Subject Definition and Selection Criteria for Stuttering Research in Adult Subjects*

Peter J. Alfonso†

The principal topics assigned to this author were to review subject definition and selection criteria reported in experiments investigating speech behaviors in adult stutterers, to determine whether the typical description of these criteria is sufficient to meet the demands of scientific investigation and replication, and to discuss a number of topics that should be considered in the development of subject definition and selection criteria. Although the organization of this paper is largely motivated by the assigned topics, I have deviated somewhat by reviewing in Section I the definitions of stutterers and stuttering that typically appear in the literature. The review is followed by a discussion of how the precise and careful use of these definitions bear directly on the development of subject definition and selection criteria. Section II represents the primary focus of the paper and includes: (a) a review of typical subject definition and selection criteria, and (b) a discussion of a number of topics that should be considered in the development of these criteria.

The argument that appropriate subject definition and selection criteria is an essential component of good experimental technique is based on the ubiquitous observations that so many of the overt characteristics of the disorder are highly variable across subjects at all levels of measurement. The argument developed here is that until the many facets of the heterogeneity of stuttering are better understood, criteria should err on the side of over-defining rather than under-defining essential details about the stutterer and his or her behaviors.

Finally, because this paper centers on experiments dealing with adults, much of what is written here presumes that the procedures employed in the experiments of interest are common to that population and are more physiologically based and invasive than those commonly used with children who stutter. Problems associated with the development of subject definition and selection criteria in the latter population group are considered in detail in a paper entitled "Childhood stuttering: What is it and who does it?" by Edward Conture (1990).

SECTION I: DEFINING STUTTERING AND STUTTERERS

I.A. An overview

Definitions of stuttering continue to evolve with our theories and abilities to measure various aspects of the disorder. Traditionally, most definitions are descriptions of behaviors. They are typically presented as a comprehensive list of behaviors that are common to all stutterers and that differentiate stuttering from normal speech. An often cited example of a descriptive definition is given by Wingate: "The term 'stuttering' means: I. (a) Disruption in the fluency of verbal expression, which is (b) characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements, namely: sounds, syllables, and words of one syllable... II. Sometimes the disruptions are (e) accompanied by accessory activities involving the speech apparatus, related or unrelated body structures, or stereotyped speech utterances... III. Also, there are not infrequently (f) indications or report of the presence of an emotional state, ranging from a general condition of 'excitement' or 'tension' to more specific emotions... (g) The immediate source of stuttering is some incoordination expressed in
the peripheral speech mechanism; the ultimate cause is presently unknown and may be complex or compound" (Wingate, 1964). Most definitions include at least the following three descriptions of the verbal behavior associated with stuttering; involuntary, repetitions, and prolongations. For example, stuttering is defined in the International Classification of Diseases as "disorders in the rhythm of speech, in which the individual knows precisely what he wishes to say, but at the time is unable to say it because of an involuntary, repetitive prolongation or cessation of a sound" (World Health Organization, 1977, p. 202). For a good discussion of the various categories of definitions see Van Riper, 1982, Chapter 2.

I.B. Needs and future directions

I.B.1. Disfluency, dysfluency, and fluency. The intent prescribed to the terms "disfluency," "dysfluency," and "fluency" varies considerably as a function of the distinctions among theoretical models of stuttering. This issue is considered in sufficient detail elsewhere, for example, Ham (1989), Perkins (1984), and Wingate (1984a,b, 1988). There should be no disagreement, however, about the critical necessity to make explicit, especially for research purposes, the descriptions of the speech behavior(s) under observation. By way of example, I will throughout this paper retain distinct definitions for the terms “disfluency,” “dysfluency,” and “fluency,” largely following the rationale reported by Wingate (1984a,b). Disfluency is used here as a general referent, pertaining to the usual and normal disruptions in the patterns of speech movements that are perceived as "fluent speech." Dysfluency, on the other hand, is used to mean abnormal disruption in the normal patterns of speech movements. Fluency is used in the perceptual sense, to mean the realization of flowing, smooth, and easily produced speech; that is, as an abstraction of the underlying articulatory gestures. Thus, a sample of speech produced by a normal talker that has been judged to be fluent, as defined here, may include, at the level of speech production, disfluent segments but not dysfluent ones. Once again, the critical point is that these terms, and the criteria developed to operationally define them, must be made explicit. It will be shown in Section II that the failure to do so can led to conflicting conclusions about a number of aspects related to stuttering.

I.B.2. Stutterer's self-identification and stuttering-identification. In defining stuttering and stutterers, infrequent attention seems to be

given to the stutterer's "self-identification" and "stuttering-identification." In particular, the stutterer's identification of a dysfluency can be very important since, as discussed below in more detail, it is often the case that disagreement will occur in fluent-dysfluent judgments that are based on data representing different accessible levels of measurement, that is, perceptual, acoustic, movement, and neuromuscular. For example, it is not unusual that an utterance is judged "fluent" at the perceptual level by an experienced listener, while analysis at deeper levels of speech, kinematic for instance, indicate inappropriate or "dysfluent" production. Until the distinction between fluent, disfluent, and dysfluent speech is better understood, the adult stutterer's judgment in the classification of him- or herself as a stutterer and in the fluency-dysfluency distinction of his or her speech should be encouraged.

I.B.3. Voluntary and involuntary speech motor output. The distinction between "voluntary" and "involuntary" in defining disfluency and/or dysfluency is not consistently made, although it may represent a critical distinction between certain types of dysfluency exhibited by stutterers and dysfluency exhibited by adults who do not stutter. The significance of "involuntary" in definitions of stuttering and stutterers is discussed in detail by Perkins (1983) in response to a review article on stuttering by Andrews and his colleagues (Andrews et al., 1983) and need not be elaborated in great detail here. By way of a brief example, Perkins (1983) writes that the presumption is that a voluntary disfluency in the adult population results from "linguistic uncertainty." This implies that a voluntary disfluency, as in the prolongation of the isolated vowel /a/ for example, is a voluntary strategy invoked by the speaker while attempting to resolve a high-level linguistic query, such as in lexical retrieval. On the hand, the mechanisms underlying involuntary speech acts are far less agreed upon. Citing Perkins (1983) once again as an example, dysfluency "presumably is a motor speech blockage." Certainly, one can argue with linguistic uncertainty and speech motor blockage models of the voluntary-involuntary distinction, but until the volition of speech motor output can be measured with validity, the stutterer's identification of his/her fluent and dysfluent speech should be encouraged. That is, stutterers appear to be in a better position to subjectively rate their utterances as involuntary dysfluent, voluntary disfluent, or fluent than are listener-judges.
These two notions, the stutterer's self-identification and the voluntary-involuntary distinction, are important because they appear to be essential parameters in distinguishing among disfluent, dysfluent, and perceptually fluent speech. The significance of developing criteria to explicitly differentiate among involuntary dysfluency, voluntary disfluency, and perceptual fluency will be discussed in greater detail in Section II.B.3.

I.B.4. Core versus secondary stuttering behaviors. Since many developmental models of stuttering consider some form of repetitions to represent the "core" of stuttering (e.g., Bloodstein, 1987; Stromsta, 1986; Van Riper, 1982) the inclusion of core behaviors in a definition appears warranted. Beyond "involuntary repetitions," and perhaps somewhat secondarily the duration and frequency of prolongations and silent pauses, there appears to be less agreement as to what is sufficient to delimit stuttering, except perhaps for the frequent remark that defining stuttering is much more complicated than some would think. Of course, there is a wide variety of so-called secondary behaviors or accessory features associated with stuttering (e.g., Wingate, 1964), the specification of which would be important especially in the consideration of severity.

I.B.5. Inclusive definitions of stuttering. Although there currently seems to be a better appreciation for psychological effects on physiological behaviors (e.g., Smith & Weber, 1988; Zimmerman, 1980c) definitions for the most part are rarely inclusive of external and internal influences on the disorder. Rather, they are either predominantly psychologically or physiologically based. Comprehensive definitions of stuttering similar to the "integrated theory" notion proposed most recently by Smith and Weber (1988) need to be better developed.

Definitions of stuttering do not always include an account of the suspected etiology of the disorder. For example, definitions could include the notion that involuntary "core" behaviors occur as a consequence of deficits, at various levels, in temporal programming (e.g., Caruso, Abbs, & Gracco, 1988; Kent, 1984), spatial or movement programming (e.g., Zimmerman, 1980c), or both temporal and spatial programming (e.g., Alfonso et al. (1986a; 1987a,b,c).

With a proper definition of stuttering, the researcher is better able to define stuttering subjects. Section II.A. shows, however, that researchers generally provide little definition of the stuttering populations that serve in their experiments, one consequence of which is that it often makes it difficult to make appropriate comparisons among experiments.

SECTION II: SUBJECT DEFINITION AND SELECTION CRITERIA

II.A. An Overview

A review of the literature indicates that little detail is given in journal articles regarding either subject definition or selection criteria. It is more often the case that stuttering severity is reported, although the means by which the severity estimate is determined is highly variable across experiments. What follows are examples of subject selection criteria (and subject definition, if given) that have been reported in recent or frequently cited research papers. The aim is to demonstrate the variability among published descriptions of experimental subjects and the criteria employed to select and define them. The following citations are in alphabetical order.

Many reports of experiments in the contemporary literature provide little information at all. For example, Freeman and Ushijima (1978) state only that their subjects were mild-moderate or severe. No other details of subject selection or definition criteria are given. Guitar et al. (1988) reports the gender of the subjects, that they were all native speakers of the same language, and had never received treatment for stuttering. The first author subjectively judged the severity of the subjects. The criteria for estimating severity were not given. Although the motivation for the Martin and Haroldson (1988) study was to experimentally increase stuttering frequency, so that the definition of stuttering, the frequency of stuttering, and the severity criteria, are crucial in this type of study, very little detail is given. The procedure employed is difficult to ascertain. "In the control room, the experimenter monitored all sessions auditorily and depressed a handswitch each time the subject stuttered. Stuttering was defined in terms of moments or instances of stuttering and not in terms of specific disfluency types." McClean (1987) reports that "informal assessment of conversations with the stutterers (7 adult male) suggested that as a group their stuttering severity ranged from mild to severe." No other details of subject selection or definition criteria are given. Zimmerman (1980a,b) and Zimmerman and Hanley (1983) used subjects who were enrolled in speech and hearing clinics at the time of data collection. The specific type of therapy is not mentioned. No details of the severity
criteria are given: “They ranged in severity from mild to severe as judged by a certified speech-language pathologist.”

Other reports of experiments provide some detail regarding subject selection criteria, though the type and amount vary in considerable degree. For example, Alfonso and his colleagues (Alfonso et al., 1986a,b; 1987a,b,c; Kalinowski & Alfonso, 1987; Story & Alfonso, 1988; Watson & Alfonso, 1982, 1983, 1987) used subjective and objective criteria to identify stutterers and group them on the basis of stuttering severity. Objective evaluations of stuttering frequency and type were completed using a combination of procedures described in the Stuttering Interview (SI) (Ryan, 1974) and the Stuttering Severity Instrument (SSI) (Riley, 1972). Subjective judgments of stuttering severity were obtained from certified speech-language pathologists and from the experimental subjects. Additional criteria are developed if severity ratings differ markedly between reading and conversational speech samples, among objective and subjective criteria (e.g., Watson & Alfonso, 1987), or if inter-test reliability, as a function of time periods or the identity of the judges, was low (Kalinowski & Alfonso, 1987). Caruso et al., (1988) assessed stutterers’ dysfluency during a conversational speech sample using two objective measures: 1), mean stuttering frequency (MSF), and 2), mean stuttering duration (MSD). Stuttering was defined as sound/syllable repetitions and sound prolongations. Conture et al. (1977) used MSF during conversation and oral reading. Stuttering severity was determined by use of the Iowa scale (Johnson, Darley, & Spriestersbach, 1963) based on the MSF. All of the subjects for Conture et al., (1985) were receiving therapy at a university speech and hearing clinic. MSF for sound/syllable repetition, sound prolongation, or within-word pause was used as a measure of severity. Metz et al. (1983) and Sacco et al. (1987) selected subjects enrolled in a residential stuttering treatment program. They used MSF based on sound and syllable repetitions, sound prolongations, and/or broken words produced during an unidentified reading sample. Severity was calculated using the SSI (Riley, 1972; 1980). The Shapiro (1980) experiment required that the locus of stuttering, type of dysfluency, and stuttering severity be reliably determined. Subjects read the Rainbow Passage five times and performed the Job and TAT Tasks (Johnson et al., 1963). Subjects were accepted if the interjudge agreement across four judges meet criteria. Operational definitions of severity of stuttering were obtained using the Stuttering Severity Scale (Johnson et al., 1963), and estimates of the specific type and locus of stuttering were subjectively made from videotape viewing.

II.B. Needs and future directions

The section above indicates that it is usually the case that one would find little detail in a journal article regarding the subject definition and selection criteria employed in the experiment. Stuttering severity estimates are given more often than subject definition and selection criteria, although the means by which the severity estimate is arrived at is highly variable across experiments. In what follows, certain issues relevant to the development of subject definition and selection criteria are discussed in detail. The issues represent broad areas of concern, and should not be construed as suggested minimal criteria. Rather, criteria should be developed as a function of the nature of the experiment at hand. The goal that all researchers should share, however, is that adequate descriptions should be given in sufficient detail so that experiments can be replicated and/or results can be appropriately interpreted and compared among experiments.

II.B.1. What identification measures should be used? The common use of a standardized perceptual test such as the SSI would lead to an obvious advantage of direct comparison of subjects across different experiments. Other identification measures should be considered, and include: (a) familial history of stuttering, (b) type(s) and duration(s) of therapies received, the distinctive characteristics of a therapy program, a description of clinical goals (e.g., slowed speech rate, gentle onsets) of these therapies, and how well they are maintained in habitual speech, (c) estimates of covert stuttering severity, for example, the stutterer’s judgment of the frequency and severity of dysfluency, secondary characteristics (e.g., Riley, 1972; Van Riper, 1982), descriptions of contextual conditions in which fluency and dysfluency are enhanced, and an estimate of the success to which the subject is able to use fluency enhancers to promote fluent speech, (d) estimates of overt stuttering severity, for example, the frequency and duration of repetitions, prolongations, and silent pauses, and (e) a description of the subject’s fluent speech, for example, rate and naturalness. Speech samples gathered as part of the evaluation should be obtained during both extemporaneous and read speech, and the differentiating stuttering behaviors should be noted. Speech samples
been shown to differentiate stutterers from nonstutterers. The former have been found to have greater respiratory criteria, based on data obtained by a Respitrace inductive plethysmograph for example, than those discussed above, and could include, for example, strain-gauge or opto-electronic movement transducers to measure lip and jaw displacement. These techniques are becoming more common in many laboratories. It would be important that supralaryngeal physiological criteria for the selection and definition of stuttering subjects be centered on organizational principles of speech motor control, that is, centered on good representatives of what are believed to be speech motor control parameters. For example, noninvasive lip and jaw displacement amplitude data could be analyzed to assess motor equivalence covariation and sequential ordering in labial gestures (Alfonso et al., 1986a, 1987b,c; Caruso et al., 1988). This issue is discussed in greater detail in Section II.B.3. It is less important to know and base group comparisons on displacement amplitudes and velocities of individual lip and jaw movements, and at what rates they move, than it is to know about the organizational principles underlying lip-jaw movements because (a) efficient, rapid, and fluent speech requires a relatively high degree of spatial and temporal coordination among the supralaryngeal speech structures, and (b) interspeaker variability at the motor control level is inherently less variable than at the phonetic level. Of course, data obtained by a combination of these techniques could be used to measure intersystem physiological parameters, for example, respiratory-laryngeal timing (e.g., Peters & Boves, 1988; Watson & Alfonso, 1987).

II.B.2. What behaviors during the moment of stuttering should be included in the definition? The type, duration, and severity of the involuntary dysfluency should be identified during the moment of stuttering. Because the relationship between linguistic structure and speech motor specification in dysfluent speech is not fully understood, a description of the linguistic context in which the dysfluency occurred should be given. At the very least, the intended fluent phonetic target should be identified, for example, stressed syllable initial voiceless aspirated stop. Other
linguistic descriptors that could be specified depending on the focus of the experiment include word position and content versus function word. A detailed consideration of psycholinguistic variables associated with stuttering is given in Wingate (1988). A number of secondary characteristics, including anxiety and emotional stress, should also be considered.

Although Shapiro (1980) concluded on analysis of EMG data that the perceptions of the judges as well as the subjects themselves regarding the identification of labial, laryngeal, and lingual predominate locus were erroneous, it might still be clinically and theoretically useful, particularly in regard to the subject's perception of his/her production of the dysfluency, to include in the definition a statement about clinician and/or client estimate of locus. We need more physiological data, in parallel with perceptual and acoustic, to determine with certainty whether or not stuttering results from a physiological disruption at one location, the larynx for example, or if the entire speech motor system fails simultaneously. Shapiro concludes that the stuttering may not be able to identify the location of the disruption, however, we don't have enough data to know this for certain, nor do we have enough data to know whether a breakdown at the larynx, for example, occurs first and leads to failures in other components of the speech system in response to the laryngeal failure.

Finally, for those experiments based on physiological data it may be useful to segment the events surrounding a dysfluency. By way of example, Alfonso and Seider (1986) examined acoustic, respiratory and laryngeal kinematic data, and laryngeal electromyographic (EMG) data during an interval beginning with the termination of fluency, followed by an inaudible dysfluent period, followed by an audible dysfluent period, and ending with a fluent period. Laryngeal EMG and movement data showed that the laryngeal configuration was clearly different during the inaudible than during the audible periods of the dysfluent episode. The configuration appeared most inappropriate during the initial inaudible period and less so during the following audible period. Segmentation also allows for the comparison of speech motor events during the dysfluent periods with events immediately following, that is, during fluent production of the phonetic target. Thus, statements about the configuration of the vocal tract during moments of stuttering can be made in reference to the vocal tract configuration during more fluent episodes immediately preceding the dysfluency, and in reference to the vocal tract configuration immediately following the dysfluency and associated with the intended fluent phonetic target.

II.B.3. Is there such a thing as “normal disfluency” and can it be reliably differentiated from stuttering? Perceptual fluency spans a wide continuum, the endpoints of which could be identified as “severely fluent” to “severely dysfluent.” There is large variability within the “normal” subsection of the continuum, ranging from something like “severely fluent” to “normal disfluent” and apparently even more variability within the “abnormal” subsection of the continuum, ranging from something like “abnormal fluent” to “severely dysfluent.” Because the “normal” and “abnormal” subsections of the continuum appear to overlap perceptually, it seems that defining abnormal fluency, particularly when produced by severe stutterers, may be more straightforward than defining normal disfluency. That is, it may be more experimentally viable and more fruitful in the long run to modify the above question and to ask: is the perceptual “fluent” speech of severe stutterers similar to the fluent speech of control subjects? The latter form of the question may be more experimentally viable because: (a) the resolution involves extreme contrasts, the continuum endpoints, represented by a severe stutterer contrasted with a fluent control subject in the example here. The extreme contrast should be easier to differentiate than more subtle contrasts; for example, certain physiological characteristics of the perceptually fluent speech of a severe stutterer compared with the corresponding characteristics of a normal subject’s fluent speech should be easier to differentiate than a mild stutterer’s fluent speech compared with a normal subject’s fluent speech, and (b) it would be easier to define a severely dysfluent stutterer than it would be to define a severely fluent control subject. The latter form of the question may be more fruitful because a better understanding of stutterers’ perceptually fluent speech would represent a relatively direct and immediate increase in our understanding of stutterers’ speech motor control. Of course, in the long run we will need to understand better the fluency variability in the normal population as well as the stuttering population. Thus, the following discussion is a modification of the originally assigned question and asks: (a) can normal disfluency be reliably differentiated from stutterers’ dysfluency, and (b) can stutterers’
perceptually fluent speech be reliably differentiated from normal fluent speech?

Considering normal disfluency first, certainly adults who do not stutter do repeat and prolong a variety of speech segments, usually words and phrases, and interject pauses between words and phrases. Generally, the speech segments in which repetitions occur differentiate the groups; nonstutterers predominantly repeat whole words and phrases whereas stutterers predominantly repeat sounds and syllables. The duration of pauses and prolongations may also distinguish normal disfluency from abnormal disfluency. However, the frequency of repetitions and the duration of prolongations, regardless of the speech segment in which these occur, are far less in magnitude in the nonstuttering population than in the stuttering population. That is, stutterers do more of everything; repetitions, prolongations, and pauses. Normal disfluencies are not usually accompanied by secondary characteristics and may be less susceptible to higher linguistic influences, for example, word order and word type (see Starkweather, 1987, Chapter 5 for a more detailed discussion of normal disfluency). An important assumption underlying the distinction between normal and abnormal speech is that normal speech movements are voluntary whereas stutterers' dysfluencies are not. Although it may be difficult, and perhaps impossible, to ascertain the volition of speech motor output, determining the extent to which repetitions, prolongations, and pauses are under voluntary motor control may be the ultimate test of the normal disfluency versus abnormal disfluency distinction.

Considering stutterers' perceptually fluent speech next, there is considerably more data regarding the contrast between stutterers' perceptually fluent speech and normal fluent speech than in the contrast discussed above. However, some experimenters conclude that stutterers' perceptually fluent speech is not different than control subjects' fluent speech, while the majority of experimenters seem to think that it is (see, for example, Van Riper, 1982, Chapter 16; Bloodstein, 1987, Chapter 1). Rather than review a relatively large literature here, it may be more useful in regards to the development of subject definition and selection criteria to discuss a few of the reasons underlying the conflicts in the results of these types of experiments. It should be noted first, however, that the conflicting results are no doubt confounded by the lack of adequate definitions and criteria for distinguishing among certain essential characteristics of stuttering discussed in Section I, for example, voluntary disfluency, involuntary dysfluency, and perceptual fluency. That is, the results of two experiments may generate conflicting conclusions whether or not stutterers' perceptually fluent speech is similar to normally fluent speech simply because the criteria (which may or may not be reported) for distinguishing the stutterers' perceptually fluent speech from their dysfluent speech differed across the two experiments.

Second, the majority of experiments are based on perceptual data alone, some are based on perceptual and speech acoustic data, and far less include kinematic and neuromuscular data in parallel with perceptual and acoustic. One source of the conflict was discussed above in Section II.B.1, that is, while a segment of a stutterer's speech may appear normal or "fluent" at the perceptual level, it may appear abnormal or "dysfluent" at the acoustic, and/or movement and muscular levels (Alfonso et al., 1984; Baer & Alfonso, 1984; Shapiro, 1980). Thus, a comparison based on perceptual data alone could indicate no difference between the groups, while a comparison of the same utterances using the same perceptual criteria for fluency and dysfluency but based on physiological data could find significant group differences. Clearly, perceptual data alone are too far removed from the source to be able to make detailed analyses of the fluent-dysfluent distinction, and as such would mark only the most obvious instances of stuttering. The lack of physiological data addressing this issue is of concern for other reasons. For example, the question of determining whether the perceptually "fluent" speech of severe stutterers is similar to the fluent speech of control subjects is important because the answer to the question will help determine whether the stutterers' speech motor system exhibits generalized spatio-temporal abnormalities regardless of the perceived fluency, or whether, alternatively, it behaves normally except during moments of dysfluency. It is generally the case that perceptual and acoustic data in the absence of simultaneously gathered physiological data are not sufficient to answer questions about speech motor control.

A second reason for the conflict in the results of reported literature on the distinction between stutterers' perceptually fluent speech and nonstutterers fluent speech is that the routinely posed form of the question, "Is stutterers' fluent speech similar to the fluent speech of adults who do not stutter," is too broad. It is possible that
certain aspects of speech do not differentiate the groups while other aspects do differentiate the groups. For example, Baken et al. (1983) found no differences between stutterers and control subjects in chest wall preposturing maneuvers immediately preceding fluent speech. Yet, significant differences in subglottic pressure (e.g., Lewis, 1975), flow rates (e.g., Hutchinson, 1975), and lung volume change and deflation for speech (e.g., Story, 1990; Watson & Alfonso, 1987) have been observed between the groups. Thus, it is less appropriate to ask whether or not respiratory control, per se, (and certainly not speech production, in general), can be differentiated between the groups. Rather, it might be that stutterers and normal speakers perform certain speech respiratory gestures in a similar fashion, namely respiratory preposturing, but that other aspects of speech respiration, namely the magnitude of the inspiratory charge, are performed differentially.

A third and perhaps most important reason for the conflict in the results of these types of experiments is that group comparisons are frequently based on physiological data that reflect relatively variant phonetic level speech gestures. Group data based on phonetic level contrasts are inherently unstable since spatial and temporal control of the speech structures to mark phonetic distinctions varies as a function of phonetic context, stress and rate, dialect, and individual speaker preferences. Rather, group comparisons should be based on relatively stable spatial and temporal characteristics of normal speech dynamics, for example those that best reflect organization principles of speech motor control. Dynamic parameters that are relatively invariant across multiple productions of an utterance are thought to be good representatives of speech motor control parameters (e.g., Gracco & Abbs, 1986). Thus, an appropriate comparison would be one based on the extent to which the magnitude of the relatively stable spatial and temporal characteristics of normal speech dynamics approximates corresponding characteristics of stutterers' perceptually fluent speech (Alfonso et al., 1986a, 1987b,c; Caruso et al., 1988). For example, motor equivalence covariation for speech is based on the observation that normal subjects control the relative displacement of individual articulators (e.g., the tongue and jaw) enlisted in a vocal tract gesture (e.g., alveolar closure) in such a way that the variability of the gesture is less than the variability associated with the individual articulators comprising the gesture. Accordingly, it is less appropriate to base group comparison on displacement amplitude and peak velocity of a single articulator, the tongue or the jaw for example, than it would be to compare the relative organization of the combined tongue-jaw displacement, because the latter comparison reflects a relatively stable component of speech dynamics while the former reflects idiosyncratic speaker preference. To summarize, the question of whether stutterers' perceptually fluent speech is similar to the fluent speech of non-stuttering adults can be appropriately addressed if criteria are developed that: (a) objectively distinguish among perceptually fluent, disfluent, and dysfluent utterances, (b) define the question more precisely, that is, provide more detail about specific aspects of speech production, (c) require appropriate types of data, and (d) require group comparisons based on appropriate speech behaviors.

II.B.4. How can a definition take into account variability in stuttering behavior? At least three types of variability should be taken into account. The first is associated with the well-known variability in the frequency and severity of dysfluent behaviors as a function of various external conditions. For example, dysfluent behaviors decrease and perceptually fluent behaviors increase with manipulation of auditory feedback; when the intensity of the feedback signal is increased, decreased, or masked by various techniques, and when the time of arrival of the auditory feedback signal to the talkers' ears is delayed. Fluency is enhanced by imposing an external rhythmical marker on the speech task, for example, a metronome signal, choral reading, and singing. Finally, fluency is enhanced by the adaptation effect. Most of these conditions will enhance fluency in non-stuttering adults as well as stutterers. The exception seems to be the delayed auditory feedback (DAF) paradigm, where a delay in the signal could result in increased dysfluency in some non-stuttering adults, although normal speakers vary widely in their susceptibility to DAF (Alfonso, 1974). Thus, alterations of the feedback signal, imposition of a rhythmical marker, and multiple repetitions of the same speech task increase fluency in most talkers, stutterers and nonstutterers alike. It appears possible that all of these conditions enhance fluency through a similar mechanism, that is, they act to highlight the control of prosody and perhaps other temporal parameters of the speech motor plan.

Likewise, there are a number of external conditions that increase dysfluent behaviors and
decrease perceptually fluent behaviors. It is well known that the contextual surround, for example, stressful situations and tasks, and specific listeners, will increase the frequency and severity of dysfluency. Certain linguistic conditions, such as long words, content versus function words, and words carrying high information loading, are known to promote dysfluency. Once again, these conditions are known to similarly affect adult non-stutterers, though not to the degree seen in stutterers. On the other hand, there are those adult non-stutterers (and apparently stutterers) who become more fluent in seemingly stressful situations. A likely causal agent for the increase in dysfluency in these conditions for both stutterers and adults who do not stutter is the effect of stress on the speech motor system (e.g., Zimmerman, 1980c).

The second type of variability is associated with changes in the type, frequency, and/or severity of dysfluent behaviors that occur across varying units of time. As an example of the variation in dysfluent behaviors that occur across the long term, it is generally observed that very young children who stutter do so primarily by repetition and prolongation, while over the course of 2 to 4 years develop a variety of other overt behaviors not present at the onset. Thus, stuttering in adults is comparatively idiosyncratic and as such varies considerably across the adult stuttering population. There are other examples of long-term changes in dysfluent behaviors, although the basis for the change is not sufficiently understood. With respect to the example given here, the source and extent of the variability of “core” behaviors, the repetitions and prolongations in very young children, compared to the variability of “secondary” behaviors in older children and adults apparently remains unresolved (Stromsta, 1986; Van Riper, 1982). Of course, variations in dysfluent behaviors occur across much smaller units of time. In fact, the perception of stuttering frequency, severity, and type within the same subject can change significantly as a function of many of the external conditions mentioned above (e.g., stressful situations and tasks) over the course of hours and days. Thus, as was pointed out in Section II.B.1., the same subject can be classified as a mild-moderate stutterer in a morning session and using the same severity rating instrument be classified as a severe-moderate stutterer in an afternoon session. This poses a particular problem in the development of subject definition and selection criteria and illustrates the importance of accounting for the variety of conditions that are known to induce short- and long-term fluency variations.

A third type of variability that could be included in subject definition and selection criteria is the type observed at the speech motor level. Greater variability for stutterers compared to control subjects at every accessible level of speech production, namely, acoustic, movement, and neuromuscular, is frequently reported and must be considered a robust observation. Relatively high levels of variability can, of course, be of serious consequence in meeting the demands of rapid, fluent speech. High variability implies that stutterers’ control of the movements of the speech structures is less precise than adults who do not stutterer. The lack of precision implies that stutterers lack the flexibility observed in normal speech, the flexibility that forms the basis for co-production of speech segments and in other human multiarticulate systems. Thus, the increased variability observed in the kinematics of many structures of the speech mechanism, and more importantly, the increased variability associated with various organizational components of fluent speech (e.g., the tongue-jaw synergies referred to in Section II.B.3.) suggest that stutterers’ control of the speech motor system lacks the flexibility and efficiency to meet the demands of rapid, fluent speech. (e.g., Alfonso et al., 1985, 1987c; Stromsta, 1986).

In summary, subject definition and selection criteria should account for the variability that is frequently observed at all levels of stuttering behaviors. In regards to treatment of the data, descriptions of data dispersion in addition to central tendency should be reported. Reporting the complete data base, perhaps as an appendix, would often be appropriate. Statements regarding the suspected source or the conditions promoting the variability, and the effect of the source or conditions on the variability, should be included; for example, fluency enhancement by rhythmic stimulation, dysfluency increase by psychological and physiological stress, accompanied by a statement regarding the magnitude of the fluency change. Finally, we need to understand better the relationship between the magnitude of the variability and stuttering severity.

II.B.5. What are the commonalities among stutterers? There are certain characteristics that appear to be common among stutterers as a group and should be considered in developing subject definition and selection criteria. For example, all stutterers seem to experience the notion of stutterer-identification discussed above in Section I.B.2. That is, regardless of severity, or the degree
to which overt or covert manifestations of stuttering are realized, stutterers seem to know that they are stutterers. If this is indeed true, then the notion of self-identification should be a central component of the definition of stuttering and stutterers. The stutterers' identification of themselves as stutterers and their identification of their moments of stuttering should focus the development of perceptual and physiological criteria.

Based on family history data, many researchers have concluded that the predisposition to stutter is genetically transmitted. For example, Van Riper (1982) concludes: "This incidence is so much greater than that found generally in the population (about 5 per cent) that it is difficult to believe that environmental factors alone could account for the results" (p. 330). Family history studies show that more males than females stutter, that there is a high incidence of stuttering within families, and that the risk of stuttering is higher if the mother, rather than the father, stutters. Certain of these studies allow the determination of risk to relatives; stuttering occurred in about 20% of male relatives and about 5% of female relatives of male stutterers, whereas stuttering occurred in about 25% of male relatives and 12% of female relatives of female stutterers (Kidd et al., 1978, 1981). However, as Pauls (1990) points out in a paper entitled "A review of the evidence for genetic factors in stuttering," there are significant limitations to the family study method. Coupled with the fact that these limitations make it impossible to understand the exact nature of the transmission, and that a disorder as common as stuttering is almost certain to be etiologically and genetically heterogeneous, the inclusion of family history data in subject definition and selection criteria should be cautiously considered.

The onset of stuttering is fairly common across stutterers. Most stuttering develops during childhood and is usually marked initially by syllable repetitions and prolongations. Adult stutterers who demonstrate a characteristic childhood onset should be distinguished from the relatively few adult stutterers who report sudden onset later in life.

Finally, certain physiological characteristics of the speech motor system are shared by most stutterers. For example, the stutterers' speech motor system is inappropriately susceptible to psychological and physiological stress. Fluency, measured at the output of the system primarily by perceptual criteria, is enhanced by rhythmical stimulation. The reaction-time of the stutterers' speech motor system is slower than normal speakers. And the stutterers' system may be characterized by a disability in spatial, temporal, or spatio-temporal coordination, particularly of the type discussed in Section II.B.3.

II.B.6. What would be the benefits to research in having a consensus definition? While it is likely that reaching the ultimate consensus definition is not possible to achieve, the alternative, that is to continue without suggestive guidelines would lead to a continuation of the variability in subject definition and selection criteria illustrated in Section II.A. One of the aims of this paper is to demonstrate the different ways in which current definitions and criteria are inadequate. With respect to definitions of stuttering, there are many instances where data have been pooled across such factors as severity and type so that a clear interpretation of the results is not possible. The lack of clearly stated subject selection criteria may have even more serious consequence on data interpretation because of the heterogeneous nature of the disorder. For example, data are frequently pooled across such variables as subject severity, varying clinical treatment influences on perceptually fluent speech patterns, and idiosyncratic dysfluent verbal and nonverbal behaviors.

There are, however, obvious and significant limitations to the consensus approach. The first is related to the wide varieties of stuttering behaviors and subjects that are often examined. Overly precise definitions may exclude critical members of a category while overly general definitions would not make the necessary distinctions between members of a category. Secondly, standardized subject definition and selection criteria would be in a constant state of reassessment as our knowledge of the disorder continues to increase.

Many of the benefits of a consensus definition can be achieved by the routine practice of reporting sufficient details of the experiment so that, at the very least, the experiment can adequately be replicated. The precise details will vary as a function of the nature of the experiment, yet it would be difficult to imagine many cases where basic descriptors, for example, severity ratings of the dysfluencies and subjects under examination, could be omitted. The paper presented here discusses a number of factors that bear on subject definition and selection criteria. Undoubtedly, other factors not listed here could also be included. Thus, the consensus opinion should be that subject definition and selection
criteria adequately deal with the heterogeneity of stuttering, that they be specifically generated as a function of the experiment at hand, and that until the significance of the heterogeneity is better understood, definitions and criteria should err on the side of over- rather than under-defining critical details of the experiment. Included in the notion of over-defining in the context of a heterogeneous behavior is the treatment of the data. Because the reported data are often used by future researchers in a way not related to the aim of the original experiment, it would be appropriate to publish certain forms of individual subject data as an appendix, in addition to the typical reporting of averaged data, especially when the data represent physiological estimates of stuttering that are technically difficult to collect and/or potentially hazardous to the subject. The consensus, then, becomes one of intent rather than the specification of consensus criteria.

REFERENCES


FOOTNOTES


†Also University of Connecticut, Storrs.

‡However, a paper entitled "Research procedures for measuring stuttering severity" by Ludlow (1990) takes a different approach to the core versus secondary behavior distinction.

§For a discussion of the problems associated with gathering familial history data, see a paper entitled "A review of the evidence of genetic factors in stuttering" by Pauls (1990).

¶For a detailed discussion of the problems associated with severity estimates see a paper entitled "Procedures for assessing the severity of stuttering for research" by Ludlow (1990).