Perceptual Judgment of Stuttering Severity:

A Comparison Between Bilingual Children Who Do and Do Not Stutter

Selma Saad Merouwe[a, b], Raymond Bertram[b], Sami Richa[e], Kurt Eggers[b, c, d]

[a] Department of Speech-Language Pathology, Saint-Joseph University, Lebanon

[b] Department of Psychology and Speech-Language Pathology, Turku University, Finland

[c] Department of Rehabilitation Sciences, Ghent University, Belgium

[d] Department of Speech-Language Pathology, Thomas More University College, Belgium

[e] Faculty of Medicine, Saint-Joseph University, Lebanon

Author note

Selma Saad Merouwe https://orcid.org/0000-0003-2062-1827

Raymond Bertram https://orcid.org/0000-0002-2709-3644

Sami Richa https://orcid.org/0000-0002-3895-1325

Kurt Eggers https://orcid.org/0000-0003-4221-2063

Correspondence concerning this article should be addressed to Selma Saad Merouwe, Dept. of Speech-Language Pathology, Saint-Joseph University, Damascus Street, B.P. 17-5208 – Mar Mikhaël, Beirut 1104 2020, Lebanon. Tel: 009613197118. Email: selma.saad.merouwe@gmail.com

Number of Tables: 5

Number of Figures: 2

Word count: 6222

Keywords: bilingualism, stuttering, stuttering-like disfluencies, other disfluencies, stuttering severity ratings

Abstract

Introduction: Recent studies conducted with bilingual populations have shown that bilingual children who do not stutter (CWNS) are often less fluent than their monolingual counterparts, which seems to affect the accuracy with which speech-language pathologists (SLPs) identify stuttering in bilinguals. That is, misdiagnosis appears frequently in bilingual children, and is more likely to occur with bilingual CWNS (false positives) than bilingual CWS (false negatives).

Methods: The goal of the current study was to gain insight in the extent of this misdiagnosis. Speech samples of 6 Lebanese bilingual CWNS and 2 CWS were rated by Lebanese SLPs in an audio-only and audiovisual presentation mode. SLPs had to identify each child as stuttering or not and subsequently rate on a 6-point scale the stuttering severity for each child. SLPs also provided background information by means of a questionnaire.

Results*:* The results showed that stuttering severity ratings (1) were on average significantly higher for CWS than for CWNS, (2) were for each CWS higher than for all but one of the CWNS, (3) varied significantly among the CWNS but not the CWS, (4) were not affected by the presentation mode, and (5) correlated positively with the percentages of stuttering-like disfluencies (SLD) and the mean number of iterations, but not with the percentages of other disfluencies (OD).

Conclusion: Misdiagnosed bilingual CWNS are perceived by the SLPs as having a mild stutter, primarily based on the frequency of their disfluencies, but can be occasionally rated at par with CWS. Further research differentiating the disfluent speech of bilingual children who do and do not stutter is needed to reach a more adequate diagnosis of stuttering.

Perceptual Judgment of Stuttering Severity:

A Comparison Between Bilinguals Children Who Do and Do Not Stutter

1. Introduction

This study investigates the perceptual judgments of speech-language pathologists (SLPs) regarding the stuttering severity of bilingual children who stutter (CWS) and children who do not stutter (CWNS). It is currently estimated that nearly 50% of the world's population is bilingual [1], and it is known that stuttering exists in all cultures and languages (see [2] for review on stuttering in several languages). Consequently, SLPs are likely to examine bilingual CWS, which implies that in their clinical decision-making process, they have to identify features unique to speech-fluency disorders. In other words, SLPs should decide whether disfluencies produced by bilinguals are indicative of stuttering or related to speaking a second language. Given the shortage of empirically based data on speech disfluencies produced by bilingual CWS and CWNS, this distinction is difficult to make. Several studies indeed indicate that the identification of stuttering in bilingual children remains challenging [3], [4], [5], [6].

There are nevertheless a number of recent studies that addressed this challenge. Byrd et al. explored the frequency and type of speech disfluencies produced by 18 bilingual Spanish-English CWNS aged between 5.06 and 6.07 years [7]. Narrative speech samples were collected in both languages. The majority of participants (78%) produced a very high frequency of stuttering-like disfluencies (SLD) in all their speech samples, ranging from 3% to 22%. Monosyllabic word and sound repetitions were the most detected SLD, and the mean number of iterations produced was 5 for sound and syllable repetitions, and 6 for monosyllabic words repetitions. Given the significance of these findings, Byrd et al.'s pilot study was replicated and expanded to other populations with different linguistic dyads. Eggers et al. investigated the

frequency and type of speech disfluencies produced by 59 bilingual Yiddish-Dutch CWNS, aged between 6.01-7.07 and 9.00-10.04 years [8]. The results showed that 46% of the participants produced more than 3% of SLD in their native Yiddish (L1). In Dutch, their second language (L2), 78% of the bilinguals produced more than 3% of SLD. Again, a high prevalence of monosyllabic word and sound repetitions was identified. However, different from Byrd et al.'s findings [7], the mean number of iterations was 1.18 for sound repetitions, 1.07 for syllable repetitions and 1.30 for monosyllabic word repetitions. In general, it has been noted that such disfluencies of bilingual CWNS are more likely to happen when they try to produce utterances at the edge of their linguistic competence, which is reached more easily than in monolinguals as bilinguals have to navigate more than one language system [7], [9], [10].

These findings on bilingual CWNS are of great clinical relevance for the differential diagnosis of stuttering in bilingual children, as they imply that speakers of more than one language produce speech disfluencies at high rates. Ambrose and Yairi classified disfluencies as SLD (including monosyllabic word repetitions, part-word repetitions, dysrhythmic phonation) and Other Disfluencies-OD (including multisyllable words repetitions, phrase repetitions, interjections, revisions and abandoned utterances). They indicated that speech with 3% SLD or more is suggestive of a stuttering disorder [11]. They also showed that CWS exhibit an average of 2 iterations (or higher). The 3% clinical threshold was established with monolinguals and is, until now, widely used in the clinical decision-making process [12]. However, as it is clear from the studies conducted by Byrd et al. [7] and Eggers et al. [8], bilingual CWNS of the explored language dyad frequently exceed this clinical threshold, with a high proportion of monosyllabic word and sound repetitions. Therefore, it appears that the 3% SLD guideline is inappropriate for bilingual children and may lead to extensive false-positive identification of stuttering.

One of the few studies that investigated the risk of misdiagnosis in stuttering assessments is the one conducted by Byrd et al. [13]. In their study, fourteen Spanish-English SLPs listened to audio-recordings of narrative re-tells in Spanish and English, produced by two bilingual children (one confirmed as a child who stutters and another as a child who does not stutter). The task was comparable to a screening test, in which diagnostic decisions are made based on the frequency and typology of speech disfluencies alone. Both children were matched for age, gender, language proficiency and frequency of disfluencies exhibited (17%). The majority of SLPs (71.4%) correctly identified the stuttering child, however, 85.7% of the SLPs falsely judged the bilingual child who does not stutter to be a child who stutters. No association was found between the identification accuracy and factors related to the SLP's background (e.g., years of experience, client caseload, educational and professional background), although this was not statistically tested given the limited number of cases and SLP participants. The false-positive identification of stuttering was -qualitatively- attributed to the high frequency of SLD (more precisely, sound, syllable and monosyllabic word repetitions) observed in the speech of the bilingual child who does not stutter. This study was replicated and expanded by Saad Merouwe et al. (unpublished data) in Lebanon, a country known for its multilingual diversity [14], [15]. The goals were to investigate whether Lebanese SLPs were accurate at identifying stuttering in bilingual children, to examine whether presentation mode affected the identification, and to explore factors affecting the SLPs' judgments. Thirty-two SLPs listened to narrative samples in Lebanese Arabic of 6 bilingual CWNS and 2 bilingual CWS, and categorized each child as having a stutter or not. In order to better understand the characteristics driving misdiagnosis in bilinguals, compared to Byrd et al. [13], the current study included children with a wider diversity in terms of bilingual status and frequency and types of disfluencies exhibited, as well as

a larger number of SLP participants. The judgments were based on audio-only recordings and subsequently on audio-visual recordings. The results showed that misdiagnosis happens frequently in bilinguals, however, it is significantly more likely to happen with bilingual CWNS (44.8% false positives) than with CWS (12.5% false negatives). The misdiagnosis rate also varied within these categories, and appeared to be driven by overlooking or misjudging physical concomitants. The use of audio-visual recordings instead of audio-only recordings did not improve SLPs' performance. Lastly, speech samples' characteristics – especially the frequency of SLD – rather than SLPs' characteristics correlated with the success rate of judgments.

The current study, which is a follow-up study of Saad Merouwe et al. (unpublished data), reports the stuttering severity ratings of CWNS and CWS rather than a binary choice (CWNS versus CWS). Using the ratings, we wanted to get further insight into the nature and the extent of false identification of CWNS. The only study that has dealt with this issue is the one conducted by Byrd et al. [13]. The SLPs ($N = 14$) were required to rate the stuttering severity of two bilingual children, one who stutters and one who does not stutter, on a 6-point scale ranging from 'no stutter' to 'severe stutter' on the basis of audio-only recorded speech samples. The differences between judging the child who stutters and the child who does not stutter were minimal. Both were rated as having a 'moderate stutter' by 43% of the SLPs. The child who stutters was even more often judged to have no stutter (29%) than the child who does not stutter (14%). The average stuttering severity ratings were approximately the same (child who does not stutter: 2.8 versus child who stutters: 3.1) and did not differ significantly from each other ($p = .61$). These results underline the difficulty in distinguishing stuttering from typical disfluencies in bilingual children. The findings are however quite exploratory given the limited number of SLP participants ($N = 14$) and the limited number of rated children ($N = 2$). Moreover, the child who

does not stutter in the Byrd et al. study had a very high proportion of total disfluencies (17%), which is likely to be beyond the typical amount of disfluencies in bilinguals [8], even though bilinguals often cross the monolingual threshold of 3%. The current study has therefore opted to select a broader range of CWNS with a total amount of disfluencies ranging from 5% to 15% and also have variation within the CWS-category by selecting a child who stutters with relatively many disfluencies (27%) and one with clearly less disfluencies (12%). We expected that the majority of the SLPs would assign – unlike what was reported in Byrd et al. [13] – higher stuttering severity ratings for CWS than for CWNS, but we wanted to explore whether this difference would be found across all children or whether it would pertain to certain children only.

In addition, we wanted to investigate to what extent stuttering severity ratings depend on presentation mode. During the past decades, a few researchers have shown interest in examining listener perceptions of speech fluency and stuttering. Some studies have focused on the identification of stuttering [16], [17], [18], [19], [20]. Others have examined the effect of stimulus presentation mode on listeners' perceptions of speakers' fluency [21], [22], [23], [24], [25]. While some studies showed that listeners did not identify a greater number of stuttering events from audio-visual versus audio-only speech samples [24], others showed a more accurate identification of stuttering moments via audio-visual recordings [21]. These equivocal findings could be related to methodological disparities (e.g., the participants' background, the stuttering severity level of the experimental samples, the presence/absence of visible or audible physical concomitants). Hence, we explored the role of presentation mode for assessment of children with a relatively broad range of SLD. From a clinical perspective, we expected that audio-only

presentation mode would yield different results than audio-visual presentation mode, as the latter allows for a better analysis of physical concomitants.

The findings of abovementioned studies by Byrd et al. [13] and Saad Merouwe et al. (unpublished data) indicate that SLPs are not accurate at correctly identifying bilingual CWNS as such. However, despite the higher percentage of false diagnosis for bilingual CWNS compared to bilingual CWS, it is still inconclusive whether all bilingual CWNS are perceived to have a less severe stutter than CWS, and whether SLPs are as sensitive to the stuttering severity in audio-only versus audio-visual presentation modes. The purpose of the current study is thus to investigate to what extent stuttering severity ratings assigned by SLPs are affected by children being CWS or CWNS, and by presentation mode. More specifically, the following research questions were addressed:

(1) Do the stuttering severity ratings differ as a function of child category (CWNS versus CWS) and presentation mode (audio-only versus audio-visual recordings)?
(2) Do the stuttering severity ratings vary within CWNS and CWS?
(3) Are the stuttering severity ratings of each child who stutters higher than those of each child who does not stutter?
(4) Do number of SLD, OD, and iterations correlate with stuttering severity ratings?

2. Materials and Methods

2.1. Participants

Thirty-two Lebanese licensed SLPs participated in this study of which 96.8% were aged between 21 and 30, and 3.2% between 31 and 40. All SLPs were bilingual, spoke Lebanese Arabic as a dominant language, and either French or/and English as a non-dominant language, according to the self-rating they reported (24 participants spoke Lebanese Arabic, French and

English, 6 participants spoke Lebanese Arabic and English, and 3 participants spoke Lebanese Arabic and French). They were recruited from 3 universities that have SLP departments in Lebanon, after obtaining the agreement of the committee of ethics of Saint-Joseph University of Beirut. An informed consent and a cover letter were emailed to the SLP graduates. They were asked to reply to the email specifying whether they wanted to participate in the study or not. Thirty-two SLPs participated in the first phase of the study, involving the completion of a survey and the judgment of 8 speech samples via audio-only presentation mode. Twenty-seven out of these 32 SLPs also participated in the second phase of the study four months later, where the same 8 speech samples were judged in the audio-visual presentation mode.

Background information about the SLPs was obtained via a Google form questionnaire. It was developed in French and English as both languages are used academically in the Lebanese educational system. The SLPs were asked to select the language of the survey according to their preference and educational background. The SLPs' characteristics obtained from the survey analyses were relative to 'work experience', 'number of bilingual CWS treated', 'self-ratings of confidence in diagnosing stuttering in a bilingual context', and 'self-ratings of difficulties in distinguishing SLD and OD'. More details about the participants are available in Table 1.

2.2. Materials

In phase 1, the participants ($N = 32$) judged audio-only recorded speech samples produced in Lebanese Arabic by 8 bilingual Lebanese children (2 with a confirmed diagnosis of stuttering and 6 confirmed as being typically developing) as being stuttered speech or not. The task was quite similar to a screening test, often used by SLPs working in schools to identify children with risk for stuttering and refer them for full assessment as early as possible [26]. In

phase 2, the participants ($N = 27$) judged the same speech samples available via audio-visual recordings.

2.2.1. Stimulus materials

In phase 1, the stimulus materials were audio-only recordings of the narrative productions in Lebanese Arabic of 2 bilingual Lebanese CWS, and 6 bilingual Lebanese CWNS matched for age, language dominance (Lebanese Arabic), and elicitation mode (narrative based on the picture book 'Frog goes to dinner' [27]). The speech samples were videotaped and audio-recorded. They were selected from a larger set of data, which was collected to analyze the speech disfluencies of bilingual Lebanese CWS and CWNS with different linguistic profiles, in order to investigate the role of bilingualism in exhibiting speech disfluencies (Saad Merouwe et al., unpublished data).

The CWS were recruited via emails sent to all Lebanese SLPs working in Lebanon. The inclusion criteria were: (a) speaking two languages of which one was Lebanese Arabic; (b) diagnosed with stuttering by their treating SLP based on a comprehensive full assessment; (c) regarded by parents as having a stuttering problem; (d) identified independently as a CWS by the first author and another Lebanese SLP (both certified multilingual SLPs, fluent in Lebanese Arabic, and both fluency specialists with extensive experience in treating fluency disorders in bilingual children); (e) age-appropriate speech-language skills based on Parents of Bilingual Children Questionnaire (PaBiQ: see description below) and the first author's evaluation on the basis of observation; and (f) no reported neurological or learning disorders based on the interview conducted by the first author with the parents.

The diagnosis of stuttering of both CWS was established independently by the first author and a second Lebanese SLP via audio-visual recording analyses of 4 speech samples per child (2 per spoken language). Both children had English as a second language. The first author

also conducted an interview with the parents over the phone to gather information about the child's stuttering, including the case history, the onset of stuttering, its development, the child's reactionary attitudes, parents' concerns and the treatment provided. Moreover, the first author used the Stuttering Severity Instrument-4th edition (SSI-4) [28] to assess the speech samples that were made available to the SLP participants. CWS1 obtained a total score of 26 (percentile rank: 61-77; severity equivalent: moderate) and CWS2 obtained a total score of 21 (percentile rank: 41-60; severity equivalent: moderate).

The CWNS were recruited via an open call sent to schools all over the country. After getting the principals' permission, letters were sent to parents to also obtain their consent. Children had to fulfill the following inclusion criteria to be considered as CWNS: (a) speaking 2 languages of which one was Lebanese Arabic; (b) no parental or teacher concern regarding stuttering; (c) no family history of stuttering; (d) no history of speech-fluency intervention; (e) no physical concomitants and normal rhythmicity of iterations (f) age-appropriate speech-language skills based on PaBiQ, and the teacher's and the first author's evaluation on the basis of observation; (g) no parental or teacher concern regarding learning abilities.

The first author and the other SLP (fluency experts with high proficiency in all languages[1]) evaluated independently all speech samples (all languages) of each child, and judged the speech as characteristic of typically developing and typically fluent children (tension-free with regular iterations) based on the analysis of 4 speech samples (a spontaneous conversation and a narrative sample collected in both dominant and non-dominant languages). Among the 6 CWNS included

[1] The first author and the other SLP are Lebanese fluency specialists, native speakers of Lebanese Arabic, exposed to both French and English languages since preschool. They use the three languages frequently and equitably in their daily life for different purposes (communication, work etc.).

in this study, two had French as a second language, two had English as a second language, and two were exposed to both English and French in addition to Lebanese Arabic.

The PaBiQ [29] is a standardized parental questionnaire that was used to define the language profile of each child. The questionnaire asks parents to provide information about language input and output in different settings. Accordingly, a language disorder risk (No risk index), a language proficiency score (in each language), and a language dominance index are calculated. The No risk index covers factors that indicate whether a child has a language disorder or not. A No risk index of 19-23 is taken as an indication of typical language development. Language proficiency is assessed by asking parents to rate the current language skills of their child in each spoken language. The 'languages used within the family and elsewhere' scores make it possible to compare the child's languages with each other in terms of frequency of use. This information on current language use is combined with information on the quantity and quality of early exposure to reach an estimation of the language dominance. The maximal score is 50 (for a more detailed description, see [29]). This questionnaire was validated for the Lebanese population, and additional analysis indicated that the No risk index highly correlated with the child's current language skills [30]. The No risk index and the language dominance index are presented in Table 2.

In the current study, 6 bilingual CWNS and 2 CWS were included with a variety of profiles in terms of frequency and type of disfluencies. To make the samples' analyses doable in terms of time needed to complete the task, we included only one full speech sample (a narrative elicited via the picture book 'Frog goes to dinner') per child in Lebanese-Arabic, which was the dominant language of 7 out of 8 children. One child (CWNS5) was a balanced bilingual with approximately equal dominance scores in Lebanese Arabic and English. As specified previously,

Lebanese Arabic was the dominant language of all SLP participants, which is important, as it was found that SLPs are better at identifying people who do and do not stutter in their native language than in a less well-known or foreign language [31], [32], [33].

The narrative speech samples were alike in number of words (M_{CWNS} = 291.67; *SD* = 120.99; M_{CWS} = 394; *SD* = 22.6), $t(6) = -1.13$; $p = .301$, and number of syllables (M_{CWNS} = 448; *SD* = 181.9; M_{CWS} = 664; *SD* = 38), $t(6) = -1.5$; $p = .164$. The total sample duration (in seconds) was significantly longer for the CWS (M_{CWNS} = 266.17; *SD* = 82.76; M_{CWS} = 444.5; *SD* = 72.83), $t(6) = -2.69$; $p = .036$, mainly due to their higher frequency of disfluencies. We calculated the percentage of disfluencies on the basis of the whole speech sample in order to obtain the most accurate disfluency estimations possible for each child. The full speech samples were included as they provide a better ecological validity [34]. Each child's speech sample consisted of the narrative of the whole picture book.

3.2.2 Speech samples

The interviewer presented the wordless book "Frog goes to dinner" [27] to each of the 8 children explaining that the story is about a boy, a frog and a dog. First each child was given time to look at each picture to understand the storyline. Subsequently, the child was asked to tell the story while going through the pictures one by one once more. The narrative samples of the 8 children were recorded and videotaped. On the basis of these recordings, the samples were transcribed, after which different types of speech disfluencies were coded into the transcriptions.

3.2.3 Transcription and coding

The first author and 3 trained SLP students transcribed and coded the speech samples. To classify the disfluencies, a similar system as that of Byrd et al. [7] was applied. Disfluencies were, cf. [11], categorized as SLD, including monosyllabic word repetitions, syllable repetitions,

sound repetitions, blocks, broken words and prolongations, and OD, including multisyllable word repetitions, phrase repetitions, interjections, unfinished words/sentences and revisions. The revisions were categorized as lexical, grammatical and phonological [35].

Before the transcription and coding, the first author trained the SLP students in both tasks and 3 speech samples, not included in the current dataset, were transcribed independently. Afterwards, transcriptions were compared, and differences among the coders were resolved through examination. The inter-judge reliability, calculated on the basis of the agreement index percentage for 10% of the dataset (point-by-point for location and type cf. [11] and [36] was 0.87. Moreover, the first author checked all the transcribed samples used specifically for the current study to ensure accuracy.

Mean SLD and OD percentages were calculated based on words as in most previous studies [37], [7], [13], [8], [38], [39], [40]. Table 3 shows the percentages for each type of disfluency in the sample, the mean number of iterations, the time it took each child to complete the narrative and the numbers of words and syllables that were produced.

3.3 Procedures

3.3.1 Survey and sample presentations

During phase 1, SLP participants were directed to a secure online password-protected platform where they completed a survey about their background information and rated 8 audio-only speech samples. They were informed that the task (survey and sample rating) would last 50 minutes and were asked to check the sound system of their laptop and to use their earphones. The audio recordings were presented in one go (stopping or repeating was not possible) and were not downloadable. After listening to each audio recording, the SLPs were instructed to rate on a 6-

point scale the stuttering severity for each child (0 = no stutter; 1 = very mild stutter; 2 = mild stutter; 3 = moderate stutter; 4 = severe stutter; 5 = very severe stutter).

During phase 2, which took place 4 months after phase 1, SLPs were directed to the same secure password-protected platform where they rated the same speech samples, this time available via audio-visual recordings, in a similar procedure as followed in phase 1. This task took about 40 minutes.

3.4 Statistical analyses

Overall analyses were conducted by an ANOVA with child category (CWNS versus CWS) and presentation mode (audio-only versus audio-visual recordings) as within (repeated) variables. Another ANOVA was performed with child (independent of child category) and presentation mode as within variables, followed by post-hoc pairwise comparisons to assess differences in severity ratings for individual children. As the data were not distributed normally, non-parametric analyses were also performed. The same effects were found with these analyses as with the ANOVAs, so only the latter were reported. Lastly, correlations between the SLP's stuttering severity ratings and samples' characteristics, i.e., % SLD, %OD and number of iterations, were explored by means of Pearson's correlation-analyses. These correlations were based on the performance in the audio-only part of the experiment, as here more participants (N = 32) were included than in the audio-visual part (N = 27). SPSS software (Version 27) was used for all analyses and multiple comparisons were performed using Bonferroni corrections.

4. Results

4.1 Stuttering severity ratings for CWS and CWNS in both presentation modes

The mean stuttering severity rating was significantly lower for CWNS (audio: 0.99; video: 0.96) compared to CWS (audio: 2.18; video: 2.24), $F(1, 26) = 102.16$, $p < .001$, $n_p^2 = .79$

(see Table 4). There was no effect for presentation mode ($F(1, 26) = 0.008$, $p = .92$, $n_p^2 = .000$). Also, the interaction between child category and presentation mode failed to reach significance ($F(1, 26) = 0.16$, $p = .69$, $n_p^2 = .006$). The effects are illustrated in Figure 1.

4.2 Differences in stuttering severity ratings between individual children

A significant effect of child (including CWNS and CWS together) was found, indicating substantial differences in ratings for individual children ($F(7, 20) = 26.82$, $p < .001$). There was no effect of presentation mode ($F(1, 26) = .001$, $p = .976$, $n_p^2 = .002$). Moreover, the interaction between child and presentation mode was not significant ($F(7, 20) = 1.2$, $p = .310$, $n_p^2 = .006$). Average ratings were the highest for CWS1 (audio: 2.44, video: 2.44) and the lowest for CWNS4 (audio: 0.16; video: 0.11), with the ratings for all other children falling in between.

For both presentation modes, post-hoc analyses revealed significant differences between both CWS and all CWNS except CWNS5 (all ps $\leq .05$). CWNS5 (audio: 2.06; video: 1.89) obtained the highest stuttering severity rating compared to other CWNS and appeared to be significantly different from CWNS2, CWNS3, CWNS4 and CWNS6 (all ps $< .05$). Moreover, CWNS4 who obtained the lowest stuttering severity rating was significantly different from all other CWNS and CWS (all ps $< .05$). For the two CWS, average ratings were higher for CWS1 (audio: 2.44, video: 2.44) compared to CWS2 (audio: 2.06, video: 2.04), however they were not statistically different ($p = 1$). More detailed information is presented in Figure 2 and Table 5.

4.3 Factors related to the stuttering severity ratings

For CWNS, at first glance, the correlations between the stuttering severity ratings and the %SLD and %OD were not significant (respectively: $r(4) = .26$; $p = .61$ and $r(4) = -.05$; $p = .91$). However, a positive strong correlation was found between the number of iterations and perceived stuttering severity ($r(4) = .95$; $p = .002$). These findings may be driven by one outlier, CWNS5,

who had relatively little SLD and OD, but the highest mean number of iterations (1.19) compared to other CWNS. This child was attributed the highest stuttering severity rating among the CWNS category (M = 2.06). After removing this outlier from the CWNS-group for further exploration, a significant positive correlation was found between the stuttering severity ratings and %SLD ($r(3) = .96$; $p = .008$), but not with OD ($r(3) = .79$; $p = .1$). Note that these correlations are exploratory, as there are only a few data points. For CWS, we abstain from correlational analyses as there are only 2 data points. As reported above, CWS1, who had the highest number of SLD and OD, and the highest mean number of iterations, was assigned a higher stuttering severity rating (M = 2.44) than CWS2 (M = 2.06), but this difference was far from statistical significance.

5. Discussion

The current study is a follow-up study conducted on the same dataset as in Saad Merouwe et al. (unpublished data), which showed that misdiagnosis is significantly more likely to occur with CWNS compared to CWS. The misdiagnosis rate also varied within both child-categories, and showed to be driven by speech samples' characteristics (e.g., %SLD) and by misinterpreting physical concomitants. The main findings of the current study exploring stuttering severity ratings assigned by Lebanese SLPs to bilingual CWS and CWNS in audio-only and audio-visual presentation modes showed that (1) stuttering severity ratings were on average significantly higher for CWS than for CWNS, (2) there was no impact of presentation mode in both categories, (3) stuttering severity ratings varied significantly within CWNS but not within CWS, (4) stuttering severity ratings were higher for the CWS than for most CWNS; one child who does not stutter was actually rated similar to both CWS, and (5) the percentages of SLD and the mean number of iterations correlated to the stuttering severity ratings.

It is intriguing that the severity ratings did not differ across presentation modes, as one would expect that audio-only judgment of disfluencies may inflate or reduce severity ratings, given that observers cannot assess the presence/absence of physical concomitants. However, earlier studies conducted with adults [22], [23], [25], and children [41] also found no differences between both presentation modes for the ratings of overall severity. Taken together, it appears that the presentation mode does not affect the perception of overall stuttering severity. One explanation might be that in the studies with both presentation modes, an overall judgment of stuttering severity has been required rather than an indication whether individual speech events are stuttered or not. With respect to the current study, SLPs had to watch the speech samples in one go, without pausing or rewinding, which is unlike common daily practice. Also, this may have caused that some SLPs missed the presence/absence of physical concomitants, which might have affected their severity ratings for both CWNS and CWS.

The average stuttering severity ratings assigned by the SLPs were found to be significantly lower for CWNS compared to those assigned to CWS. In the only directly comparable study, Byrd et al. [13] found no difference between the bilingual child who does not stutter and the one who stutters. Both were assessed equally as having a “moderate stutter” by 43% of the SLPs and were on average rated in the same way. Byrd et al. concluded that the frequency of disfluencies – often used as a guiding principle when assessing stuttering in monolinguals - could result in bilinguals being not only misdiagnosed, but could also impact the stuttering severity rating. However, Byrd et al. [13] included only one bilingual child who does not stutter and one who does stutter with both having a high percentage of total disfluencies (17% in both languages). The discrepancy between our findings and Byrd et al.’s is probably related to the number of SLP participants, cases included, and variety of profiles in terms of

speech fluency and language proficiency (see Tables 2 and 3 for details). Most importantly, all of our CWNS had a lower percentage of total disfluencies than the child who does not stutter in the Byrd et al. study, even though all but one of our CWNS exceeded the 3% SLD threshold. In general, our findings suggest that – perhaps not surprisingly - bilingual CWNS are typically perceived as less disfluent than CWS, which resonates with the studies conducted by Susca and Healey [42] and Panico et al. [23] who also found that raters are sensitive to severity of stuttering, that is, as the frequency of stuttering events increases, listeners tend to make more negative judgments about the speech samples. It is important to recognize that the bilingual CWNS in our study were mostly perceived as having a "mild stutter".

At the same time, it is important to note that the stuttering severity ratings assessing individual children revealed clear differences among CWNS children. Average ratings were the lowest for CWNS4, significantly lower than all other CWNS and CWS, which is in line with our earlier study showing that CWNS4 was also the least frequently misdiagnosed as a child who stutters (Saad Merouwe et al., unpublished data). CWNS4 also had the lowest percentage of disfluencies, which is in line with studies showing higher severity ratings as the frequency of disfluencies increases [42], [23]. The average ratings were the highest for CWNS5, which were significantly different from those for CWNS2, CWNS3, CWNS4 and CWNS6. Most importantly, the ratings for CWNS5 did not differ from the ratings assigned to the two CWS. Note that this child was also the child who does not stutter that was most often (by 70% of the SLPs) qualified as having a stutter (Saad Merouwe et al., unpublished data). In other words, this result coincides with the findings of Byrd et al. [13], showing frequent misjudgment of the CWNS and no difference in stuttering severity ratings between the CWS and CWNS. Surprisingly however, in our case, the child who does not stutter who was rated at par with the

CWS was the child with the second lowest percentage of total disfluencies (TD: 6.35%; SLD: 3.47%), unlike the child who does not stutter in Byrd et al. (TD: around 17%).

One may wonder how a child who does not stutter with a relatively low percentage of SLD and TD is misjudged so frequently. This could be related to the number of iterations as that child obtained the highest number among the CWNS-group (1.19) and the second highest number of iterations among the whole group of CWNS and CWS. We also suspect that this frequent misjudgment might be linked to her bilingual profile. One characteristic that differentiated CWNS5 from the other CWNS was her language dominance score. She had a considerably lower language dominance score in Lebanese Arabic. In fact, this score was comparable to her English dominance score (see Table 2). In addition, she was exposed to French at school. In other words, she can be considered as a trilingual, with balanced competency of Lebanese Arabic and English. This may have led to somewhat unusual linguistic characteristics (different lexical choices due to less extensive lexicon, shorter sentences), and this may have recalibrated the evaluation of her speech (mostly monosyllabic word repetitions), which may have led to a higher percentage of misdiagnosis, and a higher disfluency severity rating. As there is only one such case in this sample, this result needs additional examination, but this case does hint that linguistic profile may influence the overall perception of stuttering severity in bilingual CWNS. For the CWS-pair, average ratings were numerically higher for CWS1 than for CWS2 in line with the higher percentage of disfluencies (CWS1TD: 27.07%, SLD: 17.31; CWS2 TD: 12.16%, SLD: 6.88%). This difference was not significant though, so no-far fetching conclusions should be made. Our earlier study did show a general difference in misjudgments, with CWS2 more often being qualified as a CWNS than CWS1. This finding is expected because previous studies have shown that, as the frequency of SLD rises, assessments

of stuttering become more reliable [23]. In sum, our results are partly similar to Byrd et al. [13] as our study also includes a child who does not stutter that elicited the same ratings as the CWS, but mostly dissimilar, as most CWNS elicit stuttering severity ratings way below the CWS. Accordingly, our findings indicate that SLPs are in general perceptive of the stuttering severity among bilingual children.

The last finding of the current study is that the mean number of iterations affects the stuttering severity ratings. This implies that, in our study, a higher number of iterations led to a stronger stuttering severity perception which resonates with the findings of Ambrose and Yairi [11] showing that CWS exhibit a higher mean number of iterations compared to CWNS. With regards to the frequency of disfluencies, no correlations were found between the stuttering severity ratings and the percentages of SLD and OD when we considered the whole CWNS-group. When CWNS5 (outlier described above) was removed from the CWNS-group for further investigation, the stuttering severity ratings appeared to be positively correlated to the percentages of SLD. This finding suggests, despite the small sample, that there is a relation between the frequency of SLD and the stuttering severity ratings in general, that is, an increased percentage of SLD tend to lead to a higher stuttering severity judgment, which is in line with earlier studies that showed positive correlations between the percentage of stuttered syllables and the stuttering severity ratings attributed [42], [43], [44], [45].

6. Limitations

Although Van Borsel et al. [33] highlighted the importance of assessing speech samples in both languages in order to identify stuttering in bilingual children, in the current study the SLPs were only asked to judge one Lebanese Arabic speech sample per child, which also was the dominant language of all SLPs. This decision was based on the following considerations. We

aimed at including a larger number of children (compared to Byrd et al. [13]), with varied profiles, in order to better understand which combination of characteristics lead to a higher risk of misdiagnosis and a deviant severity perception of stuttering. We also wanted to recruit more SLP participants to ensure more certainty about the findings. Finally, we wanted to examine whether audio-visual recordings allow for better perception and judgment of speech fluency in bilinguals, given the consideration that the identification of physical concomitants may be essential to diagnose a bilingual child as a CWS [13], [46]. As the time needed to accomplish the different tasks was lengthy (90 minutes), we opted for the dominant-language only to perform the ratings upon. This decision is motivated by earlier studies showing that one can be more precise at diagnosing a child in his dominant language [46]. It would be important, however, in the future studies, to explore whether SLPs make similar stuttering severity judgments when they are provided with speech samples in the dominant and non-dominant languages.

Another limitation to the current study is the young age of the SLP participants. The recruitment process was difficult as the participation required a 90-minutes commitment overall. In general, more experienced SLPs declined participating in the study due to the demanding schedule. Recruiting more experienced SLPs should be considered for future studies.

Finally, despite the increase in the sample size compared to Byrd et al.'s study [13], the number of speech samples included in the current study remains small (n = 8). Although our findings are clinically relevant and contribute to the diagnostic criteria of stuttering in bilinguals, they should be taken with caution.

7. Conclusion

Recent studies conducted in bilingual populations have shown that establishing whether a child is stuttering or not is a difficult task, especially for typically developing CWNS who appear

to be more frequently misdiagnosed than CWS [13]. However, it was not established whether bilingual CWNS (who are highly -typically- disfluent) are perceived as 'less severely' disfluent than bilingual CWS, and whether SLPs are sensitive to severity of stuttering in both audio-only and audio-visual presentation modes. The current study shows that, although SLPs are not completely accurate at identifying bilingual CWNS as such [13], they do assign significantly higher stuttering severity ratings for bilingual CWS than bilingual CWNS in most cases. The severity ratings also correlate with the number of iterations and percentages of SLD, that is, bilingual CWNS with few SLD and little iterations are typically assigned a lower stuttering severity rating compared to those with a high number of iterations and frequency of SLD in the dominant language. However, our study also indicates that, in the case of balanced bilingualism, even a relatively low number of SLD may give a stuttering-like impression. The exact nature of this phenomenon should be further investigated, but it may well be that SLPs analyze some speech behaviors differently when paired with a lower level of linguistic competence. To avoid misdiagnosis for screening of bilingual stuttering, it thus seems important to keep an eye on the exact balance of bilingual proficiency as well as to realize that a high disfluency rate is not necessarily indicative of stuttering. Moreover, clinicians should examine thoroughly the types of SLD and the presence/absence of tension and/or arrhythmicity, as Rincon et al. established while comparing bilingual CWS and bilingual CWNS that the latter do not exhibit dysrhythmic phonation and physical concomitants [47].

Acknowledgements

The present study was supported by the University of Turku and Saint-Joseph University of Beirut. The authors would like to thank the SLPs, children, and parents who participated in the study. They would also like to thank the students Nadia Sabeh'Ayon, Reem Abi Akar, Laudy

Arida, Ghena Antoun, Christel Daaboul, Dana Barakat, Fatmé Khalife, Anna-Maria Barakat and Céline Bachalaany for their assistance in collecting and transcribing the data.

Statement of Ethics

This study protocol was reviewed and approved by the committee of ethics of Saint-Joseph University of Beirut, approval number 1272.

Written informed consent was obtained from participants. Ethical considerations for the protection of participants were considered throughout the whole process.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Author Contributions

All authors have made significant contributions to this manuscript. They have consented to the byline order, have agreed to submit it in its current form, and have no conflicts of interest.

Data availability statement

The data that support the findings of this study are not publicly available due to their containing information that could compromise the privacy of research participants but are available the corresponding author [SSM].

References

1 Bloodstein O, Bernstein Ratner N, & Brundage S. A handbook on stuttering.7th ed, Plural Publishing; 2012.

2 Saad Merouwe S, Bertram R, Richa S, Eggers K. Bilinguisme et bégaiement: Revue de la littérature et implications sur le diagnostic clinique. Glossa. 2022; 132: 67-91.

3 Byrd C. Assessing bilingual children: Are their disfluencies indicative of stuttering or the by-product of navigating two languages? Semin Speech Lang. 2018; 39(4): 324-332. DOI: 10.1055/s-0038-1667161

4 Coalson G, Peña E, Byrd C. Description of multilingual participants who stutter. J Fluency Disord. 2013; 38(2): 141-156. DOI: 10.1016/jfludis.2013.02.008

5 Shenker R. Multilingual children who stutter: clinical issues. J Fluency Disord. 2011; 36(3): 186-193. DOI: 10.1016/j.jfludis.2011.04.001

6 Tetnowski J, Richels C, Shenker R, Sisskin V, Wolk L. When diagnosis is dual. ASHA Lead. 2012; 17: 10-13. DOI: 10.1044/leader.FTR1.17022012.10

7 Byrd C T, Bedore M L, Ramos D. The disfluent speech of bilingual Spanish-English children: Considerations for differential diagnosis of stuttering. Lang Speech Hear Ser. 2015a; 46(1): 30–43. DOI: 10.1044/2014_LSHSS-14-0010

8 Eggers K, Van Eerdenbrugh S, Byrd C T. Speech disfluencies in bilingual Yiddish-Dutch speaking children. Clin Linguist Phonet. 2019; 34(6): 1-17. DOI: 10.1080/02699206.2019.1678670

9 Rispoli M, Hadley P. The significance of sentence disruptions in the development of grammar. JSLHR. 2001; 44(5): 1131-1143. DOI: 10.1044/1092-4388

10 Zackheim C, Conture E. Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: a preliminary study. J Fluency Disord. 2003; 28(2), 115-142. DOI: 10.1016/S0094-730X(03)00007

11 Ambrose N G, Yairi E. Normative disfluency data for early childhood stuttering. J Speech Lang Hear R. 1999; 42(4), 895–909. DOI: 10.1044/jslhr.4204.895

12 Yairi E, Seery C. Stuttering: Foundations and clinical applications. 3rd ed, Plural publishing; 2023.

13 Byrd C T, Watson J, Bedore L, Mullis A. (2015b). Identification of stuttering in bilingual Spanish-English speaking children. Contemporary Issues in Communication Science and Disorders. 2015b; 42: 72-87. DOI: 10.1016/s0094-730x(03)00004-4

14 Shaaban K A. Bilingual education in Lebanon. In: J. Cummins & D. Corson, editors. Encyclopedia of Language and Education. The Netherlands: Kluwer; 1997; 5: 251-259.

15 Thonhauser I. Multilingual education in Lebanon: "Arabinglizi" and other challenges of multilingualism. Mediterranean Journal of Educational Studies. 2000; 6(1), 49-2000.

16 Cordes A. Individual and consensus judgments of disfluency types in the speech of persons who stutter. J Speech Lang Hear R. 2000; 43*:* 951-964. DOI: 10.1044/jslhr.4304.951

17 Runyan C, Adams M. Perceptual study of the speech of "successfully therapeutized" stutterers. J Fluency Disord. 1978; 3: 25-39. DOI: 10.1016/0094-730X(78)90004-9

18 Runyan C, Adams M. Unsophisticated judges' perceptual evaluations of the speech of "successfully treated" stutterers. J Fluency Disord. 1979; 4: 29-38. DOI: 10.1016/0094-730X(79)90029-9

19 Wendahl R, Cole J. Identification of stuttering during relatively fluent speech. J Speech Hear Res. 1961; 4: 281-286. DOI: 10.1044/jshr.0403.281

20 Williams D, Kent L. Listener evaluations of speech interruptions. J Speech Hear Res. 1958; 1: 124-131. DOI: 10.1044/jshr.0102.124

21 Luper H. Consistency of stuttering in relation to the goal gradient hypothesis. J Speech Hear Disord.1956; 21: 336-342. DOI: 10.1044/jshd.2103.336

22 Martin R R, Haroldson S K. Stuttering and speech naturalness: Audio and audiovisual judgments. J Speech Hear Res. 1992; 35(3): 521-528. DOI: 10.1044/jshr.3503.521

23 Panico J, Healey C E, Brouwer K, Susca M. Listener perceptions of stuttering across two presentation modes: quantitative and qualitative approach. J Fluency Disord. 2005; 30(1): 65-85. DOI: 10.1016/j.jfludis.2005.01.003

24 Tuthill C. A quantitative study of extensional meaning with special reference to stuttering. Speech Monogr. 1946; 13: 81-98. DOI: 10.1044/jshd.0502.189

25 Williams D, Wark M, Minifie F. Ratings of stuttering by audio, visual, and audiovisual cues. J Speech Hear Res. 1963; 6: 91-100. DOI: 10.1044/jshr.0601.91

27 Mayer M. Frog goes to dinner. New York, NY: Dial Books for Young Readers; 1974.

28 Riley G. Stuttering severity instrument for children and adults, 4th Ed. Austin, TX: Pro Ed, Inc; 2009.

29 Tuller L. Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J.d. Jong, & N. Meir, Assessing Multilingual Children Disentangling Bilingualism from Language Impairment. 2015; 300-330, Bristol: Multilingual Matters. DOI: 10.21832/9781783093137-013

30 Hallal F, Kouba Hreich E, Messarra C M. Approche méthodologique pour l'évaluation orthophonique dans le context plurilingue Libanais: Evaluation de l'utilisation du "questionnaire pour parents d'enfants bilingues" (litmus-pabiq) dans une démarche de bilan orthophonique. [Methodological approach for speech therapy assessment in the Lebanese multilingual context: Assessment of the use of "the parents of bilingual children questionnaire" (litmus-pabiq) in a speech therapy assessment process], [Unpublished thesis]. Saint-Joseph University of Beirut; 2016.

31 Humphrey B D. Judgments of disfluency in a familiar vs. an unfamiliar language. In A. Packmann, A. Meltzer, & H. F. M. Peters, editors. Theory, research and therapy. Proceedings of the Fourth World Congress on Fluency Disorders: 424-427, Nijmegen, the Netherlands: Nijmegen University Press; 2004.

32 Van Borsel J, Britto Pereira M M. Assessment of stuttering in a familiar versus an unfamiliar language. J Fluency Disord. 2005; 30(2): 109-124. DOI: 10.1016/j.jfludis.2005.04.001

33 Van Borsel J, Leahy M M, Britto Pereira M M. Judging stuttering in an unfamiliar language: The importance of closeness to the native language. Clin Linguist Phonet. 2008; 22(1), 59-67. DOI: 10.1080/02699200701647289

34 Olness G. Structural assessment of narratives: Some issues and Illustrative Data. [Paper presentation]. Clinical Aphasiology Conference, Ghent, Belgium; 2006. https://aphasiology.pitt.edu/id/eprint/1750

35 Bedore L M, Fiestas C E, Peña E D, Nagy V J. Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. Biling-Lang Cogn. 2006; 9(3): 233-247. DOI: 10.1017/s1366728906002604

36 Suen H K, Ary D. Analyzing quantitative behavioral observation data. Mahwah, NJ: Erlbaum; 1989. DOI: 10:4324/9781315801827

37 Boey R A, Wuyts F L, Van de Heyning P H, De Bodt M S, Heylen L. Characteristics of stuttering-like disfluencies in Dutch-speaking children. J Fluency Disord. 2007; 32(4), 310–329. DOI: 10.1016/j.jfludis.2007.07.003

38 Leclercq A L, Suaire P, Moyse A. Beyond stuttering: Speech disfluencies in normally fluent French-speaking children at age 4. Clin Linguist Phonet. 2017; 32(2): 166–179. DOI: 10.1080/02699206.2017.1344878

39 Pellowski M W & Conture E G. Characteristics of speech disfluency and stuttering behaviors in 3- and 4-year-old children. J Speech Lang Hear R. 2002; 45(1): 20–34. DOI: 10.1044/1092-4388(2002/002)

40 Tumanova V, Conture E G, Lambert E W, Walden T A. Speech disfluencies of preschool-age children who do and do not stutter. J Commun Disord. 2014; 49: 25–41. DOI:

10.1016/j.jcomdis.2014.01.003

41 Rousseau I, Onslow M, Packman A, Jones M. Comparisons of audio and audiovisual measures of stuttering frequency and severity in preschool-age children. Am J Speech-Lang Pathology. 2008; 17: 173-178. DOI: 10.1044/1058-0360

42 Susca E, Healey E. Perceptions of simulated stuttering and fluency. J Speech Lang Hear R. 2001; 30: 159-165. DOI: 10.1044/1092-4388

43 Hedge M, Hartman D. Factors affecting judgments of fluency: I. interjections. J Fluency Disord. 1979a; 4: 1-11. DOI: 10.1016/0094-730X(79)90026-3

44 Hedge M, Hartman D. Factors affecting judgments of fluency: II. word repetitions. J Fluency Disord. 1979b; 4: 13-32. DOI: 10.1016/0094-730X(79)90027-5

45 Karimi K, Jones M, O'Brian S, Onslow M. Clinician percent syllables stuttered, clinician severity ratings and speaker severity ratings: are they interchangeable? Int J Lang Commun Disord. 2014; 49: 364-368. DOI: 10.1111/1460-6984.12069

46 Huffman E, Perkins W. Dysfluency characteristics identified by listeners as "stuttering" and "stutterer". J Commun Disord. 1974; 7: 89-96. DOI: 10.1016/0021-9924(74)90010-0

47 Rincon C, Johnson K N, Byrd C. An introductory examination of speech disfluencies in Spanish-English bilingual children who so and do not stutter during narratives. Perspect ASHA Spec Interest Groups. 2020; 5: 131-141. DOI: 10.1044/2019_PERSP-19-00040

Table 1. Survey-based description of the SLP participants characteristics

Questionnaire's section	**Participants' description**	
Work experience	*M* = 3.0 years	Range = 1-11
Highest educational degree	Bachelor degree 53.1%	Master degree 46.9%
Number of bilingual clients who stutter	*M* = 8.75	Range = 1-30
Experience with preschoolers who stutter	Yes (46.8%)	No (12.5%)
Courses/workshops about stuttering	Yes (87.5%)	No (12.5%)
Courses/workshops about bilingualism	Yes (59.3%)	No (40.7%)
Specialization program in stuttering	Yes (46.8%)	No (53.2%)
Self-ratings estimating difficulties in distinguishing SLD and OD (scale 1-5; 1 = not at all; 5 = extremely difficult)	*M* = 2.2	Range = 1-5
Self-ratings of confidence in accurately diagnosing stuttering in bilingual children (scale 0-9; 0 = not confident at all; 9 = extremely confident)	*M* = 6.0	Range = 0-9

Table 2. Linguistic profile of children based on Parents of Bilingual Children Questionnaire

	CWS1	**CWS2**	**CWNS1**	**CWNS2**	**CWNS3**	**CWNS4**	**CWNS5**	**CWNS6**
Sex	male	male	male	male	female	male	female	male
Age	5;3	5;5	5;9	5;9	5;5	5;4	5;6	5;8
No risk index	23	23	23	21	21	23	23	23
Language dominance Arabic	40	31.25	37.5	35.5	38.5	41	23.5	31.75
Language dominance French	/	/	11.5	16.5	/	/	15	4.5
Language dominance English	19	26.5	/	/	12.25	20	25.5	10.5

Table 3. Description of the disfluencies in the speech samples of CWS and CWNS

Disfluency type[2]	**CWS1**	**CWS2**	**CWNS1**	**CWNS2**	**CWNS3**	**CWNS4**	**CWNS5**	**CWNS6**
MonoWR	2.44	1.85	4.26	3.65	1.93	1.61	2.66	2.63
SndR	4.39	1.59	0.00	0.00	0.55	0.00	0.20	0.33
SylR	8.29	1.06	0.00	0.40	0.28	0.54	0.41	1.32
P	1.95	0.26	3.04	0.00	2.76	0.00	0.00	0.00
B	0.24	1.59	0.00	0.00	0.28	0.00	0.00	0.33
BW	0.00	0.53	0.00	0.00	0.00	0.00	0.20	0.33
MultiWR	1.71	2.38	1.21	0.00	1.38	0.00	0.82	0.00
PhR	0.49	1.06	1.82	1.62	0.55	0.00	0.61	0.00
UW/S	2.20	1.06	0.00	2.43	3.04	1.61	0.61	1.64
I	4.63	0.26	3.65	1.21	4.14	1.08	0.82	3.95
LRev	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.00
Grev	0.24	0.26	0.60	0.00	0.00	0.00	0.00	0.99
PhonRev	0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total disfluencies	27.07	12.71	14.58	9.31	14.92	4.84	6.35	11.51
%SLD	17.31	6.88	7.30	4.05	5.80	2.15	3.47	4.94
%OD	9.76	5.28	7.28	5.26	9.11	2.69	2.86	6.58
Mean Nb of iterations	1.26	1.13	1.10	1.07	1.11	1.00	1.19	1.08
Physical concomitants	yes	yes	no	no	no	no	no	no
Nb of words	410	378	164	246	362	186	488	304
Nb of syllables	691	637	233	320	660	327	650	498
Speech sample length	8’16’’	6’33’’	3’56’’	3’10’’	4’35’’	3’33’’	7’03’’	4’20’’

[2] SLD: Monosyllabic word repetitions (MonoWR), syllable repetitions (SylR), sound repetitions (SndR), blocks (B), broken words (BW) and prolongations (P)
OD: Multisyllable word repetitions (MultiWR), phrase repetitions (PhR), interjections (I), unfinished words/sentences (UW/S), lexical (LRev), grammatical revisions (GRev) and phonological revisions (PhonRev)

Table 4. Mean stuttering severity ratings per child in the audio-only and audio-visual presentation mode

	CWS1	**CWS2**	**CWNS1**	**CWNS2**	**CWNS3**	**CWNS4**	**CWNS5**	**CWNS6**
Audio-only	2.44	2.06	1.28	0.75	0.94	0.16	2.06	0.63
Audio-visual	2.44	2.04	1	0.81	0.92	0.11	1.89	1.07

Table 5. Post-hoc comparisons between rated children across presentation modes

		CWNS1	**CWNS2**	**CWNS3**	**CWNS4**	**CWNS5**	**CWNS6**	**CWS1**	**CWS2**
	Mean	1.19	0.78	0.94	0.15	1.94	0.89	2.44	2.04
CWNS1	1.19	x	0.83	1.00	**0.00**	0.56	1.00	**0.00**	**0.05**
CWNS2	0.78	0.83	x	1.00	**0.01**	**0.00**	1.00	**0.00**	**0.01**
CWNS3	0.94	1.00	1.00	x	**0.00**	**0.02**	1.00	**0.00**	**0.00**
CWNS4	0.15	**0.00**	**0.01**	**0.00**	x	**0.00**	**0.01**	**0.00**	**0.00**
CWNS5	1.94	0.56	**0.00**	0.02	**0.00**	x	**0.03**	1.00	1.00
CWNS6	0.89	1.00	1.00	1.00	**0.01**	**0.03**	x	**0.00**	**0.00**
CWS1	2.39	**0.00**	**0.00**	**0.00**	**0.00**	1.00	**0.00**	x	1.00
CWS2	2.04	**0.05**	**0.00**	**0.00**	**0.00**	1.00	**0.00**	1.00	x

Fig. 1. Average ratings for CWNS and CWS based on audio-only versus audio-visual presentation modes

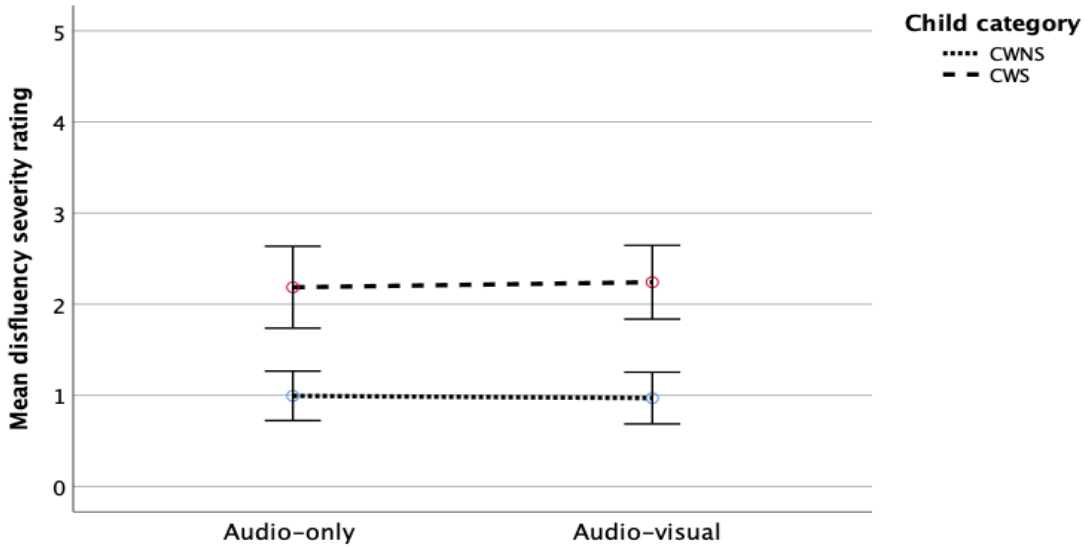

Fig. 2. Individual ratings for CWNS and CWS across presentation modes