Journal of Fluency Disorders 48 (2016) 1–15

Contents lists available at ScienceDirect

Journal of Fluency Disorders

Formant transitions in the fluent speech of Farsi-speaking people who stutter

Ali Dehqan [a,b], Fariba Yadegari [a,*], Michael Blomgren [c], Ronald C. Scherer [d]

[a] Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
[b] Health Promotion Research Center, Zahedan University of Medical Sciences, Zahedan, Iran
[c] Department of Communication Sciences & Disorders, University of Utah, Salt Lake City, UT 84112, USA
[d] Department of Communication Sciences and Disorders, Bowling Green State University, Bowling Green, OH 43403, USA

ARTICLE INFO

Article history:
Received 15 January 2015
Received in revised form 11 December 2015
Accepted 4 January 2016
Available online 11 February 2016

Keywords:
Stuttering
Coarticulation
Motor control
F2 transitions
Farsi language

ABSTRACT

Purpose: Second formant (F2) transitions can be used to infer attributes of articulatory transitions. This study compared formant transitions during fluent speech segments of Farsi (Persian) speaking people who stutter and normally fluent Farsi speakers.
Methods: Ten Iranian males who stutter and 10 normally fluent Iranian males participated. Sixteen different "CVt" tokens were embedded within the phrase "Begu CVt an". Measures included overall F2 transition frequency extents, durations, and derived overall slopes, initial F2 transition slopes at 30 ms and 60 ms, and speaking rate.
Results: (1) Mean overall formant frequency extent was significantly greater in 14 of the 16 CVt tokens for the group of stuttering speakers. (2) Stuttering speakers exhibited significantly longer overall F2 transitions for all 16 tokens compared to the nonstuttering speakers. (3) The overall F2 slopes were similar between the two groups. (4) The stuttering speakers exhibited significantly greater initial F2 transition slopes (positive or negative) for five of the 16 tokens at 30 ms and six of the 16 tokens at 60 ms. (5) The stuttering group produced a slower syllable rate than the non-stuttering group.
Conclusions: During perceptually fluent utterances, the stuttering speakers had greater F2 frequency extents during transitions, took longer to reach vowel steady state, exhibited some evidence of steeper slopes at the beginning of transitions, had overall similar F2 formant slopes, and had slower speaking rates compared to nonstuttering speakers. Findings support the notion of different speech motor timing strategies in stuttering speakers. Findings are likely to be independent of the language spoken.

Educational objectives This study compares aspects of F2 formant transitions between 10 stuttering and 10 nonstuttering speakers. Readers will be able to describe: (a) characteristics of formant frequency as a specific acoustic feature used to infer speech movements in stuttering and nonstuttering speakers, (b) two methods of measuring second formant (F2) transitions: the visual criteria method and fixed time criteria method, (c) characteristics of F2 transitions in the fluent speech of stuttering speakers and how those characteristics appear to differ from normally fluent speakers, and (d) possible cross-linguistic effects on acoustic analyses of stuttering.

* Corresponding author. Tel.: +98 2122180043.
E-mail address: fariba.yadegari@gmail.com (F. Yadegari).

http://dx.doi.org/10.1016/j.jfludis.2016.01.005

1. Introduction

Stuttering is considered a multi-dimensional speech disorder influenced by motor, linguistic, and emotional factors. The exact cause of stuttering is unknown, but it is widely accepted that stuttering is associated with deficiencies in the neural functioning that underlies speech production. For instance, studies have identified timing and functional connectivity issues between speech and language areas of the brain (Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Salmelin, Schnitzler, Schmitz, & Freund, 2000), indicating possible deficits in speech motor control strategies. Speech motor control refers to the neuromuscular organization that mediates the complex and precise movements involved in the production of speech (Kent, 2000).

A number of studies have directly and indirectly assessed speech motor control in people who stutter. Direct studies of speech motor control have used a variety of kinematic and isometric force generation tasks to compare stuttering and nonstuttering speakers. For instance, differences have been identified in the relative speed of repetitive movements involving the tongue (McClean & Tasko, 2004), the lips (Howell, Andrew, Bartrip, & Bailey, 2009; Max, Caruso, & Gracco, 2003; Namasivayam & Van Lieshout, 2004; Smith & Goffman, 2004), and the jaw (Max et al., 2003; McClean & Runyan, 2000). These findings have reinforced a long standing belief that, at least in part, the disorder of stuttering is related to a deficit in the ability to coordinate the various components of the speech-motor system (Juste et al., 2012; Van Riper, 1971; Wingate, 1969). While various direct instrumental techniques have been used to investigate speech motor skills in stuttering speakers, these approaches have some limitations including cost, ease of use, and are somewhat invasive (Howell et al., 2009).

Other studies have utilized a variety of indirect methods such as acoustic measures (e.g., Blomgren, Robb, & Chen, 1998; Civier, Tasko, & Guenther, 2010; Colcord & Adams, 1979; Klich & May, 1982; McClean & Tasko, 2004; Yaruss & Conture, 1993). Acoustic measures often lack the specificity of a direct link to an anatomical structure in the speech system; however, the benefits include the noninvasive nature of obtaining an acoustic recording and the ability to record a wide range of speech samples. While the methods vary greatly, the shared goal of all indirect measures has been to determine disordered motor control processes that may be active in people who stutter. The majority of these studies have suggested that the speech production process in adults who stutter is different from that of normally fluent peers (Alfonso, 1991; Kent, 2000). These differences are thought to be reflective, or symptomatic, of an underlying speech planning or speech motor control deficit. In many cases, these differences have been identified in the fluent speech of people who stutter. The process of comparing fluent utterances of a person who stutters to the normally fluent speech of nonstuttering speakers is referred to as the "fluent speech paradigm" (Armson & Kalinowski, 1994). The notion behind the fluent speech paradigm is that differences, however subtle, between the perceptually fluent speech of people who stutter and normally fluent speakers may provide insight into the (disordered) speech motor control strategies of people who stutter.

A specific acoustic feature that has received considerable attention in the stuttering literature is formant frequency (e.g., Blomgren et al., 1998; Howell & Vause, 1986; Klich & May, 1982; Prosek, Montgomery, Walden, & Hawkins, 1987; Robb, Blomgren, & Chen, 1998; Subramanian, Yairi, & Amir, 2003; Yaruss & Conture, 1993). Formant frequencies are spectral peaks that correspond to resonances of the human vocal tract. These formants have distinctive values depending on the shape of the vocal tract and they are particularly sensitive to changes in tongue height and advancement during production of vowels. Assessing formant frequencies, particularly measures of second formant (F2) transitions, allows for inferences to be made concerning lingual position and movement during speech production. There has been much interest regarding F2 transitions in children and adults who stutter (Chang, Ohde, & Conture, 2002; Kowalczyk & Yairi, 1995; Namita & Savithri, 2002; Robb & Blomgren, 1997; Subramanian et al., 2003; Yaruss & Conture, 1993). Viewing the research in aggregate, it appears that individuals who stutter exhibit different patterns of sound segment transitions compared to normally fluent speakers. However, different measures of F2 transitions have been used, making comparisons among studies somewhat challenging.

One method of estimating formant transition characteristics involves visual criteria to determine onset and endpoints of formant transitions. Using wideband (300 Hz) spectrograms, the formant transition has been measured from the first glottal pulse in the target vowel to the point in which a maximally steady-state is visually identified in the formants of the vowel (e.g., Klich & May, 1982; Kowalczyk & Yairi, 1995; Subramanian et al., 2003; Yaruss & Conture, 1993; Zebrowski, Conture, & Cudahy, 1985). The benefit of this method is the entire formant transition is measured in terms of duration and frequency extent. However, a challenge with this method is determining the exact point where a formant ceases to be in transition and has reached "steady state." The task of determining the end point of a transition is further complicated by the fact that the majority (up to 70%) of vowel durations in connected speech have been estimated to be in formant transition (Blomgren & Robb, 1998). The problem of subjectivity in assessing the end point of the transitions led to an additional method of assessing formant transitions that used a fixed time-point criterion (Nearey & Shammass, 1987). In the fixed time-point method, a predetermined standard timeframe is used for all speakers to evaluate rate of F2 change (slope) across tokens. For instance, F2 slopes have been measured at distinct time-points of 30 or 60 ms from the onset of the transition. This second method was used by Robb and Blomgren (1997) to compare five stuttering and five nonstuttering speakers in degree of F2 transition. Robb and Blomgren concluded that the F2 transitions in the stuttering group were characterized by greater frequency extents compared to the nonstuttering group, possibly indicating less efficient coarticulation. In summary, the primary benefit of the visual method of measuring F2 transitions is that the overall duration, overall frequency extent, and overall F2 rate of change (slope) can be assessed. The primary benefit of the fixed time-point criteria is that the initial rate of change of the transition may be compared between speakers irrespective of total transition duration.