International Journal of Pediatric Otorhinolaryngology 79 (2015) 2260–2264

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: www.elsevier.com/locate/ijporl

Speech task effects on acoustic measure of fundamental frequency in Cantonese-speaking children

Estella P.-M. Ma *, Nina L.-N. Lam

Voice Research Laboratory, Division of Speech and Hearing Sciences, The University of Hong Kong, 7/F Meng Wah Complex, Pokfulam, Hong Kong

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 8 October 2015
Accepted 14 October 2015
Available online 3 November 2015

Keywords:
Pediatric voice
Voice assessment
Speech task effect
Reliability

ABSTRACT

Objectives: Speaking fundamental frequency (F_0) is a voice measure frequently used to document changes in vocal performance over time. Knowing the intra-subject variability of speaking F_0 has implications on its clinical usefulness. The present study examined the speaking F_0 elicited from three speech tasks in Cantonese-speaking children. The study also compared the variability of speaking F_0 elicited from different speech tasks.

Methods: Fifty-six vocally healthy Cantonese-speaking children (31 boys and 25 girls) aged between 7.0 and 10.11 years participated. For each child, speaking F_0 was elicited using speech tasks at three linguistic levels (sustained vowel /a/ prolongation, reading aloud a sentence and passage). Two types of variability, within-session (trial-to-trial) and across-session (test–retest) variability, were compared across speech tasks.

Results and conclusions: Significant differences in mean speaking F_0 values were found between speech tasks. Mean speaking F_0 value elicited from sustained vowel phonations was significantly higher than those elicited from the connected speech tasks. The variability of speaking F_0 was higher in sustained vowel prolongation than that in connected speech.

1. Introduction

Speaking fundamental frequency (F_0) is a parameter frequently included in voice assessment protocol for documenting vocal performance in individuals with vocal pathologies and with neurological impairments. It is also commonly used as an outcome measure for evaluating voice before and after voice treatment. Speaking F_0 can be elicited using sustained vowel prolongation and connected speech. Sustained vowel prolongation does not require sophisticated articulatory adjustments and the task is suitable for very young children with limited speech production. Common connected speech tasks include counting, reading and monologue. Such tasks are generally carried out in children with more developed speech and linguistic abilities.

Speaking F_0 obtained can be influenced by the stimuli used for elicitation. However, the literature has mixed reports about speech task effects on speaking F_0. In a group of 12 vocally healthy young adults, Fitch [1] reported speaking F_0 obtained from sustained vowel prolongation be significantly higher than those obtained from connected speech tasks using passage reading. Similar speech task effect on speaking F_0 has been reported in the children population. Brown and Shrivastav [2] studied the comfortable speech effort level in preschoolers. Thirty preschoolers aged between 3 and 4 years were asked to produce three speech tasks including sustained vowel prolongation, repeating a sentence and repeating four English words. Significant speech task effect on speaking F_0 was found. Interestingly, contrary to Fitch's [1] results, speaking F_0 of vowel prolongation was found significantly lower than those obtained from both connected speech tasks. Baker and colleagues [3] examined the effects of task type on speaking F_0 in school-age children aged between 5.0 and 7.11 years. In their study, speaking F_0 from four speech tasks were compared including sustained vowel /a/ prolongation, sustained vowel embedded in a word at the end of a phrase, repeating a sentence, and counting from 1 to 10. They also found that speech tasks could significantly influence speaking F_0 values in young children. Counting was found to elicit a significantly higher speaking F_0 value than phrase and sentence tasks. However there is at least one report that does not find any significant speech task effects in young children and in adult males [4].

Clinically, knowing the intra-speaker variability of a voice measure can provide a reference of whether differences in client's

* Corresponding author. Tel.: +852 39177625; fax: +852 25590060.
E-mail address: estella.ma@hku.hk (Estella P.-M. Ma).

http://dx.doi.org/10.1016/j.ijporl.2015.10.018
0165-5876/

performance over time reflect true changes due to treatment effects or are simply due to variability of the measure [5]. Therefore, researchers have also examined the intra-speaker variability of speaking F_0 by evaluating whether speaking F_0 is consistent over time. Data are available on across-session (test–retest) variability of speaking F_0 for different speech tasks in normal adult and preschool speakers. In general, speaking F_0 values are found to be rather consistent across time. The literature has documented similar speaking F_0 values obtained on separate days [1,2,6] and at different times on the same day [7]. However, there are a few exceptions. Garrett and Healey [8] reported in their study that males showed significant increase in speaking F_0 across time periods from morning to early to late afternoon. Among speech tasks, speaking F_0 elicited from sustained vowels demonstrated greater test–retest variability than short phrases [9].

Speaking F_0 can vary across speakers due to linguistic, physiological and cultural differences [10–13]. Linguistically, speaking F_0 can manifest differently between tone (e.g., Mandarin Chinese) and nontone (e.g., English) languages. Mandarin Chinese speakers showed higher speaking F_0 than English speakers in reading an unemotional narrative text in their respective native language [14]. Similarly in a more recent study by Keating and Guo [15], higher speaking F_0 has been reported in Mandarin Chinese than English. Greater speaking F_0 range has been found in tone language Mandarin Chinese when compared to nontone language of English [16]. Cantonese is a dialect of Chinese and is a tone language. To date, normative voice database of Cantonese pediatric population is not yet available. The intra-speaker variability of speaking F_0 in young Cantonese-speaking children remains to be examined.

The present study was set out to achieve two objectives. The first objective was to investigate the effects of speech task on F_0 in vocally healthy, Cantonese-speaking school-age children. The second objective was to evaluate the variability of speaking F_0 elicited from different speech tasks. Two types of variability, within-session (trial-to-trial) and across-session (test–retest) variability, were compared.

2. Methods

2.1. Participants

Fifty-six school-age children (31 boys and 25 girls) participated in the study. They were recruited from a local primary school in Hong Kong. Their mean age was 8.36 years (SD = 0.98; range = 7.0 to 10.11 years) (Table 1). The lower age range was set to ensure the children have the ability to follow researcher's task instructions and have the linguistic ability to read aloud a standard passage for recording purposes. The upper age range was set to exclude puberty voice that could lead to voice instability. All children were native speakers of Cantonese. They were all confirmed to be vocally healthy by both authors. According to the parent report and class teacher report, all children had normal hearing abilities. Children who had a previous history or current oral structural abnormalities, or speech and language problems were not eligible for the study. All of them reported themselves to be free from colds or other upper respiratory tract infections on the day of testing.

Table 1
Age and gender distribution of participants.

Age (in years)	Boys	Girls	Total
7.0–7.11	6	7	13
8.0–8.11	11	6	17
9.0–9.11	8	11	19
10.0–10.11	6	1	7
Total	31	25	56

2.2. Procedures

The voice recording was conducted in a quiet room in the school, with background noise kept under 45 dBA throughout the recording session. Cautions were taken to ensure minimal fluctuations in noise level during the recording. During the recording, if there were interruptions by external noises (e.g., school bell rang before and after recess, or announcements through the school public address system), that voice sample would be discarded and another voice sample would be recorded again. Voice samples were recorded using a head-mounted professional-grade, condenser microphone (AKG Acoustics C420, Vienna, Austria) connected to a digital recorder (H4next Handy Recorder, ZOOM Corporation, Japan). The microphone was placed and maintained 5 cm from the child's mouth corner throughout the recording.

Each child performed three speech tasks including sustained vowel prolongation, reading aloud a Cantonese sentence and a standard Cantonese passage. In the vowel prolongation task, each child was asked to sustain the vowel /a/ for five seconds. In the sentence task, each child was asked to read aloud a Cantonese sentence /ba_1 ba_1 da_2 $gɔ_1$ $gɔ_1$/. In the passage task, each child was asked to read aloud a Cantonese passage. The passage had 138 characters and was selected from a local textbook of Grade One level. The passage was printed on an A4-size paper with font size 24. Two practice trials as familiarization were given to each child before actual recording. The children were instructed to complete all speech tasks using their comfortable pitch and loudness level. Five trials were recorded for each speech task. The recording session lasted for approximately 20 min.

Twenty-one of the participating children were randomly selected to perform the same speech tasks on a second occasion, around four days after the first recording was obtained. This second recording was carried out at a similar time of the day as the first recording. This procedure was to evaluate across-session (test–retest) variability of the speech tasks.

2.3. Data analysis

All voice samples were analyzed using software Praat [24]. The F_0 traces of all samples were visually checked by both authors to avoid measurement error. For sustained vowel /a/ prolongation, the middle three seconds of the sample was selected for analysis. The entire sentence from the onset of the first word (i.e., /ba_1/) to the offset of the last word (i.e., /$gɔ_1$/) was included for analysis. Similarly, the entire passage was included for analysis. All the five trials of each speech task were used for calculating mean speaking F_0. In order to better reflect what the ear perceives, speaking F_0 values were also reported in semitones in addition to in linear scale (in the unit of Hertz). The conversion to logarithmic scale was made in the unit of semitone relative to an arbitrary musical note A2 or 110 Hz.

Two data sets were used for measuring variability of speaking F_0: One data set for within-session (trial-to-trial) and the other data set for across-session (test–retest) measurement. Three measures were used to evaluate the variability of speaking F_0: Cronbach's *alpha*, coefficient of variation (CoV) and intra-class correlation coefficient (ICC). The CoV reflects the within-subject variation whereas ICC reflects the relative reliability. ICC values can range from 0 (that is, no correlation) to 1 (that is, perfect correlation). The larger the overall ICC value, the more reliable the speech task. An ICC value of greater than 0.8 suggests high degree of correlation or high reliability [17]. All statistical analyses were