COMPREHENSIBILITY ASSESSMENT: THE INFLUENCE OF FAMILIAR AND UNFAMILIAR COMMUNICATION PARTNERS

by

Jana Kim Johnston

B.A., University of Victoria, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2009

Abstract

For many years, the primary focus of assessment and intervention in the field of dysarthria rehabilitation has been the speech output of the affected speaker, as measured by intelligibility. However, this narrow focus on the acoustic adequacy of the speech signal falls short of capturing the full spectrum of information available to speakers and their partners during naturalistic interaction. Alternatively, comprehensibility measures acknowledge that everyday communicative success is greatly affected by the presence of multiple types of signal-independent information (e.g., visual cues, listener familiarity), as well as the dynamics of dyadic interaction (e.g., breakdown repair strategies). In this way, comprehensibility measures may more functionally assess dysarthria severity and the extent to which persons with dysarthria experience disability. The goals of the present study were to establish evidence for the reliability of a structured assessment of comprehensibility for dysarthric speech (developed by Visser, 2004), to investigate how the assessment reflected the effects of listener familiarity on communicative success, and to investigate how strategy use may differ under different familiarity conditions. Two speakers with dysarthria secondary to primary lateral sclerosis (PLS) were each paired with two communication partners: one familiar partner and one unfamiliar partner, to form four dyads in total. Each of these partners scored the speaker's intelligibility and participated in a dyad-based structured assessment of comprehensibility with the speaker. Results indicated that the assessment was scored with an extremely high degree of inter- and intra-rater reliability. In addition, comprehensibility scores were substantially better for both speakers with familiar versus unfamiliar partners, whereas substantial differences in intelligibility scores were noted for only one speaker. As predicted, unfamiliar dyads were found to use more strategies than familiar dyads, with some overlap across dyads in terms of the strategies used. Overall, the structured assessment of comprehensibility reliably differentiated familiar from unfamiliar dyads, suggesting it may be a clinically relevant assessment tool.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would first and foremost like to thank my research supervisor, Dr. Barbara Purves, for her incredible guidance in completing this thesis. Her insight and dedication has been invaluable and I cannot thank her enough for the countless meetings, phone calls, emails, and hours of work that have shaped this project over the past year and a half. I would also like to thank my committee members, Dr. Valter Ciocca and Tiffany Visser, for their suggestions and comments that have reminded me to adopt a fresh perspective.

My wonderful participants also deserve heartfelt thanks for inviting me into their homes and taking an interest in me and my research. Their role in this project was crucial, and I will forever be grateful to them for the time and energy they devoted to helping me in this way. I was lucky to work with such inspiring and incredibly kind individuals.

To my incredible friends and wonderful classmates: I will forever be grateful for being surrounded by such a supportive and sympathetic group of people during this challenging and busy time at UBC. Your friendships have truly kept me afloat.

Finally, to Séamus: You are the most selfless, supportive, and wonderful person I know. For all the understanding given, all the doubts assuaged, all the tantrums calmed, all the meals prepared, and all the housework done – you have been there in any and every way possible to support me and see me through grad school in one piece. You are my family, and I love you with all my heart.

1. INTRODUCTION

Rehabilitation for dysarthria has classically focused on the affected speaker alone, often effectively isolating the quality of the speaker's acoustic speech output (intelligibility) as the lone measure of severity and the central target for intervention. Considering the growing clinical emphasis on functional assessment and treatment within the field of speech-language pathology, a more comprehensive approach to dysarthria is necessary. In light of this trend, the present study proposes that a theoretical framework of comprehensibility is more appropriate than that of intelligibility in the consideration of dysarthric speech, and presents one application for the clinical assessment of communicative success. However, in order to form an appreciation for comprehensibility and its clinical applications, it is first critical to understand how the construct of intelligibility has evolved over recent years in the literature, as well as what factors have been shown to contribute to communicative success for speakers with dysarthria. In doing so, one can develop an appreciation of persons with dysarthria as not just speakers, but as communicators.

1.1. Causes and characteristics of dysarthria

The term dysarthria denotes a group of neurogenic motor speech disorders that underlie impaired control of the speech musculature. Resulting from neurological damage, dysarthria may arise from widely varying etiological sources, for example, perinatal neurological injury (e.g., cerebral palsy), acquired brain injury (e.g., stroke, traumatic brain injury), neoplastic disruption (e.g., cancer), infection (e.g., encephalitis), or progressive neurological disease (e.g., amyotrophic lateral sclerosis, Parkinson's Disease; Darley, Aronson, & Brown, 1975; Duffy, 2005). Dysarthria manifests most commonly as paralysis, weakness, or incoordination (Darley, Aronson, & Brown, 1969, 1975) affecting the motor control of one or more speech subsystems. Respiratory, laryngeal (phonatory), velopharyngeal, and/or articulatory mechanisms are commonly implicated (Darley et al., 1969, 1975; Ball, Beukelman, & Pattee, 2004).

The result of this loss of motor control is production of a degraded acoustic speech signal, causing decreased speech clarity (Comrie, Mackenzie, & McCall, 2001; Dykstra, Hakel, & Adams, 2007). In other words, individuals with dysarthria produce speech that their conversation partners may find difficult to understand. To appreciate how remediation of dysarthria is addressed from a clinical point of view requires an understanding of the theoretical constructs and terminology surrounding clinical practice and research relating to dysarthric speech. Intelligibility will be discussed first.

1.2. Intelligibility

Yorkston, Strand, and Kennedy (1996) defined *intelligibility* as "the degree to which the acoustic signal (the utterance produced by the dysarthric speaker) is understood by a listener" (p. 55). Intelligibility assessments typically isolate the speech signal via audio recordings of the speaker's single word or sentence-level utterances that are elicited using standardized tasks (Dykstra et al., 2007; Yorkston & Beukelman, 1981; Yorkston et al., 1996). An unfamiliar listener then orthographically transcribes the utterances from the audio recording to produce an intelligibility score – usually a percentage of the number of total words spoken that were transcribed correctly by the listener (Dykstra et al., 2007; Yorkston et al., 1996). This percentage has been used as an index of severity and to represent the adequacy of the speech signal produced by the dysarthric speaker (Yorkston et al., 1996). Traditional attempts to both assess and treat dysarthric speech have focused almost exclusively on the intelligibility of the acoustic signal. The factors contributing to intelligibility can be seen in Figure 1.1.

1.2.1. Intelligibility and the ICF

The World Health Organization (WHO) has developed a framework for classifying health

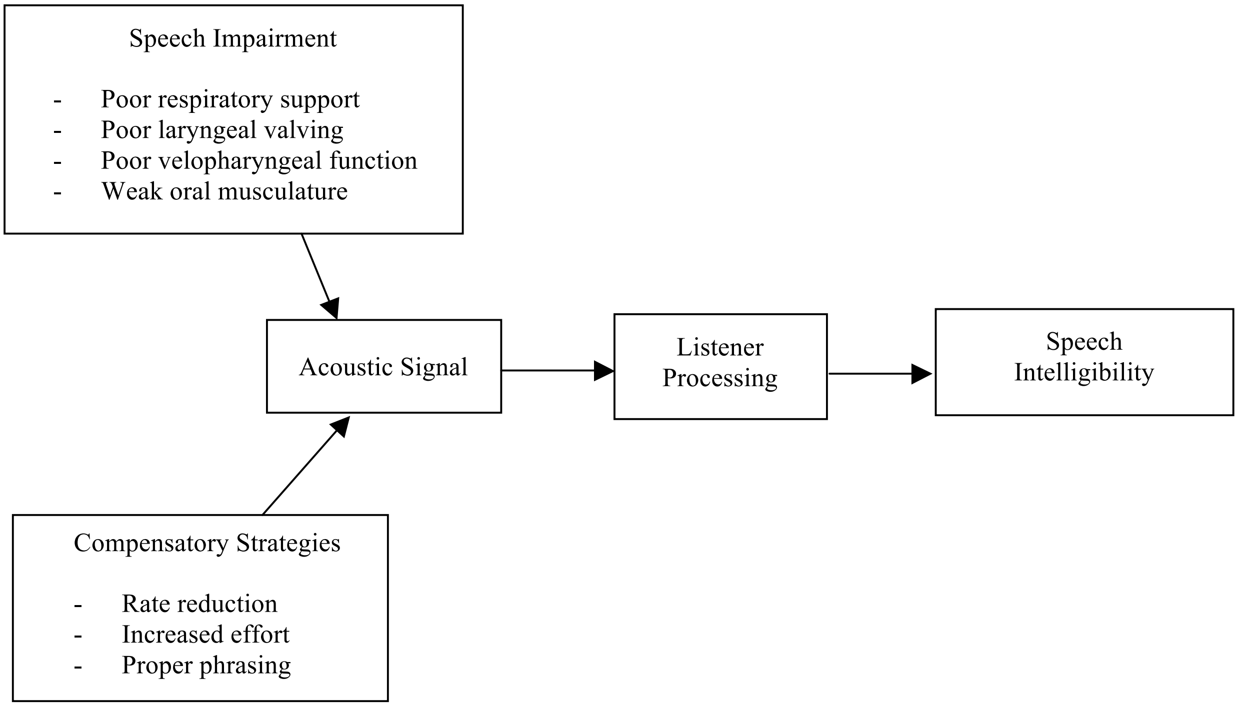

Figure 1.1. A Model of Factors that Affect Intelligibility of Individuals with Dysarthria (adapted from Yorkston et al., 1996).

conditions according to multiple influencing factors. Based on a synthesis of social and medical models of disability, the International Classification of Functioning, Disability and Health (ICF) aims to provide a measure of functioning in society (WHO, 2001). It addresses health and disability at the level of body structures and functions (anatomical and physiological systems), activity (task execution), participation (social involvement), and personal and environmental factors (contextual factors; WHO, 2001). These levels of functioning are seen within the ICF model as multi-directionally influential and highly interdependent. Essentially, this model captures the breadth of disability, from its organic cause to the physical and social environment in which it exists, including the extensive interaction between all contributing factors.

The traditional emphasis on intelligibility measures in dysarthria falls within the realm of the activity level in the ICF model (Dykstra et al., 2007; see also Yorkston et al., 1996, for a parallel reference to an older version of the WHO's model). Specifically, difficulty at the level of the speaker's anatomy and physiology (i.e., impairments caused by interactions between respiratory, laryngeal, velopharyngeal, articulatory and prosodic subsystems) results in degradation of the acoustic signal produced when speaking (Ball et al., 2004; Dykstra et al., 2007). In the language of the ICF model, the impaired body structures and functions of a person with dysarthria lead to an activity limitation, as the affected person experiences difficulty executing the activity of speaking clearly (Dykstra et al., 2007; WHO, 2001). The goal of intelligibility assessment is to measure the clarity of speech output, thus tapping how the motor speech impairment has contributed to activity limitations for the affected individual.

The emphasis on activity limitations in dysarthria has largely stemmed from the seminal Mayo Clinic studies of motor speech disorders undertaken by Darley, Aronson, and Brown (1969, 1975). These authors have contributed a substantial body of work outlining the salient perceptual features and physiological pathologies underlying the motor speech disorders, which include all types of dysarthria. While these authors describe the disordered speech patterns of the dysarthrias in extensive detail, little to no consideration is given to the nature of disability experienced by persons with dysarthria, or how impairment-level factors may interact with a multitude of other factors in the affected individual's life context. Ball et al. (2004) note that not one of Darley et al.'s 38 dimensions of assessment addressed communication performance in social contexts. Similarly, the widely used Amyotrophic Lateral Sclerosis Severity Scale (ALSSS) developed by Hillel, Miller, Yorkston, McDonald, Norris, and Konikow (1989) devotes the entirety of its speech-related items to the impairment or activity level. Until very recently, the majority of clinically available assessment procedures have largely ignored the ICF's participation level factors.

1.2.2. Limitations of intelligibility measures

Although description at the level of impairment is valuable in developing a clinical understanding of dysarthria, the intelligibility model unnaturally isolates the speech signal as the sole determinant of overall communicative ability. As the ICF model would suggest, this practice may limit the representativeness of intelligibility scores. By virtue of their tightly controlled and standardized administration and scoring procedures (Yorkston et al., 1996), intelligibility assessments produce severity scores that must be interpreted with caution. These scores may at the same time over- and under-estimate the intelligibility of those tested. Neither the quiet testing environment in which these scores are obtained (Ball et al., 2004; Dykstra et al., 2007) nor the exclusion of signal-independent information (e.g., visual cues, listener familiarity, etc.) are generally representative of typical communication environments. The assessments also typically address only word- or sentence-level speech production (e.g., Yorkston & Beukelman, 1981), controlling specifically for context and predictability, factors normally present in everyday communicative scenarios. Although assessing intelligibility based on more naturalistic, discourse-level speech samples would be more representative of the speaker's functional communication ability, such assessments are more difficult to standardize and thus rarely undertaken (Weismer, Jeng, Laures, Kent, & Kent, 2001).

Lindblom (1990) illustrates how the speech signal "is only the tip of the iceberg" (p. 228) when it comes to speech perception. He gives the example of a lecturer who asks his class to consider two written sentences (a question: "How much is two plus three?", and the answer: "Two plus three should equal five"). The lecturer plays a taped version of the answer that is masked in noise, then explains to the class that what they just heard was in fact the sentence "Poo klus free sould epwal thive". The degradation of the speech signal due to masking required that the listeners impose relevant background knowledge and context (namely, having just read the correct written sentence) on the signal in order to perceive it in a meaningful way. This example illustrates that linguistic and contextual knowledge exerts influence over what we ultimately perceive when listening to speech (Lindblom, 1990). Essentially, Lindblom's example argues that speech perception has multiple important influences, and that "our perception of speech and other communicative events is not determined by the signal alone" (p. 228). Ferrier (1991) echoes this assertion by remarking that intelligibility does not translate in any linear or simplistic manner to listener comprehension in naturalistic communication exchanges.

In light of these limitations, intelligibility scores do not adequately reflect real-life functioning, as the communicative performance they represent is not naturalistic (Ball et al., 2004; Dykstra et al., 2007). The integrity of the speech signal, while certainly a key contributor to communicative success, cannot and should not be relied upon as the lone measure of impairment for dysarthric speakers. Defining communicative ability so narrowly does not adequately capture the extent and multidimensional nature of disability experienced by people with dysarthria (Ball et al., 2004; Dykstra et al., 2007). Furthermore, it ignores the numerous tools, skills, and strategies speakers have at their disposal that can contribute to how well others understand what they say. The clarity of a dysarthric speaker's speech signal is but one of many factors affecting communicative success, and in turn, participation in social interaction.

The challenges to communicative participation are numerous for dysarthric speakers. First, faced with difficulty being understood, dysarthric speakers may avoid speaking situations. They have reported interacting less frequently with premorbid conversation partners, especially avoiding speaking with strangers (Dykstra et al., 2007; Hartelius et al., 2008). Second, a study by Comrie et al. (2001) suggests that people with dysarthria may be less active participants in conversations, contributing less major content overall to discussions and ceding the floor more readily to their unimpaired conversation partners. The results of Comrie and colleagues' study suggest that neurologically typical conversation partners may adopt a more dominant role when speaking to people with dysarthria, contributing longer and more numerous major turns. Indeed, some individuals with dysarthria have reported that they feel they are treated differently than others, possibly because their obvious speech difficulties signal to society that they are different (Hartelius et al., 2008; Yorkston, Klasner, & Swanson, 2001). It is also important to acknowledge the physical limitations that frequently co-occur with dysarthria as a result of a common underlying etiology. For example, a neurological injury may result in both dysarthria and confinement to a wheelchair. In this case, physical disability could have a considerable further impact on communicative participation in the sense that opportunities for interaction may be constrained in certain environments or contexts. Physical limitations may also contribute to differential treatment of these speakers by others in the community (Yorkston et al., 2001).

As mentioned previously, the extent of disability experienced by a person with dysarthria is not necessarily reflected by the severity of his or her speech disorder (Ball et al., 2004; Dykstra et al., 2007; Hartelius et al., 2008; Yorkston et al., 1994). Instead, it likely arises from an interaction of premorbid communication skills, complex personal factors, and emotional responses to the social consequences of disability. Individuals with dysarthria have noted that they get less enjoyment from speaking than before the causal neurological injury occurred, and many have reported feeling that the changes to their interactional pattern post-injury are generally all negative (Comrie et al., 2001; Hartelius et al., 2008).

In light of these significant barriers to communicative participation, it is integral to clinical practice that appropriate assessment and intervention programs be developed to address them in a measurable way. This includes a measure of truly functional communicative effectiveness. Currently, the tendency among clinicians is to make assumptions about dysarthric individuals' communicative effectiveness and social involvement based on intelligibility scores alone (Ball et al., 2004). It is certainly of utmost clinical importance that the extent of disability experienced by individuals with dysarthria be captured in a more quantifiable, accurate, and functionally relevant way. A truly appropriate assessment of communicative ability needs to look beyond the speech production mechanism and examine how the individual with dysarthria communicates and participates in society. As the ICF model stresses, the consideration of multiple interactions across all levels of functioning is integral for the true effects of disability to be understood. As Dykstra et al. (2007) note, as clinicians, we need to shift our attention away from the limited scope of intelligibility and towards an evaluation of how an individual's life context contributes to his or her communication disorder.

1.3. Comprehensibility

Communication is reciprocal and interactive. It is social. It is an exchange of information between individuals. And while a fundamental component of verbal communication is the ability of each speaker to produce understandable content, communication cannot occur without each party engaging socially with the other. Both the individual's participation in society and that society's response can either facilitate or hinder his or her involvement in life situations (Dykstra et al., 2007; WHO, 2001). *Comprehensibility* measures more appropriately address the participation level of the ICF by measuring communicative success in a way that goes beyond the disrupted speech mechanism and resulting speech production impairment (Yorkston et al., 1996).

Barefoot, Bochner, Johnson, and vom Eigen (1993) defined comprehensibility as "the extent to which a listener understands utterances produced by a speaker in a communication context". The influential model of comprehensibility put forth by Yorkston et al. (1996) acknowledges that multiple types of signal-independent information may supplement the acoustic signal, such as syntactic or semantic context, gestures, orthographic cues, and situational cues. Like the intelligibility model, comprehensibility in this context is seen primarily as a ‘feedforward’ model in which signal-independent information enhances the speech signal, contributing to the overall communicative output of the speaker with dysarthria. An adaptation of the Yorkston et al. model can be found in Figure 1.2.

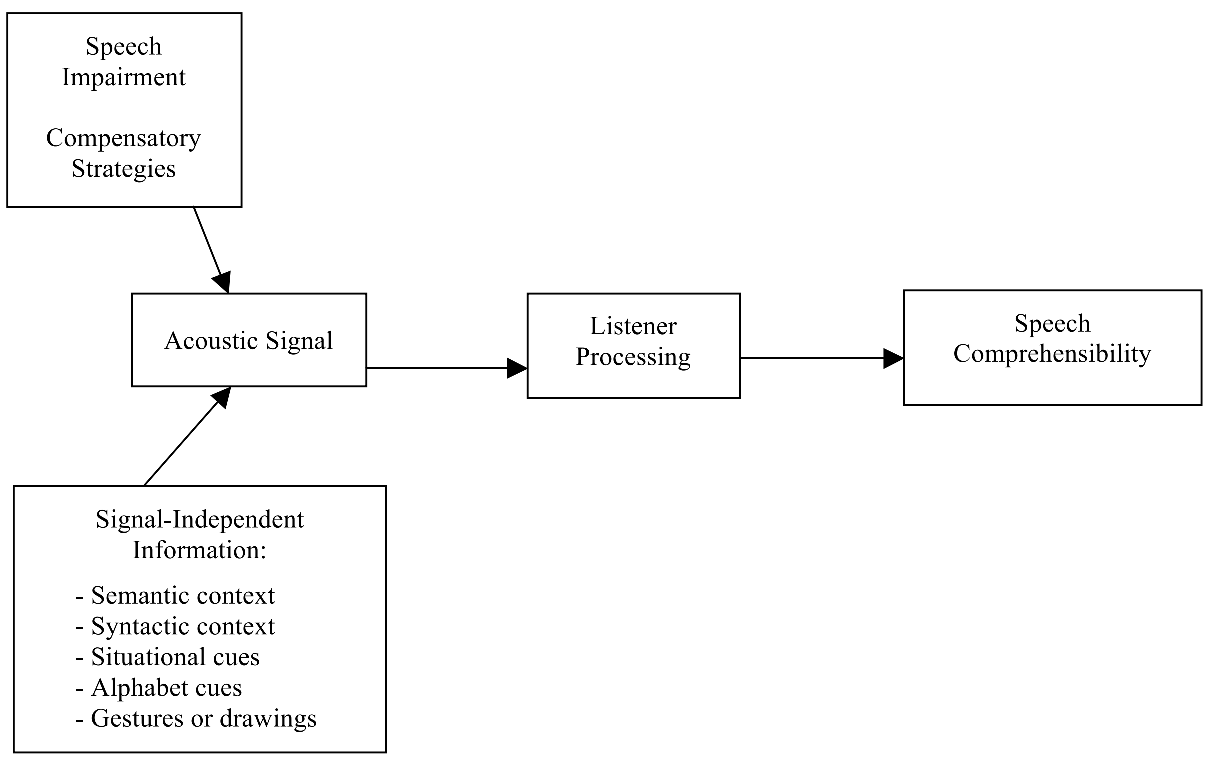

Figure 1.2. A Model of Factors that Affect Comprehensibility of Individuals with Dysarthria (adapted from Yorkston et al., 1996).

1.3.1. The feedforward model of comprehensibility

As in the model of intelligibility (Figure 1.1), the feedforward model of comprehensibility (Figure 1.2) acknowledges the contribution of the speech impairment and compensatory strategies to communicative success. However, it includes supplementary signal-independent information as an additional contributing factor.

1.3.1.1. The contribution of signal-independent information

Yorkston et al. (1996) note that although the bulk of existing research in dysarthria focuses on measures of intelligibility, many of these studies have addressed the contribution of signal-independent information to intelligibility scores. By doing so, these studies have actually tapped into aspects of comprehensibility. However, instead of calling it comprehensibility, the definition of intelligibility itself has broadened (Yorkston et al., 1996). In other words, the examination of how signal-independent information contributes to intelligibility necessitates that intelligibility itself must have a flexible definition. It should be noted that, because of this, the term may be used somewhat differently from author to author. Numerous studies have indeed shown that intelligibility, however it is operationalized, can be affected by the presence or absence of cues independent of the acoustic signal.

1.3.1.1.1. Visual information

Hustad, Dardis, and McCourt (2007) investigated the impact of visual information on intelligibility scores. Seven speakers with dysarthria secondary to cerebral palsy (CP) were audio and video recorded as they read 20 five- to seven-word sentences. Two hundred and twenty "everyday listeners" (see Klasner & Yorkston, 2000) then orthographically transcribed the speakers' utterances under either audio-only (AO) or audio-visual (AV) conditions. Listeners were separated into unique groups of 16 listeners per speaker, per condition. The number of words correctly transcribed by the listeners was then summed and divided by the total number of target words in order to derive an intelligibility score. The results of this study indicated that overall, audio-visual presentation of speech stimuli resulted in higher intelligibility scores than audio presentation alone, although this effect was only significant for three out of seven dysarthric speakers. The authors also noted a strong linear relationship (r=.85) between AO and AV difference scores and mean intelligibility (across AO and AV conditions). In other words, as mean intelligibility increased, so did the gain in intelligibility from the added visual information. These results suggest that the presence of visual information aided listeners in processing a degraded speech stream, and that this was especially true for speakers in their study with mild and moderate dysarthria. The authors also offered that the increased gains observed for speakers with higher intelligibility may indicate that there is a "minimal intelligibility level that must be attained before visual information has a meaningful impact on intelligibility" (p. 364). Interestingly, in their review of several other studies, the authors note that previous research has suggested that there may also be an upper limit to the range of intelligibility within which visual information is helpful to listeners. In other words, it appears from previous literature that visual information may contribute only nominally to improved intelligibility for speakers with mild dysarthria (see, for example, Hustad & Cahill, 2003).

While the presence or absence of visual information has been shown to affect how listeners process the degraded speech stream (e.g., Hustad et al., 2007), it is also of interest to know what visual cues in particular may be exerting this influence. In naturalistic interactions, a listener may have both articulatory and gestural information available as a supplemental source of information. An in-depth study by Garcia and Cannito (1996) investigated how, among a variety of other signal-independent variables, gesture and mode of presentation might affect intelligibility scores for a man with severe flaccid dysarthria. The authors found that for this speaker, utterances produced with accompanying gestures were more intelligible than utterances produced without gestures in an AV condition. Garcia and Cannito also noted that, interestingly, intelligibility scores for sentences with low predictability were significantly higher when listeners transcribed the speaker's utterances accompanied by gestures in an AO condition than when transcription was based on a nongestured AV presentation. Since visual information was not available to listeners in the AO condition, this finding points to the important possibility that producing gestures may alter the speaker's speech production, possibly by altering spoken rhythm or stress (Garcia & Cannito, 1996). However, this finding is difficult to interpret, as the enhancing effect of gestured AO presentation on intelligibility scores was limited to sentences with low predictability, and was not seen in a parallel condition in which sentences were highly predictable. Nonetheless, the results of this study suggest that a vast array of signal-independent variables contributes significantly (and in a highly interactive and complex way) to intelligibility. The presence of visual information, including gestures specifically, proved to be a robust source of meaning for the listeners in this study (Garcia & Cannito, 1996).

1.3.1.1.2. Cueing

In addition to maximizing naturally occurring visual or gestural information for their listeners, people with dysarthria can implement specific strategies to supplement the degraded speech signal. The literature suggests that providing extra cues to listeners (e.g., phonological, syntactic or semantic context) may modify their background knowledge and aid them in speech processing. For example, Hustad, Jones, and Dailey (2003) examined how the implementation of supplemental speech strategies affected intelligibility and speech rate for five speakers with severe dysarthria. The investigators trained the dysarthric speakers in the use of three speech supplementation strategies: topic cues (pointing to a word that indicated the topic of the proceeding utterance), alphabet cues (pointing to the first letter of each word), and mixed cues (combined topic and alphabet cues). Once the strategies were mastered, the speakers were video recorded producing various narratives while implementing each of the three strategies, or using habitual speech (non-cued condition). Intelligibility scores were then obtained from listeners who watched and transcribed the videotaped speech samples, and speech rate was measured from the video recordings. Results showed that intelligibility scores were significantly higher when listeners were provided with mixed cues or alphabet cues alone than when either topic cues or non-cued speech was used. No significant difference in intelligibility was seen between the mixed and alphabet cue conditions, or the topic and non-cued speech conditions, suggesting that alphabet cues may play a significant role in benefitting listeners as they decode dysarthric speech (Hustad et al., 2003). Results for speech rate analysis suggested that both alphabet cue and mixed cue conditions resulted in significantly reduced speech rate (approximately 70% slower) as compared to either topic cues or non-cued speech. In other words, the act of pointing to the first letter of each word spoken had an effect of significantly slowing the speaker's rate of speech. The authors note that slowed speech production alone may have contributed to the improved intelligibility scores observed in these two cue conditions. Overall, the results of this study show that people with dysarthria may benefit from training in the use of overt speech supplementation strategies. Specifically, alphabet cues seem to play a significant role in reducing speech rate and increasing intelligibility, although the mechanism behind this benefit is not clear.

Hustad et al.'s (2003) finding that topic cues did not significantly improve intelligibility compared to non-cued speech is in contrast with other literature. Providing topic information (i.e., semantic context) to listeners has been shown to enhance their intelligibility ratings for dysarthric speakers in several studies. One of the additional findings of the study by Garcia and Cannito (1996; see above discussion) was that intelligibility ratings were higher when listeners were provided with situational context (*in the house*, or *in the yard*), or when sentences were highly predictable, both of which would theoretically aid in semantic processing. Berry and Saunders (1983) also acknowledge that intelligibility can be improved when listeners are given contextual information on which to hang an often fragmented or distorted speech signal. In other words, speakers may exploit the listener's tendency to "fill in the blanks" when messages are predictable by prefacing comments with a preparatory phrase such as "I want to talk about *Thanksgiving dinner*" (p. 207). Berry and Saunders advocate for the use of such topic cueing for dysarthric speakers as part of a larger programme of educating speakers in methods to minimize situational adversity when communicating (see below). Yorkston et al. (1996) also review several other studies that highlight the benefit of both semantic and syntactic information in listener processing of dysarthric speech.

1.3.1.1.3. Situational and environmental context

Berry and Saunders (1983) suggest that, in addition to the presence of visual information and the use of specific cueing strategies, numerous environmental factors (e.g. noise and lighting levels, distance from listeners, etc.) may play a significant role in determining the success of dysarthric speakers' communication attempts. The authors advocate for an approach to rehabilitation involving what they term *environmental education*. Essentially the speaker and those in the speaker's environment should be educated about the environmental facilitators and obstacles to intelligible speech. This approach has its roots in aural rehabilitation strategies for people with hearing impairments. For example, speakers and listeners can be taught to facilitate communication through controlling environmental factors such as extraneous noise, lighting levels, distance between the speaker and listener, and resonance characteristics of the room or environment (see also Yorkston, Bombardier, & Hammen, 1994). The authors also advocate certain speaker-specific factors that can set up a communication situation for success. The speaker's posture should facilitate the maximization of visual cues by ensuring listeners have a full-face view of the speaker. Any external aids the speaker can use to maximize intelligibility (e.g., augmentative and alternative communication, voice amplification, etc.) should be present and accessible. This emphasis on minimizing environmental and situational adversity at an individual level can be considered within the context of the ICF as manipulating the Environmental Factors that contribute to the overall disability experienced by individuals with dysarthria (Dykstra et al., 2007; WHO, 2001).

1.3.1.2. Limitations

While the Yorkston et al. (1996) model of comprehensibility was a step in the right direction of measuring functional communicative abilities for dysarthric speakers, it did not acknowledge all possible contributors to successful communication. Comprehensibility is more than just the sum of all acoustic and signal-independent information produced by or available to the speaker. Dyadic communication, by definition, involves two partners interacting with each other. Yorkston et al.'s model emphasizes the role of the speaker heavily, yet fails to capture important factors related to the listener. The listener brings just as many skills and strategies to the interaction as the speaker does, and contributes significantly to communicative success depending on his or her background knowledge and experience. Furthermore, the model lacks an emphasis on the reciprocal nature of dyadic interaction. The communicative breakdowns inherent in naturalistic interaction are resolved through ongoing and bidirectional feedback mechanisms, which are simply not accounted for within the feedforward model of comprehensibility presented by Yorkston and colleagues.

1.3.2. The elaborated (feedback) model of comprehensibility.

Since Yorkston et al.'s (1996) model was first introduced, it has been acknowledged that other types of signal-independent information may also contribute to comprehensibility (e.g., facial expressions; Dykstra et al., 2007). Factors related to the listener's role in communication have also been acknowledged and emphasized to a greater degree. For example, it has been suggested that the listener's familiarity with and experience listening to dysarthric speech may impact how well he or she understands the speaker with dysarthria (Weismer at al., 2001). Furthermore, and perhaps more importantly, a greater emphasis has been placed on dyadic interaction. For instance, the extent to which the speaker and listener share background knowledge has been postulated to affect communicative success (Lindblom, 1990), and aspects of discourse such as pragmatics (Weismer et

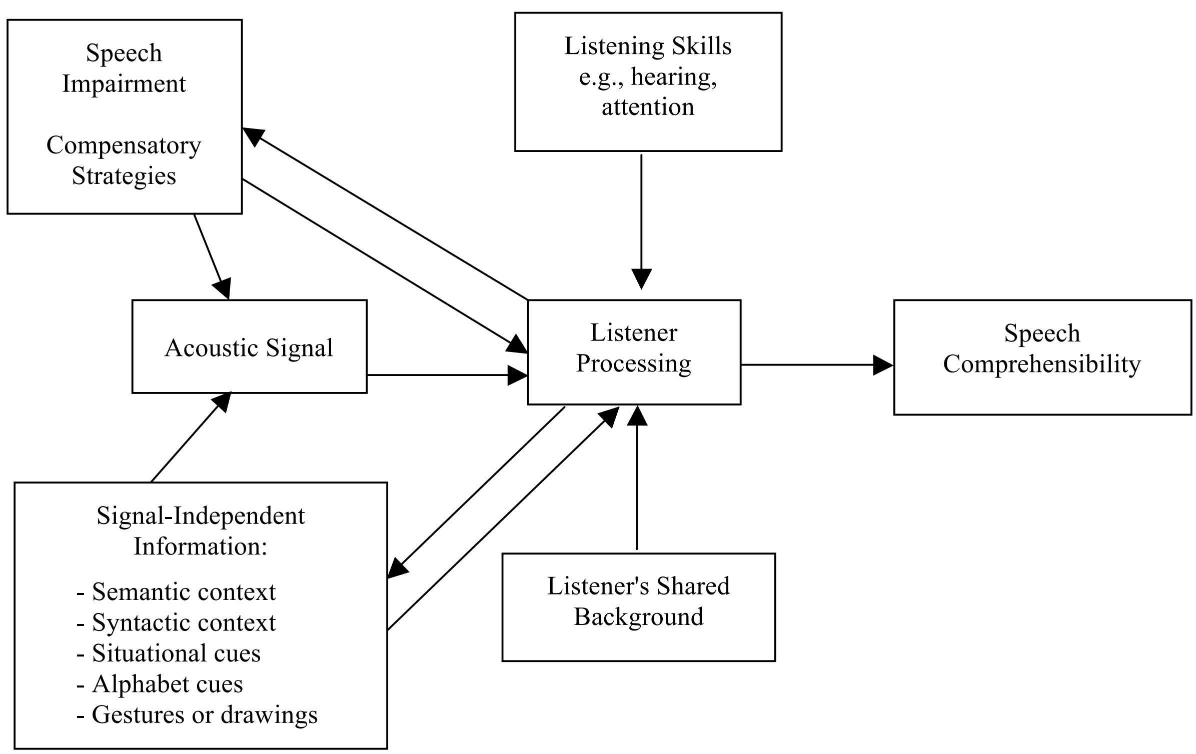

Figure 1.3. A Model of Factors that Affect Comprehensibility of Individuals with Dysarthria (adapted from Purves, 1999).

al., 2001) and breakdown repair strategies (Dykstra et al., 2007; Purves, 1999) have also been included in the consideration of comprehensibility. The key extension made to the original Yorkston et al. model is a move towards viewing comprehensibility as a product of dyadic interaction, and as such, interactive feedback mechanisms are included as significant contributors to comprehensibility (Purves, 1999). In other words, both members of the dyad are seen as contributing to communicative success, through a dynamic and ongoing combination of two mechanisms: what the speaker with dysarthria produces (feedforward) and how the listener responds (feedback). Purves' (1999) elaborated model of comprehensibility is seen in Figure 1.3.

1.3.2.1. Factors related to the listener

Of particular relevance to the present study is the observation that a listener's prior experience can affect intelligibility judgements, either through familiarity with a particular speaker or through exposure to dysarthric speech in general. Tjaden and Liss (1995b) note that investigators have used two main methods of studying familiarity effects, which both involve exposing otherwise naïve listeners to a familiarization procedure. They call the first type *general* familiarization training, which involves exposing the listener to the disordered speech of multiple speakers, or employing listeners with prior experience interacting with individuals who have a particular disorder (Tjaden & Liss, 1995b). The listener then transcribes a target speaker's speech, and the resultant intelligibility score is usually compared to those derived from listeners who were not familiarized in this manner.

Liss, Spitzer, and Caviness (2002) noted several studies that have shown that general experience listening to a degraded speech signal boosts subsequent processing of similar speech. This appears to be the case with exposure to dysarthric speech, specifically. In their own study, Liss et al. familiarized two groups of 40 listeners with dysarthric speech produced by 12 speakers with either hypokinetic or ataxic dysarthria (six speakers with each dysarthria type). The familiarization procedure involved each listener following along with a printed list of 18 familiarization phrases while listening to an audio recording of those phrases being read out loud by the dysarthric speakers. Each group was then asked to transcribe 60 phrases spoken by a speaker with the dysarthria type corresponding to that with which they had been familiarized, followed by 20 phrases spoken by a speaker with the other type of dysarthria. Their transcriptions were compared to each other as well as to a control group of listeners who had not received any familiarization. Results indicated that intelligibility scores by listeners given this brief exposure to dysarthric speech were 15%-20% higher than those scored by unfamiliarized listeners. In addition, the intelligibility scores reflected both general and specific familiarization effects. Scores from listeners transcribing dysarthric speech of the same type as that to which they were exposed in the familiarization task were significantly higher than those from speakers transcribing dysarthric speech of a different type than that with which they were familiarized (specific effect). These scores, in turn, were significantly higher than the unfamiliarized control condition scores (general effect). In fact, "the benefit of being familiarized with the specific dysarthria was roughly double the intelligibility gain for the familiarized with other condition" (p. 3028). What these results suggest is that brief experience listening to dysarthric speech from a variety of speakers may enhance the processing of degraded speech from other individuals with dysarthria, whether the perceptual features of the speech signal itself are similar or not. There does, however, appear to be a particularly substantial processing benefit gained from experience with a specific type of disordered speech (Liss et al., 2002).

Indeed, there is considerable evidence that experience with a specific speaker's speech pattern aids in comprehension. These studies have looked at the effects of what Tjaden and Liss (1995b) have called *specific* familiarization training. In this type of procedure, the listener is familiarized with the speech of one particular speaker before being asked to score his or her intelligibility.

Tjaden and Liss (1995a, 1995b) reported results of two related studies of the same speaker with moderate-to-severe mixed spastic-ataxic dysarthria secondary to CP. In the first study (Tjaden & Liss, 1995a), the authors compared intelligibility scores from listeners who transcribed sentences produced by the speaker who read them in her customary manner (control) or while implementing a breath-group strategy (treatment). A third group listened to the breath-group strategy after undergoing a familiarization procedure during which they listened twice to a recording of the speaker reading a paragraph in her usual manner (familiarization). They found that significant differences existed in intelligibility scores for each of the three conditions. The familiarization group yielded intelligibility scores that were significantly higher than those obtained from the treatment group, the scores of which were, in turn, significantly higher than those from the control group. One major limitation of this study is that there was no comparison group of familiarized listeners who listened to the speaker producing sentences habitually. Because of this, it is difficult to tease apart the precise effects of familiarization from those of treatment. Despite this limitation, these results provide some support to the notion that speaker-oriented interventions may interact with familiarity in an additive manner to improve communicative success.

In the second study (Tjaden & Liss, 1995b), listeners were familiarized in one of two ways. In the paragraph condition, they listened to the speaker reading a paragraph that contained 72 words, while in the word list condition, they listened to the speaker reading the same 72 words from the paragraph condition, but in a random order. They found that intelligibility scores based on a subsequent transcription task were higher for listeners in the paragraph condition compared to the word list condition, although this difference was not statistically significant. Both familiarization conditions resulted in significantly enhanced intelligibility scores compared to a control (unfamiliarized) condition. These results indicate that listeners may have gained some extra information from exposure to natural between-word coarticulation and prosodic sentence-level patterns (paragraph condition) compared to articulation patterns at the word level only (word list condition; Tjaden & Liss, 1995b). One should exercise caution in interpreting these data, however, as the effect was not statistically significant.

A similar study undertaken by D'Innocenzo, Tjaden, and Greenman (2006) examined the effects of familiarity and speaking condition on intelligibility scores for a speaker with mixed spastic-flaccid dysarthria (secondary to traumatic brain injury). One hundred and twenty listeners were assigned to no familiarization, word list familiarization, or paragraph familiarization conditions. Listeners in the two familiarization conditions listened to the speaker read either the paragraph or word list, and followed along with a written script. In the transcription task that followed, listeners were exposed to an audio recording of the speaker reading 15 sentences in one of the habitual, fast, slow, or loud speaking conditions. In line with the results of Tjaden and Liss (1995b), these authors found significant intelligibility gains of 8%-10% for both word list and paragraph familiarization conditions, with no differences found between the two familiarization conditions when data from all speaking conditions were pooled. The authors did observe, however, that like Tjaden and Liss (1995b), the paragraph familiarization condition yielded the highest intelligibility scores for habitual speech (D'Innocenzo et al., 2006). An additional finding from this study is that the loud speaking condition yielded a marked increase in intelligibility across familiarization groups, although there was no significant Speaking Condition × Familiarization interaction observed (D'Innocenzo et al., 2006).

Yorkston and Beukelman (1983) conducted a seminal study that involved specific familiarization procedures with multiple dysarthric speakers. Nine speakers with dysarthria of varying severity levels were recorded producing two different sets of sentences taken from the *Assessment of Intelligibility for Dysarthric Speech* (AIDS; Yorkston & Beukelman, 1981). Nine listeners were randomly assigned to one of three groups: (1) a no familiarization, (2) familiarization with the speaker, or (3) familiarization with specific feedback. Listeners familiarized with the speaker transcribed the first sentence set, then listened to it again three additional times before transcribing the second sentence set. Listeners familiarized with specific feedback transcribed the first sentence set, then listened to it again three additional times while following along with an accurate transcript, then transcribed the second sentence set. The unfamiliarized control group transcribed the second sentence set two weeks after transcribing the first. Contrary to expectations, the authors did not find any significant increase in intelligibility scores between prefamiliarization and postfamiliarization conditions. In fact, scores from the two familiarization groups were actually lower than those from the unfamiliarized group. This was observed to be the case for both sentence sets (prefamiliarization and postfamiliarization). One explanation for this is that due to random assignment, the judges' levels of experience were not balanced. Three speech pathologists were randomly assigned to the unfamiliarized group, while both the familiarization groups were made up of one speech pathologist and two less experienced student clinicians. In other words, the group that did not undergo a familiarization procedure may have actually had more general experience in listening to degraded speech overall, an effect that was not counterbalanced by the brief specific familiarization training procedure the other two groups received. Had each listening group been composed of individuals with similar experience, a familiarization effect may have emerged.

A study on perceptual learning by Buchholz (2009) examined whether specific familiarization training would increase the accuracy with which listeners identified word-initial voiced and voiceless stop contrasts. A single speaker with dysarthria secondary to CP was recorded producing a set of CVC words, and these tokens were divided into several word lists. Thirty listeners first listened to a pre-familiarization word list and were scored based on the percentage of words they correctly identified from a two-item forced-choice presentation. They were then randomly assigned to one of three familiarization conditions: feedback, no feedback, and pseudo-feedback. As in Yorkston and Beukelman's (1983) study, the feedback training consisted of presenting the listener with audio recordings of the familiarization word list along with the written words corresponding the words spoken. The pseudo-feedback condition involved the listeners listening to and receiving feedback about foils that did not contain the word-initial stop voicing contrast. After the familiarization phase, two post-familiarization lists were presented and listeners were scored again via forced-choice selection. Results showed that perceptual learning occurred after familiarization training, although the pattern of changes in the scores differed markedly for tokens with typical voice onset time (VOT; absence of prevoicing for voiceless stops and some prevoicing for voiced stops) and atypical VOT tokens (shorter prevoicing duration for voiced stops and longer prevoicing duration for voiceless stops). For typical items, the largest increase in listener accuracy was observed for the post-familiarization word list that contained novel tokens, while tokens from the post-familiarization word list containing the same items from the pre-familiarization list were actually identified with significantly less accuracy than prior to familiarization training. For atypical items, voiced stops were identified with poorer accuracy after familiarization while voiceless stops were more accurately identified, although this pattern did not generalize to the novel post-familiarization word list. It was also observed that the feedback condition did not lead to any significant increase in accuracy scores for either token type, as hypothesized. These results do suggest, however, that the familiarization procedure employed was effective in changing the accuracy with which listeners identified word-initial voiced and voiceless stops. Furthermore, the pattern of perceptual learning that occurred differed depending on the typicality of the tokens in question, suggesting that familiarization with the often atypical acoustic patterns of dysarthric speech may result in unique, speaker-specific learning for listeners.

Flipsen (1995) acknowledges that a third type of familiarity exists: *personal* familiarity.

While this type of familiarity cannot generally be achieved through an experimental manipulation or familiarization training procedure, it is of utmost clinical relevance in the study of functional communication. Personal familiarity exists between two individuals who frequently communicate with one another (and have done so for a significant time period), share a substantial amount of background knowledge, and are accustomed to each other's communicative habits. For example, the immediate family members, caregivers, or close friends of a person with dysarthria are likely to fall into this category. It is also possible for clinicians to achieve some level of personal familiarity with their clients through long-term, regular, and personalized therapeutic contact.

Anecdotal evidence suggests that these highly familiar individuals may understand more of what the speaker says than others in the community or environment (Ferrier, 1991). Self-report data reported by Hartelius, Elmberg, Holm, Lövberg, and Nikolaidis (2008) supported this observation, as dysarthric speakers were more likely to report difficulty communicating with strangers (36.3% agreed) than difficulty communicating with family members (7.3% agreed). Ball et al. (2004) have reported similar data from participants with ALS, who rated speaking in a quiet environment to a familiar person as more effective than when the listener was unfamiliar. Parents of children with significant speech delay have also reported that they understand their child's speech better than others in the child's environment (Flipsen, 1995). Flipsen examined this issue empirically by administrating an imitation-based version of the single-word portion of the AIDS (Yorkston & Beukelman, 1981) to four speech-delayed children. His results agreed with self-report data, as parents understood their child's utterances significantly more than did classroom teachers and speech pathologists, with mothers having a particular advantage in correctly interpreting their child's speech (compared to fathers). This suggests that measurable familiarity effects are robust enough that they are noticeable to both speakers and communication partners within the personal relationships affected by speech impairment.

What is the mechanism behind these familiarity effects? One possibility is that a familiar listener may process the speech signal more efficiently (D'Innocenzo et al., 2006). In other words, repeated exposure to the speaker's unique speech patterns may give familiar listeners a superior ability to 'translate' the speech signal into a meaningful stream (Flipsen, 1995). Kent and Read (1992) have referred to this perceptual learning effect within the context of normal inter-speaker variability. What they term *speaker normalization* refers to how listeners interpret the speech of multiple speakers correctly despite considerable inter-speaker variability in the input, for example, in vowel formant frequencies (Kent & Read, 1992). One explanation for how speaker normalization works is that listeners may create speaker-specific reference frames or templates through which to interpret the incoming speech signal. These reference frames could be thought of as a set of acoustic averages. Although Kent and Read referred primarily to how listeners process normal speech, this concept can also be applied to the processing of degraded speech streams, as in the case of dysarthric speech. In this case, listeners may be partaking in speaker normalization in the process of habituating to the degraded speech signal. This may be realized through a set of modified acoustic-phonetic expectations on the part of the listener that allow for the speaker's less-than-perfect productions to map onto the listener's normal underlying phonological representations (Liss et al., 2002). While Kent and Read originally presented speaker normalization as a problem to be solved, Buchholz (2009) argues that in the case of dysarthria, speaker-specific knowledge may help in the perception of degraded speech. She states, "with previous experience, the memory of talker specific acoustic-phonetic information allows listeners to customize perceptual strategies to a specific talker's speech pattern" (p. 17).

Precisely what part of this listening experience influences subsequent processing of dysarthric speech is not well understood. Some authors have hypothesized that the benefit may be attributed to information gleaned at the segmental level, or both the segmental and suprasegmental (e.g., prosodic) levels (D'Innocenzo et al., 2006). The results of D'Innocenzo et al. (2006) and Tjaden and Liss (1995b), in which paragraph familiarization failed to result in significantly higher intelligibility scores compared to a word-list familiarization, suggest that familiarization effects may be attributed primarily to enhanced segmental perception, which would result from repeated exposure to the word-level productions available in both familiarization conditions (D'Innocenzo et al., 2006; Liss et al., 2002; Tjaden & Liss, 1995b). Regardless of the mechanism behind these improvements in listener processing, it is clear that some degree of perceptual learning is taking place when listeners are exposed to dysarthric speech.

1.3.2.2. Factors related to dyadic interaction

In addition to what experience the listener brings to an interaction, the dyadic exchange itself is a significant contributor to comprehensibility. Barefoot et al. (1993), in one of the few studies to examine comprehensibility directly, investigated communicative success in a sample of 41 deaf adult speakers, a population with well-documented atypical speech and language patterns. Four adults with normal hearing watched video recorded conversation samples of the deaf speakers interacting with two normally hearing undergraduate students. The observers rated the comprehensibility of the deaf speakers' utterances on a nine-point rating scale, where a score of 1 indicated that utterances were understood "not well", a 5 indicated the utterances were understood "fairly well", and a 9 indicated the utterances were understood "very well" (Barefoot et al., 1993). Intelligibility scores were also collected for each speaker based on the number of key words correctly recognized from audio recorded reading samples. Results indicated that the mean comprehensibility score for the deaf students was 6.2, and that intelligibility scores accounted for

55.7% of the variance in comprehensibility scores. While the rating scale only addressed the comprehensibility of the deaf speakers, this study was unique in its examination of comprehensibility within dyadic interactions. Indeed, this was an important step towards acknowledging the substantial contribution of bidirectional feedback mechanisms to comprehensibility.

It is of interest what specific factors related to such dyadic interaction may contribute to comprehensibility. Several studies have examined how the characteristics of dyadic communication differ depending on the level of familiarity between the speaker and the listener. First, it is possible that communicative strategies and styles may differ due to varying familiarity levels. Supporting this notion, part of a larger study by Ferrier (1991) reported the conversational characteristics of a speaker (L.L.) with mixed spastic-ataxic dysarthria secondary to CP. The speaker engaged in two 15-minute conversations, one with an unfamiliar partner, and one with a familiar partner (a student clinician who had been working with him for approximately one year). When communicating with the familiar partner, L.L. produced far more utterances at a faster rate, produced more utterances per turn, and was more likely to use his natural speech (instead of his voice output communication device) than when communicating with the unfamiliar partner. There were also differences in the observed behaviours of the familiar and unfamiliar partners. As with the speaker, the familiar partner produced substantially more utterances at a much faster rate than did the unfamiliar partner. The authors also observed that the conversation between L.L. and the familiar partner resulted in twice as many conversational breakdowns as were observed in the unfamiliar conversation. There were, however, far more opportunities for breakdown in the familiar condition, as more total utterances were produced by both partners. Presumably as a strategy for ensuring comprehension, the familiar partner repeated L.L.'s utterances 40% of the time, while the unfamiliar partner rarely used repetition this way (8% of L.L.'s utterances were repeated). What this suggests is that the speaker's familiar partner "had evolved a system of providing feedback to L.L., after each utterance, or often after each word, to allow L.L. a check on whether he was understood" (Ferrier, 1991, p. 271). Interestingly, the authors observed that L.L. transmitted 76% of his morphemes successfully to the unfamiliar listener on the first attempt, while the familiar listener successfully received only 61% of initial presentations. This may have been due to the fact that the speaker demonstrated a strong preference for using his natural voice with the familiar partner, while his speech output device was relied on more heavily with the unfamiliar partner. What all these data suggest as a whole is that the conversation between L.L. and the familiar partner was more efficient. Each speaker produced more utterances at a faster rate, and while more communication breakdowns resulted, they were also efficiently resolved due to the repetition strategy the student clinician used.

Indeed, the evidence supporting the fact that familiar listeners may contribute to more effective and efficient resolution of communicative breakdowns is building. Findings from a study of strategy use and familiarity by King and Gallegos-Santillan (1999) support this notion. Each of seven naïve listeners participated as unfamiliar partners (UPs) in a videotaped 15-minute conversation with one of seven dysarthric speakers, while seven other participants who were personally familiar with the dysarthric speakers participated as familiar partners (FPs) doing the same task. Both UPs and FPs also completed questionnaires (adapted from Yorkston et al., 1996) about what strategies they used when speaking with the dysarthric speakers. These authors found that people with dysarthria used significantly more strategies (specifically reparative strategies) when speaking with UPs than with FPs, and that the UPs themselves also used more strategies than did the FPs. However, unlike Ferrier (1991), they attribute this observation to the fact that lower levels of listener familiarity may have led to more communication breakdowns, which in turn resulted in a greater need on the part of the dysarthric speaker to rely on additional communication strategies (King & Gallegos-Santillan, 1999). Again, these results can be interpreted to support the notion that communication between dysarthric speakers and their familiar listeners is more efficient and effective.

The second part of King & Gallegos-Santillan's (1999) study, reported in an unpublished master's thesis by Villanueva (1997), also examined the use of strategies by dysarthric speakers and their conversation partners. Following the methodology of King and Gallegos-Santillan, dysarthric speakers and their partners (FPs and UPs) completed questionnaires about what kinds of strategies they used when communicating, and these self-reported strategies were compared to those observed during 15-minute videotaped conversations. The final ten minutes of each recorded conversation were also transcribed and coded for the number of communicative exchanges that were successful (SCEs), repaired (RCEs), or unsuccessful (UCEs) as a method for examining the comprehensibility of the interactions. Villanueva operationalized comprehensibility as the percentage of total communicative exchanges that were ultimately successful, i.e., those that were successful on the first attempt (SCEs) or those that were repaired and thus eventually successful (RCEs). In other words, the number of SCEs and RCEs were summed and divided by the total number of communicative exchanges to derive a percentage score for comprehensibility. As in the King and Gallegos-Santillan study, Villanueva found that dysarthric speakers used more strategies with UPs than with FPs, and the UPs themselves also used more strategies. Furthermore, the strategies used in the unfamiliar conversations were largely reparative, signaling more communication breakdown than in the conversations with FPs. This was confirmed by examining the success of communicative exchanges, as comprehensibility percentages were higher in the familiar conversations. That is, the percentage of ultimately successful communicative exchanges was higher for interactions between dysarthric speakers and FPs. Villanueva attributes the use of fewer strategies and more successful conversation in the familiar condition to the fact that the interlocutors had likely learned how to communicate more effectively with one another, while the UPs were not accustomed to hearing dysarthric speech.

Studies from the literature on hearing impairment can help to shed light on the specific types of breakdown repair strategies dysarthric speakers and their communicative partners may use. The consideration of such studies is relevant to the current discussion due to the considerable amount of similarity in the communication problems experienced by those affected by both hearing impairment and dysarthria. In the context of dyadic interaction, both disorders have in common the processing of a degraded acoustic signal, and, as a result, a tendency towards communicative breakdown. The difference is simply in the origin of the degradation leading to breakdown. For dysarthric speakers, breakdown occurs primarily due to a disordered speech production mechanism, while for hearing impaired listeners, the signal produced is adequate, yet a degradation is imposed on it due to an impaired speech receiver. In this way, hearing impairment can be conceptualized as the reverse of dysarthria.

Caissie and Rockwell (1993) have identified several types of communicative strategies employed by both hearing impaired individuals and their conversational partners. Details of these strategies can be seen in Tables 1.1 and 1.2. The strategies listed in Table 1.1 that are typically used by hearing impaired participants (e.g., request for repetition) may be seen as analogous to those used by a normal hearing listener in a dyadic interaction with a dysarthric speaker. Similarly, the strategies typical of the communication partners of hearing impaired people (e.g., paraphrase; Table 1.2) can be compared to those that may be used by speakers with dysarthria.

Seminal research conducted by Schegloff, Jefferson, & Sacks (1977) has examined in detail the types of breakdown repair strategies used by nondisordered conversation partners. The authors

Table 1.1. Breakdown repair strategies used by people with hearing impairment (adapted from Caissie & Rockwell, 1993).

Category	Definition
Nonspecific	
Request for repetition	Neutral request for the partner to repeat the message. May be expressed either directly (e.g., "What?", "Huh?", "Could you repeat that?") or indirectly (e.g., "I didn't get that.")
Nonverbal request for clarification	Quizzical facial expression or body gesture, such as leaning in or cupping a hand over the ear, used to indicate to a partner that his or her message was misunderstood.
Specific	
Request for repetition of a specific constituent	A wh- question formulated in such a way to target the portion of the message that was misperceived (e.g., "We'll meet at 6 o'clock"; "At what time?"). Also includes requests for the partner to repeat the first, middle, or last part of the sentence.
Request for confirmation of the message	Partial or full repetition of the partner's message using a rising intonation or preceding it with "Did you say" (e.g., "We'll meet at 6 o'clock"; "did you say 6?").
Request for confirmation of the topic	Utterance that inquires about whether the partner is talking about a particular topic (e.g., "Are you still talking about tomorrow night's party?").
Request for explanation/elaboration	Utterance used to acquire additional information that might facilitate understanding. (e.g., "Could you please give me more information?").
Request for a change in the manner of presentation	Utterance that identifies the source of the difficulty and requests the partner to modify his or her speech accordingly (e.g., "Could you repeat that slowly?"; "Could you speak up?"; "Could you not obscure your mouth when speaking to me?").

observe that the attempt to initiate a repair does not necessarily result in a repair or correction being made (i.e., repair initiations can fail), and thus discuss breakdown repairs as having two component parts: an initiation and an outcome (Schegloff et al., 1977). When initiations do lead to successful repairs, it is also important to specify which party does the work of repairing the breakdown. Schegloff et al. argue that breakdown repairs can be carried out by either the *self* or the *other*, from the perspective of the speaker of the trouble source (the source of the breakdown). This dichotomy leads to four possible repair types: self-initiated self-repair (cf. spontaneous repair; Caissie &

Rockwell, 1993), self-initiated other-repair, other-initiated self-repair, and other-initiated other-repair. Schegloff and colleagues have noted there appears to be a natural preference for self-initiated self-repairs in the organization of discourse between nondisordered conversation partners.

Table 1.2. Breakdown repair strategies used by the conversational partners of people with hearing impairment (adapted from Caissie & Rockwell, 1993).

Category	Definition
Exact repetition	Full repetition of the misperceived message.
Partial repetition	Repetition of a portion of the misperceived message. Includes repeating a keyword.
Paraphrase	Rephrasing the message using words having similar meanings without providing new information.
Elaboration/explanation	Expressing additional information or defining keywords in order to provide supplementary cues.
Syntactic modification	Simplifying the syntax of the message, such as breaking the misperceived sentence into two shorter sentences.
Confirmation	Yes/no answer to the partner's request for confirmation (e.g., "Did you say 6 o'clock?"; "Yes.")
Nonverbal cue	Addition of explicit nonverbal communication to clarify the misperceived message (e.g., pointing to a referred object).
Change in prosodic features	Changes in speech patterns to make the signal more audible or clear (e.g., stressing a keyword, increasing overall loudness, or reducing speaking rate.
Spontaneous repair	Self-initiated repair when a breakdown occurs and the hearing-impaired partner does not request clarification.

Specifically, they argue that opportunities for a speaker to initiate a repair (self-initiation) naturally precede opportunities for a repair to be initiated by his or her partner (other-initiation; Schegloff et al., 1977). This is because the speaker may initiate a repair within his or her turn, while his or her partner must wait for the speaker's turn to end before gaining the conversational floor and initiating a repair (other-initiated repair). In other words, other-initiated repair generally occurs once the trouble source turn of the speaker has concluded. By waiting for the floor, the speaker's partner allows the speaker to take advantage of an opportunity to self-initiate a repair. Indeed, the most commonly observed repair type in the nondisordered discourse literature is self-initiated self-repair, which attests to its efficiency as a repair strategy. It is possible that self-initiated self-repairs tend to be conversationally favoured (Schegloff et al., 1977) because they are the least disruptive breakdown repair strategy of the four possibilities. In this way, self-initiated self-repairs may be seen as contributors to comprehensibility, while other repair types (e.g., other-initiated other-repair) may interrupt the flow of conversation more negatively.

1.4. Summary

Regardless of the model of comprehensibility one subscribes to, it is clear that naturalistic interactions inherently involve and make use of numerous kinds of signal-independent information and dynamic interactive processes, which are simply not accounted for within the limited scope of intelligibility. Thus, a model of comprehensibility more accurately represents a functional or holistic approach to the assessment and treatment of dysarthric speech than what is currently captured through an overreliance on intelligibility measures.

With the exception of a few studies (Barefoot et al., 1993; Villanueva, 1997), the majority of the available literature has not specifically addressed the notion of comprehensibility. These studies have, however, certainly ventured into its territory. What this recent exploration beyond the acoustic signal represents is a move towards viewing the perception of dysarthric speech in a more top-down manner (see Villanueva, 1997, for a discussion). In contrast to the traditional bottom-up approach where the component parts of speech and language are considered separate until added together to form a whole (e.g., as in intelligibility), a top-down approach would acknowledge the simultaneous influence of multiple factors on communicative success, but ultimately treat it as a whole. These factors would include the signal-independent information that contributes to comprehensibility, and the dynamic nature of dyadic interaction (e.g., the use of breakdown repair strategies). It is important to consider all routes to communication that may be utilized by speakers with dysarthria and their communication partners, as in the presence of a degraded speech signal, more emphasis may be placed on both partners' skill in employing them. In fact, it has been argued that, in contrast to processing normal speech, listeners may actually rely more heavily on top-down processing when decoding degraded speech (Duffy & Pisoni, 1992).

1.5. A clinical application of the comprehensibility model

While several of the aforementioned studies have investigated what factors influence comprehensibility for speakers and dyad members, none has focused on the clinical application of this construct. However, one recent study, an unpublished master's thesis by Visser (2004), ventured beyond an examination of the contributing factors and explored comprehensibility in a clinically relevant way. Visser's work described the development and application of a structured assessment of comprehensibility for a man (E.) with profound mixed spastic-hypokinetic dysarthria. The assessment process involved E. reading aloud several randomly ordered, semantically-unpredictable sentences of varying lengths (five to ten words) taken from the AIDS (Yorkston & Beukelman, 1981). Without access to the written sentence, the listener (G., a student clinician) had to try to repeat each sentence, word for word, with both members of the dyad working to resolve any communication breakdowns that occurred. These interactions were videotaped and later transcribed for analysis. The transcripts were then scored on six parameters hypothesized to represent the comprehensibility model by Purves (1999): (1) total time, (2) problem words, (3) extra speaker words, (4) extra listener words, (5) number of words guessed incorrectly, and (6) ultimate correct (i.e., the number of target words correctly identified by the listener).

Both standard intelligibility measures (the AIDS; Yorkston & Beukelman, 1981) and the videotaped comprehensibility assessments were administered at two points in time, one at the beginning of G.'s six-week clinical externship (T1), and one at the end of the externship (T2). Visser (2004) found that scores on all six measures from the structured assessment of comprehensibility were improved at T2 compared to T1, suggesting increased communicative success for the dyad over time. In addition, Visser found that although E.'s scores on a test of intelligibility were relatively unchanged over the six-week interval when based on transcriptions from an unfamiliar listener (T1: 7%, T2: 9%), the scores based on G.'s transcriptions were 15% higher at the end of her externship (T1: 7%, T2: 22%, p. 52). As both the unfamiliar scorer and G. based their intelligibility transcriptions on the same audiotapes at both time points, these findings suggest that G. had learned how to more effectively process E.'s profoundly unintelligible speech from working with him closely for six weeks (Visser, 2004). In other words, the increase in intelligibility scores (for G.) was likely due to increased familiarity.

The experience G. and E. gained from repeated interactions with each other during the six week period not only improved how well G. processed E.'s degraded speech stream (as reflected by the increased intelligibility scores), but led to some qualitative changes in their interaction as a dyad. Visser (2004) found that strategy use for both E. and G. changed slightly between T1 and T2. G.'s requests for repetition became less specific over time, and E.'s repetitions favoured partial repetitions over repetitions of the whole sentence. E. also relied more on orthographic cues (pointing to an alphabet board) at T1 than at T2. At the end of six weeks, both members of the dyad used fewer strategies and were able to repair breakdowns more effectively, which Visser attributed to increased familiarity with each other's communicative styles and breakdown repair strategies.

Visser's (2004) study presented some promising preliminary evidence that the structured assessment of comprehensibility was a reliable and valid clinical tool. However, its development was based solely on the communicative characteristics, strategies, and styles of a single dyad. Since there is such a significant natural variability in how different individuals communicate, it remains to be seen whether this assessment tool could be applied to other dyads (and to speakers with more widely varying types and severities of dysarthria) with the same high level of reliability that was found by Visser. In light of this need, the present study aimed to explore the application of the structured assessment of comprehensibility to multiple dyads, while at the same time investigating the issue of how listener familiarity contributed to intelligibility and comprehensibility scores. This investigation aimed to further establish the structured assessment of comprehensibility as a valid and reliable clinical tool that could be used as part of a more functional evaluation of severity and disability for persons with dysarthria.

2. RESEARCH QUESTIONS

The overarching goal of the present study is to evaluate the reliability and validity of the structured assessment of comprehensibility developed by Visser (2004) for measuring dyadic communicative success for dysarthric speakers. The study extends Visser's findings by exploring data from two case studies, each including assessments from two different dyads scored by multiple raters, in order to further evaluate the protocol's reliability. Further, because each case includes a familiar and an unfamiliar partner, it will be possible to examine the effect of listener familiarity on successful communication, shedding light on the validity of the structured assessment. If the assessment is valid, it should be able to differentiate between familiar and unfamiliar dyads. Specifically, the following three research questions will be addressed:

2.1. Question 1. Is the structured assessment of comprehensibility reliable?

Although Visser's (2004) thesis study established high levels of inter-rater reliability for scoring the structured assessment of comprehensibility, the assessment was created specifically for the dyad studied. Whether it is an appropriate and reliable assessment tool for multiple dyads has yet to be established. The application of the assessment to multiple dyads will help strengthen its reliability as a clinical tool, and in the process, may provide an opportunity for important modifications to administration and scoring procedures that did not emerge during its initial development.

2.2. Question 2. Will this measure of comprehensibility differentiate between familiar and unfamiliar dyads?

There is extensive research showing that communication is more successful for dysarthric speakers when communication partners are familiar than when these partners are unfamiliar (Ball et al., 2004; D'Innocenzo et al., 2006; Ferrier, 1991; Flipsen, 1995; Hartelius et al., 2008; King & Gallegos-Santillan, 1999; Liss et al., 2002; Tjaden & Liss, 1995a; Tjaden & Liss, 1995b;

Villanueva, 1997). These authors have shown that higher levels of partner familiarity benefit both intelligibility (i.e., processing of the acoustic signal) and efficiency of strategy use (e.g., King & Gallegos-Santillan, 1999). In support of these claims, Visser (2004) found that all measures of comprehensibility improved by the end of G.'s six-week externship. It is unclear how much of this improvement was due to G.'s increased ability to process E.'s degraded speech signal (as evidenced by the 15% increase in intelligibility when she scored E.'s speech at T2). Visser also noted that both G. and E. exhibited more efficient and effective communication strategies at the end of the six-week time period. Based on these findings, it is important to examine whether the results of the structured assessment of comprehensibility would reflect more successful communication in familiar dyads than unfamiliar dyads.

Hypothesis 1. Comprehensibility scores will be better for familiar dyads than unfamiliar dyads.

Hypothesis 2. Intelligibility scores will be higher when scored by familiar vs. unfamiliar listeners.

Hypothesis 3. Differences in intelligibility scores will not fully account for differences in comprehensibility scores.

2.3. Question 3. How do familiar and unfamiliar dyads differ in terms of strategy use?

Based on the findings of King & Gallegos-Santillan (1999) and Villanueva (1997) that dyads of dysarthric speakers and unfamiliar listeners use more communicative strategies than familiar listener dyads during conversation, how strategy use may differ within the context of the structured assessment of comprehensibility is of interest. As Visser (2004)'s study found qualitative changes in strategy use and efficiency between T1 (unfamiliar) and T2 (familiarized), it is reasonable to expect strategy use to differ between familiar and unfamiliar dyads.

Hypothesis 1. Unfamiliar dyads will employ a greater range of strategies than familiar dyads.

3. METHODS

The following chapter outlines the participants, materials, and procedures used in this study. Detailed explanations of the procedures for evaluating intelligibility and comprehensibility are provided, along with the method for data analysis.

3.1 Participants

3.1.1. Procedures and criteria for recruitment

3.1.1.1. Participants with dysarthria

Two participants with acquired dysarthria were recruited to participate in this study. With such a small sample size, it was important to keep the type of dysarthria constant across both speakers in order to minimize confounding variables due to differing etiology or motor speech characteristics. In addition, the speakers' level of experience with their dysarthria was an important consideration. As some types of neurological injuries may change or begin to resolve rapidly post-onset, dysarthria severity may also change considerably. That is, a relatively recent onset of dysarthria may mean that the dysarthria is unstable, making it unlikely that the affected individual would have the opportunity to develop reliable compensatory strategies. In contrast, it is logical that significant experience with a more longstanding dysarthria would lead to well-established and well-practiced communication strategies, which would be more likely to provide a reliable representation of communicative success across multiple speaking situations. Essentially, individuals who have lived with their speech impairments for a significant amount of time would be more accustomed to communicating with such a disability than would be those whose dysarthria is a relatively new development. Based on this rationale, individuals with relatively longstanding dysarthria were considered to be optimal candidates for this study.

Dysarthric participants with intact cognition were also preferred candidates for this study, so as to ensure that they would be able to make use of signal-independent information and verbal communication simultaneously, which Yorkston et al. (1996) notes is a complex cognitive activity. It was also necessary for participants to be able to understand and follow directions for completion of the intelligibility and comprehensibility tasks, which required intact and effective attention, problem-solving skills, and memory. Because of this, potential participants who presented with cognitive impairment were excluded from participation. Similarly, dysarthric participants were also excluded if they presented with other comorbidities or concurrent impairments that were judged to be significant enough to impact communicative effectiveness for the required tasks (e.g., aphasia, hearing impairment, visual impairment).

Given the above inclusion criteria, dysarthric speakers were recruited from a local clinic that offered services to clients with ALS and PLS (primary lateral sclerosis). The progressive nature of these neurological disorders ensured that the participants with dysarthria had sufficient experience with their speech disorder, and that participants could be identified within this clinical population who met the criteria for intact cognition. Once this recruitment source had been identified, further inclusion criteria were specified based on consideration of the model of comprehensibility (Purves, 1999) and the clinical tool in question (Visser, 2004), including its application to dysarthric speakers with various clinical features.

The severity of dysarthria for speakers was purposefully constrained to the moderate to severe range, effectively excluding those with mild or profound dysarthrias from participation in this study. The exclusion of mild dysarthria was partly based on Lindblom’s (1990) argument that the speech signal of speakers with mild dysarthria remains sufficiently intelligible that little added benefit would be gained through the use of signal-independent information. In other words, it was reasoned that a mildly impaired speaker’s performance might result in ceiling effects on the structured assessment of comprehensibility. Thus, this particular assessment tool was deemed to be clinically inappropriate for a population of speakers with mild dysarthria. Additionally, subjects with profound dysarthria were excluded from participation because extremely low intelligibility proved to be a significant challenge for transcription and analysis in Visser's (2004) study. For the purposes of recruitment, descriptive labels for severity were based on the perceptual judgements of those involved in recruitment (i.e., a certified speech-language pathologist and a graduate student in speech-language pathology).

While multilingual participants were not excluded from the study, English was required to be the primary language of use for both speakers across all communication settings. This requirement ensured that the speakers were able to read fluently and communicate with their partners using sentences that occasionally contained low-frequency vocabulary items (e.g., gingham, accordion, wizard).

It should be noted that this requirement did allow for the possibility of including English as a second language (ESL) speakers in the study. Considering the potential influence of an ESL speaker's prominent first language on scores of communicative success, the investigators acknowledged that monolingual English speakers would have been ideal participants for this study. However, such a strict inclusion criterion was judged to be unrealistic. Limiting recruitment to monolingual English speakers when conducting research within a diverse multicultural and multilingual community would have been a significant challenge, especially considering the clinical population under study consisted of individuals with a relatively rare degenerative disease.

3.1.1.2. Familiar and unfamiliar partners

Four neurologically intact adults were also recruited to serve as communication partners for the dysarthric speakers. For each speaker, one familiar partner and one unfamiliar partner were recruited. As with the dysarthric speakers, communication partners were required to have normal vision, hearing, and cognitive functioning. Similarly, all communication partners were required to speak English as their primary language of use when in their home environment. Relatedly, these partners were required to be able to read and write English fluently, a measure that ensured they would be able to perform the required tasks without any significant problems related to linguistic proficiency.

Familiar partners were defined as those who communicated regularly with the dysarthric speakers. Thus, the speakers' spouses, close friends, or caregivers were all considered ideal candidates. In contrast, unfamiliar partners were defined as those who did not have a close relationship or communicate regularly with someone with dysarthria. It was also necessary to ensure that unfamiliar partners did not have any significant personal or clinical experience with dysarthria or dysarthric speech. This excluded speech-language pathologists and graduate students in speech-language pathology from participation as unfamiliar partners.

Dysarthric participants and their familiar partners were recruited as dyads through a speech-language pathologist working at the clinic serving those with ALS and PLS. The speech-language pathologist approached potential participants who met the participation criteria and distributed information packages that outlined the study's purpose and objectives and included basic descriptions of the tasks involved. Unfamiliar partners were recruited via a recruitment poster located at the University of British Columbia's School of Audiology and Speech Sciences, which was directed towards audiology students in particular.

Once recruitment was complete and the inclusion and exclusion criteria had been satisfied, two triads were formed. Each triad consisted of one dysarthric speaker, his or her familiar partner, and an unfamiliar partner, who was arbitrarily assigned to the triad. Detailed information about these six participants is given in sections 3.1.3 and 3.1.4 below.

Informed consent was obtained from all participants through procedures approved by the Behavioural Research Ethics Board at the University of British Columbia (UBC). The certificate of approval to conduct the present study is available in Appendix A, and consent forms are available in Appendices B, C, and D.

3.1.2. Features of ALS and PLS

PLS is a degenerative motor neuron disease very closely related ALS. There is some debate as to whether PLS and ALS are two distinct motor neuron diseases, or whether they exist as two points on a continuum of the same disease process (Hudson, Kiernan, Munoz, Pringle, Brown, & Ebers, 1993; Tartaglia, Rowe, Findlater, Orange, Grace, & Strong, 2007). In fact, the initial presentation of either disease's symptoms may mimic the other so closely that a differential diagnosis of ALS or PLS is not possible until a few years after onset when the course of disease progression is more apparent (Tartaglia et al., 2007). Despite this difficulty in separating the two disease processes, it is known that PLS implicates degeneration of only the upper motor neurons, typically resulting in a spastic dysarthria. In contrast, ALS implicates degeneration of both central (upper) and peripheral (lower) motor neurons, leading to a mixed spastic-flaccid dysarthria (Hudson et al., 1993; Tartaglia et al., 2007). In addition, ALS progresses more rapidly than PLS (Hudson et al., 1993; Tartaglia et al., 2007).

3.1.3. Triad 1

3.1.3.1. Speaker 1 (S1)

Speaker 1 (S1) was a 75-year-old monolingual English-speaking female presenting with moderate spastic dysarthria secondary to PLS that had been ongoing for 15 years. She had experienced speech symptoms for approximately eight years prior to participation in this study.

In accordance with inclusion and exclusion criteria, S1 did not present with significant language, cognitive, visual, or hearing impairments that would affect her communication. It should be noted that S1 did have a long-standing significant unilateral sensorineural hearing loss in her left ear that had stemmed from an illness 43 years prior to her participation in this study. However, S1 reported that her hearing loss did not affect her everyday communication, and she did not notice it any longer. The investigators for this study agreed that the nature of the tasks to be performed by S1 were not such that one would expect a unilateral hearing loss to interfere with communicative success; all tasks were carried out in a quiet, well-lit room, and both members of each dyad were positioned face-to-face. In other words, since listening conditions during the tasks were optimal, S1's unilateral hearing loss was not judged to be confounding to the present study of communicative success.

3.1.3.2. Familiar partner 1 (F1)

In order to create a familiar dyad (FD), S1 was asked to identify a familiar communication partner who was willing to participate in the study. Ideal familiar partners were persons who communicated regularly with the dysarthric speakers (e.g. spouse, caregiver, etc.) and therefore had significant experience with the speakers' communicative needs, styles, and strategies.

In line with this requirement, S1's familiar partner (F1) was her live-in caregiver, a 53-year-old female from the Philippines who had worked with S1 for approximately two and a half years. She spoke English as a second language (ESL), and spoke Tagalog as her first language. F1 had spoken some English since the age of ten and had lived and worked using English in Canada for 19 years prior to participation in the present research. This, in addition to the fact that English was her primary language of use in her daily life and with S1, suggested that her skills as an English speaker were adequate for participation. F1 also reported normal vision, hearing, and cognition.

3.1.3.3. Unfamiliar partner 1 (U1)

To create a dyad with an unfamiliar listener (UD), an adult who had little or no previous experience with dysarthric speakers was also recruited and assigned to S1. The unfamiliar partner (U1) assigned to S1 was a 25-year-old male graduate student in his final year of studies in the audiology program at the University of British Columbia (UBC). Like F1, U1 reported normal vision, hearing, and cognition.

3.1.4. Triad 2

3.1.4.1. Speaker 2 (S2)

Speaker 2 (S2) was a 56-year-old monolingual English-speaking female with moderate spastic dysarthria secondary to PLS. She had been diagnosed almost three years prior to participation in the present study, yet had had dysarthric symptoms for over four years. Like S1, S2 did not present with any significant language, cognitive, visual, or hearing impairments that affected her communication.

3.1.4.2. Familiar partner 2 (F2)

S2's husband served as her familiar partner for this study. He was 57 years old and had been married to S2 for 36 years. As with the other communication partners, he reported normal vision, hearing, and cognition.

3.1.4.3. Unfamiliar partner 2 (U2)

The unfamiliar partner assigned to S2 was a 26-year-old female audiology student in her final year of studies at UBC. She had no previous personal experience in speaking with people with dysarthria, and she displayed no barriers to communicative success (e.g., impaired cognition, hearing, or vision).

3.2 Materials

3.2.1. Sentence stimuli

Printed sentences from the AIDS (Yorkston & Beukelman, 1981) were used as stimulus materials for both the comprehensibility and intelligibility assessments. Three different sentence sets (A, B, and C) were used for each assessment session to prevent practice effects. Each set consisted of 30 sentences ranging in length from five words to ten words, with five sentences of each length included in each set (225 total words). Sentences were selected using the random numbers table provided in the AIDS examination manual. Sentence length was then randomly ordered for each set.

Although the standard administration protocol of the AIDS involves sentences up to 15 words long, it was reasoned that the processing and production demands of these longer sentences may result in more articulatory breakdown than would be representative of the dysarthric speakers' real life functioning. The significant working memory demands on the listeners for these longer sentences were also judged to be unrepresentative. Furthermore, it is unlikely that people with dysarthria would typically produce sentences this long, based on the observation by Comrie et al. (2001) that dysarthric speakers' conversational contributions were shorter and more limited when compared to their non-dysarthric partners. In light of this evidence, sentences that were 11 to 15 words long were excluded from the sentence stimuli.

Sentences were randomly selected from the AIDS (Yorkston & Beukelman, 1981) in order to ensure that they were unfamiliar and unpredictable to the participants. While this is an important component of intelligibility task standardization (Yorkston & Beukelman, 1983), Klasner and Yorkston (2000) acknowledge that the use of unfamiliar sentence stimuli may negatively impact listener comprehension (and consequently, communicative success), as listeners are tasked with more online linguistic decoding than they would be if listening to a familiar or standardized passage. They argue that the simultaneous tasks of auditory processing and linguistic decoding when listening to unfamiliar material may be more cognitively taxing than listening to familiar material, which relies less heavily on linguistic decoding (e.g., the identification of word boundaries; Klasner & Yorkston, 2001). However, given the practical challenges in controlling a listener's level of experience with familiar material, unfamiliar sentence stimuli were employed in the present study.

Sentences were printed in 18-point Helvetica typeface and printed one per page on 8.5" by 5.5" sheets of paper, then bound into a booklet. In order to prevent the speaker from learning the length of the sentences, the sentences were numbered in such a way that their word length was not obvious (e.g., A1, A2, A3, etc.) rather than the standard numbering system of the AIDS, where the number of the item corresponds to the word length of the sentence (e.g., 5A is five words in length).

3.2.2. Equipment

Digital video recording equipment was used to capture speaker productions for the intelligibility assessments, and for documenting dyadic comprehensibility sessions. A Sony Handycam HDR-XR520V camera recorded digital video in high definition, and was connected to a two-channel BeachTek DXA-6A Dual Phantom power adapter. During filming for comprehensibility sessions, each speaker wore a TOA wireless microphone (model# WM-4310), the signal from each being picked up by two TOA wireless tuners (model# WT-4810). In order to split the audio input from the two speakers into right and left audio channels (one speaker per channel), each speaker's productions were transmitted from the tuners to the DXA-6A power adapter's right and left XLR inputs, and were recorded using the stereo audio input setting. For intelligibility recordings, only one microphone and tuner were used, with recordings made using the mono audio input setting.

After recording, all digital video recordings were saved digitally on computers at the Acquired Language Disorders Laboratory at the University of British Columbia. Intelligibility recordings were also saved onto an Apple MacBook for portability when having familiar partners transcribe the recordings within the speakers' homes. Intelligibility recordings were viewed using VLC Media Player on the Apple MacBook, and comprehensibility recordings were viewed using Hewlett Packard's PMB software for Windows XP. Timing measures for the intelligibility scoring were obtained using Amadeus Pro on an Apple iMac desktop computer.

3.3 Procedures

For each triad, data was collected in five separate sessions. The ordering of these events differed for each speaker. For both speakers, the first session consisted of the speaker being video recorded producing sentences from Sentence Set A for the intelligibility assessment. Next, in the case of S1, the familiar partner (F1) transcribed the intelligibility sentences and then participated in the comprehensibility task with S1. Two weeks later, the unfamiliar partner (U1) also did both these tasks, in the same order. This order was counterbalanced for the second dyad: S2's unfamiliar partner (U2) completed both tasks before the familiar partner (F2) did. In both familiar partner comprehensibility sessions, Sentence Set B was used, while both unfamiliar partners heard Sentence Set C.

With the exception of the unfamiliar partners' transcription sessions (which were conducted in a quiet room at the University of British Columbia), all sessions were conducted within the dysarthric speakers' homes. Home settings were chosen to maximize convenience for the speakers, who faced challenges in the realms of mobility and fatigue, and who both resided a considerable distance from the university. For each session, care was taken to create a suitable environment for the tasks at hand, including face-to-face seating, sufficient lighting, and minimal ambient noise.

A detailed description of the procedures for each assessment follows.

3.3.1. Intelligibility assessment

3.3.1.1. Administration

As stated above, a modified version of the sentence-level portion of the AIDS (Yorkston & Beukelman, 1981) was administered to each dysarthric speaker. Speakers were recorded reading 30 sentences one at a time from the Sentence Set A booklet. Traditionally, intelligibility assessments are based on audio recordings of the speaker's productions. However, in the present study, speaker productions were digitally video recorded. Video recording was chosen for several reasons. First, Hustad et al (2007) found that visual information aided listener processing compared to audio presentation alone. Second, the presence of visual information was also considered to be more representative of real-life functioning, as naturalistic communication is predominantly face-to-face.

The use of video recording for the intelligibility assessment also minimized differences between the intelligibility and comprehensibility assessments, as the comprehensibility task in the present study was also video recorded. Had a more traditional, audio-only approach to intelligibility assessment been pursued, it would have been more difficult to compare comprehensibility performance to intelligibility performance. Any benefit of assessing comprehensibility over intelligibility could be erroneously and solely attributed to the presence of visual information in the comprehensibility assessment. However, comprehensibility is not simply the result of adding visual information to an audio signal. It also includes the communicative strategies and bidirectional feedback employed by both communication partners. This is particularly true with regard to the structured assessment of comprehensibility used in the present study. In light of this, it was reasoned that visual information should be captured in both intelligibility and comprehensibility assessments in order to allow for a discussion of how communicative strategies and feedback may contribute to successful communication.

It is worth acknowledging that these modifications may be viewed as deviating somewhat from the standard clinical protocol for assessing intelligibility, thus potentially limiting its external validity. However, as the AIDS (Yorkston & Beukelman, 1981) is not scored using normative data, a strict adherence to its standard administration procedures was not deemed necessary for the purposes of this study. However, the basic framework of standard AIDS administration was taken into consideration and respected when possible, which helped to minimize deviation from the normative procedures. For example, although the distribution of sentence length was different, the total of number of words spoken in the sentence set was similar (225) to a standard set of stimuli from the AIDS (220).

3.3.1.2. Scoring

For each dysarthric speaker, the unfamiliar and familiar participants each watched and orthographically transcribed the video recordings of the speaker's sentence productions. When transcribing the recordings, the participants were encouraged to guess if they were uncertain, as has been done in other studies (Ferrier, 1991; Hustad et al., 2007; Klasner & Yorkston, 2000; Tjaden & Liss, 1995b). The investigators then scored these transcripts to determine the number of correctly transcribed target words for each listener's transcript. For F1's transcript, which was scored first, scoring was a collaborative effort between the two investigators, ensuring consistency in interpretation of the transcript. In contrast, each investigator independently scored U1's transcript. Later, the two versions were compared and differences were resolved through discussion until a 100% agreement was reached on the total score for each transcript. For F2's and U2's transcripts, 100% agreement was achieved after each investigator's first attempt at scoring.

Once a consensus was reached for each transcript, the agreed-upon number of correctly transcribed words was then divided by the total number of target words (225) to yield an intelligibility percentage score for each listener's transcript. Efficiency measures were also derived from these scores. As Duffy (2005) argues, an important aspect of communicative performance is the efficiency of signal transmission, or the rate of intelligible speech. Speech rate contributes significantly to the perception of speech normalcy in social contexts, and as such is a vital component to the assessment of functional communicative ability (Duffy, 2005). Duffy argues that even highly intelligible speech is perceived as unnatural when produced with a markedly slowed rate, which may limit social success.

A measure of intelligible words per minute (IWPM) was calculated for each speaker as an index of the speakers' speech efficiency. Time measures were obtained by analyzing the audio track of the intelligibility recording using Amadeus Pro. This software allowed the investigators to visualize the audio track's waveform and to take measurements of the time elapsed between the speakers' voice onset and offset for each sentence. Both investigators performed these measurements independently, then resolved discrepancies in timing measures through discussion. Measurements for each sentence were then summed to derive a measure of total elapsed time for the group of 30 sentences, which was rounded to the nearest half-second. IWPM was determined for each speaker by dividing the total number of intelligible words by the total elapsed time measure in seconds, then multiplying by 60.

3.3.2. Structured assessment of comprehensibility

3.3.2.1. Administration

After intelligibility assessments were complete, each dysarthric speaker participated in two dyadic interactions: one with a familiar partner, and one with an unfamiliar partner, resulting in four dyads total. S1 performed the comprehensibility task with her familiar partner (F1) before doing so with her unfamiliar partner (U1), while this partner ordering was reversed for S2. When participating in familiar dyads (FDs), the speakers read sentences from Sentence Set B, while in case of both unfamiliar dyads (UDs), Sentence Set C was used. Each session was recorded on digital video for later analysis.

The procedures for completing the comprehensibility task were taken from Visser (2004), and were explained to each dyad before the task began. Participants with dysarthria read the set of 30 sentences out loud, one at a time, in the order presented on the page. The listener then repeated back the sentence he or she had just heard. The speaker was encouraged to give feedback to the listener in order to clarify any misunderstood elements of the sentence, and both parties were encouraged to use whatever strategies they wanted (aside from the speaker showing the listener the printed sentence) to come to a consensus on the final version of the target sentence. The speaker was able to decide when the listener's version of the sentence was adequately accurate, and both parties mutually decided when to move on to the next sentence. Test procedures are summarized in Appendix E (adapted from Visser, 2004).

After the dyads completed the comprehensibility task, the participants were given an opportunity to discuss their impressions of the process. As Barefoot et al. (1993) noted, it is useful to document the impressions of the interlocutors themselves when considering communicative success, as communication partners may be able to provide valid comprehensibility judgments. Relatedly, one limitation of Visser's (2004) use of secondary data was that information about the experiences of the participants was unavailable. Visser noted that the impressions of the dyad members would have been a valuable source of information in the consideration of comprehensibility for the two participants studied. In regards to the present study, discussions with the participants about their experiences were not formally recorded. Instead, the investigator present recorded relevant comments via field notes.

3.3.2.2. Transcription

After data collection, all four videotaped comprehensibility assessments were transcribed. In contrast with a transcription-less approach, transcription was chosen for two reasons. First, transcription-based analysis was deemed appropriate in order to minimize the methodological differences between the proposed study and that of Visser (2004). Second, it served to reduce disputes between scorers about what each participant had said. It was recognized that when watching video recordings of dyadic interaction, it is not always clear what one or the other dyad member has said; for instance, overlapping, dysfluent, or particularly rapid speech may be difficult to decipher, resulting in different interpretations of the data across scorers. Transcription solves this potential source of dispute by providing scorers with a single interpretation of the data upon which to base their scoring. Although this practice imposes an artificially restricted interpretation of the data on scorers (as Ochs [1979] shrewdly notes, a transcript is simply someone's interpretation of the data, not the data itself), the benefit of standardizing the data in this manner was judged to exceed any limitations of using transcription.

In order to ensure the final version of each transcript was a maximally representative interpretation of the data, both primary investigators initially transcribed each of the four videos separately. After separate transcription was completed, both transcribed versions of each dyadic interaction were compared and differences in the transcripts were discussed and resolved. A final version of the transcript for each dyad was thus determined by consensus.

Appendix F contains a key to the transcription conventions used during this process. It should be noted that the interaction was separated into utterances, which were defined based on intonational breaks in the communication. That is, utterance boundaries (represented by line breaks) were inserted when either dyad member produced a falling (.), semi-rising (~) or rising intonation (?) or produced an utterance with emphatic stress (!). In addition, dashes (-) were used to denote the ends of interrupted utterances, but were also used within utterances to mark such revisions as false starts and self-initiated self-repairs. Aside from these conventions, many elements of traditional discourse analysis were not coded within the system. This is because the task itself was not conversational, and therefore the intricacies of conversation were not those under examination. Following the reasoning of Ochs (1979), who argued that selectivity is to be encouraged in transcription, the information transcribed primarily represented the behaviours of interest as they were relevant to the present research. Extraneous information was largely omitted from the transcripts in order to ensure the product was readable and contained only relevant details. Because of this, the transcription conventions used, while selected somewhat arbitrarily, were intentionally chosen for the purposes of the present investigation.

While verbal information was carefully marked in order to facilitate a thorough analysis of the words spoken by each dyad member, nonverbal information was intentionally omitted from the transcripts. This decision was based on the reasoning that recording nonverbal information inherently involved labeling it, the subjectivity of which was likely to bias the scorers regarding their interpretation of the strategies they observed the dyad members using. Instead, while scorers were instructed that nonverbal communication was relevant in their consideration of all elements of the structured assessment of comprehensibility, they were given the freedom to interpret this communication themselves and to label it accordingly. In other words, nonverbal information was omitted from the transcripts to ensure raters could make their own decisions about what they saw, what it meant, and what to call it. This ensured that the largely exploratory investigation of strategy use (see below) would be as free from constraint as possible, and that all scorers had equal opportunities to contribute their own observations of the often subtle use of nonverbal communication and strategies.

3.3.2.3. Training procedures

The primary investigator for this study met with two scorers who were unfamiliar with the assessment tool prior to the training session. The two-hour training session involved an overview of the present research, a detailed explanation of the structured assessment of comprehensibility, and an opportunity to apply it to sample data in a guided practice scoring session. The sample transcript and video data provided for practice was E. and G.'s T1 interaction from Visser (2004). While scoring the practice transcript, the raters were encouraged to ask the investigator questions if any scoring criteria required clarification. After several sentences had been scored, the raters compared total scores for each sentence, resolving any disagreements whenever possible through discussion. The number of disagreements remaining after these discussions was judged to be minimal by the investigator present and the scorers reported feeling adequately acquainted and comfortable with the scoring procedure to apply it to the new data.

3.3.2.4. Scoring

Three scorers (the principal investigator of the study and the two trained scorers) individually viewed the video recordings while following along with the transcripts, noting any disagreements with the material transcribed if and when they occurred. While scorers were asked to base their scoring on the existing transcript, any disagreements noted were investigated further in order to allow changes to be made to the finalized transcript if they were judged to be necessary. Each scorer then scored each of the four transcripts. Following Visser (2004), six measures were recorded for each sentence: (1) time to do total number of words, (2) number of problem words, (3) number of extra speaker words, (4) number of extra listener words, (5) number of words guessed incorrectly, and (6) ultimate correct (see Appendix G for detailed scoring criteria). During this process, scorers were encouraged to review any part of the recording they wished, as many times as they felt was necessary. To facilitate counting and later reliability checks, the three scorers were asked to mark on the transcript the specific words that were counted for each scoring variable using coloured pens. For example, each extra speaker word was underlined in blue, whereas each extra listener word was underlined in green, and so on.

3.3.2.5. Recording strategy use

While scoring the structured assessment of comprehensibility, each scorer also documented the use of strategies by the speaker and listener in each dyadic interaction. The focus of this study regarding strategy use was to explore the range of different strategies used in familiar vs. unfamiliar dyadic interactions, so scoring centered on the qualitative identification of different strategy types within each dyadic interaction. Scorers were given brief written instructions about how to record strategy use and were encouraged to name as a strategy anything they considered to have an enhancing or repairing effect on communication between the dyad members. They were also asked to note one example of each strategy listed by noting the location (by line number or by a range of line numbers) where it occurred in the transcript. The strategy use recording form is available in Appendix H. It should be noted that scorers were instructed to simply make note of the types of strategies they observed in use rather than tally the frequency of occurrence of particular strategies. As the aim of investigating strategy use in this context was exploratory in nature, an in-depth analysis of the frequency of strategy use in familiar versus unfamiliar conditions was beyond the scope of this study. For the purposes of the present study, an analysis of the range of strategies used by dyad members was sufficient to allow the investigators to attempt to replicate the findings of previous research in this area (King & Gallegos-Santillan, 1999; Villanueva, 1997).

3.4. Analysis

Scoring each dyad yielded three types of data: (1) quantitative data on intelligibility, (2) quantitative data on comprehensibility, and (3) qualitative data on strategy use. The types of analyses performed to test each research question are outlined below.

3.4.1. Reliability

3.4.1.1. Inter-rater agreement

To investigate whether the structured assessment of comprehensibility was reliable, inter-rater agreement was examined for each of the four dyads. This documented to what extent the raters agreed on total scores for each measure of comprehensibility for each dyad. Because each scoring category (see the six measures listed in Appendix G) yielded qualitatively different types of data, three different approaches to assessing agreement were employed. The data were divided into three groups based on the nature of each type. First, because they are continuous in nature, timing measures were analyzed separately from the rest of the measures (which are categorical). The Intraclass Correlation Coefficient (ICC) was chosen (cf. Hula, McNeil, Doyle, Rubinsky, & Fossett, 2003) to assess correlation across all three raters for timing measures. The ICC was judged to be the most appropriate correlation coefficient for capturing comparisons of a continuous measure across more than two raters at a time.

The second group (problem words and ultimate correct) were grouped together because each of these measures required the rater to make categorical judgements about a finite number of words per sentence (e.g., for a seven word sentence, only a maximum of seven words could be considered to be problem words). For this analysis, Cohen’s kappa was used. For each of the two measures, three pairwise comparisons for each dyad were calculated, indicating the agreement between each possible pairing of the three raters. The mean of these three kappa values was then calculated to derive a mean agreement value for each dyad on each of the two measures.

The remaining three measures formed the third group (extra speaker words, extra listener words, and incorrect guesses), and were grouped together because the number of possible words to fall into each of these categories was quite large and was not restricted to a maximum total based on the length of the sentence. Point-to-point agreement was chosen to address these measures. In choosing how narrowly or broadly to examine point-to-point agreement, several options were available. The two most straightforward options were word-by-word comparisons and sentence-by-sentence comparisons. In a consideration of the word-by-word method, it was recognized that a substantial portion of each transcript included words that were highly unlikely to be identified by any rater as falling into the categories of extra speaker words, extra listener words, or incorrect guesses (e.g., questions such as "did I get that right?", etc.). Thus, examining point-to-point agreement on a word-by-word basis would likely overly inflate agreement scores, as a large portion of the agreement could be attributed to raters agreeing that this significant subset of transcribed words were not scorable in any category. Such a narrow examination of agreement would have also been extremely time consuming, and so was discarded as a candidate for how to approach point-to-point agreement.

The second option for point-to-point agreement was to examine it on a sentence-by-sentence basis. However, this too was considered inadequate. Frequent communication breakdowns (and subsequent attempts at breakdown resolution) meant that many sentences involved multiple turns by each dyad member, which in turn meant that the opportunities for disagreement within a single sentence sequence were often numerous. If agreement was defined as 100% agreement on each sentence sequence (and disagreement as any disagreement within each sentence sequence, no matter how minor), this approach would fail to capture the differences between minor and major disagreements within a sentence sequence. Because of this, even very minor disagreements would cause the entire sentence sequence to be categorized as a disagreement, even if, for example, nine out of ten turns within the sequence were agreed-upon. In doing so, sentence-by-sentence comparisons would underestimate agreement. This approach to point-to-point agreement was therefore considered to be too broad, and so it too was discarded.

To find an appropriately sensitive method for investigating point-to-point agreement, a unit larger than the word and smaller than the sentence sequence was necessary for analysis. For practical purposes, each sentence sequence was divided into turn pairs. Each turn pair consisted of one or more utterances produced by the speaker followed by one or more utterances produced by the listener. That is, the entirety of the speaker's verbal turn was paired with the listener's turn that followed it to create a pair. Each new turn pair began when the speaker regained the verbal floor from the listener and produced the next utterance(s).

In the event that the speaker was the last one to speak in a sentence sequence, this final utterance was counted as a special speaker-only turn pair, which was allowable only at the end of sentence sequences. This measure was taken to ensure that any sentence final utterance(s) produced by the speaker was not counted as part of the first turn pair in the subsequent sentence sequence. Speaker-only turn pairs were included in the calculation of point-to-point agreement for the extra speaker words measure only. They were removed from the agreement calculations for the extra listener words and incorrect guesses measures in order to prevent an inappropriate inflation of agreement scores, as these two categories did not apply to utterances produced by the speaker.

Turn pairs were chosen over individual utterances as the unit of analysis because utterances containing extra speaker words were mutually exclusive with utterances containing both extra listener words and incorrect guesses, which allowed for simultaneous analysis of these three measures within the same unit. In essence, combining speaker and listener data into a single unit of analysis simplified the process of recording the data for point-to-point analysis.

Once each dyad's transcript had been divided into turn pairs, agreement was examined for each measure on a point-to-point basis for each turn pair. For each of the three possible pairings of raters (AB, AC, and BC), agreement was calculated by dividing the total number of agreements over the total number of turn pairs (total number of agreements + disagreements). The means of these three percentage scores were then calculated to determine the overall agreement for each measure for each dyad.

3.4.1.2. Intra-rater agreement

Intra-rater agreement was assessed by two of the raters in this study. To ensure this type of reliability without the time consuming process of having each rater re-score all four transcripts, the two raters re-scored a selection of sentences at least two weeks after the initial scoring. Each selection included six randomly selected sentences from each dyad (i.e., 20% of all sentences were re-scored). Measures were taken to ensure different sentences were re-scored for each FD and UD. In other words, the six sentences re-scored for the F1 transcript were different than the six selected for re-scoring from the F2 transcript. Likewise, the re-scored sentences from the U1 transcript differed from those from the U2 transcript. It was also ensured that the two raters each re-scored a different selection of sentences, to ensure the maximum number of sentences was re-scored for each dyad. These comparisons of two scores given by the same rater at different time points examined to what extent each measure of comprehensibility received the same score by the same rater between T1 and T2.

For the categories of extra speaker words, extra listener words, and incorrect guesses, intra-rater agreement was calculated for each rater by dividing the number of agreements by the number of items re-scored (number of agreements + disagreements) with turn pairs again as the unit of analysis. As was the case for inter-rater agreement, Cohen's kappa statistic was calculated to measure intra-rater agreement for the categories of problem words and ultimate correct. Finally, simple correlation coefficients were calculated to measure intra-rater agreement for timing measures. The mean of all these intra-rater agreement values was then calculated for each of the two raters to give an overall measure of intra-rater reliability.

3.4.3. Listener familiarity and comprehensibility

It is important to note that high scores on each of the six measures of the structured assessment of comprehensibility would not necessarily reflect communicative success. For instance, total time would be low in an efficient, successful communication, whereas ultimate correct would be high in a successful exchange. Thus, it is inappropriate to hypothesize that comprehensibility scores would be higher for one type of dyad compared to another, as the individual elements of comprehensibility need to be addressed individually. Comprehensibility was instead operationalized by comparing FDs and UDs on three measures: (1) total time, (2) ratio of extra listener to extra speaker words (a synthesis of two measures from the assessment), and (3) number of problem words (following the model of comprehensibility put forth by Purves, 1999, see Figure 1.3). The ratio of extra listener to extra speaker words was chosen based on the finding in Visser (2004) that this ratio increased from T1 to T2, along with a noticeable shift in communicative responsibility onto the listener. Since G. was a more familiar listener at T2 compared to T1, it was expected that the FDs in the present study would also display greater ratios of extra listener to extra speaker words than the UDs. Therefore, superior comprehensibility, which one could expect from the FDs, could be defined as taking less time, having a higher ratio of extra listener to extra speaker words, and involving fewer problem words. To test the first hypothesis (that comprehensibility scores would be better for familiar dyads than unfamiliar dyads), the dyads were compared on these three measures.

To test the second hypothesis (that intelligibility scores would also be higher when scored by familiar versus unfamiliar listeners), intelligibility scores based on the transcriptions of unfamiliar and familiar listeners were compared to investigate whether the data aligned with findings from previous studies (e.g., Visser, 2004). Two measures of intelligibility were explored, following the general protocol of the AIDS (Yorkston & Beukelman, 1981): the percentage of correctly transcribed target words, and intelligible words per minute (IWPM), the latter of which sheds light on the efficiency of speech produced by the speaker with dysarthria. Descriptive statistics were employed for these comparisons.

To address the third hypothesis (that differences in intelligibility scores would not fully account for differences in comprehensibility scores), intelligibility scores were compared with results of the comprehensibility assessment to investigate whether prominent associations between the two measures could be found. Each of the three chosen measures in the structured assessment of comprehensibility were also examined to determine if any particular measure was strongly associated with intelligibility, and whether or not this relationship was consistent across dyads of similar familiarity levels. This shed light on how each component contributed to increased communicative success. A descriptive approach was taken in the analysis of this data.

3.4.4. Strategy use

To examine whether, as hypothesized, unfamiliar dyad members employed a greater range of strategies than familiar dyads (due to more frequent communication breakdowns), qualitative data from the records of participants' strategy use were examined. As a first step in synthesizing the raw data, the strategies recorded by the scorers were grouped into three categories: strategies that had been identified by all three raters, strategies that had been identified by two out of three raters, and strategies that had been identified by only one rater. This categorization was challenging because of the inherent variability of the raw data, as each rater may have named or interpreted a given strategy in a unique way. Therefore, combining three different interpretations into one synthesized category was not straightforward. Indeed, the highly exploratory nature of this analysis should be emphasized for two reasons. First, the scorers themselves had a considerable amount of freedom in terms of how to interpret the strategies they observed (see Appendix H for the instructions given to scorers on how to record strategy use). Second, the process of summarizing the recorded strategies into a single list also involved numerous subjective judgements on the part of the primary investigator. The aim of this highly subjective synthesis was simply to collapse the raw data across scorers to allow for further analysis.

The summarized list of strategies was then used to tally the total number of strategies employed by each dyad member in each interaction. This was done in order to examine the range of strategies used by members of dyads of different familiarity levels. In doing so, it was possible to compare the results of the present study to the findings of previous studies on this topic (King & Gallegos-Santillan, 1999; Villanueva, 1997).

4. RESULTS

The primary goal of the present study was to evaluate the reliability of the structured assessment of comprehensibility, and in so doing, investigate how two speakers with dysarthria and their communication partners differed with regard to intelligibility, comprehensibility, and strategy use when the partners were either familiar or unfamiliar. Shedding light on these research questions, the following chapter reports the results of these investigations, detailing the inter- and intra-rater reliability findings for the assessment tool, as well as quantitative results of the intelligibility and comprehensibility assessments, and qualitative results from the investigation of strategy use. Because many of these findings should be considered in light of the fact that one participant (F1) spoke English as a second language, a brief description of how this unique case was addressed from a methodological standpoint is presented first.

4.1. The case of F1

As has already been noted, S1's familiar partner was an ESL speaker. This presented some unique challenges for scoring her intelligibility and comprehensibility transcripts. In examining her intelligibility transcriptions, it was observed that certain errors on F1's transcript were unlikely to have been committed by a native speaker of English (e.g., errors in grammatical morphology), suggesting that F1 lacked some knowledge about the predictability and grammaticality of certain constructions. Although F1's intelligibility scores were not purposefully adjusted to account for her language level, some of these issues were addressed in scoring. If F1's written sentence contained an error in grammatical morphology the investigators judged with confidence to be atypical of a native English speaker, the item was not counted as an error. For example, in F1's transcription 'you're not *suppose* to be talking during the performance', the word 'suppose' was credited as correct despite it lacking the obligatory past tense *–ed* morpheme.

In regards to the comprehensibility assessment, similar ‘errors’ in F1’s speech were observed. For instance, during the transcription process it became apparent F1 commonly omitted certain grammatical morphemes (most notably the past tense *–ed* and plural *–s*) from her speech. These grammatically incorrect guesses at the target words were almost invariably accepted by S1 as correct. That is, S1 did not appear to regard them as incorrect guesses at the targets that merited initiating a repair strategy.

While the scoring protocol for the structured assessment of comprehensibility does not specify how to approach such error patterns, it was judged that these omissions may not be errors per se, but dialectal differences due to the fact that F1 was not a native English speaker. As such, it was judged to be inappropriate to penalize F1 for incorrect guesses when in fact her guesses were, at least on a lexical level, correct. Indeed, there appear to be different levels of correctness. For a target word such as ‘apples’, most would agree that the guess ‘apple’ is more correct than say, a guess such as ‘porcupine’. Thus, differentiating between grammatically correct forms and lexically correct forms may be necessary in a consideration of communicative performance for an ESL speaker. Certainly, from a clinical standpoint, this is a valuable distinction that would allow clinicians flexibility when working with diverse populations.

To address this issue during scoring, two sets of comprehensibility scores were calculated for the interaction featuring S1 and F1. First, all raters were instructed to score F1’s transcript similarly to the other three, by adopting a strict (or grammatically correct) criterion for correctness. That is, to be considered correct, the speaker had to repeat the target word exactly as the speaker had presented it, including all obligatory grammatical morphemes. However, the raters were also alerted to the fact that some errors in F1’s transcript could be interpreted as lexically correct. For these cases, the raters were asked to mark on the transcript those guesses where they judged the error to be lexically correct, and thus a clinically acceptable response that was simply due to the influence of F1's ESL status. Once scoring was complete, the errors marked as lexically correct were consolidated across all raters' transcripts, and then analyzed on a case-by-case basis by the primary investigator. Based on this analysis, an additional set of adjusted scores was assigned to F1 based on a less strict (or lexically correct) criterion for correctness.

In deriving the lexically correct scores (F1-L) from F1's grammatically correct scores (F1-G), all comprehensibility measures except total time and extra speaker words were adjusted. Although it was possible that F1's ESL status affected these two measures, there was no way to evaluate this potential impact systematically. It should be noted, too, that for the F1-L score set, the extra listener words that were excluded from scoring were a subset of the excluded incorrect guesses. The reason for this, as was previously mentioned, was that all grammatically incorrect but lexically correct guesses were marked as incorrect for the F1-G scores, and were all considered to be correct for F1-L scoring. In contrast, while these lexically correct guesses were all also considered to be extra listener words (because they were incorrect guesses) in the F1-G score set, only the first presentation of the guess in question was excluded in the F1-L score set. This is because all subsequent repetitions of this lexically correct guess were still classified as extra listener words by definition. Therefore, the difference in incorrect guesses seen between the F1-G and F1-L score sets was more pronounced for incorrect guesses than for extra listener words.

4.2. Intelligibility

The results from the modified version of the AIDS (Yorkston & Beukelman, 1981) are displayed in Table 3. As hypothesized, both speakers' intelligibility percentage scores were higher in the familiar condition (F1: 64.4%; F2: 92.0%) compared to the unfamiliar condition (U1: 62.7%; U2: 60.4%). From a descriptive standpoint, the difference between familiarity conditions for S1 was minimal (1.7%) compared to the difference between scores for S2 (31.6%).

Following the operational definitions of severity outlined by Hustad (2008; see Table 4.1 below), these percentage scores for intelligibility confirm that S1's dysarthria would be classified as moderate when scored by either F1 or U1. For S2, while U2's scores suggest a classification of moderate dysarthria, F2's scores fall within the range of mild dysarthria. Since intelligibility is typically scored by unfamiliar listeners, the descriptive labels associated with the scores from U1 and U2 (i.e., moderate dysarthria for both speakers) are most clinically representative. These labels also confirm that the perceptual judgements of those involved in recruiting these two speakers were accurate.

Table 4.1. Operational definitions for severity ratings (adapted from Hustad, 2008).

Intelligibility range	*Descriptive label*
76%-95%	Mild
46%-75%	Moderate
25%-45%	Severe
5%-24%	Profound

Also noteworthy is the total time each speaker took to produce the same set of 30 sentences. While S1 required 188.5 seconds (3 minutes and 8.5 seconds) to produce all the sentences in Sentence Set A, S2 took substantially longer to complete this task (312 seconds, or 5 minutes and 12 seconds). This time difference indicates that S2's rate of speech was significantly slower than that of S1.

Speech efficiency scores, represented by the number of intelligible words per minute (IWPM), are also displayed for both speakers in Table 4.2. As with intelligibility percentages, this measure produced similar results for both of S1's partners (F1: 46.2 IWPM; U1: 44.9 IWPM). While the intelligibility percentage was extremely high for F2 (92.0%), the IWPM score from this rater was 39.8 IWPM, which was nonetheless slightly lower than the IWPM score for both raters in the case of S1. This lower value reflects the high number of correctly transcribed words by F2 (207 out of 225) in relation to S2's relatively slow rate of speech production. For U2, the number of correctly transcribed words (136 out of 225) more closely approximated the raw scores for S1 given by both F1 and U1. Given the amount of time S2 took to complete the task, however, the IWPM

Table 4.2. Intelligibility scores for both speakers with dysarthria, as scored by familiar and unfamiliar raters.

Speaker	*Rater*	*Raw score*	*% Intelligible*	*Time to complete*	*IWPM*
S1	F1	145/225	64.4%	188.5 seconds	46.2
	U1	141/225	62.7%		44.9
S2	F2	207/225	92.0%	308.5 seconds	40.3
	U2	136/225	60.4%		26.5

score derived for U2 was noticeably lower (26.2 IWPM). In other words, as IWPM measures the efficiency of speech production, it can be inferred from these data that S1's speech efficiency was approximately equal in both familiarity conditions, while S2's speech efficiency was clearly higher in the familiar condition compared to the unfamiliar condition.

4.3. Comprehensibility

4.3.1. Revision to scoring protocol

Throughout the process of transcription, the two primary investigators of this study were faced with several challenges regarding how to represent the video recorded data in a transcript format. These challenges arose due to difficulty, both anticipated and realized, in applying the scoring protocol originally developed by Visser (2004). While evaluating the applicability of Visser's original protocol to multiple dyads was a central aim of the present research, the number of actual transcription and anticipated scoring challenges meant that applying it without modification would have been inappropriate. Failing to update and expand the original protocol would have led to excessively poor inter-rater reliability and erroneously demonstrated that the structured assessment of comprehensibility was an invalid clinical tool; essentially, it would have been an exercise in futility. Instead, it was decided that changes to the scoring criteria were necessary to address these challenges.

4.3.1.1. Process of revision

Beginning with the process of transcription, several revisions were made to Visser's original scoring protocol to account for novel occurrences in the data that had not presented themselves during the protocol's initial development. At the same time, the transcription conventions were refined to ensure the protocol could be applied as clearly as possible. In essence, knowing that certain questions would arise during scoring allowed the investigators to make operational decisions about how to address them.

Unfortunately, not all scoring challenges could be anticipated. After the initial updates were made to the protocol, the primary investigator undertook the task of scoring each of the four transcripts using the updated criteria. During this process, further difficulties in applying the scoring protocol to the present data were noted. After discussing these issues together, the two investigators decided what further changes to the protocol were necessary and a second revision to the protocol was drafted.

The revised protocol was then given to a pilot scorer who was unfamiliar with the assessment tool, along with one of the four transcripts. She was asked to apply the tool to the transcript and note any questions, difficulties, or inconsistencies that presented themselves during the scoring process. These observations were then shared with the primary investigator and, based on this feedback, minor changes and clarifications were made to create a finalized version of the test protocol with which to train scorers. This final updated version of the scoring criteria was what the three raters used to score each dyad's comprehensibility.

4.3.1.2. Changes made to the scoring protocol

It is important to note that the changes made to the scoring protocol were not intended to account for every single scoring challenge encountered. For example, no guidelines regarding how to deal with periods of laughter were included in the revised criteria. Instead, when possible, general rules were added to account for a broader category of potential situations. Only issues that presented themselves multiple times were addressed. This was done in an effort to discourage the addition of a multitude of situation-specific rules in any future revisions as well. That is, instead of creating one rule for every scenario with which each dyad presented, the added rules were intended to be broad enough to address all the dyads in the current study and any future dyads to which the protocol may be applied. This also minimized the density of the document the scorers had to use, reducing the demands on their working memories.

The most significant updates to the scoring protocol were (1) an elaboration of the criteria for identifying the timing boundaries (especially final agreement), along with the specification that all other measures must be scored based on timing decisions, and (2) the addition of a rule that repeated words or part words within the context of a self-initiated self-repair or false start were not to be counted in any category. This latter rule served to avoid penalizing dysarthric speakers on the extra speakers words measure for the frequent and numerous part word, full word, and phrase-level repetitions they produced due to reduced motor speech control. Since these repetitions differed qualitatively from full and/or purposeful repetitions of the target words due to communicative breakdown and subsequent attempts at clarification, treating them as the same type of repetition would have been unfair. Labeling these behaviours as self-initiated self-repairs and false starts allowed them to be treated as if the speaker had only produced one iteration of the target sentence. It was reasoned that the impact of these motorically based delays would be borne out in the timing measures instead.

Several other rules were also updated or elaborated in order to capture repeated occurrences in the present study's data that had not emerged during Visser's (2004) initial development process (e.g., how to score part words depending on context, how to handle words added to the target sentence, etc.). Compared to the major changes regarding timing and repairs, however, this latter group of changes had a relatively minor effect on how transcripts were scored.

4.3.2. Inter-rater agreement

The results of the inter-rater agreement calculations can be seen in Table 4.3. As is evident in the first column, the ICC measures for total time were extremely high for each dyad (mean $\hat{\rho}$ = .9985), representing near perfect agreement for time. In the second two columns, problem words and ultimate correct, mean values for Cohen's kappa were .9043 and .9713, respectively. According to the descriptive labels for kappa values given by Landis and Koch (1977), these numbers represent almost perfect agreement between raters. For the final three measures (extra speaker words, extra listener words, and words guessed incorrectly), the mean percentage agreement values between raters were .9687, .9258, and .9607, respectively. As was discussed in chapter three (Methods), these values were derived from each pairwise point-to-point comparison by dividing the total number of agreements between the two raters by the total number of turn pairs in the transcript (i.e., the total number of agreements + disagreements). The mean of these three percentages (one per pairwise comparison) was then calculated. While not specifically corrected for chance, these final three values nonetheless represent extremely high inter-rater reliability. Indeed, across all four dyads, inter-rater agreement was extremely high (often near perfect) for all six measures from the structured assessment of comprehensibility.

Table 4.3. Inter-rater agreement values (ICC, Cohen's kappa, and means from pairwise comparisons) for the six comprehensibility measures for each dyad.

Measure	*Total time*	*Problem words*	*Ultimate Correct*	*Extra speaker words*	*Extra listener words*	*Words guessed incorrectly*
Agreement value	ICC ($\hat{\rho}$)	Cohen's kappa (k)		Means of % agreement values from pairwise comparisons		
F1	.9980	.9342	.9676	.9781	.9524	.9832
U1	.9986	.8254	1.0000	.9658	.9143	.9500
F2	.9981	.9130	1.0000	.9692	.8678	.9253
U2	.9993	.9444	.9175	.9615	.9688	.9844
Mean	**.9985**	**.9043**	**.9713**	**.9687**	**.9258**	**.9607**

4.3.3. Intra-rater agreement

Table 4.4 displays the results of the intra-rater reliability calculations for raters A and B. As was discussed in chapter three (Methods), each rater re-scored a unique set of six sentences per dyad, resulting in 20% of the all sentences being re-scored. Overall agreement for both raters was near perfect (range: 0.90 to 1.00), as was agreement for each individual measure of comprehensibility. These high levels of intra-rater reliability suggest that the structured assessment of comprehensibility resulted in nearly identical scores when administered at two different points in time. In other words, within individual raters, scoring maintained an extremely high level of consistency over time.

4.3.4. Scoring challenges and sources of disagreement

Once scoring was completed and the raw data from each of the three scorers had been reviewed, it became evident that some scoring challenges persisted. Common themes emerged across scorers regarding certain scoring pitfalls, which may have been due to some combination of

Table 4.4. Intra-rater agreement values for raters A and B.

	Total time		*Problem words*		*Ultimate correct*		*Extra speaker words*		*Extra listener words*		*Words guessed incorrectly*	
	Correlation coefficient		Cohen's kappa (k)				% Agreement for within-rater turn pair comparisons					
	A	B	A	B	A	B	A	B	A	B	A	B
F1	0.9996	0.9541	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000	1.0000	1.0000
U1	0.9986	0.9982	1.0000	0.8993	1.0000	1.0000	1.0000	1.0000	1.0000	0.8235	1.0000	1.0000
F2	0.9913	0.9954	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.8750	1.0000	0.9375
U2	0.9996	0.9996	1.0000	1.0000	0.9073	1.0000	0.9836	1.0000	0.9831	0.9048	1.0000	0.9524
Mean	0.9972	0.9868	1.0000	0.9748	0.9768	1.0000	0.9959	1.0000	0.9958	0.9008	1.0000	0.9725

unique events within the data and/or a lack of specificity in the scoring protocol. The most common source of difficulty (and thus, disagreement) was identifying timing markers for each sentence. This was especially true for identifying the end of the timed sequence, or where the precise moment of agreement (or abandonment) occurred. This variability in how scorers identified where to stop timing necessarily led to some disagreement about which final guesses or comments to include or exclude from scoring for extra listener words and incorrect guesses. Indeed, timing decisions affected all five of the other comprehensibility measures, as each included a rule that words may only be counted if they were included within the timing for that sentence. Relatedly, determining which parts of the interaction to exclude was a source of difficulty and disagreement, as whether or not communication was directed towards the examiner or the listener (or both) was not always clear. Another minimal source of disagreement was the inexact manner in which timing was recorded. Raters simply observed the time on the video recording at the point in the interaction where they decided a timing boundary should exist. This method led to routine disagreements of one to two seconds between scorers.

Another prominent trouble source encountered by all raters was the identification and handling of self-initiated self-repairs and false starts. As has already been noted, the addition of the rules surrounding these occurrences was a key revision to the scoring protocol. However, the distinction between a self-initiated self-repair or false start and a repetition or revision that was explicitly called for by the other partner was inconsistent across raters. Whether to classify a repetition as a self-initiated self-repair or not affected agreement for the categories of problem words and extra speaker words most prominently. However, because the application of the rules regarding self-initiated self-repairs and false starts for dysarthric speakers was extended to the speech of their nonimpaired partners, extra listener words and incorrect guesses were also affected by these decisions, albeit to a lesser degree.

Additional sources of disagreement were less universal across categories (e.g., how to deal with part words given in the context of an interruption or overlapping speech). However, agreement for all comprehensibility measures suffered to some extent due to simple errors on the part of the raters. These errors included rule violations (e.g., counting a placeholder as an incorrect guess) and a variety of obvious mistakes (e.g., failing to count incorrect guesses as such). In considering these sources of disagreement, a distinction should be drawn between the notion of system error (a problem with the scoring paradigm) and that of human error (mistakes made by scorers).

4.3.5. Summarized comprehensibility scores

Because inter-rater agreement was not perfect, it was necessary to synthesize scores across all three raters prior to reporting the results of the comprehensibility assessment. In other words, a consensus version of the scores for each dyad was created. In general, if two or three raters assigned the same score to the same sentence, this score was used. However, an exception to this rule was made for cases where raters B and C agreed but rater A (the primary investigator) did not, because of rater A's status as an expert rater. In each of these cases, the source of disagreement was investigated carefully to ensure that it was not due to a frank violation of a scoring rule or a misinterpretation of the assessment protocol. If a rule violation was discovered during this process, the primary investigator consulted with a second expert rater to resolve the disagreement and choose the most appropriately representative score.

If all three raters' scores were different for a given sentence, the differences were investigated and a representative score was determined on an item-by-item basis. The method of achieving this differed somewhat for the individual measures of comprehensibility. For timing measures, the timing markers noted in the margins of the transcripts were compared to derive a consensus about when timing started and stopped. Similarly, word-by-word comparisons were used to solve any scoring discrepancies for problem words and ultimate correct, and turn pair comparisons were used for extra speaker words, extra listener words, and incorrect guesses. If no consensus could be reached at the item level and all three raters assigned different scores to a given item, the score assigned by rater A (the primary investigator) was chosen as representative for practical purposes. Synthesizing scores across raters this way yielded a single version of the total scores for each sentence, which made possible a discussion of the implications of these scores in terms of how comprehensibility may have been affected by listener familiarity.

Scores for each measure of comprehensibility, which were synthesized across raters, are displayed in Table 4.5. The adjusted scores for F1 reflecting the lexical level criterion for correctness (F1-L) are displayed below her grammatically correct scores (F1-G) for ease of comparison.

As previously discussed, the four dyads were compared on three measures to test the

Table 4.5. Comprehensibility scores for all dyads (synthesized across raters). Measures that were not adjusted for the F1-L score set are shown in brackets.

		Total time	*Problem words*	*Extra speaker words*	*Extra listener words*	*Words guessed incorrectly*	*Ultimate correct*	*Extra-list to extra-spkr words ratio*
S1	**F1-G**	714.5	66	117	179	48	202	1.53
	F1-L	(714.5)	57	(117)	165	25	215	1.41
	U1	1025	71	204	375	106	219	1.84
S2	**F2**	563	25	79	79	22	224	1.00
	U2	1191.5	70	346	306	98	213	0.88

hypothesis that comprehensibility scores would be better for familiar dyads than unfamiliar dyads: (1) total time, (2) ratio of extra listener to extra speaker words, and (3) number of problem words. As was expected, the amount of time taken to complete the entire 30-sentence task was substantially longer for unfamiliar dyads (U1: 1025 seconds; U2: 1191.5 seconds) than for familiar dyads (F1: 714.5; F2: 563). Also aligning with expectations, the extra listener words to extra speaker words ratios for the two S2 dyads were larger in the familiar condition (F2=1.00) than in the unfamiliar condition (U2=0.88). However, this pattern was reversed in the case of S1, where the ratio was smaller for the familiar condition (F1-G=1.53, F1-L=1.41; U1=1.84). In other words, the ratios of extra listener words to extra speaker words across dyads were not consistently larger for unfamiliar dyads. Finally, for both S1 and S2, the number of problem words was higher in the unfamiliar condition (U1=70; U2=71) than the familiar condition (F1-G=66, F1-L=57; F2=25), although the difference was only noteworthy for S2.

It was further predicted that differences in intelligibility scores would not fully account for differences in comprehensibility scores. Indeed, some support for this notion can be seen by examining Tables 3 and 7. While only a small difference in intelligibility scores was observed between F1 (64.4%; 46.2 IWPM) and U1 (62.7%; 44.9 IWPM), the differences in comprehensibility scores were substantial between these two familiarity conditions. In the case of S2, F2's intelligibility scores (92.0%; 40.3 IWPM) were substantially higher than U2's (60.4%; 26.5 IWPM), a difference that was mirrored in the comprehensibility scores. Indeed, all S2's comprehensibility scores were significantly better for F2 when compared with U2.

Upon analysis of the comprehensibility data, it was observed that for each dyad, a trend existed for scores to improve over the course of the sentence set. This was supported anecdotally by S2 and U2, who commented after they completed the comprehensibility task that they noticed the task got easier over time. To investigate this issue empirically, each sentence set was divided into approximate halves, and the total scores for each half were compared. For each sentence set, the first half contained 111 words (sentences #1 to #13 from Sentence Set B; sentences #1 to #15 from Sentence Set C), and the second half contained 114 words (sentences #14 to #30 from Sentence Set

Table 4.6. Comparison of first and second halves of comprehensibility scores for each dyad. Light grey = improved scores; dark grey = scores that did not improve.

		Total time	*Problem words*	*Extra speaker words*	*Extra listener words*	*Words guessed incorrectly*	*Ultimate correct*
F1	1st half	404.5	35	69	102	28	99
	2nd half	310	31	48	77	20	103
U1	1st half	551	50	98	221	49	107
	2nd half	474	21	106	154	57	112
F2	1st half	308	16	51	58	17	110
	2nd half	255	9	28	21	5	114
U2	1st half	799	39	257	198	72	101
	2nd half	392.5	31	89	108	26	112

B; sentences #16 to #30 from Sentence Set C). The results of these comparisons are shown in Table 4.6. Scores that improve from the first to the second half of the sentence set are shown in light grey while those that do not are shown in dark grey. Note that for the overwhelming trend was for scores to improve between the first and second halves of the sentence set, for all dyads. There were only two exceptions to this trend, and both were for U1. For the S1U1 dyad, extra speaker words increased from the first to the second half of the sentence set, as did the number of words guessed incorrectly by U1. However, these findings can be accounted for by the presence of a particularly challenging sentence for S1 and U1 that occurred in the last half of the transcript. This sentence took a significant amount of time to complete due to considerable breakdown and numerous unsuccessful attempts at repair. When this sentence's data was excluded from the comparison of the two halves of the data set (1st half: 111 words, 2nd half: 106 words), a decrease in both extra speaker words (1st half: 98; 2nd half: 68) and incorrect guesses (1st half: 49; 2nd half: 40) over time was evident, indicating that overall, the S1U1 dyad also experienced improved comprehensibility over time.

4.4. Strategy use

As can be seen in Table 4.7, the total number of strategies identified by one or more raters for each triad was consistently higher in both unfamiliar conditions than in the familiar conditions. Both the dysarthric speakers and their partners used more strategies when the partner was unfamiliar than when the partner was familiar, which supports findings from previous research on this subject (King & Gallegos-Santillan, 1999; Villanueva, 1997, Visser, 2004).

The number of strategies identified generally increased as the number of identifying raters decreased. Because of this, the majority of the total strategies identified were observed by only a

Table 4.7. Strategies observed in use by dysarthric speakers and their communication partners.

	Speaker Strategies				*Partner Strategies*			
	3 Raters	2 Raters	1 Rater	**Total**	3 Raters	2 Raters	1 Rater	**Total**
F1	2	6	15	**23**	3	5	11	**19**
U1	4	6	20	**30**	4	7	22	**33**
F2	2	4	10	**16**	1	2	10	**13**
U2	6	8	19	**33**	3	3	10	**16**

single rater. For practical purposes, instead of presenting them in the text, they are available in Appendix I. Strategies identified by two or all three raters differed from dyad to dyad, although some overlap between dyads existed. Tables 4.8 and 4.9 list these strategies, along with which dyad members were observed using them. The number of check marks in each cell indicates whether the strategies in question were observed by either two or three raters. In order to facilitate a comparison of strategies across dyads, a single check mark in brackets is also displayed for the strategies that also were identified by a single rater for a different dyad member.

As can be seen in Table 4.8, certain speaker strategies were in fact identified across all four dyads. For instance, both speakers, regardless of which partner they were interacting with, used the strategies of repeating a specific constituent (e.g., a word or phrase) and nodding to indicate 'yes' (or general agreement), with both these strategies being identified by two or more raters in each case. Additional strategies common to both speakers (in all dyads) were breaking up a sentence into individual phrases or words, shaking the head for 'no' or general disagreement, self-initiated self-repairs, and clarifying a trouble word by using emphasis. In addition, oral spelling was also common to all dyads except S1F1.

Table 4.8. Communication and repair strategies used by speakers with dysarthria (observed by at least two raters).

Strategy	*S1F1*	*S1U1*	*S2F2*	*S2U2*
• Repetition of a specific constituent (word or phrase)	✓✓✓	✓✓	✓✓	✓✓
• Breaks up sentence into word/phrase level "chunks"	✓✓✓	(✓)	(✓)	✓✓✓
• Contrasts incorrect and correct guesses	✓✓	(✓)		
• Shakes head for 'no'	✓✓	(✓)	✓✓	✓✓✓
• Nods for 'yes' (or general agreement)	✓✓	✓✓✓	✓✓	✓✓
• Rephrases the message	✓✓	(✓)		
• Requests clarification of what listener guessed/said	✓✓	✓✓✓		
• Self-correction/self-initiated self-repair	✓✓	✓✓	✓✓	(✓)
• Request for partner to speak more loudly		✓✓✓		
• Orally spells word for partner		✓✓✓	✓✓✓	✓✓✓
• Splitting contractions (e.g. "weren't" to "were not"), later clarifying target was contracted version		✓✓		
• Explains the personal strategies found to be helpful		✓✓		
• Clarifies or corrects partner using emphasis on trouble words	(✓)	✓✓	(✓)	✓✓
• Warns partner that sentence is awkward or atypical (verbally or nonverbally)		✓✓		
• Gesture/pantomime			✓✓✓	✓✓✓
• Finger spelling using hand shapes				✓✓✓
• Traces letter shapes in the air				✓✓✓
• Breaks target word into syllables or part words	(✓)			✓✓
• Giving an example of a more general term				✓✓
• Giving a synonym for a target word				✓✓
• Gives partner feedback (e.g. "good!" or "you're close!")	(✓)			✓✓
• Holds up fingers to indicate number words (e.g. 3 fingers for "three")				✓✓

When comparing speaker strategies in familiar versus unfamiliar dyads, no visible trends emerged regarding commonly identified strategies across speakers. Instead, strategies appeared to be more speaker-specific. Strategies such as contrasting correct guesses with incorrect ones, rephrasing messages, and requesting clarification of the listener's guess were used by S1 with both F1 and U1. In contrast, S2 tended to use gesture and pantomime with both F2 and U2. Within the set of strategies used by each speaker, however, there was evidence that certain strategies were uniquely employed when partners were unfamiliar. For example, when communicating with U2 in particular, S2 used finger spelling techniques, gave examples and synonyms for words, broke target words into syllables, and gave her partner specific feedback. Similarly, S1 used strategies while interacting with U1 that included splitting contractions (e.g., *weren't* to *were not*), giving warnings

Table 4.9. Communication and repair strategies used by the partners of speakers with dysarthria (observed by at least two raters).

Strategy	*S1F1*	*S1U1*	*S2F2*	*S2U2*
• Verbally requests repetition	✓✓✓	✓✓✓		✓✓
• Asks speaker to speak more slowly	✓✓✓			
• Identifies which word is source of trouble (e.g. "the second word is?")	✓✓✓	✓✓		✓✓✓
• Uses body language to request repetition (e.g. leaning in with quizzical facial expression)	✓✓	(✓)	(✓)	
• Holds up hand to stop speaker (in order to request repetition)	✓✓			
• Lack of guess/silence to prompt a repetition	✓✓	(✓)	(✓)	
• Verbally states doesn't know/has no guess (e.g. "I can't…")	✓✓			
• Repeats previously guessed (correct) words for confirmation	✓✓	✓✓✓	(✓)	(✓)
• Requests oral spelling		✓✓✓		✓✓✓
• Requests semantic information about the target (e.g. category membership)		✓✓✓		
• Uses placeholder phrase (e.g. "something" or "blank")	(✓)	✓✓		
• Uses gestures to help describe guess		✓✓		
• Contrasts two plausible guesses at target (e.g. "is it X or Y?")		✓✓		
• Specific request for feedback (e.g. "is that right?")		✓✓		✓✓
• Guesses	(✓)	✓✓	(✓)	✓✓
• Rephrases the message or guess		✓✓		
• Pauses mid-guess to indicate need for repetition	(✓)	(✓)	✓✓✓	(✓)
• Repeats as much of the sentence as was understood	(✓)		✓✓	✓✓✓
• Mouthing along with the speaker's presentation		(✓)	✓✓	

about the awkwardness and atypicality of sentences, and explaining her personal communication strategies. For both speakers, none of these strategies were observed in use with the familiar partners.

Strategies were also identified as common to all four communication partners (see Table 4.9). All four partners used a guessing strategy when attempting to repeat target words, and repeated previously guessed (correct) words when guessing new ones. They also used the strategy of pausing mid-guess to indicate they required another repetition from the dysarthric speaker, although this strategy was only identified by all three raters for F2. Strategies identified for three out of four dyads were specifying which word was problematic (e.g., “what is the second word?”), using body language (e.g., leaning in) to indicate a need for repetition, using silence to indicate a repetition was needed, and repeating as many words in the sentence as possible. The only evidence that unfamiliar partners used different strategies than familiar partners was that both unfamiliar partners requested oral spelling and specifically requested feedback (e.g., “did I get that right?”) while neither familiar partner used either of these strategies.

5. DISCUSSION

The purpose of the present study was to address the following three questions. First, it aimed to examine the reliability of the structured assessment of comprehensibility (originally developed by Visser, 2004) in order to ascertain its feasibility for the clinical assessment of communicative success for dysarthric speakers and their partners. Second, it aimed to determine whether the structured assessment of comprehensibility would differentiate between familiar and unfamiliar dyads. It was predicted based on previous research that both intelligibility and comprehensibility scores would be superior under conditions where partners were familiar. It was also anticipated that there would not be a one-to-one relationship between intelligibility and comprehensibility scores. Third, this study aimed to explore how strategy use might differ based on the level of familiarity between dyad members, with the prediction that unfamiliar dyads would use a greater range of strategies than familiar dyads. The following chapter provides a discussion of the results from each of these investigations. Findings are discussed as they are relevant to the specific research questions and hypotheses laid out above. The chapter concludes with a discussion of the theoretical and clinical implications of these findings, along with a review of the limitations concerning both the structured assessment of comprehensibility itself and the methodological approach taken to evaluate its feasibility as a clinical assessment tool.

5.1. Reliability of the structured assessment of comprehensibility

The primary purpose of the present study was to examine whether the structured assessment of comprehensibility, originally developed by Visser (2004), is a reliable clinical tool for evaluating the communicative success of speakers with dysarthria and their communication partners. To this end, four dyadic interactions were video recorded and transcribed, then scored by three raters. A comparison of scores across the three raters revealed that inter-rater agreement was extremely high for all six measures of comprehensibility. In addition, two raters re-scored a subset of sentences at least two weeks after the initial scoring in order to determine the intra-rater reliability of the structured assessment of comprehensibility. Intra-rater agreement for both raters was nearly perfect for all measures, indicating that scores assigned by a given rater were stable over time. These findings provide further evidence (in addition to that of Visser, 2004) that this particular method of assessing comprehensibility holds promise in its ability to reliably differentiate the communicative characteristics of multiple dyads.

No matter how meticulously any scoring protocol is reviewed during the development process, and no matter how comprehensive the product of that development is, there will always be scoring challenges and sources of disagreement faced by those who implement it. In this case, the two most prominent additions to Visser's (2004) original scoring protocol were also the most frequent sources of disagreement. The challenges associated with these two situations (the identification of timing boundaries associated with the agreement process, and how to score self-initiated self-repairs and false starts) appeared generally to be due to difficulty interpreting the subtleties of communication, especially nonverbal communication.

The main issue contributing to disagreement for timing measures was identifying the end of each timed sentence sequence. This was often challenging, as the precise moment of agreement (or, in a few cases, abandonment) was often not particularly obvious. The slightest change in eye gaze, or a nearly imperceptible nod of the head were sometimes the only indication that the speaker had accepted the listener's guess as correct. Even when agreement was signaled verbally, repetitions of the target sentence (which were to be counted as part of the agreement process) sometimes seamlessly transitioned into editorial comments about the sentence itself (which were to be excluded). A common difference in scores often resulted from one scorer counting the listener's final reiteration of the entire target sentence within the timed portion of the sequence, while another scorer perceived this repetition as superfluous, since agreement had been reached in the previous utterances. Indeed, these final repetitions often appeared to be a way for the listener to confirm the entirety of the sentence to him- or herself. The potential for confusion is obvious in this case, given the requirement that timing should be stopped when agreement has been reached *by both dyad members*. Was this self-confirmation valuable for the listener in transitioning into a state of agreement? This question is not easy to answer, making any attempt at clarifying the wording of the scoring criteria difficult as well.

The other scoring issue that proved challenging was the identification and treatment of self-initiated self-repairs. Like timing boundaries, self-initiated self-repairs were often rather nebulous. Again, the subtleties of dyad members' nonverbal communication contributed to disagreement over whether or not repetition of a constituent was self-selected or other-selected. In other words, there was disagreement surrounding whether the speaker decided herself to repeat part or all of the target sentence based on having just committed a speech error or experienced a motoric breakdown, or whether some indication was given by the listener that repetition was necessary. Being that different scorers interpreted these events differently, this meant that decisions about whether target words were problematic (i.e., counted as problem words), and whether the repetitions that followed were countable as extra speaker words were also left open to interpretation.

Since both of these issues center on very subtle phenomena, there may not be concrete ways to address them in terms of changing the scoring protocol to systematically account for these occurrences. Instead, it may be more prudent to acknowledge that, while certain aspects of comprehensibility can be measured systematically, others are far less tangible. From a scorer's perspective, it is impossible to know the thoughts or intents of a dyad member unless they are visibly expressed. What goes unexpressed (or ambiguously expressed) must thus be absorbed into a certain expected amount of inter-rater disagreement. In fact, perfect agreement is an inappropriate goal for an assessment of comprehensibility based on semi-naturalistic interaction.

Despite these two repeated sources of disagreement, both intra- and inter-rater agreement were extremely high. It should be acknowledged, however, that this assessment tool is still in its initial stages of development. Future applications of the structured assessment of comprehensibility to more dyads will verify whether the most recent version of the assessment is indeed as reliable as the current data suggests. The two dyads in the present study are, after all, technically the first dyads to which the revised protocol has been applied. However, the reliability remains promising. First, in considering all sources of disagreement, it is helpful to draw a distinction between the notion of system error (a problem with the scoring paradigm) and that of human error (mistakes made by scorers). Therefore, it should be kept in mind that only a portion of the disagreement between raters in this study was due to ambiguity in the scoring system itself. Were the pure human errors to be removed from the reliability calculations, agreement would have been even higher than what was reported. Furthermore, given Visser's (2004) finding that reliability was also strong for the single dyad she studied, and given the fact that the bulk of the scoring criteria remained unchanged between Visser's original version and the updated version used in the present study, these reliability data suggest that the structured assessment of comprehensibility holds promise.

5.2. How intelligibility and comprehensibility are affected by partner familiarity

The results from the structured assessment of comprehensibility indicated that the vast majority of the six measures of comprehensibility were better in the familiar condition compared to the unfamiliar condition. That is, the total amount of time, number of problem words, number of extra speaker words, number of extra listener words, and number of words guessed incorrectly were all lower for familiar dyads, reflecting superior comprehensibility. Results of the ultimate correct measure reflected that, in accordance with expectations, the familiar partner in the second dyad (F2) repeated more words correctly on final presentation than did the unfamiliar partner (U2); however this pattern was reversed in the first dyad, (with the unfamiliar partner, U1, repeating more total words correctly). The ratios of extra listener words to extra speaker words also aligned with expectations for only one of the two dyads, as the ratio was higher for F2 compared to U2, but lower for F1 compared to U1. Nonetheless, the overall trend for superior comprehensibility scores in familiar dyads suggests that the assessment tool has the ability to differentiate between the qualitatively different communicative characteristics of familiar and unfamiliar dyads. Furthermore, a post-hoc comparison of the first and second halves of each sentence set revealed that overall, comprehensibility scores improved across the duration of the assessment, especially for unfamiliar dyads.

As expected, intelligibility scores for both speakers also supported findings from previous research that familiar listeners more effectively process the speech signals of speakers with dysarthria (D'Innocenzo et al., 2006; Flipsen, 1995; Liss et al., 2002; Tjaden & Liss, 1995a; Tjaden & Liss, 1995b; Villanueva, 1997, Visser, 2004). For each speaker, percentage scores indicated that more words were correctly identified by the familiar listener compared to the unfamiliar listener. Furthermore, the number of intelligible words per minute was higher for familiar partners than unfamiliar partners, indicating that speech was more efficiently transmitted in the familiar conditions. It should be noted, however, that difference between familiarity conditions for both of the intelligibility measures was more marked in the case of S2 than with S1.

Broadly speaking, for the four dyads who participated in the present study, both comprehensibility and intelligibility were better for familiar dyads compared to unfamiliar dyads.

Comprehensibility scores in particular reflected substantial differences in communicative success for both speakers between the two familiarity conditions and their unfamiliar counterparts. However, one measure of comprehensibility that was hypothesized to be higher in familiar dyads (ratio of extra listener to extra speaker words) displayed contrasting patterns for each speaker. Inconsistent with Visser's (2004) finding that the ratio of extra listener to extra speaker words increased markedly from T1 to T2, (equated here with increased familiarity), only one triad in the present study followed the predicted pattern (F2>U2), while the other did not (F1<U1). In other words, Visser's finding that the communicative burden on the familiar partner was higher than when the partner was unfamiliar was not totally supported by the results of this study, since this pattern was only observed for one of two dyads.

There are a few possible explanations for these findings. First, it is possible that a similar study with a larger number of participants may have found an overall trend of support for Visser's (2004) findings. In other words, the small sample size in the present study may have simply led to the equal weighting of two sets of data, where one of which may have otherwise been considered an outlier in a larger study. Some factors other than familiarity (e.g., individual differences in personality, communicative style, cultural background, etc.) may have contributed to the atypical findings from the two S1 dyads. It is also possible that Visser's (2004) findings were the anomalous ones, and that the ratio of extra listener to extra speaker words is not actually a particularly relevant measure of comprehensibility. Differences in scores on this measure between familiar and unfamiliar dyads could be dyad specific, which would explain why they did not correlate as strongly with partner familiarity as was predicted.

When examining the large majority of comprehensibility measures that did differentiate between familiar and unfamiliar dyads, the advantage of familiarity was the most apparent in the comprehensibility scores for the S2F2 dyad. The data from F2 may have differed from F1 for a few reasons. First, F1 was a speaker of ESL, while F2 was a native English speaker. The differing language abilities of these two participants meant that direct comparison of their data was challenging. While some effort was made to adjust F1's scores to be more directly comparable to those of a native English speaker (as demonstrated in the F1-L, or lexically correct score set), it was impossible to account for every communicative nuance that may have resulted from her non-native language proficiency. Indeed, even these adjusted scores, when compared to those of F2, showed a far more modest advantage of familiarity for F1.

Another related possibility is that F1 communicated differently than F2 for sociocultural reasons. Among numerous possible contributors, expectations for communicative accuracy and rule following may have been influenced by F1's culture of origin. Furthermore, F1 was employed by S1 as her live-in caregiver. This position of authority held by S1 may have contributed to unique patterns in her communication with F1 that were not present in the spousal relationship between S2 and F2. Because of the small number of participants in this preliminary study, personality differences may also have resulted in different communicative styles and patterns between the two familiar dyads. For example, the number of attempts taken by either dyad member to resolve a communicative breakdown could have been influenced by cultural views regarding politeness, the nature of the social relationship between the person in question and his or her partner, perceived expectations to perform, personal characteristics such as perfectionism, and so on. While these factors may be seen as having the potential to confound the data (or at least significantly complicate the interpretation of findings), they may be beneficial to consider, as dealing with multilingual and multicultural clients and families is a reality for many practicing clinicians.

A third important difference between the two familiar dyads was the length of time the dyad members had been communicating together on a regular basis. While S2 and F2 had been married for 36 years, F1 had been working for S1 for only two and a half years and had known S1 primarily in a professional capacity. Both the nature and length of these relationships are cause for speculation that the marked advantage F2 had in communicating successfully with S2 was due to his being more personally familiar with her.

One interesting trend noted in all four dyads was that comprehensibility scores improved for all measures between the first and second halves of each transcript. This was especially true for the unfamiliar partners, who exhibited particularly noticeable improvements over time. This effect may have been partially due to the fact that sentence sets were not controlled for the distribution of length or complexity, making it possible that the sentence set used by both unfamiliar listeners may have had a more challenging first half. However, another important possibility is that communication partners experienced some degree of learning during the relatively short comprehensibility task, and were better able to decipher the speaker's communicative efforts during the second half compared to the first. If so, mechanisms of habituation to the degraded speech signal itself, such as speaker normalization (Lindblom, 1990) or perceptual learning (Buchholz, 2009), may have played a role in such adjustment, although familiarization with the speakers' use of signal-independent information (e.g., strategies, visual cues, etc.) would have been just as likely to influence comprehensibility scores. These are interesting possibilities, since the unfamiliar partners displayed greater improvements over the course of the task than did the familiar listeners. Given this finding, it would appear that unfamiliar partners experienced a greater change in familiarity over the course of the task itself. This may not be surprising, since the impact of a twenty minute session on the familiar partners (who had known the speakers for years) would intuitively have less impact than the same session might for someone who previously did not know the speaker at all. In other words, the learning curve for unfamiliar partners appears to be steeper than for familiar partners.

Another important finding was that for S1 and her partners, a very minimal difference in intelligibility scores between the two familiarity conditions coincided with a relatively large difference in comprehensibility scores. In other words, while the intelligibility results for S1 suggested that familiar and unfamiliar listeners may understand approximately equal amounts of her speech, comprehensibility scores suggested that her communication with unfamiliar listeners would involve substantially more time, effort, and breakdown than would communication with more familiar listeners. Assuming that the present study's investigation of comprehensibility captured communicative success in a truly functional way (i.e., tapping into the Participation level of the ICF; WHO, 2001), these scores support the work of previous authors suggesting that dysarthria severity (as it is classically defined based on intelligibility scores alone) and the experience of disability for afflicted speakers are not in perfect correlation (Ball et al., 2004; Dykstra et al., 2007; Hartelius et al., 2008; Yorkston et al., 1994). Essentially, this demonstrates that intelligibility measures fail to capture the range of communicative skill across individuals, and that comprehensibility scores may be a more sensitive metric for understanding how successfully persons with dysarthria communicate with others.

5.3. Strategy use

The results of the strategy use analysis showed that both the speakers with dysarthria and their communication partners used more strategies in the unfamiliar conditions than in the familiar conditions. This supported our hypothesis, as it was in accordance with previous literature on this topic (King & Gallegos-Santillan, 1999; Villanueva, 1997). Several strategies were used by both speakers in all four dyads (e.g., repetition, using emphasis, etc.), and one (oral spelling) was used in three of the four dyads. Speaker strategies appeared to hinge on a combination of speaker preferences and the familiarity of the communication partner. In other words, certain strategies were unique to one or the other speaker and were used with either both her familiar and unfamiliar partners (e.g., S1 use of rephrasing messages, S2's use of gesture and pantomime), or with her unfamiliar partner alone (e.g., S1's use of explaining her strategies, S2's use of finger spelling). Strategies common to all four communication partners were also observed (e.g., guessing), as were those identified for three out of four dyads (e.g., using body movements to signal the need for a repetition). Two strategies were common to both unfamiliar partners (requesting oral spelling and requesting feedback), but were not observed for either familiar partner, indicating that unfamiliar and familiar communication partners may use different types of strategies.

Certainly, the fact that more strategies were used by unfamiliar dyad members was at least in part due to the larger amount of communicative breakdown that occurred in these dyads (as evidenced by comprehensibility scores, see Table 4.5). The mechanism behind why unfamiliar dyads experience more breakdown may be as follows. The fact that an unfamiliar partner lacks experience with both processing the dysarthric speaker's degraded speech stream and interpreting the relevant signal-independent information efficiently tends to lead to more frequent communicative breakdowns. These breakdowns, in turn, lead to more attempts at repair from both dyad members. As these attempts fail (again because of a lack of familiarity) and it becomes obvious that a given strategy is not successful at repairing the breakdown in question, the dyad members move on to different strategies, repeating this process until a repair is achieved. This means that an unfamiliar dyad will use a more diverse array of strategies than will a familiar dyad, the members of which presumably know which strategies are the most successful based on their previous experience communicating together. Familiar dyads, therefore, operate in a less 'hit-andmiss' manner with regard to strategy use.

The investigation of strategy use was by far the most exploratory part of the present study. Prior to collecting data, the only hypothesis regarding strategy use was the one outlined above, that unfamiliar dyad members would use more strategies than familiar dyad members. While the data did confirm findings from previous studies (King & Gallegos-Santillan, 1999; Villanueva, 1997), a post-hoc qualitative analysis was also deemed necessary to explore the types of strategies that were identified. What resulted was the observation that certain strategies by both the speakers with dysarthria and their communication partners were identified in all four dyadic interactions. This may have been because certain strategies are more universal than others, at least for the comprehensibility task in question. For example, the nature of the comprehensibility task lends itself most obviously towards repetition, so it is not surprising that repetition was used by all participants across all dyads. For listeners in particular, guessing was a strategy all four partners used, which also was an inherent component of the task itself. Certain strategies are also not specific to the task itself, but commonly used in daily communication by most people (e.g., nodding. in agreement), so their appearance within the context of the comprehensibility task were not surprising or systematic.

In terms of strategies that differed based on familiarity, there were no strategies common to both speakers with dysarthria that were used exclusively in either familiarity condition. Instead, speaker strategies appeared to be largely unique to each speaker. However, there appeared to be some interaction between these speaker preferences and familiarity, as several of each speaker's strategies were observed in use only with unfamiliar partners. Two patterns relating to familiarity were also seen when examining the strategies used by unfamiliar partners. Both unfamiliar partners used the strategies of requesting oral spelling and requesting feedback, while neither familiar partner did so. While it is tempting to assume these two strategies might be representative of how unfamiliar communication partners on a larger scale might approach the comprehensibility task differently than familiar partners, caution should be taken in interpreting this finding. Both U1 and U2 were final year graduate students in a small audiology program at UBC. Because of this, these two participants had far more in common than did F1 and F2 (who did not know each other and worked in very different fields). Furthermore, it is possible that any training in aural rehabilitation the two unfamiliar partners had received throughout the course of their studies might have influenced the way they interacted with the dysarthric speakers. When questioned about this after data collection, U1 reported that he never tapped consciously into that skill set during the comprehensibility task. It is still possible, however, that U1 and U2's clinical experiences in audiology led them to subconsciously employ certain communication strategies they had learned to be useful in repairing the communicative breakdowns commonly experienced when working with people with hearing impairment. As was observed previously, many of the strategies targeted to the deaf and hard of hearing and their communication partners (Caissie & Rockwell, 1993) can be applied to the population of people with dysarthria (see section 1.3.2.2. for a discussion).

It is tempting to assume that a strategy observed by only a single rater must be less valid or less prominent than those identified by more than one rater. However, it is important to note that it is not only the strategies identified by two or three raters that merit attention. Obvious strategies may have been missed by two out of three raters for various reasons. First, the raw data collected from all three raters suggested that the level of attention given to strategy use differed significantly across raters. That is, some raters recorded very few strategies where others recorded many. Second, the definition of what a strategy was may have differed from rater to rater, as the scorers were purposefully given very minimal guidelines to frame their observations of strategy use. Due to the highly subjective nature of how raters recorded and labeled strategy use in this study, it would be inappropriate to give significantly more weight to the strategies identified by two or three of the raters. It is likely that many of the strategies identified by only one rater were perfectly valid enhancements of communication, but were simply less obvious. In other words, it is possible that they were there, but they were just harder to detect. Or, if they weren't difficult to detect, the raters who did not record these behaviours as strategies may simply have not viewed them as enhancing communication (e.g., laughter, comments regarding the awkwardness of wording, etc.). It is possible that pointing out such strategies to the raters who missed them may have resulted in the raters acknowledging their (now more obvious) impact on communication. Finally, given the small number of raters in this study, all identified strategies should be treated with equal merit.

It is also of interest how the strategies employed by the participants in this study may have affected the comprehensibility scores discussed earlier. To this end, two observations were made regarding strategy use by the speakers with dysarthria. First, while filming the S2U2 dyad, the investigator noted that S2 relied heavily on enthusiastic nodding to guide U2 to the correct target words. Her use of nodding, however, was not particularly discriminating. She tended to nod after each guess where U2 had correctly guessed any part of the sentence, regardless of whether the rest of her guess was correct or not. This appeared to be a source of confusion for U2, who received positive feedback for the incorrect parts of her guesses along with the correct parts. Because of this, U2 would often persist with incorrect guesses until S2 eventually corrected them. This dynamic undoubtedly contributed to the substantial amount of time this dyad took to complete the comprehensibility task (theirs was the longest of the four interactions), as well as other high values for measures such as extra speaker words, extra listener words, and incorrect guesses.

Second, it was observed that S2's strategy use was far more obvious than that of S1, especially in the unfamiliar conditions. While S2's strategies with U2 often involved the incorporation of various iconic gestures, S1 relied more on commenting (giving warnings, explaining, rephrasing, etc.). This difference may have been due, in part at least, to the mobility challenges S1 experienced, which appeared to affect her hands. In contrast, S2 was highly mobile and exhibited dexterous use of both hands. Regardless of the underlying reasons behind this difference, it is possible that the visibly salient iconic gestures S2 employed frequently with U2 (and less frequently with F2) in some way contributed to the large difference in comprehensibility scores between familiarity conditions. If so, S1's far less obvious use of strategies may have contributed to the more modest comprehensibility differences between the two conditions.

While the types of strategies observed with these four dyads may be representative of what would be seen from these same participants in other communicative contexts (e.g., conversational discourse), it is also possible that the contrived nature of the task affected what information was produced by each partner. Perhaps certain gestural or emotional-affective information was less likely to emerge than with naturalistic conversation. As Hustad et al. (2007) notes, scripted sentence stimuli may have the effect of limiting the production of some gestures or facial cues that signal emphasis and emotion. While the structured assessment of comprehensibility is currently based on the use of scripted sentences, this potential for results that may not fully reflect naturalistic communication must be kept in mind, especially when interpreting nonverbal cues and strategies used by dyad members.

5.4. Theoretical and clinical implications

The present study found that the structured assessment of comprehensibility not only successfully differentiated familiar dyads from unfamiliar dyads, but did so with extremely high levels of reliability within and across raters. In other words, the assessment tool reliably indicated that familiar dyads' comprehensibility scores were superior to those of unfamiliar dyads. This supported findings from numerous other studies that have indicated that communication is more successful for speakers with dysarthria when their communication partners are familiar (Ball et al., 2004; D'Innocenzo et al., 2006; Ferrier, 1991; Flipsen, 1995; Hartelius et al., 2008; King & Gallegos-Santillan, 1999; Liss et al., 2002; Tjaden & Liss, 1995a; Tjaden & Liss, 1995b; Villanueva, 1997, Visser, 2004). These findings suggest that the structured assessment of comprehensibility holds promise as a valid and reliable clinical tool for evaluating communicative success for dysarthric speakers and their partners.

As has already been noted, defining severity for speakers with dysarthria through measuring intelligibility alone falls short of addressing the true nature of disability. In terms of the ICF (discussed in Chapter 1), while intelligibility measures only address the Activity level, comprehensibility measures, through a more holistic analysis of communicative success, address the Participation level (WHO, 2001). The shift in focus that is moving away from intelligibility and towards comprehensibility has important clinical implications. First, the structured assessment of comprehensibility dealt with in the present study may provide a more representative view of dysarthric speakers' ability to make themselves understood within naturalistic communication contexts. In this way, the assessment tool may surpass traditional measures of intelligibility in addressing the participation level of the ICF, which is a step toward understanding the speaker's level of functioning in society (WHO, 2001; Yorkston et al., 1996). Furthermore, it has been acknowledged by others that severity of dysarthria does not reliably predict perceived disability or participation level, as even mild forms of dysarthria can have a significant impact on the affected person's life participation (Ball et al., 2004; Dykstra et al., 2007; Hartelius et al., 2008; Yorkston et al., 1994). Since intelligibility measures are the current clinical standard in measuring severity, it follows that they would inappropriately predict the speaker's disability level. This was confirmed by the findings of the present study, which indicated that, for one speaker, a very minimal difference in intelligibility scores between the two familiarity conditions occurred despite a substantial difference in comprehensibility scores. This finding provides strength to the argument that a comprehensibility measure such as the structured assessment of comprehensibility may be a more appropriate index of severity, and thus more holistically evaluate the impact of dysarthria as a disability.

There is a need in the field of dysarthria rehabilitation for changing the way we evaluate disability to include factors across all of the ICF's levels of functioning (WHO, 2001). In particular, systematically evaluating the impact of an individual's life context on his or her experience of disability is an important goal. While the present study did not specifically address the life context of either of the dysarthric speaker or their partners, it did move towards assessing functional communicative participation through acknowledging the contribution of more than just the speech signal they were able to produce. Compared to the highly limited nature of intelligibility assessment, the structured assessment of comprehensibility captured the cumulative effects of how each speaker and her partners utilized numerous types of signal-independent information to enhance communication. Furthermore, the task itself, which involved dynamic face-to-face interaction, allowed the dyad members a certain amount of creative freedom in repairing breakdown and resolving misunderstanding. The fact that dyad members were permitted to use any strategies they wished was itself a significant step towards an understanding of naturalistic communication for speakers with dysarthria and their partners. Indeed, the comprehensibility model promotes the consideration of all facets of the communicative context as key contributors to successful communication.

The present findings about familiarity align with what previous authors have found: that there are qualitative differences in the ways people with dysarthria communicate depending on who their partners are (Ball et al., 2004; D'Innocenzo et al., 2006; Ferrier, 1991; Flipsen, 1995; Hartelius et al., 2008; King & Gallegos-Santillan, 1999; Liss et al., 2002; Tjaden & Liss, 1995a; Tjaden & Liss, 1995b; Villanueva, 1997, Visser, 2004). That is, due to differing levels of familiarity, communication between a speaker with dysarthria and a store clerk may be far more difficult and frustrating compared to communication between the same speaker and her spouse or close friend. It may be prudent to apply this information to the assessment process. Specifically, the structured assessment of comprehensibility could be administered to a speaker with dysarthria and a variety of his or her communication partners. Even a single comparison of two different partners (preferably with maximally different levels of familiarity) may allow clinicians to evaluate a particular speaker's severity using a range of scores rather than a single metric.

This is an important observation in the sense that it is reasonable to expect communication to improve if familiarity is increased between the partners. Procedures that familiarize a listener or partner with the dysarthric speaker may greatly benefit comprehensibility, and thus communicative success. Approaches to intervention that focus on improving comprehensibility through repeated dyadic interaction and partner instruction may be important for improving communication on a functional level for dysarthric speakers. Treatments involving instruction in how to maximize the use of signal-independent information and repair strategies may also prove beneficial at improving communicative success (Ball et al., 2004; Ferrier, 1991; Hustad et al., 2007; Tjaden & Liss, 1995a; Villanueva, 1997; Visser, 2004; Yorkston & Beukelman, 1983) as well as social participation and quality of life (Visser, 2004). However, several factors (e.g. dysarthria severity, level of cognitive functioning, time since onset, etiology, etc.) may interact to determine whether a speaker is an appropriate candidate for treatment targeting the use of signal-independent information (Yorkston et al., 1996). In the case that a dysarthric speaker is deemed appropriate for this type of intervention, the incorporation of familiar conversation partners into therapy sessions would be important to improve the speaker's communication skills within natural, functional contexts, and to emphasize the role of dyadic interaction in communicative success (Yorkston et al., 1996). This inclusion may also encourage generalization to the environments in which the dysarthric speaker habitually communicates (Visser, 2004). As Yorkston et al. (1996) note, "bringing about functional changes is at the heart of our clinical practice" (p. 60).

Determining how dyad members utilize signal-independent information and strategies, and how these behaviours contribute to both successful and unsuccessful communication is an important first step in planning treatment targeting improved comprehensibility (Dykstra et al., 2007; King & Gallegos-Santillan, 1999). In other words, the information gleaned from the structured assessment of comprehensibility certainly can be seen as having important implications for guiding treatment. In fact, even the task itself may be useful in enlightening dyad members about which particular aspects of their communication are the most problematic, or even which are the most helpful. After completing the comprehensibility task, one of the participants in the present study (F2) noted that he felt the process had helped him identify which specific sound patterns S2 had the most trouble producing. This comment is encouraging, and suggests that engaging in this activity provided newfound insight for at least one participant that may have helped him navigate future interactions with his partner. Indeed, after 36 years of marriage to S2, the fact that F2 learned something new from this experience speaks to its potential for useful clinical application.

5.5. Limitations

5.5.1. Limitations of the structured assessment of comprehensibility

While some of the challenges that arose during scoring have already been discussed, additional limitations of the structured assessment of comprehensibility merit attention. First, as assessment tools go, the structured assessment of comprehensibility is relatively time consuming. There are several steps which cumulatively require significant time and attention on the part of the clinician. This is especially true in the case of both transcription and scoring. Of course, the time required to complete each of these portions of the assessment varies widely based on how successful the dyad was at communicating efficiently. However, as a rough estimate, scoring each transcript for the purposes of the present study took approximately four hours for an expert rater, and approximately five hours for the recently trained raters. This is certainly a significant amount of time to spend on an assessment, from a clinical perspective.

One of the two most time consuming parts of completing the assessment was the process of transcribing the four dyadic interactions. Since transcription was so cumbersome, it raises the question of whether written transcripts were the most optimal media on which to base data analysis. Because it was sometimes difficult to understand what had been said in the video recordings, it was recognized that different scorers may have different interpretations of what they observed. Therefore, transcription was chosen for this study to allow scorers to base their decisions on a standardized interpretation of the data. In other words, that transcription was chosen to represent the video data was simply a pragmatic choice on the part of the investigators. Certainly, it was not the only option. In fact, Armstrong, Brady, Mackenzie and Norrie (2007) have recently shown that a transcription-less approach could be a promising method for the analysis of aphasic discourse. Given the significant time commitment transcription requires, the development of a transcription-less version of the structured assessment of comprehensibility may be a goal worth pursuing.

Something else that undoubtedly contributed to the length of the assessment process was the density of the document detailing the scoring criteria (see Appendix G). Despite attempts to avoid the addition of highly specific rules to the test protocol, the revised scoring criteria ended up being relatively long and dense, and as such may have been taxing on the working memory of its users. In fact, the pilot scorer in the present study commented that she found this to be the case. If further revisions were to be made, perhaps a focus of the revision process should be to consolidate similar rules into more succinct statements, reducing the demands on the reader.

In addition to the above issues regarding the time required to complete it, the structured assessment of comprehensibility has limitations in terms of the population to which it can be applied. As was discussed in Chapter 2 (Methods), it is probably most useful when applied to speakers with dysarthria that is moderate or severe. Ceiling effects limit its application to populations with milder forms of dysarthria, while scoring challenges are too numerous with more profound dysarthric speakers. Profound speakers, particularly those with diagnoses of ALS or PLS, may experience a more notable deterioration of intra-articulator coordination and timing (Yunusova, Weismer, Westbury, and Lindstrom, 2008), which may lead to conflicting acoustic and visual-articulatory information for their communication partners, which in turn may lead to more significant breakdown. For these reasons, the structured assessment of comprehensibility is only ideal for speakers with dysarthria that fall within a specific window of severity. Furthermore, as this tool has only been applied to speakers with dysarthria resulting from parkinsonism (Visser, 2004) or PLS, it remains to be seen whether it may be applied (with equally promising results) to speakers with dysarthria resulting from different etiologies.

The application of this assessment tool is further restricted because it is (at least currently) inappropriate for dysarthric speakers with comorbid cognitive and/or language impairments. Certainly, the nature of neurological injury is such that these comorbidities often co-occur with dysarthria. Essentially, excluding speakers with these issues significantly limits the population of neurologically impaired people who could potentially benefit from the structured assessment of comprehensibility. Yet another limitation is that questions remain regarding how (or whether) to adjust procedures for dyad members with varying levels of English proficiency, although a multilingual participant was accommodated in the present study.

Finally, some areas for improvement remain if the structured assessment of comprehensibility is to be further refined for clinical use. First, as has been noted, the task itself is somewhat artificial, owing to the scripted nature of the sentences. In the future, perhaps more personally relevant sentences could be collected from the dysarthric speakers in order to maximize the functional nature of the task and collect the most representative data possible from dyad members. Second, refining certain wording in the scoring criteria may be necessary. For example, the way scoring was conducted according to the revised test protocol would mean that a speaker who used an alphabet board would receive a lower (better) score for extra speaker words than a speaker spelling orally. This is inconsistent with the fact that a speaker using his or her voice should, logically speaking, be considered to be communicating more efficiently, and thus receive a better comprehensibility score.

5.5.2. Limitations in design

Some limitations in terms of the methodological decisions made throughout the course of this study should be addressed. First, certain issues existed regarding the characteristics of the participants selected. Perhaps most obviously, F1's status as an ESL speaker presented a challenge. Because F1 had spoken English while living and working in Canada for 19 years, it was assumed that she was proficient enough in English to participate in the study. While this was not untrue, the influence of her first language was nonetheless evident in the resulting data. If F1's English proficiency had been formally screened, the resulting data may have been useful in framing the interpretation of her performance on both intelligibility and comprehensibility tasks.

Also potentially affecting the scores for the comparison of the two S1 dyads was S1's level of hearing loss. Prior to participation, S1 reported her hearing loss to be unilateral and having no noticeable effect on her communication. However, on the day of data collection with U1, S1 reported that her other (normally hearing) ear was feeling 'plugged'. She also reported that her voice was 'worse' that day than it had been in the previous session with F1. Despite these issues, data collection continued. During the process of data analysis, however, it was apparent that S1 had not heard some of U1's guesses, which visibly contributed to subsequent breakdown. This may have inflated the difference in comprehensibility scores between the F1 and U1 dyads to some degree. However, it can also be argued that this may have made scores in the two conditions more similar. While S1's transient hearing loss usually led to breakdown between the two dyad members, it also sometimes resulted in S1 accepting a wrong guess from U1 as correct. Since S1 habitually accepted some of F1's grammatically incorrect guesses as correct, it is possible that breakdown was avoided for both dyads in a few cases, albeit for very different reasons. Regardless of the direction of influence here, it is clear that this hearing loss may have affected comprehensibility scores and thus limited direct comparisons between the F1 and U1 dyads.

The present study did not systematically control for the degree of familiarity of each familiar partner recruited. This led to the formation of two familiar dyads with very different histories. Indeed, it is quite likely that S2's husband of 36 years was far more personally familiar with her than S1's live-in caregiver of two and a half years was with her. In contrast, both unfamiliar partners were very similar. In fact, they were classmates and in the same graduate program who had had many shared experiences in the preceding two years. Naturally, they also had been exposed to the same education (e.g., aural rehabilitation training) that dealt with how best to communicate with hearing impaired individuals. While neither unfamiliar partner reported consciously tapping into these skills, it is quite possible that these experiences had made them very similar communicators on a basic or subconscious level.

Other issues that may have limited the findings of this study concerned the selection of sentence stimuli used for each dyad. While sentences were controlled for length, the sentences were not counterbalanced for length or complexity across the sentence sets. That is, the way sentences were randomized did not prevent clusters of longer, more complex sentences from occurring at certain points in a given set. What resulted was one set (Sentence Set C) that contained a disproportionate number of long (9- and 10-word) sentences in the very beginning of the set. Since this set was used by both unfamiliar dyads, it may have led to the greater difficulty at the beginning of the comprehensibility task that both unfamiliar dyads appeared to experience (especially U2). More broadly speaking, it may have contributed to the poorer comprehensibility scores the unfamiliar dyads displayed overall. In fact, S1 commented after her session with U1 that it was a far more frustrating experience than the interaction with F1 had been, and that she felt it was due in part to the sentences being harder.

It should also be noted that transcribing the dysarthric speaker's sentences during the intelligibility session may have served as a sort of brief specific familiarization training procedure for both unfamiliar partners. While the effects of this were likely minimal since each unfamiliar partner performed the transcription task prior to meeting and interacting with the speaker with dysarthria for the comprehensibility task, it is worth acknowledging that their comprehensibility scores may not be wholly representative of truly unfamiliar people, for example, strangers in the community.

The exploratory findings on strategy use are also limited, as the frequency or effectiveness of their use was not examined in this study. Information about which strategies were most frequently employed as well as which were successful versus unsuccessful would have considerable clinical utility in terms of goal selection and approaches to intervention. The identification of useful communication strategies for speakers with dysarthria is absolutely essential in supporting and encouraging successful communication between clients and those with whom they interact.

Finally, the extent to which the findings of this study are generalizable is unclear. As has been noted, only a very restricted portion of the greater population of dysarthric speakers has been examined. The two speakers in this study both had moderate dysarthria secondary to PLS, were cognitively intact, were native English speakers, and had no comorbidities that otherwise impaired their communication (e.g., hearing loss, aphasia, etc.). Furthermore, the small sample size of two speakers meant that the full range of communicative characteristics, habits, and strategies that are representative of the greater population of people with dysarthria was unlikely to be captured in the data.

Despite the various limitations of this research, it is clear that the structured assessment of comprehensibility has the potential to be a valuable assessment tool in the context of an ever-increasing clinical focus on functional assessment and intervention. With further evaluation and refinement, it holds promise for evaluating comprehensibility in a way that captures the nature of disability as it is experienced by speakers with dysarthria.

6. CONCLUSION

Two speakers with dysarthria secondary to PLS were video recorded interacting with both familiar and unfamiliar communication partners in the context of a structured comprehensibility task. The central finding of this study was that the structured assessment of comprehensibility, originally developed by Visser (2004), could be reliably applied to these four dyads. In the process, evidence was found to support previous research that showed familiar dyads to have superior comprehensibility compared to unfamiliar dyads.

This research builds on the findings of many previous studies, which have cumulatively demonstrated the existence of a gradual shift in how communicative success is defined in the field of dysarthria research and clinical practice. Instead of measuring a dysarthric speaker's intelligibility, which considers only the impact of the disrupted acoustic signal, comprehensibility measures account for a wide range of signal-independent information, including communication enhancing strategies; furthermore, they necessarily demand more specification of the communication context, including communication partners. Increasingly, attention is being paid to how people with dysarthria communicate in more functional contexts. In light of these new directions in dysarthria research, the structured assessment of comprehensibility evaluates communicative success dyadically. As an assessment tool, the structured assessment of comprehensibility aims to play a key role in evaluating the participation of speakers with dysarthria in society.

7. REFERENCES

Armstrong, L., Brady, M., Mackenzie, C., & Norrie, J. (2007). Transcription-less analysis of aphasic discourse: A clinician's dream or a possibility? *Aphasiology, 21*(3-4), 355-374.

Ball, L. J., Beukelman, D. R., & Pattee, G. L. (2004). Communication effectiveness of individuals with amyotrophic lateral sclerosis. *Journal of Communication Disorders, 37,* 197-215.

Barefoot, S. M., Bochner, J. H., Johnson, B. A., & vom Eigen, B. A. (1993). Rating deaf speakers' comprehensibility: An exploratory investigation. *American Journal of Speech-Language Pathology, 2*(3), 31-35.

Berry, W. R., & Saunders, S. B. (1983). Environmental education: The universal management approach for adults with dysarthria. In: W. R. Berry (Ed.), *Clinical Dysarthria.* San Diego: College-Hill Press.

Buchholz, L. K. (2009). Perceptual Learning of Dysarthric Speech: Effects of Familiarization and Feedback. Unpublished master's thesis. University of British Columbia, Vancouver, British Columbia, Canada.

Caissie, R., & Rockwell, E. (1993). A videotape analysis procedure for assessing conversational fluency in hearing-impaired adults. *Ear & Hearing, 14*(3), 202-209.

Comrie, P., Mackenzie, C., & McCall, J. (2001). The influence of acquired dysarthria on conversational turn-taking. *Clinical Linguistics & Phonetics, 15*(5), 383-398.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12,* 472-496.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor Speech Disorders.* Philadelphia: W. B. Saunders Company.

D'Innocenzo, J., Tjaden, K., & Greenman, G. (2006). Intelligibility in dysarthria: Effects of listener familiarity and speaking condition. *Clinical Linguistics & Phonetics, 20*(9), 659-675.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (2nd ed.). St. Louis, MO: Elsevier Mosby.

Duffy, S. A., & Pisoni, D. B. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language and Speech, 35*(4), 351-389.

Dykstra, A. D., Hakel, M. E., & Adams, S. G. (2007). Application of the ICF in reduced speech intelligibility in dysarthria. *Seminars in Speech and Language, 28*(4), 301-311.

Ferrier, L. J. (1991). Clinical study of a dysarthric adult using a Touch Talker with Words Strategy. *Augmentative and Alternative Communication, 7,* 266-274.

Flipsen, P., Jr. (1995). Speaker-listener familiarity: Parents as judges of delayed speech intelligibility. *Journal of Communication Disorders, 28*, 3-19.

Garcia, J. M., & Cannito, M. P. (1996). Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. *Journal of Speech & Hearing Reseaerch, 39*(4), 750-760.

Hartelius, L., Elmberg, M., Holm, R., Lövberg, A-S., & Nikolaidis, S. (2008). Living with dysarthria: Evaluation of a self-report questionnaire. *Folia Phoniatrica et Logopaedica, 60,* 11-19.

Hillel, A. D., Miller, R. M., Yorkston, K., McDonald, E., Norris, F. H., & Konikow, N. (1989). Amyotrophic lateral sclerosis severity scale. *Neuroepidemiology, 8,* 142-150.

Hudson, A. J., Kiernan, J. A., Munoz, D. G., Pringle, E., Brown, W. F., & Ebers, G. C. (1993). Clinicopathological features of primary lateral sclerosis are different from amyotrophic lateral sclerosis. *Brain Research Bulletin, 30,* 359-364.

Hula, W. D., McNeil, M. R., Doyle, P. J., Rubinsky, H. J., & Fossett, T. R. D. (2003). The inter-rater reliability of the story retell procedure. *Aphasiology, 17*(5), 523-528.

Hustad, K. C. (2008). The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech, Language, and Hearing Research, 51,* 562-573.

Hustad, K., & Cahill, M. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology, 12*(2), 198-208.

Hustad, K. C., Dardis, C. M., & McCourt, K. A. (2007). Effects of visual information on intelligibility of open and closed class words in predictable sentences produced by speakers with dysarthria. *Clinical Linguistics & Phonetics, 21*(5), 353-367.

Hustad, K. C., Jones, T., & Dailey, S. (2003). Implementing speech supplementation strategies: Effects on intelligibility and speech rate of individuals with chronic severe dysarthria. *Journal of Speech, Language, and Hearing Research, 46,* 462-474.

Kent, R. D., & Read, C. (1992). *The Acoustic Analysis of Speech* (pp. 87-104). San Diego, CA: Singular Publishing Group.

King, J. M., & Gallegos-Santillan, P. (1999). Strategy use by speakers with dysarthria and both familiar and unfamiliar conversational partners. *Journal of Medical Speech-Language Pathology, 7*(2), 113-116.

Klasner, E. R., & Yorkston, K. M. (2000). Dysarthria in ALS: A method for obtaining the everyday listener's perception. *Journal of Medical Speech-Language Pathology, 8*(4), 261-264.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Lindblom, B. (1990). On the communication process: Speaker-listener interaction and the development of speech. *Augmentative and Alternative Communication, 6*(4), 220-230.

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *Journal of* *the Acoustical Society of America, 112*, 3022-3030.

Ochs, E. (1979). Transcription as theory. In Ochs, E. & Schieffelin, B. (Eds.), *Developmental Pragmatics* (pp. 43-72). New York: Academic Press.

Purves, B. (1999). From theory to practice. Workshop: Fredericton, NB.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361-382.

Tartaglia, M. C., Rowe, A., Findlater, K., Orange, J. B., Grace, G., & Strong, M. J. (2007). Differentiation between primary lateral sclerosis and amyotrophic lateral sclerosis: Examination of symptoms and signs at disease onset and during follow-up. *Archives of Neurology, 64,* 232-236.

Tjaden, K. K., & Liss, J. M. (1995a). The influence of familiarity judgments of treated speech. *American Journal of Speech-Language Pathology, 4,* 39-48.

Tjaden, K. K., & Liss, J. M. (1995b). The role of listener familiarity in the perception of dysarthric speech. *Clinical Linguistics & Phonetics, 9*(2), 139-154.

Villanueva, P. J. (1997). Comprehensibility in communicative exchanges: Dysarthric speakers and both familiar and unfamiliar conversation partners. Unpublished master's thesis. University of Nevada, Reno.

Visser, T. (2004). Comprehensibility: A potential measure for improvement in communication. Unpublished master's thesis, University of British Columbia, Vancouver, British Columbia, Canada.

Weismer, G., Jeng, J, Laures, J. S., Kent, R. D., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica, 53*(1), 1-18.

World Health Organization. (2001). Towards a common language for functioning, disability and health. In *International Classification of Functioning, Disability and Health* (pp. 1-22). Geneva, Switzerland: WHO.

Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of intelligibility of dysarthric speech.* Tigard, Oregon: CC Publications.

Yorkston, K. M., & Beukelman, D. R. (1983). The influence of judge familiarization with the speaker on dysarthric speech intelligibility. In W. Berry (Ed.), *Clinical Dysarthria* (pp. 155-164). San Diego, CA: College-Hill Press.

Yorkston, K. M., Bombardier, C., & Hammen, V. L. (1994). Dysarthria from the viewpoint of individuals with dysarthria. In J. A. Till, K. M. Yorkston & D. R. Beukelman (Eds.), *Motor Speech Disorders: Advances in Assessment and Treatment* (pp. 19-35). Baltimore: Paul H. Brooks Publishing.

Yorkston, K. M., Klasner, E. R., & Swanson, K. M. (2001). Communication in context: A qualitative study of the experiences in individuals with multiple sclerosis. *American Journal of Speech-Language Pathology, 10*(2), 126-137.

Yorkston, K. M., Strand, E. A., & Kennedy, M. R. T. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology, 5,* 55-66.

Yunusova, Y., Weismer, G., Westbury, J. R., & Lindstrom, M. J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research, 51,* 596-611.

APPENDICES

Appendix A: Certificate of approval from UBC behavioural research ethics board

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - FULL BOARD

PRINCIPAL INVESTIGATOR:	INSTITUTION / DEPARTMENT:	UBC BREB NUMBER:
Barbara A. Purves	UBC/Medicine, Faculty of/Audiology & Speech Sciences	H09-00407

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Vancouver (excludes UBC Hospital)
Vancouver Coastal Health (VCHRI/VCHA)	GF Strong Rehabilitation Centre

Other locations where the research will be conducted:

Research sessions may be conducted at participant's home, if participant with dysarthria prefers this, i.e., participant with dysarthria will have choice of two locations (UBC research lab or his/her own home).

CO-INVESTIGATOR(S):
Jana Johnston

SPONSORING AGENCIES:
N/A

PROJECT TITLE:
Comprehensibility assessment: The influence of familiar and unfamiliar communication partners.

REB MEETING DATE:	CERTIFICATE EXPIRY DATE:
March 12, 2009	**March 12, 2010**

DOCUMENTS INCLUDED IN THIS APPROVAL:	DATE APPROVED:
	April 16, 2009

Document Name	Version	Date
Consent Forms:		
Consent Form - Dysarthric Speaker	N/A	March 30, 2009
Consent Form - Familiar Partner	N/A	March 30, 2009
Consent Form - Unfamiliar Partner	N/A	March 30, 2009
Advertisements:		
Recruitment Poster - Unfamiliar Partners	N/A	February 26, 2009
Letter of Initial Contact:		
Information Letter - Unfamiliar Partner	N/A	March 26, 2009
Information Letter - Dysarthric Speaker	N/A	March 26, 2009
Initial Contact Letter & Clinical Information Package	N/A	February 26, 2009
Information Letter - Familiar Partner	N/A	March 26, 2009
Other Documents:		
Comprehensibility Assessment - Scoring Protocol	N/A	February 24, 2009
Comprehensibility Materials - Sample Sentence Set 1	N/A	February 26, 2009
Comprehensibility Task - Instructions for Participants	N/A	February 24, 2009
Other:		
N/A		

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair
Dr. Ken Craig, Chair
Dr. Jim Rupert, Associate Chair
Dr. Laurie Ford, Associate Chair
Dr. Anita Ho, Associate Chair

Appendix B: Consent form for speaker with dysarthria

THE UNIVERSITY OF BRITISH COLUMBIA

School of Audiology and Speech Sciences
Faculty of Medicine
2177 Wesbrook Mall
Vancouver, B.C., Canada V6T 1Z3
Tel: 604-822-5591 Fax: 604-822-6569

CONSENT FORM (PARTICIPANT WITH DYSARTHRIA)

COMPREHENSIBILITY ASSESSMENT: THE INFLUENCE OF FAMILIAR AND UNFAMILIAR COMMUNICATION PARTNERS

Principal Investigator:
Barbara Purves, M.Sc., Ph.D., S-LP(C), Assistant Professor

Co-Investigator:
Jana Johnston, B.A., Graduate Student, Speech-Language Pathology

PURPOSE

The purpose of this study is to examine communicative success (comprehensibility) between speakers with dysarthria and communication partners who are either familiar or unfamiliar with them. You have been invited to take part because you (1) have dysarthria, and (2) have a partner with whom you communicate frequently who would also be willing to participate in this study. This research is being conducted as part of a Master's thesis project.

STUDY PROCEDURES

This study involves three participants: You, a person with whom you communicate frequently (e.g., a member of your immediate family, close friend, etc.), and a person who is unfamiliar to you. If you agree to be in this study, we will ask you to participate in three tasks. First, we will ask you to read a set of 30 sentences out loud while you are video recorded. This will take approximately 30 minutes to complete. In another session on a different day, we will ask you to read another list of 30 sentences out loud, one at a time, this time to your familiar partner. Your partner will repeat back to you each sentence he/she hears, and if any communication breakdowns occur, you and the listener will resolve them as you see fit. Finally, on a third day, we will ask you to perform the same task, this time to an unfamiliar listener (a university student you have never met) while reading a third set of 30 sentences. Again, you will work with this listener to resolve any breakdowns that occur so that he/she can repeat back the sentences you have read out. Each of these joint sessions will be video recorded, and will take up to one hour each. In any session, you can feel free to take breaks at any time and you can stop at any time. All three of these sessions will be conducted at a location convenient to you (either in your home or at UBC). Following the final session, you will be given the opportunity to discuss your participation experience with the researcher. A summary of the overall findings of the project will be provided to you upon request.

CONFIDENTIALITY

Any information obtained from you will be kept strictly confidential. All documents and recordings will be identified by a code number and will be kept in a locked filing cabinet or on a password-protected computer. Access to these data will be restricted to the Principal Investigator and Co-Investigator (named in this consent form) as well as two paid research assistants (to be appointed at a later date). The data will be kept securely by the Principal Investigator for at least five years after completion of this study. You will not be identified by name in any reports or publications resulting from this study. The data from this study may be used in future analyses related to this study and/or for educational purposes. In this case, confidentiality will be strictly maintained. You will have the opportunity to look at your video data and give additional signed consent before any video recordings of you can be used for research or educational purposes.

RISKS AND BENEFITS

Other than the possibility of fatigue or frustration, there are no known risks associated with participating in this study. The potential benefits include learning more about your communication with unfamiliar and familiar partners. You will be given the option of learning your individual results and/or the general results of the study once it is complete.

Be advised that your participation in this research is voluntary and will not jeopardize your access to health care or speech-language pathology services in any way. The speech-language pathologist distributing this information to you will not be advised as to whether you decide to pursue involvement in this study or not.

HONORARIUM

You will receive $10.00 per session in compensation for any inconvenience and up to $10.00 for any travel expenses incurred.

CONTACT

If you have any questions or if you want further information about the study, you may contact:

Barbara Purves
Tel: 604-822-2288 E-mail: purves@audiospeech.ubc.ca
Jana Johnston
Tel: 604-827-3042 E-mail: janaj@interchange.ubc.ca

If you have any concerns about your treatment or your rights as a research subject, you may contact the Research Subject Information Line at the University of British Columbia at 604-822-8598.

CONSENT

Your participation is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy to your access to services or health care. Your signature below indicates that you have received a copy of this consent for your own records. Your signature indicates that you consent to participate in this study.

_________________________ _________________________

Participant Name (please print) Date

_________________________ _________________________

Participant Signature Date

CONSENT TO USE VIDEO RECORDINGS FOR RESEARCH AND EDUCATIONAL PURPOSES (to be completed after video recording)

Your signature indicates that you have had the opportunity to look at video recordings made of you in the course of the study. Your signature indicates that you consent to the use of these video recordings for the research purposes described above, (i.e., to analyze comprehensibility, and for related research questions that may arise in the course of this research), and that you consent to viewing by both investigators and research assistants, by members of Ms. Johnston's thesis committee and by those attending her thesis defense, and, in the event of further analysis, by others working under the supervision of Dr. Purves.

_________________________ _________________________

Participant Signature Date

Your signature indicates that you consent to the showing of these video recordings for educational purposes, including dissemination of research findings at professional and scientific conferences, and education for graduate students in speech-language pathology or related health professions, and that you understand that no additional identifying information will be given with these recordings during such exhibition.

_________________________ _________________________

Participant Signature Date

Appendix C: Consent form for familiar partner

THE UNIVERSITY OF BRITISH COLUMBIA

School of Audiology and Speech Sciences
Faculty of Medicine
2177 Wesbrook Mall
Vancouver, B.C., Canada V6T 1Z3
Tel: 604-822-5591 Fax: 604-822-6569

CONSENT FORM (FAMILIAR PARTNER)

COMPREHENSIBILITY ASSESSMENT: THE INFLUENCE OF FAMILIAR AND UNFAMILIAR COMMUNICATION PARTNERS

Principal Investigator:
Barbara Purves, M.Sc., Ph.D., S-LP(C), Assistant Professor

Co-Investigator:
Jana Johnston, B.A., Graduate Student, Speech-Language Pathology

PURPOSE

The purpose of this study is to examine communicative success (comprehensibility) between speakers with dysarthria and communication partners who are either familiar or unfamiliar with them. You have been invited to take part because you have a close relationship and communicate frequently with a person who has dysarthria (a motor speech disorder that affects the clarity of speech production). This research is being conducted as part of a Master's thesis project.

STUDY PROCEDURES

This study involves three participants: You, your partner, and an unfamiliar person. Should you agree to participate, your role in this study will consist of two parts. In one session, we will ask you to complete a task together with your partner. This will involve him/her reading a set of 30 sentences out loud, one at a time, while you repeat back to him/her each sentence you hear. If any communication breakdowns occur during this process, you and your partner will resolve them as you see fit. This session will be video recorded, and will take up to one hour to complete. It will be conducted at a location convenient to your partner (either in his/her home or at UBC). In a separate session, we will ask you to watch a video recording of your partner reading a different set of 30 sentences, and you will be asked to write down the sentences that you hear as accurately as you can. This will take approximately 30 minutes. In any session, you can feel free to take breaks at any time and you can stop at any time. This session will take place in the Acquired Language Disorders research lab on the UBC campus or in your own home, whichever you prefer. Following the final session, you will be given the opportunity to discuss your participation experience with the researcher. A summary of the overall findings of the project will be provided to you upon request.

CONFIDENTIALITY

Any information obtained from you will be kept strictly confidential. All documents and recordings will be identified by a code number and will be kept in a locked filing cabinet or on a password-protected computer. Access to these data will be restricted to the Principal Investigator and Co-Investigator (named in this consent form) as well as two paid research assistants (to be appointed at a later date). The data will be kept securely by the Principal Investigator for at least five years after completion of this study. You will not be identified by name in any reports or publications resulting from this study. The data from this study may be used in future analyses related to this study and/or for educational purposes. In this case, confidentiality will be strictly maintained. You will have the opportunity to look at your video data and give additional signed consent before any video recordings of you can be used for research or educational purposes.

RISKS AND BENEFITS

Other than the possibility of fatigue or frustration, there are no known risks associated with participating in this study. The potential benefits include learning more about the communication between you and your partner. You will be given the option of learning your individual results and/or the general results of the study once it is complete.

Be advised that your participation in this research is voluntary and will not jeopardize your (or your partner's) access to health care or speech-language pathology services in any way. The speech-language pathologist distributing this information to you will not be advised as to whether you decide to pursue involvement in this study or not.

HONORARIUM

You will receive $10.00 per session in compensation for any inconvenience and up to $10.00 for any travel expenses incurred.

CONTACT

If you have any questions or if you want further information about the study, you may contact:

Barbara Purves
Tel: 604-822-2288 E-mail: purves@audiospeech.ubc.ca
Jana Johnston
Tel: 604-827-3042 E-mail: janaj@interchange.ubc.ca

If you have any concerns about your treatment or your rights as a research subject, you may contact the Research Subject Information Line at the University of British Columbia at 604-822-8598.

CONSENT

Your participation is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy to your access to services or health care. Your signature below indicates that you have received a copy of this consent for your own records. Your signature indicates that you consent to participate in this study.

_________________________ _________________________

Participant Name (please print) Date

_________________________ _________________________

Participant Signature Date

CONSENT TO USE VIDEO RECORDINGS FOR RESEARCH AND EDUCATIONAL PURPOSES (to be completed after video recording)

Your signature indicates that you have had the opportunity to look at video recordings made of you in the course of the study. Your signature indicates that you consent to the use of these video recordings for the research purposes described above, (i.e., to analyze comprehensibility, and for related research questions that may arise in the course of this research), and that you consent to viewing by both investigators and research assistants, by members of Ms. Johnston's thesis committee and by those attending her thesis defense, and, in the event of further analysis, by others working under the supervision of Dr. Purves.

_________________________ _________________________

Participant Signature Date

Your signature indicates that you consent to the showing of these video recordings for educational purposes, including dissemination of research findings at professional and scientific conferences, and education for graduate students in speech-language pathology or related health professions, and that you understand that no additional identifying information will be given with these recordings during such exhibition.

_________________________ _________________________

Participant Signature Date

Appendix D: Consent form for unfamiliar partner

THE UNIVERSITY OF BRITISH COLUMBIA

School of Audiology and Speech Sciences
Faculty of Medicine
2177 Wesbrook Mall
Vancouver, B.C., Canada V6T 1Z3
Tel: 604-822-5591 Fax: 604-822-6569

CONSENT FORM (UNFAMILIAR PARTNER)

COMPREHENSIBILITY ASSESSMENT: THE INFLUENCE OF FAMILIAR AND UNFAMILIAR COMMUNICATION PARTNERS

Principal Investigator:
Barbara Purves, M.Sc., Ph.D., S-LP(C), Assistant Professor

Co-Investigator:
Jana Johnston, B.A., Graduate Student, Speech-Language Pathology

PURPOSE

The purpose of this study is to examine communicative success (comprehensibility) between speakers with dysarthria and communication partners who are either familiar or unfamiliar with them. You have been invited to take part because you do not have a close relationship or communicate frequently with a person with dysarthria (a motor speech disorder that affects the clarity of speech production). This research is being conducted as part of a Master's thesis project.

STUDY PROCEDURES

This study involves three participants: You, a person with dysarthria, and a close contact of the person with dysarthria. Should you agree to participate, your role in this study will consist of two parts. In one session, we will ask you to complete a task together with the dysarthric speaker. This will involve him/her reading a set of 30 sentences out loud, one at a time, while you repeat back to him/her each sentence you hear. If any communication breakdowns occur during this process, you and the speaker will resolve them as you see fit. This session will be video recorded, and will take up to one hour to complete. It will be conducted at a location convenient to the speaker with dysarthria (either in his/her home or at UBC). In a separate session, we will ask you to watch a video recording of the speaker reading a different set of 30 sentences, and you will be asked to write down the sentences that you hear as accurately as you can. This will take approximately 30 minutes. In any session, you can feel free to take breaks at any time and you can stop at any time. This session will take place in the Acquired Language Disorders research lab on the UBC campus. Following the final session, you will be given the opportunity to discuss your participation experience with the researcher. A summary of the overall findings of the project will be provided to you upon request.

CONFIDENTIALITY

Any information obtained from you will be kept strictly confidential. All documents and recordings will be identified by a code number and will be kept in a locked filing cabinet or on a password-protected computer. Access to these data will be restricted to the Principal Investigator and Co-Investigator (named in this consent form) as well as two paid research assistants (to be appointed at a later date). The data will be kept securely by the Principal Investigator for at least five years after completion of this study. You will not be identified by name in any reports or publications resulting from this study. The data from this study may be used in future analyses related to this study and/or for educational purposes. In this case, confidentiality will be strictly maintained. You will have the opportunity to look at your video data and give additional signed consent before any video recordings of you can be used for research or educational purposes.

RISKS AND BENEFITS

Other than the possibility of fatigue or frustration, there are no known risks associated with participating in this study. The potential benefits include learning more about communication with persons with dysarthria. You will be given the option of learning your individual results and/or the general results of the study once it is complete.

HONORARIUM

You will receive $10.00 per session in compensation for any inconvenience and up to $10.00 for any travel expenses incurred.

CONTACT

If you have any questions or if you want further information about the study, you may contact:

Barbara Purves
Tel: 604-822-2288 E-mail: purves@audiospeech.ubc.ca
Jana Johnston
Tel: 604-827-3042 E-mail: janaj@interchange.ubc.ca

If you have any concerns about your treatment or your rights as a research subject, you may contact the Research Subject Information Line at the University of British Columbia at 604-822-8598.

CONSENT

Your participation is entirely voluntary and you may refuse to participate or withdraw from the study at any time without jeopardy to your access to services or health care. Your signature below indicates that you have received a copy of this consent for your own records. Your signature indicates that you consent to participate in this study.

__________________________ __________________________

Participant Name (please print) Date

__________________________ __________________________

Participant Signature Date

CONSENT TO USE VIDEO RECORDINGS FOR RESEARCH AND EDUCATIONAL PURPOSES (to be completed after video recording)

Your signature indicates that you have had the opportunity to look at video recordings made of you in the course of the study. Your signature indicates that you consent to the use of these video recordings for the research purposes described above, (i.e., to analyze comprehensibility, and for related research questions that may arise in the course of this research), and that you consent to viewing by both investigators and research assistants, by members of Ms. Johnston's thesis committee and by those attending her thesis defense, and, in the event of further analysis, by others working under the supervision of Dr. Purves.

__________________________ __________________________

Participant Signature Date

Your signature indicates that you consent to the showing of these video recordings for educational purposes, including dissemination of research findings at professional and scientific conferences, and education for graduate students in speech-language pathology or related health professions, and that you understand that no additional identifying information will be given with these recordings during such exhibition.

__________________________ __________________________

Participant Signature Date

Appendix E: Test procedures for structured assessment of comprehensibility
(adapted from Visser, 2004)

1. Select sentences from the Assessment of Intelligibility of Dysarthric Speech (AIDS) of varying length (e.g. 5-10 words). Randomize length.
2. Have speaker read each sentence to the listener.
3. Listener must repeat each sentence back to the speaker.
4. Speaker gives feedback to the listener.
5. Speaker and listener mutually decide when to move on to the next sentence.

Conditions:

1. Speaker cannot show the listener the printed sentence.
2. Speaker can decide how accurate listener's response needs to be.
3. Speaker and listener can use any strategies they choose.

Appendix F: Transcription conventions for structured assessment of comprehensibility

(abc) Transcriber comments
#abc# Extra-textual information (e.g., noting beginnings and endings of recordings)
<abc> Overlapping speech
abc- Abandoned or interrupted utterance
abc. Falling intonation (e.g., sentence-final)
abc? Rising intonation (e.g., questioning)
abc~ Semi-rising intonation (e.g., listing, pausing expectantly)
abc! Emphasis
0:00 Time marker for video recording*
X Unintelligible syllable

Participants

E: Experimenter

S1: Speaker 1
F1: Familiar Partner 1
U1: Unfamiliar Partner 1

S2: Speaker 2
F2: Familiar Partner 2
U2: Unfamiliar Partner 2

*NOTE: Time markers on transcripts are simply a guide for the transcriber to locate the beginning of the sentence. They do NOT denote where scorers should base the beginnings of their timing measures.

Appendix G: Structured assessment of comprehensibility: Scoring criteria and procedures – Revised (based on Visser, 2004)

(1) Time to do total number of words

- INCLUDE:
  - ALL comments and feedback from speaker's first presentation to the final agreement.
    - Begin timing from where speaker initiates phonation (may include long, audible inhalations).
    - Stop timing when agreement has been reached by both dyad members. Agreement can be signaled verbally (e.g., "correct", etc.) or nonverbally (e.g., head nod, moving on to the next sentence, etc.). Final agreement must be determined between the speakers in the dyad and not by a third party listener.
      - If repetition of all or part of a sentence (even after one dyad member has signaled agreement) is judged to be part of the agreement process itself, include the repetition in the timing measure for that sentence.
      - If there is NO clear moment of agreement, stop timing at the moment where the sentence is abandoned.
    - In cases where there is a recording break within a single sentence sequence, begin counting time again from the first presentation of the interrupted target sentence by either dyad member in the subsequent recording.
  - Reinforcers
    - E.g., "excellent", "good", "right"
- DO NOT INCLUDE:
  - Any interaction with the examiner, including false starts where the examiner needs to redirect the dyad.
  - Any external interruption
    - E.g., phone ringing
  - Any editorial comments about the sentence itself that contain part or all of the target words
    - E.g., "the book has a red cover, what a funny sentence!"

(2) Number of problem words

- INCLUDE:
  - ONLY communication that occurs within the timing measures identified in criteria (1) above.
  - Every word of the target sentence (no more than once) that, after the first presentation, needs subsequent repetition OR is incorrectly guessed by the listener.
  - When the listener responds to presentation of the sentence with, e.g., "I didn't get ANY of that" or indicates nonverbally that a repetition is needed (e.g., with quizzical facial expression), count EVERY target word as a problem word.
  - When the listener got it right the first time (i.e., seemed to agree with speaker as to the target on first presentation) but subsequently forgets it OR omits it OR changes it to an incorrect guess.
    - Note: forgotten words may or may not be indicated by their use of a placeholder phrase, e.g., "something", "some kind of". Count placeholder phrases based on the number of words they are used to substitute:
      - e.g., for the target "blue chair", if the listener guessed "some kind of chair", the placeholder phrase "some kind of" would indicate that "blue"

is one problem word.

- DO NOT INCLUDE:
  - Words in the target sentence where the speaker repeats parts or all of the sentence more than once in the context of a self-initiated self-repair or a false start (e.g., c-c-coffee). In other words, if the speaker produces part or all of the sentence once, then self-selects to repeat it again for increased clarity (i.e., breakdown is not acknowledged by the listener, either verbally or nonverbally), consider this a self-initiated self-repair.
    - Note: false starts and self-initiated self-repairs may or may not span several utterances (even multiple recordings).
  - Words produced by the listener with a slip of the tongue or an obviously accidental misarticulation (e.g., enormlous) but that are otherwise correct.
  - Words that are accidentally omitted from or added to the sentence by the speaker.
  - Where the listener guesses two words that sound identical to the single word target (e.g., horseback/horse back, cannot/can not). Consider this to be a correct guess (i.e., no problem words).

(3) Number of extra speaker words

- INCLUDE:
  - ONLY communication that occurs within the timing measures identified in criteria (1) above.
  - EVERY repetition of the target word, even if the word is embedded in a sentence (e.g., "you got 'he' right": count 'he'). See exception below regarding false starts and self-initiated self-repairs.
  - Part words (i.e., if the speaker says a syllable purposefully and in isolation) and each letter given verbally (e.g., H… A…S…T…).
  - Where the speaker separates a contraction (e.g., "weren't" → "were not"), count the two parts of the contraction as two extra speaker words.
- DO NOT INCLUDE:
  - The speaker's first production of the target word in the sentence.
  - Other comments, even if related to the task (e.g. "you're close").
  - Where the speaker repeats parts or all of the target sentence more than once in the context of a self-initiated self-repair or a false start (e.g., c-c-coffee). In other words, if the speaker produces part or all of the sentence once, then self-selects to repeat it again for increased clarity (i.e., breakdown is not acknowledged by the listener, either verbally or nonverbally), consider this a self-initiated self-repair.
    - Note: false starts and self-initiated self-repairs may or may not span several utterances (even multiple recordings).

(4) **Number of extra listener words**

- INCLUDE:
  - ONLY communication that occurs within the timing measures identified in criteria (1) above.
  - EVERY repetition of a target word (even if for confirmation at the very end, e.g., repetition of the entire sentence), and even if the word is embedded in a sentence (e.g., "did I get 'he' right?": count 'he').
  - Every incorrect guess from (5) below, including nonwords and extra words added by the listener. See below for an exception regarding self corrections.
  - Placeholder words (e.g., "something something", to indicate two words not understood and not guessed at) and placeholder phrases (e.g. "some kind of") are counted as one

extra listener word each, regardless of how many target words they are used to represent.
    - E.g., "He… something… island" : count the placeholder word *something* as one extra listener word
    - E.g., "He bought… some kind of tractor" : count the placeholder phrase *some kind of* as one extra listener word.
  - Part words (e.g., if the speaker says a syllable purposefully and in isolation) and each letter given during oral spelling (e.g., H…A…S…T…).
  - Where the listener separates a contraction (e.g., "weren't" → "were not"), count the two parts of the contraction as two extra listener words.
  - Where the listener guesses two words for a single word target, count two extra listener words (see exception below)
- DO NOT INCLUDE:
  - The listener's first production of each target word in the sentence, even if it comes quite far along in the sequence.
  - Other comments, even if related to task (e.g. "am I close?").
  - Where the listener repeats parts or all of the target sentence more than once in the context of a self-initiated self-repair or a false start (e.g., c-c-coffee). In other words, if the listener produces part or all of the sentence once, then self-selects to repeat it again for increased clarity (i.e., breakdown is not acknowledged by the speaker, either verbally or nonverbally), consider this a self-initiated self-repair.
    - Note: false starts may or may not span several utterances (even multiple recordings).
  - Errors that are immediately self-corrected (to the correct target word).
  - Where the listener guesses two words that sound identical to the single word target (e.g., horseback/horse back, cannot/can not). Consider this to be a correct guess (i.e., no extra listener words).

(5) Number of words guessed incorrectly
- INCLUDE:
  - ONLY communication that occurs within the timing measures identified in criteria (1) above.
  - EVERY incorrect guess at a target (even if it is a repetition of an earlier guess OR where the listener got it right the first time but subsequently changes it to an incorrect guess later in the sequence).
    - See exception below regarding self-corrections.
  - Incorrect guesses documented by the listener him/herself:
    - e.g., "it's not X, is it?": count X as an error.
  - Incorrect part word guesses (e.g., if the listener says an incorrect syllable)
  - Nonword guesses (e.g., impoachers).
  - Extra words added to the target sentence.
  - Each incorrect oral letter.
  - Where the listener separates a contraction (e.g., "weren't" → "were not"), count the two parts of the contraction as two incorrect guesses.
  - Where the listener guesses two words for a single word target, count two incorrect guesses.
- DO NOT INCLUDE:
  - Clear placeholder, (e.g., "something X") where the listener does not seem to be offering "something" as a guess at the target.
  - Incorrect or nonword guesses made in the context of a slip of the tongue (e.g.,

enormlous), a false start.
    - Errors that are immediately self-corrected (to the correct target word) in the context of a self-initiated self-repair.

(6) Ultimate correct:

- ONLY consider communication that occurs within the timing measures identified in criteria (1) above.
- Consider only the correct number of words in FINAL presentation. Even if a word was correctly produced earlier by the listener, but in final repetition, it is incorrect or omitted, if the speaker accepts it, it is counted as an error word (giving a score of n-1/n, where n is the total number of words in the sentence).
- Each word added to the target sentence that is accepted in final presentation results in one point deducted from the score for ultimate words correct.
    - E.g., the listener correctly repeats 7 out of 7 words of a sentence, but adds in an extraneous word that is accepted by both dyad members: ultimate correct score would be 6 out of 7.
- If a contraction is expanded and accepted in final presentation (e.g., “we will” for “we’ll”), deduct one point for the target contraction being incorrect.

Appendix H: Strategy record form for scorers

Dyad: _________________ Rater: __________________

Recording Strategy Use:

In the course of watching and reading through these interactions, you will notice that both the speaker with dysarthria and her communication partners use various strategies to improve their communication.

Strategies can be thought of as any method by which the sender **or** receiver of a message enhances, clarifies, or repairs communication (e.g., requesting repetition, requesting feedback, gesture, etc.)

Please list below the strategies you observe used **at least once** during the interaction. You do not need tally how many times each strategy is used. Simply indicate the type of strategy and the line number on the transcript where **one** example can be located. (Note: If the strategy is nonverbal and not represented on the transcript with a line number, please indicate between which transcript lines the strategy occurs.)

Speaker with Dysarthria	Line #	Partner	Line #

Appendix I: Communicative strategies identified by a single rater

Dyad: S1F1

S1	F1
• Specific verbal confirmation (e.g. "that's right") • General verbal confirmation (e.g. "okay") • Corrects/clarifies with listener regarding the correct target word • Slowed speech • Takes breaths between words • Emphasis on trouble word • Laughter • Moving on to the next sentence (indicates agreement indirectly) • Breaks word into syllables • Explains a particular word may be unfamiliar to listener • Indicates that wording is awkward or atypical • Emphasis on humourous target word • Indicates it's time to move on • Indicates task is almost done • Indicates one last try is necessary	• Leans in in preparation for listening/attending to speaker's production • Uses placeholder phrase (e.g. "blank") • Laughter • Explains general difficulty with English pronunciation • Guessing • General nonverbal agreement (e.g. nodding) • Repeats any words understood to inform speaker of which words were not understood • Pausing mid-guess to indicate need for repetition • Indicates difficulty with the task (e.g. "it's hard") • Gives general feedback about how task is going/how well the speaker is doing • Gives feedback about whether a clarification helped or not (e.g. "it's clearer now")

Dyad: S1U1

S1	U1
• Laughter • Takes breaths between words • Shakes head to indicate having difficulty • Warning partner that certain constructs are difficult to understand (e.g., contractions), therefore giving the partner insight about future sources of communication breakdown • Breaking up the sentence into "chunks" (word or phrase) • Providing problematic adjective within a noun phrase to give partner clue about part of speech of the target • Specific confirmation of message ("yes") • Interrupting partner's incorrect guess • Quizzical facial expression indicating speaker has not heard partner's guess correctly • Leans in or turns head to hear partner better • Puts emphasis on targets that are funny/enjoyable • Contrasts incorrect and correct guesses • Rephrases the message • Moving on to the next sentence (indicates agreement indirectly) • Makes general comments on the difficulty/nature of the task materials • Gives specific verbal feedback (e.g., "good!") • Comments on the truth value of a sentence in	• Watches speaker's mouth (lip reading?) • Mouths along with speaker's presentation • General verbal agreement (e.g., "okay") • General nonverbal agreement (e.g., nodding) • Deictic gestures showing one beat per word in the sentence • Pausing mid-guess to indicate need for repetition • Lack of guess to indicate need for repetition • Breaking oral spelling into single-letter turns • Verbally acknowledging that part of sentence was missed • Reconstructing sentence that has been broken into "chunks" to confirm entire message at the end of a sentence sequence • Uses rising (questioning) intonation to indicate uncertainty • Comments on logic of the sentence (whether it makes sense or not) • Acknowledges wrong guess • Commenting on sentence content in relation to speaker's life context (to build rapport?) • Leans in to listen/attend carefully or request repetition • Commenting on difficulty of the sentence • Nonverbally and verbally indicates sentence judged to be odd/atypical

relation to physical context
- Jokes and comments about the partner's performance
- Verbally signals task is over ("that's it")
- Shakes head to indicate 'no'

- Comments on/acknowledges speaker's communication strategy (splitting contractions)
- Jokes and comments about humorous sentence content
- Confused facial expression
- Apologizes for incorrect guess
- Compares current sentence to previous sentence

Dyad: S2F2

S2

- Speaks slowly
- Laughter (to indicate wrong guess)
- Requesting a break/"time-out"
- Specific verbal confirmation ("uhhuh"/"right")
- Break sentence into "chunks" (word or phrase)
- Clarification of the correct target word
- Beat gestures for each orally spelled letter
- Wiping/clearing saliva before speaking
- Nonverbally signals task is over ("ta-da!" gesture)
- Emphasis on trouble word

F2

- Self-initiated self-repair
- Guessing
- Confused/unsure facial expression signals need for repetition
- Laughter (acknowledging wrong guess)
- Requesting a break/"time-out" for speaker
- Repeating the target sentence after agreement has been reached (as self-confirmation?)
- General verbal agreement/confirmation ('kay)
- Lack of guess/silence signals need for repetition
- Initiates breaking sentence into "chunks" by repeating each phrase as it is delivered
- Leans in to signal need for repetition

Dyad: S2U2

S2

- Specific confirmation (e.g., "yes")
- Pointing at partner while indicating confirmation (e.g., while nodding and saying "that's right")
- Interrupting a wrong guess
- Breaks oral spelling into single letter turns (or turns with few letters)
- Speaks slowly
- Gestures plus sign (+) for 'and'
- Deictic beat gestures marking syllables
- Chooses to abandon a frustrating sentence, using gesture to signal "forget it"
- Laughter when sentence becomes too frustrating
- Requests a break/"time-out"
- Indicates that a guess was correct despite one added word (e.g., "yes, but no X")
- Repeats the full sentence after agreement for confirmation that partner understands
- Indicates positive or negative tone of sentence or target word by nodding or shaking head during presentation
- Points to self/body to indicate target word
- Self-initiated self-repair
- Shakes head to indicate difficult sentence
- Wiping/clearing saliva before speaking
- Drinking water before speaking
- Nonverbally signals task is over ("ta-da!" gesture)

U2

- Self-initiated self-repair
- Apologizing
- General verbal agreement (e.g., "'kay")
- Questioning/rising intonation to indicate uncertainty
- Using placeholder to indicate which word not understood
- Whispering/rehearsing to self before making a guess
- Repeating oral spelling in single letter turn format
- Commenting on difficulty of a sentence
- Pausing mid-guess to indicate need for repetition
- Repetition of previously guessed (correct) targets as frame for new guesses