Published in final edited form as:

J Acoust Soc Am. 1999 January ; 105(1): 481–490.

Utterance rate and linguistic properties as determinants of lexical dysfluencies in children who stutter

Peter Howell, James Au-Yeung, and Lesley Pilgrim

Department of Psychology, University College London, Gower Street, London WC1E 6BT, England

Abstract

Two important determinants of variation in stuttering frequency are utterance rate and the linguistic properties of the words being spoken. Little is known how these determinants interrelate. It is hypothesized that those linguistic factors that lead to change in word duration, alter utterance rate locally within an utterance that then gives rise to an increase in stuttering frequency. According to the hypothesis, utterance rate variation should occur locally within the linguistic segments in an utterance that is known to increase the likelihood of stuttering. The hypothesis is tested using length of tone unit as the linguistic factor. Three predictions are confirmed: Utterance rate varies locally within the tone units and this local variation affects stuttering frequency; stuttering frequency is positively related to the length of tone units; variations in utterance rate are correlated with tone unit length. Alternative theoretical formulations of these findings are considered.

INTRODUCTION

Traditionally, calculation of speech rate has included pauses and stuttering events in their measurement of the total speech time, and speech rate has been expressed as the number of words or syllables uttered in that period of time divided by the total speech time. Recent researchers have advocated the use of articulation rate (see Kalinowski et al., 1993, 1995, 1996; Kelly and Conture, 1992; Logan and Conture, 1995; Yaruss and Conture, 1995). This rate excludes dysfluent words or syllables from the rate calculations and time spent pausing is left out of the total speech time. These two methods of calculation of rate are termed here “global speech rate” and “global articulatory rate,” respectively.

Rate measures may need further refinement to establish the relationship between speech rate and stuttering. One reason is that speakers vary rate at different points in time. One aim of this article is to investigate whether such “local” variations occur and to see if such local measurement of articulation rate is appropriate in elucidating the relationship between speech rate and stuttering. The local rate is measured by calculating the articulation rate of speech material that is grouped together prosodically (within a tone unit) which often only spans a few words. The second reason for employing local rate measures is in order to allow the relationship between speech rate and linguistic factors to be studied (for the reasons described below).

When global speech rate measures have been made, researchers have been in almost unanimous agreement that reducing speech rate decreases stuttering frequency. Johnson and Rosen (1937), for example, found that the amount of stuttering decreased during slow speech compared with normal speech rate. Other authorities who concur with this conclusion include Perkins et al. (1991), Starkweather (1985), and Wingate (1976). However, a qualifying rider has been made by Kalinowski and his colleagues concerning articulatory rate. They have reported results employing altered auditory feedback that led

them to question whether a slow articulatory rate is always necessary for fluency to improve (Kalinowski et al., 1993, 1995, 1996). It is frequently reported that increased global speech rate has a complementary effect on stuttering frequency. Thus Johnson and Rosen found that subjects had greater difficulty in speaking at fast rates compared to slower rates. Bloodstein (1987) also noted that high speaking rates can result in stuttering. In an experiment on the effects of speech rate on stuttering frequency, Kalinowski et al. (1993) found that when reading at a faster than normal articulatory rate, seven out of nine subjects showed an increase in stuttering frequency. In a later study, however, Kalinowski et al. (1995) found no difference in stuttering frequency between normal and fast rates. In the light of this discrepancy, Kalinowski et al. (1995) concluded that an increase in articulatory rate does not determine stuttering frequency with the same consistency as does a decrease in rate.

The other established determinants of frequency of stuttering are those associated with the linguistic properties of the words. Identification of the linguistic factors that are indicative of the likelihood of stuttering on particular words started with the work of Brown (reviewed in Brown, 1945). In this review, he concluded that four linguistic factors were paramount: (a) Words occurring in early positions in sentences are more likely to be stuttered than those nearer the end. (b) Content words are more likely to be stuttered than function words. (c) Long words are more prone to be stuttered than short ones. (d) Words starting with consonants are stuttered more often than those starting with vowels. There have been three main areas of investigation on linguistic determinants of stuttering since Brown's seminal work: (a) Differences between the way these factors operate in children who stutter compared with adults. (b) Identification of additional factors that need to be added to Brown's list. (c) Attempts to provide a unifying theoretical framework to explain the operation of some or all of the linguistic factors that affect the frequency of stuttering (Au-Yeung et al., 1998; Brown, 1945; Howell et al., in press; Wingate, 1988).

The best documented linguistic indicator that is sensitive to speakers' age is word type. Bloodstein and Gantwerk (1967) and Bloodstein and Grossman (1981) have found that children who stutter experience more difficulty on function words than on content words. Recently, Howell et al. (in press) compared frequency of stuttering directly for different age groups. They found that frequency of stuttering of function words was significantly higher than on content words for age groups up to nine years but not for older age groups.

One of the additional linguistic factors that has been investigated in detail since Brown's work is sentence length (Gaines et al., 1991; Logan and Conture, 1995; Tornick and Bloodstein, 1976; Weiss and Zebrowski, 1992). Logan and Conture (1995), for example, have investigated this factor in children who stutter. They found that frequency of stuttering was higher on long sentences than on short ones. This factor was noted to occur for adults who stutter in an early study by Brown and Moren (1942) but did not appear in his 1945 list of principal determinants.

Previous attempts at supplying a unifying framework have concentrated on explaining the linguistic determinants listed by Brown (1945). Brown himself considered that the four linguistic factors he identified were all associated with semantic difficulty. Wingate (1988) argued that stress could explain how all these same factors operate in these speakers. Although these two explanations may account for stuttering in adults, it is less clear how they would apply to children who stutter. Thus Brown's argument for a semantic association cannot explain why children have difficulties on the semantically simpler function words observed in these speakers. Similarly, Wingate's emphasis on the role of stress is less applicable to children's speech due to the prevalence of dysfluency on function words which are rarely stressed.

The principle reason why speech rate has not been considered along with linguistic determinants within such unifying frameworks is that analyses of the influence of rate have usually been made globally whereas linguistic determinants are assessed at a local level in an utterance. Thus the greatest extent investigated in Brown's (1945) work was when he examined word position effects using sentences as the unit of analysis. Restricting attention to short linguistic segments has the advantage that it allows the words or linguistic contexts that lead to stuttering to be specified with precision. Increased precision would be a sufficient reason on its own to motivate investigating whether speech rate influences occur at a local level within extended utterances. However, there are also reasons to suppose that it is misleading to base conclusions about how rate affects stuttering frequency on long utterances, even utterances the length of a sentence. The grounds for this follow.

As stated earlier, in continuous spontaneous speech, rate variations occur between constituent stretches even though speakers receive no instructions to do this. There is surprisingly little information how such local rate variations are made in speech. One thing tested in the below experiment is whether words within fast spoken groups have the highest tendency to be stuttered. When speakers make global rate changes, this might or might not be expected to lead to changes in frequency of stuttering. This can be illustrated by considering two ways which contrast in what happens to stuttering frequency depending on how global rate increases are effected at a local level. The first possibility is that a speaker increases rate proportionately for all the utterance. The parts of the utterance that are spoken at fast, medium, and slow rates would tend to change with slow local stretches becoming medium rate, medium rate local stretches becoming fast, and fast stretches becoming faster still. In this case, the global rate increase would be expected to increase the frequency of stuttering due to the increased number of the problematic fast local stretches. The second possibility arises if speakers only have flexibility to increase rate in local stretches that are spoken at a slow rate. When the rate of these slow stretches is increased, they would become medium rate. In this situation, there would be no increase in frequency of stuttering with increased speech rate as no additional local stretches are spoken sufficiently fast. The different findings about the effects of speech rate increases on stuttering frequency noted in the studies of Kalinowski et al. (1993, 1995) may have occurred because speakers in the two studies used these different ways of speeding speech up. It is possible that, by chance, the speakers in Kalinowski et al. (1993), who showed an increase in stuttering frequency with rate, may have increased rate proportionately throughout the utterance, while the subjects in Kalinowski et al. (1995) may have just accelerated the slow stretches.

Alternative ways of changing speech rate are not investigated in the current study but they do have methodological implications for the way the study was conducted. They show that it is necessary to investigate the association between stuttering frequency and rate changes at local levels in utterances as misleading conclusions can be drawn if only global measures are investigated. That is, global rate changes can occur without affecting stuttering frequency even though local rate changes influence the frequency of stuttering. The current experiment shows that the variations in stuttering frequency that occur in groups of words within the same utterance depend on local speech rate.

In the experimental procedure adopted below, the recorded speech is segmented into regions spoken at different rates. The segmental unit that is employed is the tone unit. These segments, by definition, reflect the prosodic grouping imposed on an utterance by a speaker whether the material is read or spoken spontaneously (Crystal, 1969, 1987). Tone units consist of a minimum of one syllable and its boundaries are marked by the prosodic properties of a change in pitch and a junctural feature (usually a very slight pause). Every tone unit has a coherent intonation and one accented nucleus syllable (Kreidler, 1997). An example of tone unit segmentation within a sentence would be, "If the delivery man comes ll

give me a shout” (where || marks the boundary between tone units). A typical reading of this sentence would be to speak the words in the first tone unit relatively rapidly and those in the second more slowly which a speaker might adopt so as to emphasize the instruction. This example shows that tone units can be produced at different rates from one unit to the next. It was assumed in the preceding paragraph that words grouped together prosodically that are spoken at a fast rate would be the ones prone to contain stutterings. To establish this, tone units are classified into different rates and examined to see whether the groups of tone units spoken at different rates have any influence on frequency of stuttering.

Spontaneous speech was used in the current study while most previous studies on speech rate employed read material. Choice of this type of material may appear surprising as read material is more convenient to deal with experimentally. For example, it is straightforward to get speakers to change their speech rate by telling them to read faster. Read speech also has the advantage that it allows sentences to be used which are also commonly employed in studies of linguistic factors. Consequently, comparisons between rate and linguistic factors could be made directly if sentences were used. A procedure which involved reading sentences was avoided here for the primary reason that the rate changes that occur in read material are unnatural. The reason for this is that reading does not involve certain on-going processes like utterance formulation which are likely to have important roles in determining speech rate. Thus it is not clear whether variation in speech rate enforced by an experimenter has similarities to the mechanisms of rate variation that speakers use spontaneously in everyday speaking situations. It is also known from empirical evidence that the two modes of delivery (read and spontaneous) lead to differences in prosodic organization. For instance, Howell and Kadi-Hanifi (1991) showed speakers do not produce the same tone unit boundaries in speech they produced spontaneously compared with those they produce where they were required to read the same material at a later date.

The speech of children around nine years old is used in the analyses. Nine years was chosen for two reasons. First, children who are stuttering at this age are more likely to persist than younger ones (Andrews and Harris, 1964). Thus at this age, it has been more definitely established that these speakers have a stuttering problem. Second, secondary idiosyncratic influences characteristic of long-term persistent stuttering that may affect utterance rate measures at older ages are less likely to have appeared in speakers of this age. The types of stuttering incidents that are counted are what Howell et al. (1997) refer to as lexical dysfluencies. These consist of word and part-word repetitions, prolongations, and broken words (the latter category is referred to by the more neutral term “other” dysfluencies). Supralexical dysfluencies that extend over groups of words (such as phrase repetitions) can be split up over more than one tone unit.

The hypothesis about how rate changes within spontaneous speech and how these rate changes are associated with the linguistic determinants that affect the likelihood of stuttering is as follows. The linguistic characteristics associated with stuttering change speech rate. When speech rate increases, stuttering is more likely to occur. In the current experiment, the length of tone units is the linguistic property under investigation which would affect the utterance rate. As noted earlier, it has been reported previously that stuttering frequency increases when sentence length increases (Brown and Moren, 1942; Gaines et al., 1991; Logan and Conture, 1995; Tornick and Bloodstein, 1976; Weiss and Zebrowski, 1992). In addition, it is known that utterance rate increases when utterance length increases in fluent speakers (Ferreira, 1993; Levelt, 1989). A check is first made as to whether utterance rate and stuttering frequency covary within tone units in spontaneous speech in the same way that they do in sentences in read speech. Next, analyses are made to ascertain whether tone unit length in spontaneous speech influences stuttering frequency in a similar way to that which occurs in sentences in read speech. The two preceding analyses do not rule out the

possibility that utterance rate and tone unit length operate independently. Therefore, the data are examined in the final analysis to assess whether there is a correlation between tone unit length and utterance rate. To obtain a satisfactory measure of utterance rate for this purpose, rate is measured only on nonstuttered syllables in a tone unit (Hargrave et al., 1994).

I. METHOD

A. Subjects

The data used for speech assessments are from eight children who stutter. The children ranged in age from 9 to 11 years old. There were seven boys and one girl. The children had been assessed by therapists and admitted to a 2 week intensive therapy course. Detailed information on individual speakers is provided in Table I.

B. Materials and apparatus

Recordings were made at the commencement of the intensive therapy course (i.e., before they received treatment). The children were asked to speak in monologue for about 2 min on any topics they liked, for example, hobbies, friends, holidays. The children talked continuously during this period, which is probably due to them having free choice about topic. Recordings of the children were made in an Amplisilence sound-attenuated booth (Howell et al., 1998). The speech was transduced with a Sennheiser K6 microphone positioned 6 in. in front of the speaker in direct line with the mouth. The speech was recorded on DAT tape and transferred digitally to computer for further processing. The speech from the DAT tapes was down-sampled to 20 kHz.

C. Transcription and segmentation procedures

The speech was first transcribed in orthographic form with lexical dysfluencies not marked. These were made independently by two judges. The experimental judge had eight years experience transcribing stuttered speech. The second judge was used for assessing the reliability of this and all other aspects of assessment and had two years experience transcribing stuttered speech but was naive as to the purpose of the experiment. Agreement between the two judges for the orthographic transcriptions was high at 96%, giving a Kappa coefficient of 0.92 which is much higher than chance (Fleiss, 1971). The version of the transcription produced by the more experienced judge was used in the next stages of the procedure.

The end-points of the words in this transcription were located. Using end-points ensured that any pauses or word-initiation attempts were included at the start of the subsequent word (for the purpose of this experiment, word repetitions were also treated as word-initiation attempts). The position where each word ended was located with the aid of two traveling cursors that were superimposed on the oscillographic display of the speech waveform. Each cursor position was independently adjusted with a mouse. The first cursor was initially placed at the start of the speech. The second cursor was positioned slightly before where the end of the first word was thought to be located. The section of speech between the first and second cursors was then played and listened to over an RS 250-924 headset, to check whether the end of the word had been correctly located. The second cursor was then adjusted further forward and the word replayed again. Once the judge had adjusted past the end of the word, he adjusted the second cursor back in time into the file until the end of the word was not heard. By continually crossing backward and forward across the boundary at the end of the word, he became satisfied that he had located it accurately. This marker was stored and the subject proceeded to locate the boundary between the next words in the sequence in the same way (this procedure is similar to that used by Osberger and Levitt, 1979). Once the judge was happy about the position of the end-point cursor, this marker was

stored and the boundary between the next words in the sequence was located in the same way. Silent pauses not associated with stuttered words were located using the same procedure. Filled pauses were treated as words both here and when the words were categorized. Both the experimental judge and the judge used for assessing reliability worked independently through the recordings of each speaker sequentially. Judge 1 was about as likely to place his marker in advance of Judge 2 (49.2%) as the reverse (48.7%). When this happened, the differences were small (3.9% of the duration of a fluent word and 4.6% of a dysfluent word). Thus the placement of the word end-point markers is consistent between judges and, therefore, the markers of the experienced judge were used in the next phase of the procedure.

In this next phase, the two judges used a similar procedure to that employed to locate word end-points to establish words at tone unit boundaries. The inter-judge agreement about tone unit boundaries was calculated as the average of tone units identified by the first judge and agreed by the second, and vice versa. Agreement about tone-unit segmentation was 83% and the Kappa coefficient was 0.66 which is considered as a good reproducibility (Fleiss, 1971).

D. Assessment procedures for dysfluencies?

Fluency assessments were made using words as defined previously (that is, extending from the interval from the end of one word to the end of the next) using the word end-point markers produced by the more experienced judge again. Each word in the speech of all eight subjects was assessed. A randomized presentation order was used so that the global context in which all judgments were made was as constant as possible (thus minimizing range effects; Parducci, 1965). The random sequence was obtained computationally. The first randomly selected word was heard in isolation. After a short pause the test word was heard along with the word that had preceded it (the two words had the same timing and were in the same order as in the original recording). Consequently, pauses were apparent when they occurred between the context and test word. The test word alone and this word with the word that preceded it could be heard as often as the judge required by hitting the return key on the computer keyboard. Thus presentation of the test and context words was initiated by the judge, and the current trial ended and the next trial commenced after the judge entered his responses (detailed in the following paragraph). The full context was not available at the beginning of the recording (in these cases, the word to be judged was still played in isolation beforehand).

Assessments obtained about each word were a rating about how comfortably the word “flowed” and a categorization of the word as fluent, prolongation, repetition of word or part-word, or other dysfluency. The flow judgments were included as a classificatory response does not allow a judge to indicate anything about the relative fluency or dysfluency within a category. The flow judgments were made after the categorization on a 5-point scale. The 5-point scale employed a Likert scale format (Likert, 1932) in which the judge indicated the extent to which he agreed with the statement that “the speech is flowing smoothly” (1=agree, 5=disagree with intermediate values showing levels of agreement). Thus it represents a judge's assessment of the speaker's ability or inability to proceed with speech (Perkins, 1990). It was stressed that the rating scale was not a finer-grained indication of whether a word is fluent or dysfluent: A word high in “flow” might nevertheless be categorized as fluent or vice versa. Howell et al. (1997) have shown that inter-judge agreement is high for repetitions. Agreement about prolongations is also satisfactory provided that words called prolongation and given ratings of 4 or 5 are selected (as in the below analyses). The inter-judge agreement level for location of dysfluent words where both judges agreed that a word was dysfluent was 85.6% out of all words identified dysfluent by either judge (Kappa's coefficient is 0.71 which is considered as a good level of agreement).

E. Classification of utterance rate of each tone unit

As just described, fluency assessments were based on words. Utterance rate measures, on the other hand, were always made on syllables as they are less variable in duration than words. First, the duration of each tone unit was computed by removing all dysfluent words and pauses between words (Hargrave et al., 1994). The number of syllables in all fluent words in a tone unit was then divided by the duration of the tone unit after dysfluent words had been removed to produce the stutter-free utterance rate (syllables/second) of that tone unit. Utterance rates of 4–5 syllables/s were classified as medium [based on fluent speakers' speech rate found by Pickett (1980) and concurred by Kalinowski and his colleagues]. Utterance rates above and below the 4–5 syllables/s were classified as fast and slow, respectively. In a second analysis, the utterance rate was classified as fast, medium, or slow depending on the individual speaker's utterance rate. The mean and s.d. of utterance rate of all words were calculated for each speaker. The fast rate was defined as the utterance rate more than one s.d. above the mean rate while the slow rate was defined as lower than one s.d. below the mean utterance rate. The rest were classified as medium rate.

II. RESULTS

The hypothesis that there is a greater likelihood of stuttering on tone units which are spoken fast compared to those spoken at a medium or slow rate was tested. In the first analysis, the division of utterance rate was carried out using the 4–5 syllables/s criteria. For each speaker, all the tone units at a selected rate were located and the number of words across all these tone units was calculated. The number of these words that were stuttered was then determined. Stuttering rate was expressed as the percentage of the number of stuttered words out of the total number of words in tone units of a particular utterance rate. This provides a measure of the per word stuttering rate of words in that rate division of tone units. It is unaffected by differences in tone unit length between the utterance rate divisions. A one-way ANOVA was performed on the stuttering rates at different utterance rate (with the overall stuttering rate of the speakers taken out as covariate) to see if there was a relationship between utterance rate divisions and, the dependent variable, stuttering rate. Words in fast tone units had a significantly higher chance of being stuttered than tone units at slower utterance rates, $F(2,20)=4.36$, $p<0.05$. The adjusted stuttering rates are plotted as boxplots for each utterance rate in Fig. 1 (the caption to the figure describes the convention used in boxplots). *Post hoc* Tukey test showed that the tone units of fast utterance rate had significantly higher stuttering rate than those of the slow utterance rate.

In order to test if fast speech rate alone increased the likelihood of stuttering, the medium and slow tone units were grouped together and tested against the fast tone units. Words in the fast tone units were found to be stuttered more when the stuttering rate of individual speakers were taken out as covariate, $F(1,13)=15.03$, $p<0.005$, in a one-way ANOVA test with speech rate as the independent variable. This confirmed that fast speech rate did have higher level of stuttering than the remaining rate divisions.

In the second analysis, the utterance rate was defined for individual speakers using the mean and s.d. of utterance rate of all words as described above. A one-way ANOVA was performed on the stuttering rates at different utterance rates (with the overall stuttering rate of the speakers taken out as covariate). Words which were classified as fast had a significantly higher chance of being stuttered than tone units at slower utterance rates, $F(2,20)=4.07$, $p<0.05$. The adjusted stuttering rates are plotted as boxplots for each utterance rate in Fig. 2. A *post hoc* Tukey test showed that the tone units of fast utterance rate has significantly higher stuttering rate than those of the medium utterance rate. Once again, the medium and slow tone units were grouped together to test against the fast tone units similar to the first analysis. Words in the fast tone units were found to be stuttered more when the

stuttering rate of individual speakers was taken out as covariate, $F(1,13)=6.58$, $p<0.05$, in a one-way ANOVA test with speech rate as the independent variable. This again confirmed that fast speech rate had a higher level of stuttering.

Next, the hypothesis that words in longer tone units have a higher chance of being stuttered was tested. Tone unit length was measured both in words and syllables. The first step in obtaining the length of tone units in words was to divide tone units into those that contained one, two, three, four, and more than four words. For each speaker, all the tone units of a given length were located and the words in them summed to obtain the total. The number of words out of this total that were stuttered was then determined. The stuttering rate for tone units of the specified length was expressed as the percentage of stuttered words out of the total number of words for the tone units of that particular length. A one-way ANOVA (with the overall stuttering rate of the speakers taken out as covariate) showed that stuttering rate differed significantly over tone units of different word lengths [$F(4,34)=3.04$, $p<0.05$]. The stuttering rates against tone unit length in words are presented as boxplots in Fig. 3. Tone unit length was measured on syllables in a similar way as with words. Stuttering rate varied across tone units of different length measured in syllables, in a similar way to that found with words [$F(4,34)=8.00$, $p<0.001$]. These data are plotted in Fig. 4 in the same way as the word data (which appeared in Fig. 3).

The third analysis examined whether utterance rate and tone unit length were correlated. Tone unit length and utterance rate correlated (linearly) significantly both when tone unit length was measured in words ($r=0.500$, $p<0.001$) and in syllables ($r=0.568$, $p<0.001$). The quadratic correlations were also significant ($r=0.557$, $p<0.001$ and $r=0.633$, $p<0.001$, respectively). Cubic regressions were also carried out and not much improvement was found from the quadratic ones, $r=0.558$, $p<0.001$, and $r=0.643$, $p<0.001$, respectively. This suggests that rate asymptotes in long tone units. Boxplots showing the correlation between raw tone unit length and utterance rate scores are given for the case where tone unit length was measured in words in Fig. 5 and when measured on syllables in Fig. 6.

III. DISCUSSION

In the analysis using fixed criteria for dividing utterances into fast, medium, and slow rate, the fast rate produced a higher stuttering rate than the slow rate. Medium rate utterances had lower frequency of stuttering than fast rates, but not significantly so. When criteria based on the individual's speech rate were employed, fast rate speech produced a significantly higher rate of stuttering than medium rate speech. Slow rate speech had lower rates of stuttering than fast rate, although not significantly so. The common factors in these different rate analyses are: (a) The significant effect in these analyses both involve the fast rate division which produced higher rates of stutters than other rate partitions. (b) In rate partitions that do not differ significantly in frequency of stuttering from the fast one, stuttering rates were always lower. We conclude, therefore that the hypothesis that there is a greater likelihood of stuttering on tone units which are spoken rapidly was supported. These tone units are interspersed with other tone units spoken at different rates. The findings suggest that not all stretches of speech in an utterance are equally problematic for speakers. A local utterance rate measure allows the stretches that are likely to be difficult to be specified in part. The two subsequent analyses addressed the question whether linguistic factors, known from previous studies to operate at a local level in an utterance, bring about rate changes that then lead to stuttering. One linguistic factor that could change utterance rate is utterance length. It was first shown that long tone units have a greater likelihood of stuttering than short ones, as has been reported to occur in speech read by children (Logan and Conture, 1995). It was then shown that the length of nonstuttered tone units was correlated with speech rate. A similar correlation has been established between speech rate and tone unit length for fluent

speakers of English (Ferreira, 1993). Two questions need considering with respect to these findings. (a) Why should utterance rate increases lead speakers to be dysfluent? (b) Why should long utterances be spoken more rapidly than short ones? Some proposed answers to these questions will be developed dealing, in particular, with the level at which the effects arise. This will be followed by an attempt at specifying how utterance length and utterance rate interact leading to dysfluency.

First, two potential ways of explaining why utterance rate increases frequency of stuttering are considered. In both these accounts, a distinction is drawn between planning a sound and its execution (Levelt, 1989). The two views are differentiated by whether rate-dependent speech dysfluencies occur through the operation of the planning or the execution processes. The first possibility is a development of Postma and Kolk's (Kolk and Postma, 1997; Postma, 1997; Postma and Kolk, 1992, 1993) Covert Repair Hypothesis (CRH). CRH is a model that maintains stuttering arises out of repair processes (Levelt, 1983) that occur in fluent speakers' speech as well as in the speech of people who stutter. Repairing appears to be taking place in an utterance like "turn left at the, no, turn right at the crossroads" where the speech indicates that the speaker realizes he or she has given the wrong direction and alters it. CRH maintains that stuttering is a by-product of repairing covert, rather than such overt, phonemic errors in the speech plan before it is executed. CRH proposes that the errors occur as a result of a defect in the phonological planning process. In Postma and Kolk's account, speakers replan the speech when they detect that such an error has occurred. Such a replanning process is hypothesized by Postma and Kolk and remains to be proven. The model could be applied to the rate data reported in the current experiment. It would simply be necessary to make the reasonable assumption that the covert errors that require replanning are more likely to occur when utterance rate is high.

Blackmer and Mitton (1991) have questioned whether dysfluencies that fluent speakers produce during self-repairs allow sufficient time for replanning. Their data were from recordings of speech errors and repairs made by adults to late night radio shows. They found that speakers were able to interrupt themselves extremely rapidly and would then repeat words they had spoken immediately prior to the interruption. Elsewhere, Clark and Clark (1977) have proposed that interruptions and repetitions, similar to those in the examples Blackmer and Mitton give, occur because the plan for a subsequent word or words is not ready. Speakers have the plan for words spoken previously and these can be automatically re-executed (i.e., without replanning them). The time during which the repetition occurs allows further time for planning to be completed. It would be expected that the process just described would be prone to happen more when speech rate is high, as rapid rates would increase the chance that the plan for the next word is unavailable. Note that the role of repetition is to allow time for completing the plan of a later word and is a stalling tactic, not a malfunction in central nervous system structures responsible for planning speech. Also, an erroneous plan has not been produced for the word that initiates the stalling, more time to complete the plan is just needed. Thus neither the sequence of repetitions nor the later word involve planning errors. The occurrence of dysfluencies arises out of strategic responses speakers make during execution. Blackmer and Mitton's (1991) account offers a critique of the notion that replanning takes place during rapidly produced repetitions associated with speech repairs produced by fluent speakers. As well as the critical side of their argument, they propose an execution level strategy to explain the surface form of rapid repetitions. Au-Yeung and Howell (1998), Au-Yeung et al. (1998), and Howell et al. (in press) have shown how a process based on Blackmer and Mitton (1991) and Clark and Clark (1977) could explain features of stuttering behavior.

Two further comments about the execution level account of rate-induced dysfluencies are necessary: First, when applied as an account of all classes of dysfluency, the execution level

explanations make predictions about activity in different sites in the brain. The execution level account proposes that there are no fundamental planning-level differences between fluent speakers and people who stutter though there would be at the levels involved in execution. Consequently, this view would predict that activity differences between fluency groups are more likely to be seen in the motor centers than in the areas attributed a role in language formulation. Even in the strong form of the execution level account, errors during planning would be allowed to occur in the speech of people who stutter. However, the prediction that there are no brain process differences between speaker groups due to replanning utterances would still apply providing that the propensity for errors that require replanning, occurs equally often between the speaker groups. The second comment concerns what descriptors are appropriate to account for difficulties arising from plan unavailability. CRH was developed from a linguistic perspective and describes difficulty at an abstract phonological level. The execution level account has continued this tradition using, for example, terminology from the repair literature and linguistic-based typologies of stuttered dysfluencies. However, since the problem of stuttering is deemed to be tied up with execution-level strategies, motoric descriptions would be more appropriate. The translation is straightforward as linguistic characterizations which specify phonological difficulty are directly related to motoric descriptions.

The second question mentioned was why should long utterances be spoken faster than short ones? The more specific question will be addressed about the level at which utterance length dependencies on stuttering arise. One possibility is that all the linguistic factors that lead to variations in frequency of stuttering emerge during planning. Since, as argued above, rate variations associated with stuttering appear through the operation of execution levels, linguistic factors that appear during planning would have supremacy. According to this account, the first step in utterance formulation is to decide on the words that will appear in a tone unit. Once the lexical content of a tone unit has been decided, the duration of individual words and syllables could arise through contextual influences operating in the tone unit. Empirical evidence suggests that planning that leads to long tone units being produced changes the duration of the constituent syllables. Thus if the stressed syllable duration in a word list is taken as the normal (or baseline) duration for a syllable in that particular word, compression of the syllable occurs when the syllable is uttered in the same word in connected speech (referred to by Levelt, 1989, as the word-in-isolation effect). Rate changes and an increase in frequency of stuttering would then follow from the change in duration.

A factor that commends this account (referred to as duration compression theory) is that it could also account for how some of Brown's (1945) other linguistic factors operate. A circumstantial interpretation can be made that duration compression occurs for different word positions in tone units. Thus Nakatani et al. (1981) have shown that, in fluent speakers, word durations are shortened when a stressed syllable occurs in tone unit-initial position than when it appears at later positions (Nakatani et al., 1981). Syllable compression could also potentially explain the word length effect. Thus Lehisté (1970) has reported that the duration of a syllable in a long word is significantly shorter than when the same syllable occurs in a shorter word in fluent speakers' speech. Nooteboom (1972; cited by Levelt, 1989) showed that the duration of a syllable decreases the more syllables that follow it in a word. This could potentially explain why stuttering occurs almost exclusively on phones in early positions in words (Wingate, 1988). There does not seem any easy way of accounting for the word-initial consonant effect by duration compression. If this theory is correct, this factor may have a different basis to the others discussed probably because difficulty arises due to words having this factor being phonologically more difficult.

The duration compression account developed above has the major drawbacks that it relies on observational data and circumstantial arguments to account for the relationship between

linguistic factors, speech rate and frequency of stuttering. Saying that the effects of duration are planned has little explanatory value. Moreover, the durational effects could well arise during the execution of the plan rather than being inherent specifications of the plan itself. The duration-compression approach would be more convincing if some superordinate factor could be found which explains why duration compression is associated with the different linguistic characteristics where the level at which it operates (planning or execution) can be deduced. One possibility is that duration compression factors in the plan could be mediated by longer lexical access time (time to retrieve the mental lexicon) in people who stutter (Prins et al., 1997) and that lexical access time is influenced by Brown's factors. Thus lexical access would have to take longer when words are in, for instance, long tone units than short ones. This view can be contrasted with one in which duration compression is mediated by latency differences attributed to lexical access time but which derive from different times required to execute words with different forms and in different contexts known to affect frequency of stuttering.

Data to support the point of view that people who stutter show differences in lexical access time are not extensive and there are no data of speakers at the ages of 9–11 we tested (Prins et al., 1997; Van Lieshout et al., 1996a, 1996b). Furthermore, there is disagreement between the two principal groups of researchers about how to interpret these results. The main dispute between the groups is whether lexical access time differences can be accounted for at central (Prins et al., 1997) or peripheral levels (Van Lieshout et al., 1996a, 1996b). We have examined Prins et al.'s study elsewhere and shown that the latency differences interpreted as revealing differences in lexical access time of verbs between fluent speakers and people who stutter, could occur during speech-execution rather than while these words are being planned (Au-Yeung and Howell, submitted). The latency differences may reflect, then, more about execution times of phonologically difficult words rather than lexical access time per se.

Considering the evidence on these two questions together, some tentative conclusions can be drawn about how speech rate and linguistic factors interact. First, increases in stuttering frequency associated with increased utterance rate appear to arise during execution rather than planning. Although it is possible that duration differences associated with Brown's factors occur during planning, there is no compelling reason for this conclusion unless the process through which these effects arise can be specified. Lexical access time differences associated with words with these linguistic characteristics is one possible way of supplying such a process. Although latency differences that are supposed to indicate that people who stutter have slow access to verbs (instances of content words) and conceivably also for words that have other of Brown's factors, the latency differences between fluency groups can be explained by differences arising during execution of these word forms. Taking all this evidence together, it appears that people who stutter can be differentiated from fluent speakers in the way they execute speech, in particular, because they speak too fast in specified contexts this then precipitates stuttering. The contexts that are likely to be stuttered when spoken too fast are those in which the linguistic or motor load is high. At present, no metric has been definitively established to measure this. EMG work on fluent speakers (Van Lieshout et al., 1995) and work identifying phonological properties indicating high motor load in people who stutter (Howell et al., submitted) may provide such indices in the future. Linguistic factors and utterance rate operate interactively at the execution level leading to dysfluencies. This interaction is strategic rather than hard-wired as shown by the fact that fluency is reversible in various ways for a time when speech rate is artificially controlled. Two important questions that remain to be answered are why does fluency not stick when it is induced and why do some children stay dysfluent if dysfluent speech patterns are merely strategic?

In this study, only two factors were considered that trigger stuttering: local utterance rate and tone unit length. There are other linguistic factors that interact with tone unit length such as syntactic, prosodic, phonetic structures. The interaction of these factors with utterance length will be investigated in further studies. These observations mean that the explanation offered here is by no means exhaustive. We consider we have supplied a possible way that the linguistic utterance being planned leads to a high motor load at the execution/motor level. Further work is needed to separate between the motor/execution account versus accounts that maintain the problem arises in speech planning processes (such as the proposal that stuttering arises out when covert repairs to plans are made).

Acknowledgments

We wish to express our thanks to the reviewers, Dr. Joe Kalinowski, Dr. Pascal van Lieshout, and Dr. Anders Lofqvist, the associate editor of JASA, for their comments on this paper. This research was supported by a grant from the Wellcome Trust.

References

- Andrews, G.; Harris, M. *The Syndrome of Stuttering*. London: Heinemann; 1964.
- Au-Yeung J, Howell P. Lexical and syntactic context and stuttering. *Clinical Linguistics and Phonetics*. 1998; 12:67–78.
- Au-Yeung J, Howell P. Lexicalization and stuttering: Comments on Prins. Main and Wampler (1997). (submitted).
- Au-Yeung J, Howell P, Pilgrim L. Phonological words and stuttering on function words. *J. Speech, Language and Hearing Research*. 1998; 41:1019–1030.
- Blackmer ER, Mitton JL. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*. 1991; 39:173–194. [PubMed: 1841032]
- Bloodstein, O. *A Handbook on Stuttering*. 4th ed.. Chicago: National Easter Seal Society; 1987.
- Bloodstein O, Gantwerk BF. Grammatical function in relation to stuttering in young children. *J. Speech Hear. Res.* 1967; 10:786–789. [PubMed: 5586944]
- Bloodstein O, Grossman M. Early stutterings: Some aspects of their form and distribution. *J. Speech Hear. Res.* 1981; 24:298–302. [PubMed: 7265947]
- Brown SF. The loci of stutterings in the speech sequence. *J. Speech Disorders*. 1945; 10:182–192.
- Brown SF, Moren A. The frequency of stuttering with relation to word length during oral reading. *J. Speech Disorders*. 1942; 7:153–159.
- Clark, HH.; Clark, E. *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt; 1977.
- Crystal, D. *Prosodic Systems and Intonation in English*. Cambridge, MA: Cambridge U.P.; 1969.
- Crystal, D. *Clinical Linguistics*. London: Edward Arnold; 1987.
- Ferreira F. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory Language*. 1993; 30:210–233.
- Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol. Bull.* 1971; 76:378–382.
- Gaines N, Runyan CM, Meyers SC. A comparison of young stutterers' fluent versus stuttered utterances on measures of length and complexity. *J. Speech Hear. Res.* 1991; 34:37–42. [PubMed: 2008079]
- Hargrave S, Kalinowski J, Stuart A, Armson J, Jones K. Effect of frequency-altered feedback on stuttering frequency at normal and fast rates. *J. Speech Hear. Res.* 1994; 37:1313–1319. [PubMed: 7877290]
- Howell P, Au-Yeung J, Sackin S. Exchange of stuttering from function words to content words with age. *Journal of Speech, Language and Hearing Research*. (in press).
- Howell P, Au-Yeung J, Sackin S. Internal structure of content words leading to lifespan differences in phonological difficulty in stuttering. (submitted).

- Howell P, Kadi-Hanifi K. Comparison of prosodic properties between read and spontaneous speech. *Speech Commun.* 1991; 10:163–169.
- Howell P, Sackin S, Glenn K. Development of a two-stage procedure for the automatic recognition of dysfluencies in the speech of children who stutter. I. Psychometric procedures appropriate for selection of training material for lexical dysfluency classifiers. *Journal of Speech, Language and Hearing Research.* 1997; 40:1073–1084.
- Howell P, Staveley A, Sackin S, Rustin L. Methods of interval selection, presence of noise and their effects on detectability of repetitions and prolongations. *J. Acoust. Soc. Am.* 1998; 104:3558–3567. [PubMed: 9857514]
- Johnson W, Rosen L. Effect of certain changes in speech pattern upon frequency of stuttering. *Journal of Speech Disorders.* 1937; 2:105–109.
- Kalinowski J, Armson J, Mieszkowski M, Stuart A, Gracco VL. Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech.* 1993; 36:1–16. [PubMed: 8345771]
- Kalinowski J, Armson J, Stuart A. Effect of normal and fast articulatory rates on stuttering frequency. *Journal of Fluency Disorders.* 1995; 20:293–302.
- Kalinowski J, Stuart A, Sark S, Armson J. Stuttering amelioration at various auditory feedback delays and speech rates. *European Journal of Disorders of Communication.* 1996; 31:259–269. [PubMed: 8944848]
- Kelly EM, Conture EG. Speaking rates, response time latencies, and interrupting behaviors of young stutterers, nonstutterers, and their mothers. *J. Speech Hear. Res.* 1992; 35:1256–1267. [PubMed: 1494271]
- Kolk, H.; Postma, A. Stuttering as a covert repair phenomenon. In: Curlee, RF.; Siegel, GM., editors. *The Nature and Treatment of Stuttering: New Directions.* 2nd ed.. New York: Allyn & Bacon; 1997.
- Kreidler, CW. *Describing Spoken English: An Introduction.* London: Routledge; 1997.
- Lehiste, I. *Suprasegmentals.* Cambridge, MA: MIT; 1970.
- Levelt WJM. Monitoring and self-repair in speech. *Cognition.* 1983; 14:41–104. [PubMed: 6685011]
- Levelt, WJM. *Speaking: From Intention to Articulation.* Cambridge, MA: MIT; 1989.
- Likert R. A technique for the measurement of attitudes. *Archives of Psychology.* 1932; (140)
- Logan KJ, Conture EG. Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. *Journal of Fluency Disorders.* 1995; 20:36–61.
- Nakatani LH, O'Connor JD, Aston CH. Prosodic aspects of American English speech rhythm. *Phonetica.* 1981; 38:84–106.
- Nooteboom, S. *Production and perception of vowel duration.* Utrecht University; 1972. Unpublished doctoral dissertation
- Osberger MJ, Levitt H. The effect of timing errors on the intelligibility of deaf children's speech. *J. Acoust. Soc. Am.* 1979; 66:1316–1324. [PubMed: 500969]
- Parducci A. Category judgement: a range-frequency model. *Psychol. Rev.* 1965; 17:9–16.
- Perkins WH. What is stuttering? *J. Speech Hear. Res.* 1990; 55:370–382.
- Perkins WH, Kent R, Curlee R. A theory of neuropsycholinguistic function in stuttering. *J. Speech Hear. Res.* 1991; 34:734–752. [PubMed: 1956181]
- Pickett, JM. *The Sounds of Speech Communication: A Primer of Acoustic Phonetics and Speech Perception.* Baltimore: University Park; 1980.
- Postma, A. On the mechanisms of speech monitoring. In: Hulstijn, W.; Peters, HFM.; van Lieshout, PHHM., editors. *Speech Production: Motor Control, Brain Research and Fluency Disorders.* New York: Elsevier Science; 1997.
- Postma A, Kolk H. Error monitoring in people who stutter: Evidence against auditory feedback defect theories. *J. Speech Hear. Res.* 1992; 35:1024–1032. [PubMed: 1447913]
- Postma A, Kolk H. The Covert Repair Hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *J. Speech Hear. Res.* 1993; 36:472–487. [PubMed: 8331905]
- Prins D, Main V, Wampler S. Lexicalization in adults who stutter. *Journal of Speech, Language and Hearing Research.* 1997; 40:373–384.

- Starkweather, C. The development of fluency in normal children. In: Gregory, H., editor. *Stuttering Therapy: Prevention and Intervention with Children*. Memphis, TN: Speech Foundation of America; 1985. p. 9-42.
- Tornick GB, Bloodstein O. Stuttering and sentence length. *J. Speech Hear. Res.* 1976; 19:631-654.
- Van Lieshout PHHM, Hulstijn W, Peters HFM. Speech production in people who stutter: Testing the motor plan assembly hypothesis. *J. Speech Hear. Res.* 1996a; 39:76-92. [PubMed: 8820700]
- Van Lieshout PHHM, Hulstijn W, Peters HFM. From planning to articulation in speech production: What differentiates a person who stutters from a person who does not stutter. *J. Speech Hear. Res.* 1996b; 39:546-564. [PubMed: 8783133]
- Van Lieshout PHHM, Starkweather CW, Hulstijn W, Peters H. Effects of linguistic correlates of stuttering on EMG activity in nonstuttering speakers. *J. Speech Hear. Res.* 1995; 38:360-372. [PubMed: 7596101]
- Weiss AL, Zebrowski PM. Dysfluencies in the conversations of young children who stutter: Some answers about questions. *J. Speech Hear. Res.* 1992; 35:1230-1238. [PubMed: 1494268]
- Wingate, ME. *Stuttering: Theory and Treatment*. New York: Irvington-Wiley; 1976.
- Wingate, ME. *The Structure of Stuttering: A Psycholinguistic Analysis*. New York: Springer-Verlag; 1988.
- Yaruss JS, Conture EG. Mother and child speaking rates and utterance lengths in adjacent fluent utterances: Preliminary observations. *Journal of Fluency Disorders*. 1995; 20:257-278.

FIG. 1.

Boxplots of the adjusted stuttering rates on words (ordinate) for the three utterance rate divisions based on syllables. Fast, medium, and slow utterance rates are labeled on the abscissa. Fast is any rate above 5 syllable per second; medium as between 4 and 5 syllable per second; slow as under 4 syllable per second. The boxes at each rate indicate the interquartile range of stuttering rates and the mean is shown by the solid horizontal line within the box. The full range of stuttering rates is shown by the whiskers.

FIG. 2.

Boxplots (formatted as in Fig. 1) of the adjusted stuttering rates on words (ordinate) for the three utterance rate divisions based on the speech rate of individual speaker. Fast, medium and slow utterance rates are labeled on the abscissa. Fast is defined as speech rate higher than the level at one s.d. above the mean speech rate of all fluent words. Slow rate is defined as less than one s.d. below the mean rate. Medium is the rate in between +1 and -1 s.d. levels.

FIG. 3. Boxplots (formatted as in Fig. 1) of stuttering rates on words (ordinate) for different tone unit lengths measured in words (abscissa).

FIG. 4. Boxplots (formatted as in Fig. 1) of stuttering rates on words (ordinate) for different tone unit lengths measured in syllables (abscissa).

FIG. 5. Boxplots (formatted as in Fig. 1) showing the correlation between utterance rate (ordinate) measured on syllables and utterance length measured in number of words (abscissa).

FIG. 6. Boxplots (formatted as in Fig. 1) showing the correlation between utterance rate (ordinate) measured on syllables and utterance length also measured in number of syllables (abscissa).

TABLE I

Speaker information including speaker initials, age, gender, overall stuttering rate, and the number of words uttered by individual speakers and used in this study.

Speaker	Age	Gender	Stuttering rate	Number of words
cn	10	M	15.08	166
dc	10	M	4.92	230
dm	9	M	12.07	246
jg	9	M	7.69	98
mt	10	M	25.00	199
pm	9	M	20.35	200
rp	11	M	35.83	145
sh	11	F	23.27	164