Folia Phoniatrica et Logopaedica

Research Article

Folia Phoniatr Logop
DOI: 10.1159/000518138

Received: February 4, 2021
Accepted: June 26, 2021
Published online: June 30, 2021

Disfluencies in English Speech Produced by Spanish-English Bilinguals

Sara Ashley Smith[a] Ai Leen Choo[b] Stephanie R. Seitz[c]

[a]Department of Language, Literacy, Ed.D., Exceptional, and Physical Education, University of South Florida, Tampa, FL, USA; [b]Department of Communication Sciences and Disorders, Georgia State University, Atlanta, GA, USA; [c]Department of Management, California State University East Bay, Hayward, CA, USA

Keywords
Fluency · Disfluency · Bilingualism · Spanish-English · Speech

Abstract
Background: Fluency and disfluency exist on a continuum of speech production. Typically fluent speakers produce varying numbers of disfluencies; this number increases in stressful or cognitively demanding situations. Prior research indicates that adult second language learners produce more disfluencies in their weaker, second language, however, this has not been explored among heritage bilinguals who developed in both languages during childhood. There is a lack of foundational knowledge regarding disfluencies among typically fluent adult bilinguals; typical fluency patterns are likely influenced by bidirectional relationships between languages. These patterns may be viewed as disfluencies by listeners who generally perceive disfluencies unfavorably. ***Objectives:*** The current study explores the quantity and types of disfluencies produced by bilinguals. ***Methods:*** Twenty Spanish-English bilinguals took part in a simulated job interview. Responses were transcribed and the total number and percent of disfluencies were calculated. ***Results:*** The findings indicated that typically fluent Spanish-English bilingual adults produce 6.99 typical (nonstuttered) disfluencies per 100 words and are therefore within the range of normative data on monolingual adults (5.1–10.99 per 100 words). The 2 most common disfluencies were superfluous verbal behaviors and pauses. ***Conclusions:*** The findings revealed that typically fluent Spanish-English bilingual adult participants produced more fixed postures than previously reported among monolingual English speakers.

Introduction

Fluency and disfluency exist on a continuum of speech production [1]. Typically fluent speakers produce varying numbers of disfluencies in normal speech [2]. Research indicates that listeners perceive disfluent speech negatively [3] and disfluencies increase in stressful situations [4, 5]. As such, exploring disfluencies among typically fluent adults in real-life speech scenarios, particularly those in which speech performance is being evaluated (e.g., job interviews), is highly relevant.

Prior research indicates that second language (L2) speaker adults have a higher prevalence of disfluencies in the second language (L2) [6–9]. Using multiple languages

karger@karger.com
www.karger.com/fpl

Correspondence to:
Sara Ashley Smith, sarasmith3@usf.edu

Table 1. Examples of typical disfluencies

Disfluency	Example
Syllable repetitions	*There's…there's*
Incomplete syllable repetitions	*W-w-w-w-would*
Multisyllable unit repetitions	*We were… we were…*
Fixed postures with audible airflow	*Sssssssssome*
Fixed postures without audible airflow	*A*––––– (no sound) *asking*
Superfluous verbal behaviors	*Um, uh*
Pauses	Marked by an absence of speech (>250 ms in duration between words)

may be innately more cognitively demanding due to monitoring and suppression of the stronger nontarget first language (L1) [10]. Conditions with a greater processing load and/or increased linguistic complexity have been associated with increased disfluency [11, 12]. At present, however, comparatively little research has examined disfluency among typically fluent bilingual adults (i.e. bilinguals who do not have speech disorders) who acquired both languages in childhood.

Evidence from Spanish-English bilingual children further indicates that there may be disfluencies in English specific to Spanish speech patterns. Typically fluent Spanish-English bilingual children produce more disfluencies in English, including more mazes and syllable repetitions, than monolingual English speaking children [13]. Researchers have posited that this is evidence of typically fluent Spanish speech patterns occurring in English [5, 14]. Spanish contains more mazes and syllable repetitions than English and these patterns are not perceived as "disfluent" by Spanish listeners [14]. However, this has not been examined among typically fluent adult Spanish-English bilinguals.

Understanding the nature of speech disruptions among bilingual speakers is valuable for informing a broader perspective on speech and language interactions and how Spanish-English bilinguals may be perceived by English monolinguals. This is relevant in the US context, where approximately 15% of the population speaks Spanish at home [15]. The current study explored disfluencies produced in English by typically fluent Spanish-English bilingual adults in a stressful speaking condition, i.e., a simulated job interview. We predicted that Spanish-English bilinguals would produce more disfluencies than prior normative data among English monolinguals.

Fluency and Disfluency

Fluency is commonly defined as the absence of disruptions in the forward flow of speech. On the opposite end of this spectrum are disfluencies, or interruptions in the forward flow of speech. Fluency, in the context of the current study, is different from *proficiency*, the knowledge of a specific language. Disfluencies could be categorized into superfluous behaviors, repeated movements, and fixed postures [16]. Superfluous behaviors include verbal behaviors such as fillers and nonverbal behaviors such as grimacing [16]. Repeated movements consist of syllable repetitions, incomplete syllable repetitions, and multisyllable unit repetitions [16]. Fixed postures consist of prolongations of speech sound and stoppage of speech, which could be accompanied by audible or inaudible airflow [16]. Table 1 presents examples. Although rates of disfluencies vary across individuals, certain types are more common compared to others. Multisyllable unit repetitions and superfluous verbal behaviors are common in English [17, 18]. Other disfluencies such as pauses and mazes (series of words or word fragments that do not contribute to the meaning of the utterance) are also commonly reported [13, 14, 17, 18]. Disfluencies such as syllable repetition, incomplete syllable repetition, and fixed postures with and without airflow are also present in everyday English speech, although excessive levels and increased tension may indicate a fluency disorder, e.g., stuttering [16]. The presence of muscular tension and a longer duration of disfluencies also distinguishes stuttered disfluencies from typical (nonstuttered) disfluencies [16, 17]. However, it should be noted that clinical thresholds for stuttering in English are based on monolingual norms [13].

The literature generally reports that the spontaneous English speech of nonstuttering monolingual adults contains between 5.1 and 10.9 total disfluencies per 100 words [19–22]. Fox Tree [19] estimated that 6 out of every 100 words are disfluencies. Roberts et al. [21] found 5.1 per 100 words for monologues from English speaking males. Shriberg [22] reported a mean of 6 per 100 words ($n = 6$) for telephone conversations with an unfamiliar person on

partner on a chosen topic. Among monolingual English speakers, superfluous verbal behaviors and revisions are the most frequently occurring [21].

Disfluencies that are commonly found in speech are not necessarily errors. Typical or nonstuttered disfluencies are associated with increased planning and processing loads and are more likely to occur where the planning effort is higher [4]. Thus, typical or nonstuttered disfluencies may be related to the speaker coping with the increased processing load [4]. Further, typical disfluencies can improve the discourse content and form [6] and can serve communicative purposes [4, 23]. Hieke [6] proposed that a pause followed by a repeat can serve to stall speech during a hesitation. Superfluous verbal behaviors (um, er, and uh) have been posited to facilitate interpersonal communication; proposed examples include: helping to manage turn-taking by blocking a listener from interrupting, creating a space for the listener to help with information, indicating to the listener that the speaker is hesitant, and indicating a speaker's own recognition when they misspeak [4]. Further, superfluous verbal behaviors have been shown to speed up response times during word/monitoring tasks [24]. Repetitions of single words or multiple words can provide a speaker with more planning time [9]. Repetitions can also serve as a self-initiated repair, allowing the speaker to correct a misspoken word within the same turn [9].

Stressful speaking tasks are associated with greater disfluency. Buchanan et al. [26] found that participants produced more disfluencies during more stressful tasks. McDougall and Duckworth [20] reported 10.9 disfluencies per 100 words for simulated police interviews conducted with British males – likely a stressful task. This number is higher than the previously reported means of 5.1 and 6 [21, 22]. Roberts et al. [21] found that monologues about work had higher rates of disfluencies (i.e., 78) compared to monologues on neutral topics (i.e., 23). Disfluency rates may be higher for abstract topics due to planning difficulty [4]. However, see also the findings of Merlo and Mansur [25] of no relationships between topic familiarity and disfluency.

Certain speaking scenarios, e.g., interviews, have specific stress-inducting features such as social evaluation, motivated performance, and unpredictability [26]. Being evaluated as a speaker is stressful and anxiety-producing, implying concern over one's perception. Listeners may view certain types of speech more favorably (i.e., fluent speech) and others more negatively (i.e., disfluent speech). Interviews, i.e., formal meetings in which an individual is questioned for the purpose of evaluation, present a common high-stakes speech performance scenario.

Bilinguals

The existing research regarding bilinguals has yielded mixed insights into typical disfluencies. Adult L2 speakers produce more mazes than monolinguals in their weaker L2 than in their stronger L1 [7–9]. Hieke [6] reported comparable numbers of stalls between L1 English and L2 English speakers but found that L2 speakers had higher rates of repairs. Enxhi et al. [27] reported that superfluous verbal behaviors were the most common type of disfluency among Malaysian speakers of L2 English, similar to findings among typically fluent monolinguals.

Less research has addressed typically fluent bilingual adults who acquired both of their languages in early childhood [28]. Prior research has examined stuttering among bilingual children [for a review, see 29]. This research has focused on comparing bilinguals to monolinguals and comparing fluency rates between the bilinguals' 2 languages to determine whether there is a discrepancy [5]. Bilingual children may produce more syllable repetitions [5]; however, bilingual children who stutter produce comparable rates of stuttered disfluencies in both languages [13].

We posit that typical patterns of disfluencies among bilinguals are likely unique to the combination of languages being spoken. Research on childhood bilingualism demonstrates bidirectional relationships between linguistic features in each developing language, including vocabulary, phonology, and morphology, such that each language being acquired influences the other [see, e.g., 29–32]. These findings can be interpreted as evidence of a unified language-processing system in which multiple languages actively interact and influence each other. Further, evidence of bidirectional relationships between developing languages has also been observed in fluency research. Spanish-English bilingual children produce a wider range of disfluency rates in English than in Spanish [5] and they produce more mazes than monolinguals [13]. Byrd [14] suggests that Spanish may induce disfluencies (so-called, as defined by standards based on monolingual English research) due to the grammatical structure of the language, independently of language dominance. As such, typical Spanish-English bilingual adults who acquired both languages in childhood may produce more so-called disfluencies in English.

Current Study

There have been increasing calls for documentation of typical disfluent speech among bilinguals [14, 29]. Approximately 15% of the US population (roughly 40 million individuals) speak Spanish at home [15]. The term

heritage bilingual describes an individual who speaks a language due to personal and/or historical connection to that language, such as indigenous or immigrant languages [34]. English is often the dominant language for heritage bilinguals in the US by the time they reach adulthood [35]. The comparative paucity of research addressing disfluency among typically fluent bilingual adults does not match the growing population of heritage bilinguals in the US. If typically fluent Spanish-English bilinguals produce distinct speech patterns in English, there may be implications for how monolingual English listeners perceive the speech of Spanish-English heritage bilinguals.

Much of the research on disfluencies has focused on fluency disorders and, to our knowledge, few prior studies have examined disfluencies among typically fluent Spanish-English bilingual adults. The aim of the current study was to explore English disfluencies produced during a simulated job interview among heritage Spanish-English bilinguals. Using an experimental task designed to simulate a real-world stressful speaking activity, the current study addressed the following research question: What is the nature and distribution of disfluencies (including types and frequency) in English speech produced by heritage Spanish-English bilinguals during a simulated interview?

Prior research points to higher rates of typical disfluencies among bilinguals [9] and higher rates of typical disfluencies in situations with a higher processing load [4] and with high communicative stress [26, 36]. Thus, we expected that interview responses would contain typical disfluencies. Findings can provide insights on the nature of English disfluencies among typically fluent Spanish-English heritage bilinguals.

Materials and Methods

Participants

Twenty Spanish-English bilinguals [14 females, mean age = 24.1 years, SD = 6.45] took part in the current study. The participants were self-reportedly typically developing with no diagnosed speech or language disorders. All spoke Spanish from birth and had learned English in school between the ages of 3 and 7 years. Five participants were born in a Spanish-speaking country and subsequently moved to the USA and began English language schooling (the age at time of moving ranged from 1.5 to 7 years; mean age = 4.5 years, SD = 1.78 years). All of the participants were students at a US university and had English-only education backgrounds. No participant had received Spanish language instruction or attended schooling in Spanish, and none worked in jobs that used Spanish as an explicit job task. Participants self-identified as having higher proficiency in English than Spanish and reported daily use of English and Spanish in home and social settings.

Participants self-identified as "Spanish-English bilinguals" and completed a language environment questionnaire and a language proficiency self-report. Participants all indicated their current use of Spanish and English as a percentage ratio (e.g., 40% Spanish and 60% English) in the home and social settings across all 4 modalities, with a more "balanced" use of Spanish and English for speaking and listening and more English than Spanish in reading and writing (percentage of English used for: speaking, mean = 63.30, SD = 16.74; listening, mean = 59.25, SD = 28.13; reading, mean = 91.4, SD = 7.25; writing, mean = 89.25, SD = 19.9; watching TV, mean = 65.35, SD = 25.6; and listening to music, mean = 53.55, SD = 24.78). Participants self-reported their own language abilities for English and Spanish across the 4 modalities on a scale of 1–10, with 10 being "native-like." Overall, the mean self-ratings for English (speaking, mean = 9.4, SD = 0.75; listening, mean = 9.65, SD = 0.49; reading, mean = 8.9, SD = 1; and writing, mean = 8.9, SD = 1.07) were higher than for Spanish across all modalities (speaking, mean = 8.2 , SD = 1.3; listening, mean = 9.45, SD = 0.69; reading, mean = 7.6, SD = 1.64; and writing, mean = 6.4, SD = 1.76), although self-reported listening abilities were near comparable in English and Spanish.

The participants were administered 2 language proficiency measures – one was a measure of Spanish single word vocabulary (Bilingual Verbal Ability Tests; BVAT) [37] and the other was a standardized measure of English single word vocabulary (WAIS-IV) [38]. Spanish vocabulary raw scores out of 58 were used (mean = 22.5, SD = 4.22). WAIS-IV English vocabulary scores were scaled at 10 (representing the 50th percentile; mean = 8.4, SD = 2.09).

Procedures

The participants attended 2 testing sessions of approximately 1 h each as part of a larger study [39], and they were given a battery of background assessments on language (Spanish and English vocabulary, self-reported daily use) and cognition (executive function, working memory, and nonverbal problem solving) to ensure that they were typically developing, of a comparable language background, and eligible to participate (i.e., heritage Spanish-English bilinguals with comparable language and education history). Eligible participants returned to complete the experiment activity in the second session. The research took place in a quiet room with minimal distractions. The participants were compensated with USD 25 in campus-specific currency for each testing session (USD 50 in total). Compensation was given immediately following the session; participants who did not schedule or show up for the second testing session were still compensated for the first testing session.

Measures

Experimental Task

The participants took part in a simulated video conference interview (comparable to using Skype or Zoom) in which various "interviewers" would call in to the meeting and ask a question. The participants were informed prior to the beginning of the activity that they were going to take part in a mock interview though a virtual meeting platform. They were told that various interviewers would call in to the meeting one by one and ask a question. The participants wore headphones and accessed the activity on a laptop. Using prerecorded videos, interviewers "called in" and joined the meeting such that the interviewee was only looking at and

speaking to 1 person at a time. The interviewers introduced themselves and their role (as potential bosses, coworkers, or subordinates) and then asked a question adapted from interview preparation materials. Each interviewer asked 1 question, which the participant answered. All of the participants "interacted" with each of the interviewers. Interaction with individuals from different roles was consistent across participants. Interview videos were filmed in a recording studio using professional recording equipment by professional videographers. The question order was randomly assigned. Interview questions and simulation were piloted with 6 participants. Examples of questions included: "Can you tell me about a difficult experience you had in a job or class and how you overcame it?", "What would you do if right before a deadline you realized that a report you wrote for your boss or professor was not very good?", and "Tell me about a time when you had too many things to do and had to prioritize. How did you organize your time?" The participants had 2 min and 30 s to respond, uninterrupted (the time was shown by a timer on the screen), and then a new actor would request that the speaker wrap up and ask the next question. While the participant spoke, a prerecorded video of the actor "listening" remained on the screen. When the participant finished responding, he or she pushed a button to indicate that he or she was ready for the next interviewer. The participants' responses were audio recorded. A research assistant remained in the physical room to ensure task fidelity and equipment function.

Audio Speech Samples and Transcription

Six English speech samples were obtained from each participant (i.e., full participant responses to 6 questions) from the larger corpus of interviewee question responses; see Hood [36]. The mean total number of words per response was 152.76 (SD = 93.33). Utterances were defined as a stream of speech under a single intonation contour bounded by pauses and/or constituting a single semantic unit, as per Crookes [40]. Eight hundred three utterances were analyzed, and each interview question response contained an average of 6.5 utterances. Speech samples were transcribed and coded by 4 trained researchers, who were all native English speakers, using CHAT annotation software – a program under the Computerized Language Analysis (CLAN), part of the CHILDES (Child Language Data Exchange System) project [41]. The transcription procedure followed the guidelines for conversational units (c-units); a c-unit is a string of words followed by a pause of 1 or more seconds, concludes with intonation, or contains a grammatically correct structure [42].

Data Analyses

Disfluencies were calculated for each speech sample using the language features in CLAN. For each speech sample, the following disfluencies were coded using CHAT annotation software: syllable repetitions, incomplete syllable repetitions, multisyllable unit repetitions, fixed postures with and without audible airflow, superfluous verbal behaviors, and pauses (Table 1) [17, 41]. For pauses, only durations ≥ 250 ms between words were counted [43]. FLUCALC, a program under CLAN, was then used to generate the number and percent of each type of disfluency from each CHAT transcript. The percent of disfluencies was obtained by dividing the total number of each disfluency type over the total number of words.

Disfluency coding reliability was achieved through consensus reliability. A total of 3 passes (transcription and coding, coding check, and reliability) were performed as follows: (1) trained transcribers transcribed and coded the complete transcripts, (2) another transcriber (who was not involved in the initial transcription) checked the disfluencies codes, and (3) a transcriber randomly selected 10% of the transcriptions for reliability checks. To evaluate the reliability of coding for disfluencies, the randomly selected transcripts were recoded and analyzed. The interrater reliability was high (ICC = 0.933). When disagreements in speech disfluencies occurred, the coders listened to them again and then came to a consensus. A fluency specialist resolved a few remaining disagreements. All statistical analyses were conducted using SPSS version 24 [44].

Table 2. Mean percentage of the different types of disfluencies

Type of disfluency	% (SD)	Range
Superfluous verbal behaviors	4.44 (2.92)	0–20.00
Multisyllable unit repetitions	2.02 (1.96)	0–8.70
Pauses	1.35 (1.70)	0–12.00
Fixed postures with audible airflow	0.88 (1.42)	0–9.76
Syllable repetitions	0.73 (1.04)	0–4.69
Incomplete syllable repetitions	0.28 (0.48)	0–2.63
Fixed postures without audible airflow	0.11 (0.26)	0–1.14
Total	9.81 (6.08)	0–48.00

Results

The research question was: What is the nature and distribution of disfluencies (including types and frequencies) in English speech produced by heritage Spanish-English bilinguals during a simulated interview?

Table 2 presents the types and frequencies of all types of disfluencies. Among the different types of disfluencies, the most common were superfluous verbal behaviors (also known as filled pauses) [16, 17], multisyllable unit repetitions, and pauses. Pauses ranged between 1,000 and 3,000 ms in duration, with an average of 1,283 ms. The current study found an average of 6.99 typical disfluencies per 100 words among Spanish-English heritage bilinguals speaking English. The findings revealed that there were 2.09 stuttered disfluencies per 100 words, including fixed postures with audible airflow (0.88) and fixed postures without audible airflow (0.11).

Discussion/Conclusion

The current study furthers our understanding of typical disfluencies among Spanish-English bilinguals when speaking English. We presented the frequency and type

of disfluencies in English speech produced by heritage Spanish-English bilinguals in a stressful condition, i.e., a mock job interview. The simulated interview had features that are known to induce stress, including social evaluation, motivated performance, and unpredictability [26]. The findings indicate that typically fluent Spanish-English bilingual adults produce 6.99 typical disfluencies per 100 words, which is within the range of normative data for monolingual adults (5.1–10.99 per 100 words). As expected, bilinguals in our study who are typically fluent showed higher percentage rates of typical versus stuttered disfluencies. The findings revealed that typically fluent Spanish-English bilingual adult participants produced more fixed postures (with and without audible airflow) than previously reported among monolingual English speakers.

Characterizing Typical Bilingual Disfluencies

The results showed that typically fluent Spanish-English bilinguals are within the range of normative data on spontaneous English speech in nonstuttering monolingual adults (5.1–10.99 per 100 words). This finding is slightly higher than some previous estimates, such as those of Fox Tree [19] (6 per 100 words), Roberts et al. [21] (5.1 per 100 words), and Shriberg [22] (a mean of 6 per 100 words). Our findings are lower than those of McDougall and Duckworth [20], who reported that monolingual English-speaking adult males produced 10.9 typical disfluencies per 100 words; however this was during a potentially even more stressful activity, i.e., simulated police interviews in which the participants were "accused" of shoplifting.

The current study found that superfluous verbal behaviors, multisyllable unit repetitions, and pauses were among the most common typical disfluencies. This is aligned with previous findings among monolinguals which that showed superfluous verbal behaviors and revisions are the most frequently occurring [21]. Prior research among bilingual, Malaysian university students similarly found that superfluous verbal behaviors were the most common TD [27].

Researchers have discussed the possibility of a higher prevalence of mazes among bilinguals [9] and Spanish speakers specifically [13]. Multisyllable unit repetitions were indeed common (i.e., 73), while syllable repetitions (i.e., 17) were the least common. Our findings are comparable to those of Roberts et al. [21] among monolingual English-speaking males (albeit with some beginner to lower intermediate French as a foreign language experience). The study of Roberts et al. [21] showed more multisyllable unit repetitions than the current study (i.e., 1.08–1.63, depending on the topic).

The number of syllable repetitions found in the current study was comparable to prior findings among monolinguals. Roberts et al. [21] reported syllable repetitions between 0.58 and 0.78 (again, depending on the topic), and incomplete syllable repetitions ranged from 0.15 to 0.23); the current study found values of 0.73 and 0.28, respectively. However, the current study employed a stressful task (i.e., a simulated job interview), and this may have resulted in more disfluencies. Buchanan et al. [26] found that participants produced more disfluencies during more stressful tasks. Roberts et al. [21] also found that monologs about work had higher rates of disfluencies (0.78 and 0.23), comparable to those in the current study which also focused on talking about work.

The current study's findings on stuttered disfluencies differ from prior work among monolinguals. The current study found 2.09 stuttered disfluencies per 100 words, which is higher than in prior work among monolingual English speakers, such as the study of Roberts et al. [21]. Specifically, a fixed posture with audible airflow were more prevalent in the current study (i.e., 0.88 compared to only 0.14–0.24 in the study of Roberts et al. [21]). The current study also reports fixed postures without audible airflow [11] while Roberts et al. [21] report none across all topics. Thus, the current study presents preliminary evidence that typically fluent Spanish-English bilingual adults might produce more stuttered disfluencies, specifically fixed postures, when speaking in English than monolingual English speakers. This finding is consistent with prior work among Spanish-English bilingual children [13].

Bilingual speakers likely show distinct typical patterns of disfluencies that are specific to the *combination* of the 2 languages being spoken. Prior research demonstrates that there are bidirectional relationships between a developing bilingual's 2 languages; these have been found for vocabulary, phonology, and morphology [see, e.g., 30–33]. The fluency patterns observed here are therefore likely distinct to Spanish-English bilinguals.

Considering the Role of Disfluencies

Disfluencies can reveal insights into speech planning, demonstrate how speech production can break down, and provide information on the speech production system's constraints [4]. Disfluencies also provide valuable information to the listener, informing the listener that the speaker is experiencing planning difficulties, allowing turn taking, or communicating metalinguistic informa-

tion, such as recognition of a mistake [4]. Brennan and Williams [45] posit that typical disfluencies also reveal information about the speaker's confidence, or lack thereof, to the listener.

In the current study, participants answered an interview-themed prompt that required reflection and often a retelling of past events (e.g., "Tell me about a time you had too many things to do and had to prioritize. How did you organize your time?"). The task likely required planning and the interview format, though simulated, may have negatively impacted confidence. Findings revealed high numbers of superfluous verbal behaviors and pauses. Superfluous verbal behaviors serve important communicative purposes (2-way communication); however, the speaking task in the current study did not involve turn taking [4] given that there were no opportunities for the other speaker, in this case the interviewer, to provide a missing word [4]. It is likely that the superfluous verbal behaviors observed in the current study were expressions of hesitation, conveying metalinguistic information about the speakers' confidence and/or indicating the speakers' own recognition of having misspoken. Pauses may serve to stall speech during a hesitation [6]. Repetitions of single words or multiple words can also provide the speaker with more planning time [9]. Repetitions can also serve as a self-initiated repair, allowing the speaker to correct a misspoken word within the same turn [9]. The current study also found more fixed postures than prior work among monolinguals, which could similarly indicate that participants were experiencing planning difficulties or were communicating uncertainty or other metalinguistic information to the listener. However, our interview task is comparable to tasks used in prior research with monolinguals (e.g. monologs about work) [25]; these previously used tasks also required retelling of information or work-related events.

Limitations

Our results should be interpreted within the context of the study's limitations. First, 20 participants took part in the current study. Six speech samples (over 15 min of speech) from each participant were analyzed. With a larger number of participants or more speech samples from each participant, additional trends in the data could have reached statistical significance. Our study revealed trends that need to be confirmed by additional research. Further, our small sample size could have constrained our ability to detect key trends. A larger sample size may reveal more trends and result in statistical significance. Finally, future research with other samples, in terms of speakers and speech circumstances, could broaden our understanding. Results from the present study may not be comparable to results obtained from bilingual speakers engaging in a nonstressful speaking task.

It is possible that the duration of time during which participants were required to speak (2.5 min) in response to each question played a role in the results. However, the findings of Buchanan et al. [26] revealed significantly more disfluencies even when comparing the first 2 min of the respective tasks. It is possible that, if participants had engaged in a face-to-face simulated interview and/or had to speak longer during the simulated interview, the activity would have become more stressful and/or cognitively demanding and there would have been more or different patterns of disfluencies. While each response lasted only 2.5 min, the simulated interview activity did require more than 30 min in total (listening and speaking) from each participant.

A simulated interview, by nature, does not have the same stakes as a real job interview. Buchanan et al. [26] found increased disfluencies associated with stress during a simulated activity; participants engaged in a stressful 5-min simulated speech activity (defending themselves from an imaginary shoplifting accusation) and produced more disfluencies (pauses) than when performing a less stressful 5-min speech activity (describing a travel article). The simulation of Buchanan et al. [26] was face-to-face, however, while our speech activity involved a virtual meeting. Although we attempted to simulate a remote interview, our interlocutors were prerecorded. It was beyond the scope of the current study to have a team of actors for each participant to converse with in real time.

The current study compared novel data to previously published normative data in English speakers. While this method provides new insights, some variables remain uncontrolled. A future research study could build on the current findings by comparing typical disfluencies in both English and Spanish and compare findings to a group of monolingual English speakers matched for age, gender, and educational background.

Conclusions and Suggested Future Directions

The current study examined disfluencies produced during a common speaking situation that involved judgment of speech performance (a job interview) among a growing US population (Spanish-English bilinguals). Relative to monolinguals, little research has been conducted to understand the character of disfluencies among bilinguals. Foundational research among typically fluent heritage bilinguals is an essential first step. Our study also has implications for those conducting job interviews.

There is a potential for interviewers to have biases with regard to typically fluent bilingual speech patterns (i.e., slightly elevated rates of disfluencies compared to monolinguals) that do not reflect the bilinguals' ability as a potential employee, which in turn has implications for both equity and loss of valuable talent.

Findings from the current study should not be generalized to other populations. We presented data detailing disfluencies in English, which is both the stronger language and, sequentially, the L2. All of the participants reported that English is their stronger language and the dominant language for work and school. This is a common profile for adult heritage bilinguals in Anglophone settings such as the USA [34]. The English fluency and disfluency patterns observed in the present study may deviate from those seen among Spanish-English bilinguals for whom Spanish is stronger and/or more frequently used. Our findings are specific to the languages of our participants (Spanish and English), since bidirectional relationships exist between linguistic features in each of a bilingual's 2 languages. These disfluencies may be related to the fact that the speakers' first language was Spanish. Future research studies could compare our findings with disfluency patterns among Spanish-dominant Spanish-English bilinguals or other bilingual samples, as this could provide further insight into fluency and bilingual language development.

Acknowledgement

The authors would like to thank all of the participants who took part in the current study. The authors would also like to thank the research assistants who worked on collecting data the project, including Dalia Beccara, Katie Brewer, Jorge N. Garcia Jr., Kelly H. Koutnik, and Jane Lee.

Statement of Ethics

The participants provided written informed consent and the study protocol was approved by the California State University East Bay Institutional Committee on Human Research (01062016). All of the procedures were given ethical approval by the institutional research board.

Conflict of Interest Statement

The authors have no conflict of interests to declare.

Funding Sources

Funding for this study was received through a Faculty Support Grant for Collaborative Research from California State University East Bay to the first and third authors.

Data Availability Statement

All of the data generated or analyzed during this study are available upon request. Further enquiries can be directed to the corresponding author.

Author Contributions

All 3 authors made substantial contributions to the conception or design of this work. The first and third authors were responsible for acquisition of the data. The first and second authors were responsible for analysis and interpretation of data for this work. All 3 authors were responsible for drafting this work and revising it critically for important intellectual content, gave their final approval of the version to be published, and are accountable for all aspects of this work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1 Starkweather CW, Givens-Ackerman J. Stuttering. Austin: Pro Ed; 1996.
2 Schegloff EA, Jefferson G, Sacks H. The preference for self-correction in the organization of repair in conversation. Language. 1977;53(2):361–82.
3 Martin RR, Haroldson SK. Stuttering and speech naturalness: audio and audiovisual judgments. J Speech Hear Res. 1992 Jun;35(3):521–8.
4 Bortfeld H, Leon SD, Bloom JE, Schober MF, Brennan SE. Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. Lang Speech. 2001 Jun;44(Pt 2):123–47.
5 Brundage SB, Rowe H. Rates of typical disfluency in the conversational speech of 30-month-old Spanish–English simultaneous bilinguals. Am J Speech Lang Pathol. 2018 Oct;27 3S:1287–98.
6 Hieke AE. A content-processing view of hesitation phenomena. Lang Speech. 1981 Apr;24(2):147–60.
7 Lennon P. Investigating fluency in EFL: A quantitative approach. Lang Learn. 1990 Sep;40(3):387–417.
8 Poulisse N. Slips of the tongue: Speech errors in first and second production. Amsterdam: Benjamins; 1999.
9 Rieger CL. Repetitions as self-repair strategies in English and German conversations. J Pragmatics. 2003 Jan;35(1):47–69.
10 Green DW, Abutalebi J. Language control in bilinguals: the adaptive control hypothesis. J Cogn Psychol. 2013 Aug;25(5):515–30.
11 Bock K, Levelt WJ. Language production: Grammatical encoding. Cambridge: Academic Press; 1994.
12 Roche JM, Arnold HS. The effects of emotion suppression during language planning and production. J Speech Lang Hear Res. 2018 Aug;61(8):2076–83.
13 Byrd CT, Bedore LM, Ramos D. The disfluent speech of bilingual spanish-english children: considerations for differential diagnosis of stuttering. Lang Speech Hear Serv Sch. 2015 Jan;46(1):30–43.
14 Byrd CT. Assessing bilingual children: Are their disfluencies indicative of stuttering or the by-product of navigating two languages? Lepzig: Thieme; 2018. p. 324–32.

15 United States Census American Community Survey: Demographic and housing estimates – United States Census. 2018 [cited 2021 Jan 21]. Available from: https://data.census.gov/cedsci/table?d=ACS%205-Year%20Estimates%20Data%20Profiles&table=DP05&tid=ACSDP5Y2018.DP05.
16 Teesson K, Packman A, Onslow M. The Lidcombe behavioral data language of stuttering. 2003. Available from: https://doi.org/10.1044/1092-4388(2003/078).
17 Ambrose NG, Yairi E. Normative disfluency data for early childhood stuttering. J Speech Lang Hear Res. 1999 Aug;42(4):895–909.
18 Cordes AK, Ingham RJ. Stuttering includes both within-word and between-word disfluencies. J Speech Hear Res. 1995 Apr;38(2):382–6.
19 Fox Tree JE. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. J Mem Lang. 1995 Dec;34(6):709–38.
20 McDougall K, Duckworth M. Profiling fluency: an analysis of individual variation in disfluencies in adult males. Speech Commun. 2017 Dec;95:16–27.
21 Roberts PM, Meltzer A, Wilding J. Disfluencies in non-stuttering adults across sample lengths and topics. J Commun Disord. 2009 Nov-Dec;42(6):414–27.
22 Shriberg E. To 'errrr' is human: ecology and acoustics of speech disfluencies. J Int Phon Assoc. 2001 Jan;31(1):153–69.
23 Brown-Schmidt S, Tanenhaus MK. Real-time investigation of referential domains in unscripted conversation: a targeted language game approach. Cogn Sci. 2008 Jun;32(4):643–84.
24 Fox Tree JE. Listeners' uses of um and uh in speech comprehension. Mem Cognit. 2001 Mar;29(2):320–6.
25 Merlo S, Mansur LL. Descriptive discourse: topic familiarity and disfluencies. J Commun Disord. 2004 Nov-Dec;37(6):489–503.
26 Buchanan TW, Laures-Gore JS, Duff MC. Acute stress reduces speech fluency. Biol Psychol. 2014 Mar;97:60–6.
27 Enxhi SY, Hoon TB, Fung YM. Speech disfluencies and mispronunciations in English oral communication among Malaysian undergraduates. Int J Appl Linguist Engl Lit. 2012 Nov;1(7):19–32.
28 Liu H, Cao F. L1 and L2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. Brain Lang. 2016 Aug;159:60–73.
29 Choo AL, Smith SA. Bilingual children who stutter: Convergence, gaps and directions for research. J Fluency Disord. 2020 Mar;-63:105741.
30 Hayashi Y, Murphy VA. On the nature of morphological awareness in Japanese-English bilingual children: A cross-linguistic perspective. Biling Lang Cogn. 2013;16(1):49–67.
31 Luo YC, Chen X, Geva E. Concurrent and longitudinal cross-linguistic transfer of phonological awareness and morphological awareness in Chinese-English bilingual children. Writ Lang Lit. 2014 Jan;17(1):89–115.
32 Marinova-Todd SH, Siegel LS, Mazabel S. The association between morphological awareness and literacy in English language learners from different language backgrounds. Top Lang Disord. 2013 Jan;33(1):93–107.
33 Ramírez G, Chen X, Pasquarella A. Cross-linguistic transfer of morphological awareness in Spanish-speaking English language learners: the facilitating effect of cognate knowledge. Top Lang Disord. 2013 Jan;33(1):73–92.
34 Valdés G. Bilingualism, heritage language learners, and SLA research: opportunities lost or seized? Mod Lang J. 2005 Sep;89(3):410–26.
35 Montrul SA. Is the heritage language like a second language? EuroSLA Yearbook. Amsterdam: Benjamins; 2012. p. 1–29.
36 Hood SB. Effect of communicative stress on the frequency and form-types of disfluent behavior in adult stutterers. J Fluency Disord. 1974 Jan;1(3):36–47.
37 Muñoz-Sandoval AF, Cummins J, Alvarado CG, Ruef ML. Bilingual verbal ability tests. Rolling Meadows: Riverside Publishing Company; 1998.
38 Wechsler D, Coalson DL, Raiford SE. WAIS-IV: Wechsler adult intelligence scale. San Antonio: Pearson; 2008.
39 Smith SA, Seitz SR, Koutnik KH, Mckenna M, Garcia JN. The "work" of being a bilingual: exploring effects of forced language switching on language production and stress level in a real-world setting. Appl Psycholinguist. 2020 May;41(3):701–25.
40 Crookes G. The utterance and other basic units for second language discourse analysis. Appl Linguist. 1990;11(2):183–99.
41 MacWhinney B. The CHILDES Project: Tools for analyzing talk. transcription format and programs. London: Psychology Press; 2000.
42 Ratner NB, Brundage S. A clinician's complete guide to CLAN and PRAAT. 2018 [cited 2020 Apr 14]. Available from: https://talkbank.org/manuals/Clin-CLAN.pdf.
43 Bosshardt HG. Effects of concurrent cognitive processing on the fluency of word repetition: comparison between persons who do and do not stutter. J Fluency Disord. 2002;27(2):93–113.
44 IBM. IBM SPSS Statistics for Windows, version 24.0. Armonk: IBM; 2016.
45 Brennan SE, Williams M. The feeling of another's knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. J Mem Lang. 1995 Jun;34(3):383–98.