ARTICULATORY CONTROL IN SPEECH PRODUCTION

Chen Shen, Esther Janse

Centre for Language Studies, Radboud University Nijmegen, the Netherlands

ABSTRACT

Articulatory control can be quantified in various ways. Clinical studies frequently use maximum performance measures (diadochokinesis or DDK) to elicit speakers' maximum rate of repeating syllable sequences. Psycholinguistic studies, on the other hand, often use tongue twister phrases to elicit speech errors in healthy populations. Although both tasks require speakers to rapidly alternate between similar syllables, no direct comparison has been made to investigate the expected overlap between speakers' performance in these two tasks. We collected speech data from 78 healthy young adults, testing their maximum performance on syllable repetitions and tongue twister sentences, and their habitual reading rate. Our results show that individual maximum speech rate in tongue twister sentences was predicted by maximum DDK rate, illustrating that both tasks contain elements of articulatory control. Speakers' habitual sentence reading rate was, however, not correlated to their maximum rate, highlighting a dissociation between maximum and actual performance in speech rate.

Keywords: speech production; articulatory control; maximum performance

1. INTRODUCTION

Speech is practiced daily by most of us, yet speech production can be complex and challenging. The complexity of speech production is illustrated by the experience of 'twisting one's tongue' when trying to produce sentences like 'She sells seashells on the sea shore', in which speakers have to alternate between /s/ and /ʃ/ at word onsets. What articulatory control abilities enable speakers to produce this fluently?

Clinical evaluation of articulatory control often uses repetitive syllable sequences for assessing speech motor capacity in persons with speech disorders (e.g., dysarthria [1]). In this so-called diadochokinesis (henceforth DDK) task, speakers are asked to repeat the same syllable as fast and as accurately as possible (e.g., 'papapapa…') or to alternate between syllables (e.g., to produce 'pataka' repeatedly). The latter task thus asks speakers for their maximum performance (in terms of rate and accuracy) in quickly alternating between syllables that only differ in place of articulation of the onset consonant: labial-alveolar-velar. Likewise, production of tongue twister sentences also requires speakers to alternate between similar onset phonemes, between similar onset clusters, or between singleton onset phonemes and onset clusters. Evidently, production of a meaningful sentence such as a tongue twister sentence entails more linguistic processing than repeating nonsensical syllable sequences. First, a tongue twister sentence requires reading or memorising of a longer fragment than a DDK stimulus. The longer fragment naturally has more variegated alternation between similar syllable onsets than a DDK stimulus. Second, sentence production entails grammatical and semantic processes that are absent in sequence repetition.

Although both DDK and tongue twister tasks contain elements of articulatory control, to our knowledge, no study so far has investigated the relationship between speakers' maximum performance on these two tasks. This is most likely due to the former (DDK) task being typically used in a clinical setting [1-3], and the latter task being mainly used in psycholinguistic studies [4-6] as a means to elicit speech errors from healthy speakers. To quantify articulatory control from different angles, we examined variation within and associations between maximum performance in the two speech tasks, in a healthy adult population.

Reference rates for healthy control speakers already exist for DDK in multiple languages, including Dutch [7]. Additionally, several studies have investigated rate differences between DDK performance on repetitions of non-words versus real words in native speakers of multiple languages [8, 9]. As speakers have access to stored motor programmes for real words, but not for non-words, maximum performance can be expected to be better for word than non-word repetition. Indeed, school-aged children as well as healthy older adults achieved faster repetition rates in producing real word relative to non-word stimuli in DDK tasks [8, 9].

In addition, several clinical studies have addressed the question of whether patients' DDK performance is actually representative of their 'normal' speech behaviour, operationalised as their habitual speech rate in sentence reading. Some have stressed the discrepancy between patients' maximum performance on DDK stimuli and their sentence reading rate [10], thereby questioning the utility of DDK as a clinical measure. Others have observed that habitual rate in healthy adults is associated with their maximum articulation rate, but note that they have used the very same reading materials for eliciting both habitual and maximum rate [11].

In this study, we aimed to quantify articulatory control using two maximum performance speech tasks (a DDK and a tongue twister task) that require fast and accurate alternation between similar syllables. Through the novel combination of these two speech tasks, we aimed to achieve the following three objectives. First, through the maximum performance speech rate and accuracy measures, we investigated the variability in a sample of young healthy adult speakers on stimuli that differ in the level of linguistic content (ranging from non-words to real words to tongue twister sentences). Second, we examined whether speakers' tongue twister performance is related to their maximum articulatory (DDK) performance, as measured with words and non-words, to explore the underlying articulatory control mechanisms these measures may reflect. Third, we investigated whether speakers' maximum rate measures (in DDK and tongue twister tasks) are associated with their habitual sentence reading rate.

2. METHODS

2.1. Participants

In total, 78 participants (age: M = 23 years, SD = 3; 61 females) completed the speech tasks in the Centre for Language Studies lab at Radboud University Nijmegen. They were reimbursed for their time through course credits or gift vouchers. Participants were all native speakers of Dutch, with no speech, hearing, or reading disabilities, nor past diagnosis of speech pathology or brain injury. Normal or corrected-to-normal vision was also required. All 78 participants gave informed consent for their audio recordings to be analysed.

2.2. Description and analysis of the speech tasks

Two speech tasks were used to elicit participants' maximum performance (rate and accuracy) as indices of their articulatory control. An additional sentence reading task was used to gather data for participants' habitual speech rate. Stimuli of all three tasks were presented using PowerPoint slides on a 24'' full HD monitor placed on a table in front of the participant. Recordings were made using a Sennheiser ME 64 cardioid capsule microphone through a pre-amplifier (Audi Ton) onto a steady-state 2 wave/mp3 recorder Roland R-05 in a sound-attenuating recording booth.

The first author monitored participants' task progress and controlled the changing of stimulus slides outside the recording booth on the stimulus computer (Dell Precision T3600).

2.2.1 DDK task description and analysis

Clinical DDK task normally contains repetitions of mono- and tri-syllabic nonsense words such as 'pa' and 'pataka'. Given the focus of this study on alternating articulatory movements, we only selected the commonly used tri-syllabic non-word 'pataka' /pataka/, and added the reversed syllable-order variant 'katapa' /katapa/. In addition, two common real Dutch words that were closest to the nonsense words 'pataka' and 'katapa' were added: 'pakketten' /pɑˈkɛtə(n)/ (packages) and 'kapotte' /kaˈpɔtə/ (broken). Whereas no stress pattern was available for the non-words, both real words had lexical stress on the second syllable. The mono- and di-syllabic nonsense stimuli ('pa', 'ta', 'ka', 'pata', 'taka') were presented as practice trials. All of the nonsense words used here were phonotactically legal in Dutch.

During the task, each DDK stimulus was presented in the centre of a full-screen PowerPoint slide. To elicit repetitive production of the stimulus, multiple (nonsense) words were presented next to each other, for instance "patakapatakapataka...". Participants were instructed to repeatedly produce the presented stimulus as accurately and as fast as possible. A pre-recorded example was played prior to the practices to familiarise the participants with the task. A brief line of text reminding them about accuracy and speed of repetition was constantly on-display at the top of each slide. A 2-second pause (preparation time) followed by a 75-millisecond beep-tone was used to mark the start of articulation. Each stimulus was to be repeated for around 10 seconds. Mean DDK task duration was three minutes.

Participants' maximum performance in terms of articulation rate and accuracy was analysed acoustically in *Praat* [12]. Most participants were already making some errors in a 3-second time window, but errors generally increased in longer time windows. We therefore opted for a relatively long time-window (7s) to capture accuracy and rate in a reliable and representative way.

Individual DDK articulation rate (syllables/sec) was calculated by multiplying the total number of correct-and-full (non)words produced by each participant in a 7-second time window (or as close to 7-second as possible for the repetition counts to be an integer) by three (syllables), and divided this number of total syllables by the actual production time (total-duration minus error-duration, in-breaths, and pauses longer than 200 ms between repetitions).

Individual DDK accuracy (fraction) was calculated as number of correct and full repetitions divided by number of all repetitions in the same 7-second time window. A repetition was only counted as correct if it did not contain any form of error or pauses longer than 200 ms within the sequence.

2.2.2 Tongue twister task description and analysis

Following Wilshire's (1999) tongue twister paradigm [13], we selected four tongue twister sentences that contain a combination of repetition and alternation of word-initial consonants (e.g., **p**oes **k**otst **p**ostzak, and **fr**its **v**indt **v**is **fr**ietjes). Below are the four Dutch tongue twister sentences that were used as test stimuli with their literal English translations in parentheses:

- De **p**oes **k**otst in de **p**ostzak
  (*The cat puked in the mail bag*)
- **Fr**its **v**indt **v**is**fr**ietjes **vr**eselijk **v**ies
  (*Frits finds fish-fries terribly gross*)
- Ik **b**ak een **pl**ak **b**ak**bl**oedworst
  (*I fry a slice of blood-sausage*)
- **P**apa **p**akt de **bl**auwe **pl**atte **b**akpan
  (*Daddy grabs the blue flat frying pan*)

Prior to the task stimuli, two additional tongue twister sentences were presented as practice stimuli:

- **Sl**imme **Sj**aantje **sl**oeg de **sl**ome **sl**ager
  (*Smart Sjaantje hit the slow butcher*)
- **B**akker **B**as **b**akt de **b**olle **br**oodjes **br**uin
  (*Baker Bas bakes the round buns brown*)

Participants were instructed to repeat the tongue twister sentences minimally five times as accurately and as fast as possible. As in the DDK task, tongue twister stimuli were each presented in the centre of a full-screen PowerPoint slide with a reminder of the accuracy and speed of repetition. A picture related to the meaning of each tongue twister sentence (e.g., a blue frying pan) was shown on the same slide, and disappeared after about two seconds (preparation time). Participants were instructed to start repeating the tongue twister as soon as the picture disappeared. Mean tongue twister task duration was four minutes.

Maximum performance (rate and accuracy) was analysed acoustically in *Praat* [12]. Individual tongue twister rate (syllables/sec) was calculated by averaging the articulation rate of the correct repetitions of the four tongue twister sentences. Rate of each correct stimulus was measured by dividing the number of syllables in a tongue twister sentence by the time used for that repetition.

Similar to accuracy measures in the DDK task, individual tongue twister accuracy (fraction) for the first five repetitions per sentence was calculated by number of correct and fluent repetitions divided by five. A repetition was counted as fluent if it did not contain any form of error or pause longer than 200 ms in the tongue twister sentence.

2.2.3 Sentence reading task description and analysis

In addition to the two maximum performance speech tasks, participants also performed a sentence reading task. The reading task contained 48 meaningful Dutch sentences that are between 12 and 16 syllables in length (e.g. *De grote kat heeft de vaas per ongeluk gebroken 'The big cat has accidentally broken the vase'*). Participants were instructed to read the sentences fluently in a natural way. Habitual articulation (HA) rate (syllables/sec) of each speaker was averaged across all 48 sentences.

3. RESULTS

3.1. Quantifying variability in speech performance

Table 1: Speech task performance.
CV corresponds to coefficient of variation as a variability index ((SD/mean)*100%).

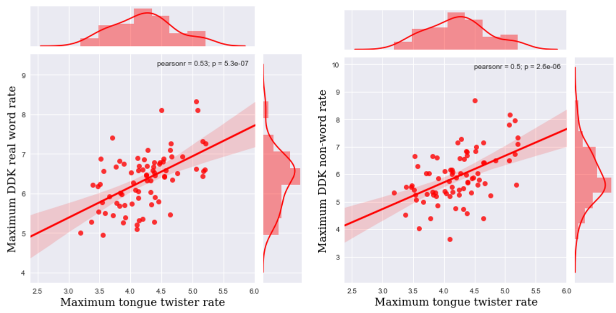
	Rate (syll./sec)			Accuracy (fraction)		
	Mean	*SD*	*CV%*	*Mean*	*SD*	*CV%*
DDK (real word)	6.33	0.71	11.2	0.94	0.06	6.2
DDK (non-word)	5.91	0.93	15.8	0.89	0.10	10.7
Tongue Twister	4.22	0.49	11.5	0.59	0.16	26.8
Habitual Articulation	5.62	0.61	10.9	/	/	/

Table 1 presents the descriptive statistics of the two maximum performance speech tasks and the sentence reading task. Rate and accuracy measures averaged over task stimuli were entered as dependent variables in two models for rate and accuracy respectively. Task (DDK real word, DDK non-word, and Tongue Twister) was entered as the fixed effect of interest, with participant as random effect [14]. Results from linear mixed-effects analysis, using the lme4 package [15], showed that real word DDK performance is significantly better than non-word DDK performance for both rate (t = 5.45, p < .001) and accuracy (t = 2.71, p < .01). Maximum performance in the tongue twister task is significantly worse than in DDK non-word repetition (t = -25.17, p < .001 and t = -18.24, p < .001 for rate and accuracy respectively). This indicates that the difficulty level of repetitively producing tongue twister sentences is relatively high for healthy young adult speakers, possibly also due to the fact that the tongue twister sentences contain syllables of varying complexity (e.g., some have consonant clusters) and voicing alternation in consonants. Furthermore, the more difficult the speech task, the higher the variability in accuracy between speakers, as evident from the coefficient of variation values (cf. Table 1).

3.2. Correlations between maximum performance measures and between maximum and habitual rate

Our second question was whether individual's maximum performance in tongue twister and DDK tasks are associated. Figure 1 below shows the between-task correlations for maximum rate.

Figure 1: Correlations between maximum rate measures in tongue twister and DDK (real word to the left and non-word to the right) tasks.



Rates in the two maximum performance speech tasks correlated significantly ($r = .53$*** for tongue twister and DDK real word rate, $r = .50$*** for tongue twister and DDK non-word rate). Accuracy of tongue twister production was not correlated with DDK accuracy: neither for DDK word stimuli ($r = .13$), nor for DDK non-word stimuli ($r = .16$). This lack of association between accuracy levels may be due to limited variability in DDK accuracy (cf. Table 1).

Our last question was whether speakers' maximum rate measures are associated with their habitual sentence reading rate. None of the correlations between rate performance measured in the two maximum performance speech tasks on the one hand and habitual articulation rate on the other reached significance (all r values < .14), suggesting that speech rates which speakers can maximally obtain alternating between similar syllables are not clearly reflected in their habitual sentence reading.

4. DISCUSSION

In this study, we investigated articulatory control in a young adult speaker sample through examining their maximum performance (rate and accuracy) in two speech tasks as indices of articulatory control. More specifically, we used a repetitive syllable-sequence production (DDK) task, which is often used in clinical settings, and a tongue twister task, which is typically used as an experimental means to elicit speech errors in non-clinical populations.

The descriptive statistics show that maximum rate in DDK non-word production and maximum accuracy in tongue twister production were highly variable, even in our homogeneous young and non-clinical speaker group. This variability illustrates that speakers differ considerably in their articulatory control ability.

Our observation of faster DDK performance on real words than nonsense sequences is in line with findings for other languages with school-aged children and older adults [8, 9]. This may suggest that speakers were better able to rapidly move their articulators in the correct manner when they are more familiar with the required motor programmes. Alternative explanations, however, cannot be ruled out. For instance, confounded with lexicality, words in Dutch have lexical stress patterns (and hence involve unstressed syllables that are reduced acoustically) that are lacking in meaningless sequences like 'pataka' or 'katapa'. Additionally, the word sequences also contained short vowels whereas the non-words only consisted of long vowels, which might have contributed to the rate differences observed between real and nonsense words too.

Our second aim was to examine whether speakers' tongue twister performance is related to their maximum articulatory (DDK) performance, given that both tasks require rapid alternation between similar syllables. Maximum speech rates but not accuracy measures in the two speech tasks were correlated. The rate correlation suggests that both tasks contain elements of speakers' ability to plan and execute similar articulatory programmes, despite differences between tasks in terms of difficulty level, length of the speech stimuli, and the amount of linguistic processing involved. This then provides evidence for both tasks tapping articulatory control.

Our third and last aim was to assess whether speakers' habitual articulation rate, as measured with a sentence reading task, is associated with their speech rate on (either of) the two maximum performance measures. In line with patient data [10] and with rate measures of speakers' semi-spontaneous speech [16], maximum rates obtained with neither DDK nor tongue twister production were predictive of speakers' habitual articulation rate. These results highlight a dissociation between maximum and actual performance in speech rate, likely due to differences in task demands.

5. AKNOWLEDGEMENT

This project has received funding from the EU's H2020 research and innovation programme under MSCA GA 675324. We also thank Mirjam Ernestus for her valuable input and comments.

6. REFERENCES

[1] Duffy, J. R. 2013. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. 3rd Edition. St. Louis: Elsevier.

[2] Fletcher, S. 1972. Time-by-count measurement of Diadochokinetic syllable rate. *Journal of Speech Language and Hearing Research*, 15, 763-771.

[3] Wang, Y. T., Kent, R. D., Duffy, J. R., Thomas, J. E., Weismer, G. 2004. Alternating motion rate as an index of speech motor disorder in traumatic brain injury. *Clinical Linguistics and Phonetics*, 18, 57-84.

[4] Goldrick, M., & Blumstein, S. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, 21, 649-683.

[5] McMillan, C. T., Corley, M. 2010. Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117, 243-260.

[6] Acheson, D. J., Hagoort, P. 2014. Twisting tongues to test for conflict-monitoring in speech production. *Frontiers in Human Neuroscience*, 8, 1-16.

[7] Knuijt, S., Kalf, J., van Engelen, B., Geurts, A., de Swart, B. 2017. Reference values of maximum performance tests of speech production. *International Journal of Speech-Language Pathology,* 21, 56-64.

[8] Icht, M., Ben-David, B. M. 2015. Oral-diadochokinetic rates for Hebrew-speaking school-age children: Real words vs. non-words repetition. *Clinical Linguistics and Phonetics,* 29, 102-114.

[9] Ben-David, B., M, Icht, M. 2017. Oral-diadochokinetic rates for Hebrew-speaking healthy ageing population: non-word versus real-word repetition. *International Journal of Language and Communication Disorders,* 52, 301-310.

[10] Ziegler, W. 2002. Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language,* 80, 556-575.

[11] Tsao, Y., Weismer, G. 1997. Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech Language and Hearing Research,* 40, 858-866.

[12] Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/.

[13] Wilshire, C. 1999. The “Tongue Twister” paradigm as a technique for studying phonological encoding. *Language and Speech,* 42, 57-78.

[14] Baayen H. R., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

[15] Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4 (version 1.1-19). *Journal of Statistical Software*, 67, 1-48.

[16] De Jong, N., Mora, J. 2017. Does having good articulatory skills lead to more fluent speech in first and second languages? *Studies in Second Language Acquisition,* 1-13.