the study.

2.1. Materials and Design

We selected 144 common nouns from CELEX [11] in which the initial consonants and their adjacent vowels were carefully chosen taking into account the following factors: *place* (labial vs. velar/glottal) and *manner* of the initial consonant (plosive, fricative, nasal, glottal stop), vowel backness (front: /ʏ, ɪ, ɛ, yː, iː, ɛː/, central: /a, aː/, back: /ɔ, ʊ, oː, uː/) and *vowel length* (short vs. long). *Transition:* three transitional distances (small, medium, large) were defined as the distance in position between the word initial consonant and its adjacent vowel: A small distance is found between a labial consonant and a frontal vowel or a velar/glottal consonant and a back vowel. Medium transitions were defined as the distance of the labial or velar/dorsal consonants to a central vowel. Large transitions reflect the distance between a labial consonant and a back vowel or a dorsal/glottal consonant and a frontal vowel. We controlled stress by choosing words with stressed first syllables, while no word contained consonant clusters or diphthongs in any position. The words were matched between the manipulated conditions according to their number of syllables (one or two syllables). Because there are no German words starting with velar nasal /ŋ/ and words orthographically starting with vowels are realized with glottal stops, manipulations resulted in an incomplete factorial design with 2+1+1 (manner) x 2 (place) x 3 (transition) x 2 (vowel length) = 36 possible CV-combinations (see Table 1 for examples of potential initial CV-structures). We selected four mono- or disyllabic words for these 36 CV-combinations resulting in 144 words and constructed 144 corresponding nonsense-words by exchanging one phoneme, keeping the initial CV-structure intact and also obeying constraints of German phonotactics and orthography (i.e., Gummi [gum] – Gubbi), resulting in 288 targets. In addition to these manipulations, we retrieved syllable frequency of the first syllable and bigram frequency of the initial CV-structure from CELEX as potential covariates.

Experimental lists alternated between word and nonsense-word blocks with a break between blocks. The order of items within word blocks was pseudorandomized to ensure that consecutive items were neither phonologically nor semantically related. The matching block of nonsense-words was ordered accordingly. The order of experimental blocks was counterbalanced across experimental lists.

	labial			dorsal		
	small	medium	large	small	medium	large
vowel				ʔɔ,ʔoː	ʔa, ʔaː	ʔɪ, ʔiː
nasal	mɪ, miː	ma, maː	mɔ, moː			
plosive	pɪ, piː	pa, paː	pɔ, poː	kɔ, koː	ka, kaː	kɪ, kiː
fricative	fɪ, fiː	fa, faː	fɔ, foː	hɔ, hoː	ha, haː	hɪ, hiː

Table 1: Overview of the word-initial CV-structure manipulations: *place* and *manner* of C, *transition* to V and *vowel Length*

2.2. Method and Procedure

In a pretesting session, participants reported their stuttering history and took part in a video-recorded diagnostic session in which the stuttering severity instrument (SSI-4 [12]) was administered. This test uses samples of spontaneous and read speech and an evaluation of speech-accompanying physical concomitants to calculate standardized severity scores ranging from 10 (mild stutter severity) to a maximum score of 56. The SSI-4 was evaluated afterward based on the video-recordings.

During the reading experiment, participants sat in front of a laptop, wearing a headset. They were videotaped from two different perspectives: a close-up of the face and from a further distance to also capture speech-accompanying hand gestures. The experimental procedure was controlled by the DMDX software [13]. After a short introduction presented on the screen, participants read the word and nonsense-word lists following the order described above. Participants used the space key to prompt the next word on the screen. The experiment lasted about 20 minutes resulting in a total session length of 45 to 60 minutes.

2.3. Participants

24 native German adults who stutter (AWS) who were independently diagnosed with persistent developmental stuttering participated in our study. Six participants were excluded from the analysis, as their total SSI-4-score was below the critical limit of 10. The SSI-4-scores of the remaining 18 participants (two women, 16 men; ranging from 22-59 years of age) ranged from 12 (very mild) to 32 (severe).

2.4. Annotation

The second author, a certified speech therapist, manually annotated each stuttered event during the reading experiment using Praat [14]. A word was labelled as stuttered if at least one of the symptoms (prolongation, repetition, or block) occurred during its production.

2.5. Modelling procedure

Regression modelling was conducted on our binary response variable *stuttered* (stuttered vs. not stuttered) using generalized logistic mixed-effects models (GLMMs [15]). We conducted separate analyses for manipulations of labial place of the initial consonant and for velar/glottal place (cf. Materials: labial: nasal, plosive, fricative vs. velar/glottal: glottal stop, plosive, fricative, see Table 1).

All manipulated predictors (*manner, transition, vowel length, word status*) and their two-way interactions as well as the psycholinguistic covariates (*word length, biphone frequency* of initial CV-structure, *syllable frequency* of initial syllable) and their interactions with *word status* were used as independent variables, and AWS was included as random factor. Starting with a fully specified model [16], we used backward selection procedures [17] to first reduce the random effects structure, keeping only significant random slopes and intercepts in the model. In a second step, we reduced the fixed effects structure, using a stepwise backward elimination procedure, dropping all non-significant interactions and main effects until the model contained only terms that contributed to the overall model. All model comparisons were conducted using LRT [18]. We report type three ANOVA chi-square statistics of main effects and interactions from package *car* [19] only for factorial predictors with more than two levels.

3. RESULTS

After excluding those AWS who did not score as clinically relevant in the SSI-4, three trials due to technical errors, and 61 additional trials containing pronunciation errors, the data set consisted of 5120 data points from 18 AWS. 1363 (27%) of these cases were stuttered with a high variability between speakers. Two AWS, although noticeably dysfluent in spontaneous conversation, rarely stuttered during the experiment. We report two separate analyses, modelling stuttered events of firstly, targets with initial labial consonants and secondly targets with initial glottal or velar consonants.

3.1. Labial onsets

AWS stuttered on 724 (28%) out of 2550 possible targets with labial consonants at word onset, ranging from 0.7% to 94.2% stuttering per AWS. For the modeling procedure as described above, we used Helmert contrasts for *manner*, differentiating effects of fricatives from effects of plosives in the first contrast and compared both of them to nasal /m/ in the second contrast. *Transition* (small, medium, large), as an ordered factor, was modelled using polynomial contrasts, and we used sum contrasts for *word status* and *vowel length*. Table 2 shows the summary model statistics of targets with an initial labial consonant. The β-estimates represent the log odds ratios with positive values reflecting a tendency towards higher stuttering rates and vice versa.

We replicated a significant positive relationship between *word length* and stuttering. Analysis also showed a significant linear component of *transition*: larger distances between consonants' place and vowels' backness are related to higher stuttering probabilities. Furthermore, nonsense-words were significantly more often stuttered than words. *Word status* (word vs. nonsense-word), *manner* of the word initial consonant and *vowel length* significantly contributed to the random effects. All other predictors and their interactions did not significantly contribute to the model.

Random effects *Groups Name*				*Variance*
(Intercept) | AWS				7.68
Word Status | AWS				0.18
Manner (plosive vs. fricative) | AWS				0.41
Manner (plos./fric. Vs. nasal) | AWS				0.15
Vowel Length | AWS				0.59
Number of obs.: 2550, groups: AWS: 18				
Fixed effects *Predictor*	β	*SE*	*Z*	*p*
(Intercept)	-2.15	0.66	-3.24	**.001**
Word Length	0.71	0.13	5.29	**<.001**
Transition linear	0.24	0.11	2.15	**.032**
Word Status (Words)	-0.29	0.13	-2.25	**.024**

Table 2: Outcome of the GLMM analysis of targets with labial CV-onsets

3.2. Velar/glottal onsets

Targets with an initial velar/glottal place of articulation were stuttered in 639 of all 2570 cases (25%), ranging from 0% to 74.6% per AWS. We again used Helmert contrasts for *manner* differentiating effects of glottal stops from the effect of plosives in the first contrast and compared both of them to fricative /h/ in the second contrast. All other factors (*transition*, *word status* and *vowel length*) were modelled as in the labial onset condition. Table 3 displays the summary statistics of the model.

Analysis revealed a main effect of *manner* of the initial consonant ($\chi^2(2) = 15.29$, $p < .001$). Although there was no significant difference in stuttering between words starting with glottal stops and words starting with plosives, words starting with glottal fricative /h/ were far less often stuttered. Interestingly, we also found a significant interaction between *manner* and *vowel length* ($\chi^2(2) = 8.48$, $p = .014$). Figure 1 displays the huge stuttering difference on glottal stops before short vowels compared to stuttering on glottal stops before long vowels. This difference is slightly more prominent than the difference between plosives before short vs. long vowels. However, with glottal fricative /h/ vowel length did not affect stuttering frequency. Again, we replicated the effect of *word length* on stuttering frequency. The structure of relevant random effects resembled that of the labial model. Again, no other effects contributed to the model.

Random effects *Groups Name*	*Variance*
(Intercept) | AWS	8.70
Word Status | AWS	0.19
Manner (glottal stop vs. plosive) | AWS	0.61
Manner (g.s./p. vs. fricative) | AWS	0.21
Vowel Length | AWS	0.93
Number of obs.: 2570, groups: AWS: 18	

Fixed effects *Predictor*	*β*	*SE*	*z*	*p*
(Intercept)	-2.74	0.71	-3.84	**<0.001**
Manner (g.s. - p.)	-0.27	0.22	-1.23	0.219
Manner (g.s./p. - f.)	-0.51	0.14	-3.72	**<0.001**
Word Length	0.83	0.16	5.12	**<0.001**
Vowel Length (short v.)	0.35	0.26	1.38	0.163
M (g.s. - p.): VL	-0.16	0.08	-1.94	0.052
M (g.s./p. - f.): VL	-0.12	0.05	-2.29	**0.022**

Table 3: Outcome of the GLMM analysis of targets with velar/glottal CV-onsets (M = Manner, VL = Vowel Length, g.s. = glottal stop, p. = plosive, f. = fricative)

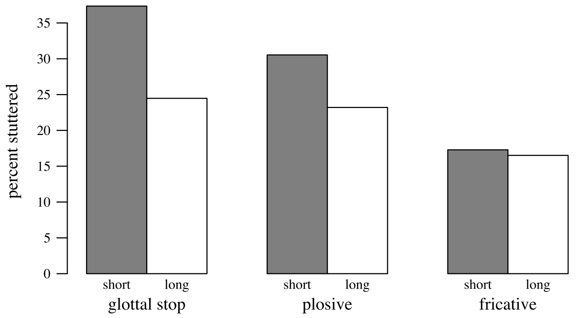

Figure 1: Percent of stuttered targets displaying the interaction of *manner* and *vowel length* of CV-structures with velar/glottal onset consonants

4. DISCUSSION

This study investigated whether a specific combination of phonological-phonetic factors influences speech fluency in experimental reading lists consisting of words and phonotactically legal nonsense-words in German. We replicated the word length effect in line with many other studies on stuttering [20, 21, 22]. Disyllabic items require a higher effort of coordinating two syllables into one metrical unit. The significant interaction between manner of articulation (of first consonants) and vowel length hints at a special status of the fricative /h/ offering itself as a softer onset often targeted in fluency enhancing techniques. In line with Natke and colleagues [10], in all other conditions (glottal stops and velar plosives), short vowels are more susceptible to stuttering than long vowels (see Fig. 1) which is also supported by the significant contribution of vowel length to the random effects in both models. The specific combination of consonantal onsets that cannot be lengthened followed by short vowels might be the most critical structure as it offers no flexibility to (i) smoothen the onset or to (ii) lengthen the vowel. Additionally, short vowels in our disyllabic items always occurred within ambisyllabic structures, making syllable boundaries highly opaque, thereby even increasing the likelihood of stuttering. Syllables with long vowels, in contrast, offer more flexibility to handle vowel production (i.e., by lengthening them) and are associated with more transparent syllable boundaries. The sooner a syllabary boundary is identified at the interface between phonological and phonetic encoding, the faster a phonological syllable can be transformed into phonetic units or motor programs to be overtly articulated. The significant transition effect, i.e., higher stuttering rates due to a larger distance between initial consonants and adjacent vowels, reflects additional articulatory effort. In Japanese, similar transition effects have been found from syllable nuclei to codas [7, 8] suggesting that language-specific articulatory properties drive the different findings (also regarding manner) in cross-linguistic studies on stuttering. On a more general level, the finding that nonsense-words increased stuttering rates while there was no interaction with other factors might again point toward fluency enhancing techniques which are bound to familiar word structures.

In sum, our results show that both lexical variables and motor-articulatory variables play a crucial role in stuttering. A common framework that spans across the phonology-phonetics interface as well as motor programming and articulation is needed to understand the contribution of the single levels and the interplay of linguistic and motor-level variables (for a recent discussion see also [22, 4]). To date, existing models are too narrowly focused on their respective areas (i.e., psycholinguistics, phonetics). Integrating data from different speaker groups will allow for the incorporation of individual differences of fluent and dysfluent speakers in such models. This seems a necessary asset of such models as the high variability between speakers as well as language specific differences particularly play out at the interface between (psycho-)linguistic and motor processing.

5. REFERENCES

[1] Bloodstein, O., Ratner, N. B., & Brundage, S. B. 2021. *A handbook on stuttering*. Plural Publishing.

[2] Wingate, M. E. 2002. *Foundations of stuttering*. San Diego, CA: Academic Press.

[3] Sasisekaran, J. 2014. Exploring the link between stuttering and phonology: A review and implications for treatment. *Seminars in Speech and Language*, *35*(2), 95–113.

[4] Brundage, S. B., & Bernstein Ratner, N. 2022. Linguistic aspects of stuttering: Research updates on the language–fluency interface. *Topics in Language Disorders*, *42*(1), 5–23.

[5] Levelt, W. J., Roelofs, A., & Meyer, A. S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.

[6] Wingate, M. E. 1988. *The Structure of Stuttering: A Psycholinguistic Analysis*. New York: Springer.

[7] Shimamori, S., & Ito, T. 2009. Difference in frequency of stuttering between light and heavy syllables in the production of monosyllables: From the viewpoint of phonetic transition. *The Japanese Journal of Logopedics and Phoniatrics*, *50*, 116–122.

[8] Shimamori, S., & Ito, T. 2010. Does the transition from the onsets to the core vowels affect the frequency of stuttering in Japanese? *Japanese Journal of Special Education*, *48*, 23–29.

[9] Wingate, M. E. 1979. The first three words. *Journal of Speech, Language, and Hearing Research*, *22*(3), 604–612.

[10] Natke, U., Sandrieser, P., van Ark, M., Pietrowsky, R., & Kalveram, K. T. (2004). Linguistic stress, within-word position, and grammatical class in relation to early childhood stuttering. *Journal of Fluency Disorders*, *29*(2), 109-122.

[11] Baayen, R. H., Piepenbrock, R., & Gulikers, L. 1995. The CELEX lexical database (release 2). *Distributed by the linguistic data consortium, University of Pennsylvania*.

[12] Riley, G.D. 2008. *Stuttering Severity Instrument (SSI-4): Examiner Manual and Picture Plates.* Austin, TX: Pro-Ed.

[13] Forster, K.I., & Forster, J.C. 2003. DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers*, *35*(1), 116–124.

[14] Boersma, P. & Weenink, D. 2018. Praat: doing phonetic by computer [Computer program]. Version 6.0.40 retrieved 22 May 2018 from https://www.praat.org.

[15] Bates D., Mächler M., Bolker B., Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

[16] Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

[17] Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. 2015. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

[18] Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. 2017. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

[19] Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., ... & Monette, G. 2012. Package 'car'. *Vienna: R Foundation for Statistical Computing*, *16*.

[20] Brown, S. F. 1945. The Loci of Stutterings in the Speech Sequence. *Journal of Speech Disorders*, *10*(3), 181–192.

[21] Wingate, M. E. 1967. Stuttering and word length. *Journal of Speech and Hearing Research*, *10*(1), 146–152.

[22] Max, L., Kadri, M., Mitsuya, T., & Balasubramanian, V. 2019. Similar within-utterance loci of dysfluency in acquired neurogenic and persistent developmental stuttering. *Brain and Language*, *189*, 1–9.