Stutter-free and stutter-filled speech signals and their role in stuttering amelioration for English speaking adults

Joseph Kalinowski*a,*, Vikram N. Dayalua, Andrew Stuarta, Michael P. Rastattera, Manish K. Ramib

aDepartment of Communication Sciences and Disorders, School of Allied Health Sciences, East Carolina University, Greenville, NC 27858-4353, USA
bDepartment of Communication Sciences and Disorders, University of North Dakota, Grand Forks, ND-58202-8040, USA

Received 1 June 2000; received in revised form 30 August 2000; accepted 31 August 2000

Abstract

This study examined the power of an exogenously generated stuttered speech signal on stuttering frequency when compared to an exogenously generated normal speech signal. In addition, we examined the specific components of the second speech signal, which might be responsible for the inducement of fluency in people who stutter. Eight males and two females who stuttered participated in this study. Experiment I involved meaningful speech: normal continuous speech, normal interrupted speech, stuttered continuous speech, and stuttered interrupted speech, whereas Experiment II involved vowels and consonants: /a/, /a-i-u/, /s/, /s-sh-f/. The results indicated that stuttered and normal speech signals were equally effective in reducing stuttering frequency. Further, the vowels were more powerful than consonants in inducing fluency for people who stutter. It is suggested that acoustic manifestations of stuttering, rather than a problem, may be a natural compensatory mechanism to bypass or inhibit the ‘involuntary block’ at the neural level.

0304-3940/00/$ - see front matter © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Stuttering; Auditory signals; Auditory feedback; Speech gesture; Choral speech

A second speech signal comes from another speaker in unison, choral, semi-choral, or shadow forms, where the speech signal from the other speaker is temporally either ahead, behind, or in perfect synchronization with the person who stutters. In terms of probability, perfect synchronization of two speech signals is relatively rare. In its electronic manifestations, a second speaker is ‘derived’ via electronic means to produce a signal, which is temporally out of phase and oftentimes shifted in frequency. A substantial number of replicated empirical reports coupled with the near total absence of opposing reports indicate that endogenous alterations of the speech signal output (e.g. prolonged or slowed speech, rhythmic speech, singing, and lipped speech) or exogenous alterations of speech signal input induces relatively fluent speech in people who stutter [6]. Exogenous auditory speech signals in the form of a ‘second speech signal’ (e.g. chorus reading, shadow speech, delayed auditory feedback, frequency altered feedback) [2–6,8,14,15], or speech signals via the visual modality (e.g. visual choral speech) [16], produce more powerful and natural-sounding reductions in stuttering than incongruous non-speech auditory (e.g. masking noise, clicks) [1,8,10,19] or visual (e.g. flashing lights) inputs [17,18]. All exogenous speech signals, either produced by a second speaker or produced via electronic means, should be considered as ‘second speech signals’. In other words, when persons who stutter are speaking in the presence of an exogenous second speech signal, they always receive two speech signals due to the impossibility of removing the bone conducted signal generated by the endogenous speech signal (i.e. the person’s own speech), and the incomplete masking of the air-borne speech signal.

Generally, the reduction in stuttering frequency under alterations of the second speech signal has been attributed to such factors as entrained rhythm [14], distraction [5,6], modified vocalization, and rate reduction [20]. These explanations are based on the assumption that the peripheral manifestations of stuttering (e.g. repetitions, prolongations, and audible struggle behaviors) are the problem. There has been, however, no examination as to the possibility that
stuttering behaviors may be a natural compensatory mechanism for an ‘involuntary block’ at the central level, rather than a manifestation of a mechanistic problem at the peripheral level. If so, then the immediacy and naturalness of the fluency enhancement via the second speech signal may be explained in a dramatically different manner.

In order to investigate these notions, we examined the power of an exogenous stuttered speech signal on stuttering frequency when compared to an exogenous normal speech signal. We used incongruent second speech signals in order to compare the inherently incongruent nature of exogenous stuttered speech to that of the disconsonant fluent speech (in incongruent or disconsonant speech, the second speech signal contains different phonemic material than that read aloud by the participants). If fluency reduction was observed as would be expected in the aforementioned compensatory mechanism model, it would be essential to determine the specific components of the incongruent second speech signal, which might be responsible for the inducement of fluency. Specifically, the natural classification scheme of vowels and consonants were examined in both dynamic and relatively static vocal tract positions as this provides an appropriate continuum for examination.

Ten normal-hearing adults who stutter (eight males, two females, mean = 27.9 years, SD = 9.4) participated in both experiments. Participants did not present with any other speech and language disorders. All participants had a history of therapy, but were currently not receiving any formal therapeutic intervention. Participants read different junior-high level passages of 300 syllables with similar theme and syntactic complexity in both experiments. Based on informal clinical measures, severity of stuttering for the participants in this study ranged from moderate to severe. The two experiments were counterbalanced while the experimental conditions and the passages were randomized. They were instructed throughout the experiment to read at a normal rate and not to use any controls to reduce or inhibit stuttering. In both experiments participants listened to the exogenous auditory speech signals via supra-aural earphones (headphones that completely cover the external ear) at a most comfortable listening level. All participants spoke into a lapel microphone (Sony model 8819) 10 cm from their mouths with an approximate orientation of 0° azimuth and −120° altitude. The microphone output was fed into a video camera (Sony model CCD-TV 75).

Stuttering episodes were calculated from the participants’ videotape recorded passages. Stuttering was defined as word repetitions, part-word prolongations, and/or inaudible postural fixations. Interjudge syllable-by-syllable agreement, as indexed by Cohen’s kappa [9] was 0.79. Intrajudge Cohen’s kappa syllable-by-syllable agreement was 0.81. kappa values above 0.75 represent excellent agreement beyond chance [12].

The means and standard deviations for stuttering frequency as a function of the exogenous auditory speech signals delivered in Experiment I are as follows: (a) NAF (mean = 35.4, SD = 25.11), (b) fluent interrupted (mean = 15.4, SD = 17.11), (c) stuttered interrupted (mean = 14, SD = 15.59), (d) stuttered continuous (mean = 10.4, SD = 10.09), and (e) fluent continuous (mean = 8.6, SD = 13.2). A one-factor-repeated-measure-analysis-of-
variance revealed a significant main effect of exogenous auditory speech signal on stuttering frequency ($F_{(4,36)}=14.35$, Greenhouse-Geiser $P = 0.0004$, $\eta^2 = 0.62$). A post hoc single-d.f. comparison revealed there was a significant reduction in stuttering frequency for all forms of exogenous auditory speech signals relative to NAF ($P < 0.0001$). No statistically significant differences were observed between fluent and stuttered speech signals ($P = 0.76$), or continuous and interrupted speech signals ($P = 0.10$).

The means and standard deviations for stuttering frequency as a function of exogenous auditory speech signals delivered in Experiment II – are as follows: (a) NAF (mean = 37.5, SD = 27.23), (b) /a/ (mean = 10.3, SD = 15.98), (c) /a-i-u/ (mean = 8.2, SD = 11.92), (d) /s/ (mean = 18, SD = 20.49), and (e) /s-sh-f/ (mean = 24.9, SD = 25.53). A one factor repeated measure analysis of variance revealed a significant main effect of the exogenous auditory speech signal on stuttering frequency ($F_{(4,36)} = 17.77$, Greenhouse-Geiser $P = 0.0001$, $\eta^2 = 0.66$). A post hoc single-d.f. comparison revealed there was a significant reduction in stuttering frequency for all forms of exogenous auditory speech signals relative to NAF ($P < 0.0001$). There were also statistically significant fewer stuttering episodes when the exogenous auditory signals were vowel(s) vs. consonant(s) ($P < 0.0001$). Non-significant differences in stuttering frequencies were found between single versus trains of speech signals ($P = 0.40$).

Before this set of experiments, to the best of our knowledge, there has been no empirical documentation that both exogenously generated stutter-filled and stutter-free incongruous speech signals could induce same levels of fluency enhancement in people who stutter. Our results indicated that stuttering frequency was substantially reduced, irrespective of whether the exogenous auditory speech signal were stuttered or stutter free. What are the possible mechanisms that could be involved in the generation of true fluency via the incongruous auditory speech signal presented in the form of a second speech signal? The fluency enhancement produced by a congruent but temporally shifted speech signal might suggest a timing or entraining support mechanism (i.e. choral speech, DAF, FAF). However, asynchronous speech signals (i.e. signals used in this study) produce the same level of fluency enhancement as the signals mentioned above, suggesting some form or type of ‘inhibitory response’. That is, the second speech signal provides the required compensation to aid the system in effectively executing the motor plan and generates fluent speech in persons who stutter. Our data suggests that a sonorant, voiced speech signal, stuttered or fluent, intermittent or continuous, is an effective acoustic mechanism to release the system from the involuntary block at the neural level. However, these second speech signals do not totally eliminate the involuntary block as they only occur in compensation to the formation of a ‘neural block’ that has already occurred. For a second signal to be a prophylaxis to the overt manifestations of stuttering, it must impact the neural system before the block occurs- as is evidenced in identical choral speech.

We also suggest that this data might illuminate the purpose of producing oscillations and sustained vowel gestures by a person who stutters (i.e. repetitions and prolongations). It is hypothesized that the person produces the overt manifestations of stuttering (i.e. repetitions and prolongations), in an attempt to generate an auditory release mechanism from an ‘involuntary block’ in speech execution at the neural level. Simply put the overt manifestations of stuttering are an attempt to compensate at the peripheral level for a loss of control at the central level-albeit a conspicuous compensation. Thus the overt stuttering manifestations are hypothesized to be a form of compensation rather than a problem in itself. The overt manifestations of stuttering appear to be analogous to the role of fever in an infectious disease state. Oftentimes, the fever has been confused as the problem, when in fact it is an inherent compensatory mechanism of the body to a disease state.

Thus, a lack of an appropriate fluency enhancing gesture is hypothesized to be the predominate etiological factor that is exhibited or manifested due to a lack of inhibition on the part of the auditory cortex in assimilating the appropriate plan required for the smooth execution of the speech act. Recent brain imaging procedures have employed choral speech condition to induce fluent speech in adults who stutter and have compared the brain images obtained, to those attained during the stuttering behavior [13,21]. A lack of activation in the auditory areas during the motor planning of stuttered speech was observed, but an essential normalization under the choral speech condition was noted, indicating its inherent fluency enhancing capabilities. Simply put, the second speech signals provide the required inhibition that is essential for the system of a person who stutters, to smoothly execute the speech act. The resultant inhibition via the second speech signal eliminates the system from requiring the production of oscillations or stuttering manifestations (i.e. repetitions and prolongations) that was previously required to produce speech. Thus the resultant speech is almost free of the overt manifestations of stuttering.

It should be noted that a parallel between the manifestations of stuttering and that of expressive aphasia has been identified previously and appears to be support, at least in part in the data revealed in these experiments [7,11]. That is, the stuttering and expressive aphasia may exist on a functional-morphological continuum and have similar central explanations. The power of auditory cueing or priming in both disorders provides an impetus for further investigation of the possible parallels of these disorders.


