Agreement during identification of /k/ and /g/ in cleft lip and palate at speech stimuli

Concordância durante a identificação de /k/ e /g/ na fissura labioapalatina em diferentes estímulos de fala

Concordancia durante identificación de /k/ y /g/ en la fisura palatina en estímulos del habla distintos

Tayrine Sousa Marques Borges[*]

Jeniffer de Cassia Rillo Dutka[**]

Thais Alves Guerra[***]

Gabriela Zuin Ferreira[****]

Aveliny Lima-Gregio Mantovan[*****]

Viviane Cristina de Castro Marino[******]

Paper presented at the Third International Symposium of Orofacial Cleft and Related Anomalies promoted by the Rehabilitation Hospital for Craniofacial Anomalies, USP, Bauru - SP, held on November 08th and 09th 2013.

Abstract

Introduction: Children with cleft lip and palate (CLP) can use compensatory articulation during speech. The stimuli used for classifying atypical articulatory productions can affect examiners' agreement during identification of

[*]*Resident for the Multidisciplinary Residency Program in Health of Children and Adolescents from the Federal University of São Paulo – UNIFESP- Brazil.*[**]*Assistant Professor Doctor at the Speech Pathology and Audiology course at the Dental School in Bauru and the Graduate Program from the Rehabilitation Hospital for Craniofacial Anomalies from the University of São Paulo - USP, Bauru, SP, Brazil.* [***]*Master Student at the Rehabilitation Sciences Program from the Rehabilitation Hospital for Craniofacial Anomalies (Hospital de Reabilitação de Anomalias Craniofaciais - HRAC), University of São Paulo (USP), Bauru, SP, Brazil.* [****] *Associate Professor in the Department of Speech Pathology and Audiology at the Ceilândia College from the University of Brasilia (UNB), Brasília, DF, Brazil.* [*****] *Assistant Professor Doctor at the Philosophy and Sciences College, Universidade Estadual Paulista - UNESP, Marilia, SP, Brazil*

Authors´ contributions: *TSMB - involved in the development and study design, data collection, editing of speech samples, preparation of samples and audio presentation to the judges, analysis and interpretation of data and preparation of the article. JCRD - participated in the preparation and design of the study, the analysis and interpretation of data and also in preparing and revising it. TAG - involved in data collection, editing of speech samples, preparation of samples and audio presentation to the judges and also in data analysis and preparation of this article. GZF - involved in data collection, editing of speech samples, preparation of samples and audio presentation to the judges and also in data analysis and preparation of this article. AMLG - participated in the preparation and design of the study, the analysis and interpretation of data and also in preparing and revising it. VCCM - involved in the development and design of the study, in data collection, analysis and interpretation of data and also in the preparation and review of this article.*

Correspondence Address: *Viviane Cristina Castro Marino. Rua Santa Helena, 1967, casa 07, Jardim Estoril, Marília. CEP 17514.410 - Brasil*
E-mail adress*: vivianemarino2@yahoo.com.brr*
Received: *17/07/2014* **Accepted***: 23/10/2014*

glottal stops. **Purposes**: To establish the correlation between Auditory-Perceptual Evaluation by speech-language-pathologists (SLP) and a golden-standard rating of use of /k/ and /g/ produced by children with operated CP during production of single words, carrier phrases and phrases with recurrence of targets. **Method:** This prospective study involved ratings of 420 recordings of speech samples produced by six children, three with operated cleft lip and palate or operated cleft palate and three controls. The speech stimuli consisted of single words (SW), carrier phrases (CP) and phrases with recurrence of velar consonants (PR). Five SLPs performed auditory-perceptual ratings to identify if they heard use of glottal stop, omission of target consonants or adequate production. Agreement between the SLPs and a golden-standard rating was established. **Results:** The degree of agreement between the SLPs and the golden-standard rating varied according to the speech stimuli (SW = 69%, CP= 79% and PR = 98%, with Kappa statistics ranging from substantial to almost perfect). When taking into account the type of response the agreement was greater for "presence of glottal stop", particularly for PR. **Conclusion:** Levels of agreement for perceptual judgments of speech recordings varied according to the stimuli. The use of phrases with recurrence of the target consonant had the best agreement.

Keywords: Articulation Disorders; Cleft Palate; Child.

Resumo

Introdução: *Crianças com fissura labiopalatina podem fazer uso de articulações compensatórias durante a produção da fala. Os estímulos de fala utilizados para identificação das produções atípicas podem afetar a concordância durante os julgamentos das oclusivas glotais.* ***Objetivo:*** *Estabelecer a concordância entre julgamentos perceptivo-auditivos por fonoaudiólogos e um julgamento padrão ouro de /k/ e /g/ produzidos por crianças com fissura de palato operada durante produção de palavras isoladas, frases-veículo e frases com recorrência de som-alvo.* ***Método:*** *Estudo prospectivo envolvendo o julgamento de 420 gravações de amostras de fala produzidas por 6 crianças, três com fissura labiopalatina operada ou de palato operada e três controles. Os estímulos de fala foram constituídos por palavras isoladas (PI), frases-veículo (FV) e frases com consoantes velares em recorrência (FR). Cinco fonoaudiólogas realizaram o julgamento perceptivo-auditivo das gravações com a tarefa de selecionar se ouviram o uso de oclusiva glotal, omissão do som-alvo ou fala típica. A concordância entre esses julgamentos e um julgamento considerado padrão ouro foi estabelecida.* ***Resultados:*** *O grau de concordância entre os fonoaudiólogos e o julgamento padrão ouro variou de acordo com os estímulos de fala (PI= 69%, FV= 79% e FR= 98%; Kappa variando de substancial a quase perfeito). Levando-se em conta o tipo de resposta, houve maior concordância para "presença de oclusiva glotal" em FR.* ***Conclusão:*** *Níveis de concordância durante julgamentos perceptivo-auditivos de gravações de fala variaram de acordo com os estímulos apresentados. O uso de frases com recorrência da consoante /k/ e /g/ resultou em melhor concordância.*

Palavras-chave: *Transtornos da Articulação; Fissura Palatina; Criança.*

Resumen

Introducción: *Niños con paladar hendido (PH) pueden utilizar articulación compensatoria durante el habla. Los estímulos utilizados para evaluar las producciones articulatorias atípicas pueden afectar la concordancia durante la identificación de golpe de glote.* ***Objetivo:*** *Establecer la correlación entre las juzgamiento perceptivo-auditivo por Terapistas de Habla (TH) y un índice de oro-estándar durante identificación de /k/ y /g/ en niños con PH operado durante producción de palabras aisladas, frase- vehículo y frases con la recurrencia de /k/ y /g/.* ***Método:*** *Este estudio prospectivo envolvió evaluación de 420 grabaciones de habla producidas por seis niños, tres con LPH o PH operado y tres controles. Los estímulos del habla consistían en palabras aisladas (PI), frases-vehículo (FV) y frases con repetición de consonantes velares (FR). Cinco TH evaluaron las grabaciones cuanto el uso de golpe de glote, omisión o producción adecuada. Se estableció el nivel de concordancia entre los juzgamientos de las TH y un índice de oro-estándar.* ***Resultados:*** *El nivel de concordancia entre los TH y la calificación oro-estándar varió de acuerdo con los estímulos del habla (PI=69%, FV=79% y FR=98%), con estadísticas Kappa variando desde sustancial a casi perfecta. Si se tiene en cuenta el tipo de respuesta, la concordancia fue mayor para la "presencia de golpe de glote" para las FR.* ***Conclusiones:*** *Los niveles de acuerdo durante juzgamiento perceptivo-auditivo de las grabaciones de hable variaran de acuerdo con los estímulos. El uso de frases con la recurrencia de la consonante de interés tenía el mejor acuerdo.*

Palabras clave: *Trastornos de la Articulación; Fissura del paladar; Niño.*

Introduction

The speech-language-pathologists who work with individuals with cleft lip and palate (CLP) or velopharyngeal dysfunction (VPD) must be able to identify the use of compensatory articulations (CA) which can occur in the speech of these subjcts[1] in order to establish the appropriate treatment planning. The perceptual judgment is essential in speech evaluation for subjects with CLP as the final decision on their production is based on this judgment[2,3].

It is known that the identification and characterization of compensatory productions are based on perceptual evaluation of speech with focus on place, manner of articulation and the sound observed during attempts to target sound production[4]. Thus, it is important that health professionals use standardized clinical procedures that favor the collection and analysis of information from these production[1,2-5].

For this reason, researchers have been concerned about possible factors that can influence the interpretation of results obtained in the perceptual evaluation of speech changes[2-5,6]. The literature suggests several aspects that may particularly influence the evaluation of CA[1,7,8], including agreement among the judges in the records of the clinical reviews[5,6,9-12] and the judge's experience in performing the auditory perception judgment of CA[2,8,13]. In order to minimize the effects of variable "experience" in perceptual judgments of speech it is suggested to constitute a team of trained listeners who can offer reliable judgments, once multiple judges conducting the trials are preferable to those performed by a single judge[1, 5].

Besides the reviewer's experience, literature has highlighted the need to select samples of speech sounds that allow comparisons of results for different centers and subjects who speak different languages[1,2]. Some scholars pointed the importance of using speech stimuli with different levels of complexity (syllables, words and sentences) to evaluate the different aspects of speech in subjects with VPD[3]. Others recommended the use of a combination of "naming words and repeating phrases" stimuli when only two procedures may be selected for the evaluation of phonological aspects and production of the sounds of children with CLP[14]. Particularly, in regards to the CA, the literature reports variability in selecting them, namely: spontaneous conversation, texts, sentences, directed words and speech (counting and days of the week). Different studies have used such stimuli alone or in combination, with more descriptions for the combined use of spontaneous and isolated words[2]. A study, in particular, used controlled words formed by consonant + vowel + consonant (CVC), which were inserted in a vehicle sentence8. In this study, only the plosive, fricatives and affricates consonants were included, as they are consonants which the CA are commonly identified by requiring greater amount of intraoral pressure in their production.

Overall, in the last decade literature has recommended, for the purpose of multicenter research, the use of controlled speech samples consisted of words and especially by phonetically elaborate sentences, identification of CA[1,5]. Some researchers have proposed the use of phrases formed by a single consonant (target) to be investigated, inserted at different positions in the word, in order to avoid the influence of other consonants in the evaluation of the consonant (target), or even to prevent the occurrence of assimilation[15]. The use of sentences consisting of the occurrence of the target consonant in the starting and the final position of the word, in phonetically controlled environments, has been proposed as part of a standardized protocol developed for professional training in the evaluation of speech (*Cleft Audit Protocol for Speech – CAPS-A: a comprehensive training package for speech analysis*)[11]. As emphasized by Sell[5], for the task of analysis of samples for multiple judges, besides the necessary training reference samples, it becomes important to standardize the sampling, that is, what the speaker should produce.

In Brazil, initiatives for the creation of phonetically standardized phrases for clinical and research purposes have been made by initiatives from *BrasilCleft* task force which aimed a standardization and systematization of the collection and analyses of the primary post-intervention results in CLP fissure, from the integration of Brazilian services that work with craniofacial anomalies. However, targeted studies to verify compliance during the perceptual judgments of speech stimuli involving occlusive /k/ and /g/ are restricted. The interest in investigations involving occlusive /k/ and /g/ is due to the high occurrence of the glottal occlusive (also known as "glottal stop") in occlusive consonants[16] and, particularly, in /k/ and

/g/, when this CA is present in the speech of the population with CLP.

The glottal stop (GS) is a transient articulation of consonant nature, whose correlate acoustic is a sound essentially abrupt[17]. This production cannot be often auditory perceptually distinguished from the omission of a consonant segment. However, this distinction seems to be favored when the speech stimulus consists of recurrence of the investigated sound target. The hypothesis adopted in this study is that speech stimuli involving more opportunities recurrence of stops /k/ and /g/ can ease the identification of the GS.

The aim of this study is to establish the correlation between auditory-perceptual judgments by speech-language-pathologists and a gold standard judgment of the /k/ and /g/ produced by children with cleft lip and/or operated palate during the production of isolated words, phrase vehicles and phrases to target sound recurrence. It also aims to establish the relationship between the degree of agreement of the judgments for the three used stimuli (single words, sentences and phrases with vehicle-target sound recurrence).

Materials and methods

Prospective study of perceptual judgments of occlusive consonants /k/ and /g/ produced in different speech samples of children with cleft palate and/or operated palate. These samples were the stimuli that are the focus of this study. The speech samples were collected in a high complexity center in the management of craniofacial anomalies (Rehabilitation Hospital for Craniofacial Anomalies - Hospital de Reabilitação de Anomalias Craniofaciais: HRAC, USP, Bauru). Reference judgments ("gold standard") were established by three speech-language-pathologists from this center. Five other speech-language-pathologists with experience in the speech evaluation of children with CLP participated as judges of the speech samples recorded. The study was reviewed by the local Ethics Committee (HRAC- USP - Bauru) and was approved on 02.26. 2013, under number 207 833.

Casuistry

Three children, both genders (a boy with operated unilateral incisive transforaminal cleft, a girl with operated incomplete post foramen fissures and another with operated bilateral transforaminal cleft) with an average age of 10.5 years old (SD = 1.23), were selected for this study. The three children had VPD and GS in occlusive /k/ and /g/, during the repetition of words and spontaneous speech, as recorded in the medical records of these children and confirmed in clinical evaluation (real time). None of the children had severe dysphonia and/or language disorders which could impair the production of speech samples of interest and hence the auditory-perceptual identification of these samples.

Three other children were also selected (average age = 11 years old, SD = 1.72), matched on gender, who showed typical production of velar consonants (control). Those responsible for these children did not report complaints of speech disorders, voice and/or hearing. No abnormal speech was identified during the recording of the speech samples for the study.

Procedure of speech stimuli

Speech stimuli selected to constitute the judged speech samples in the study refer to single words and short sentences involving occlusive /k/ and /g/ in recurrence, prepared in accordance with the recommendations proposed by Henningsson et al.[1]. They also refer to the words entered in phrase vehicle that meet the criteria for acoustic analysis, but which can also be used in auditory-perceptual judgments to compare performance of judges and/or stimulus effect/context in speech production.

As for single words (SW), they met the following criteria: (a) being two syllables, with recurrence of the velar occlusive (/k/ or /g/) or formed by the combination of occlusive /k/ or /g/ and liquid sounds and (b) the velar occlusive should be associated with high vowels /i/, /u/ and /a/ and in stressed position in the word. Thus, six words have been prepared ("caca, quilo, cuca" and "gago, guizo, gula"), with the velar occlusive /k/ and /g/, respectively.

As for the words used in vehicle phrases (VP), these consisted of the same six disyllable paroxytone words that matched the velar occlusive consonants in initial position of words with vowels /a/, /i/, /u/ in stressed position. These words were inserted into the vehicle phrase: "Fale ____bem bonito" ("Say____ pretty well)."

As for short sentences (SS), these were made of words with recurrence of a single velar consonant of interest (/k/ or /g/). As a result, the following sentences were obtained: "A cuca correu e caiu"

(the bogery ran and fell down) and "O Gugu é gago" (Gugu stutters).

Collection procedure of stored speech samples

The speech samples produced by children with GS and children with typical speech (control) were collected following the same procedures. At the time of the recording of the samples it was requested that each child repeated five times the set of stimuli of interest (i.e., 5 x 6 single words = 30, 5 x 6 words in vehicle sentence = 30 and 5 x 2 short sentences = 10), summing up 70 repetitions per child. Recordings that showed noise or audio with weak vocal intensity that do not permit identification of recorded audio stimuli were excluded. In these cases, new recordings were made and only recordings with high quality were kept in the study. In total, 420 samples (70 repetitions x 6 children) were used to provide the auditory-perceptual judgment of the variables of interest in this study.

All recordings were made in the same acoustically treated room, using digital equipment of high definition (digital recorder MARANTZ, Shure unidirectional microphone). The microphone was positioned 10 cm from the lips of the children, in order to allow enough intensity of the audio signal favoring later auditory-perceptual analysis. The recordings were digitized and stored on a computer.

Preparation of samples for analysis by the judges

The 420 speech samples obtained were edited for the auditory-perceptual analysis of the judges of these speech samples through Praat[18] software. First, the samples were separated into audio files per patient. Later, editing was carried out according to the speech stimulus (e.g., single words, phrases and short sentences vehicle with target sound recurrence) and these stimuli were separated into folders. On the final phase, speech samples were the three sets of speech stimuli and they were organized randomly.

In addition to the recordings of interest to the study, speech samples belonging to the three other children were also recorded, in order to serve as a reference (control) for the judges. These samples represent each of the three assessment possibilities (presence of occlusive /k/ and /g/ as typical (expected) production, presence of GS in occlusive and absence of consonants /k/ and /g/ (omission), and the edited by means of the same Praat[18] software and stored in a separate folder.

After editing every audio-recorded material, the same was filed on CD-ROM, along with the answer sheets. In each recorded CD there were two main folders, one containing the speech samples that should be used for reference for judges and another containing all the samples of interest that should be judged. The instructions printed on the answer sheets were also included on the CD-ROM. The materials (CD and answer sheets) were delivered to the three judges who establish reference judgments (the gold standard) and to the other five judges who participated in the study.

Establishing reference judgments: "gold standard"

Consensual judgments of the 420 speech samples recorded were obtained from three judges with experience in assessing the speech of children with CLP. These speech-language-pathologists (SLP) work for a high complexity center for the treatment of craniofacial anomalies and act in the same center for over five years, with extensive experience in speech evaluation for this population. These professionals reported having normal hearing, had no contact with the children who have had their recorded statements and did not have information about the purpose of the study.

Prior to judgments, instructions and reference samples (audio recorded) representative of each type of production were offered. When submitting samples of references, it was appointed which was the kind of production that should be judged by the judges directing them that they should use this information as a parameter to make their judgments. The judges were instructed to only judge the presence of typical speech, the absence of occlusive /k/ or /g/ (omission) or the presence of GS in different speech stimuli (single words, vehicle phrase and short sentence with target sound recurrence) regardless hearing other speech disorders in the displayed sentence. For example, to hear the phrase "Fala caca bem bonito" ("Say *caca* pretty well"), the judge should decide between the presence of the consonant /k/ (typical), the absence of occlusive /k/ (omission), or the presence of GS.

After the instructions, the judges had the opportunity to hear simultaneously in the same room the

speech samples using individual headphones. They were allowed to listen to the samples as often as they thought necessary, and they could also adjust the volume of recorded samples. After listening to (once or more) each speech sample, practitioners have written down their response option (presence/absence of the occlusive /k/ and /g/ or presence of GS) on answer sheets made for this purpose and then found their answers. In the case of disagreement with their judgments, judges heard again until they reached a consensus judgment (one judgment for each one of the 420 samples heard). The consensual judgment was named in this study "gold standard" and these judgments were reported to verify compliance with the judgments of the other participants (i.e., the other five SLP- judges).

A total of 420 "gold standard" judgments were obtained. From these, 180 referred to single words (30 repetitions x 6 children), 180 referred to the vehicle phrase (30 samples x 6 children) and 60 referred to the short sentences with target sound recurrence (10 samples x 6 children).

Auditory-Perceptual judgment of the speech samples by the judges

Five judges with experience of at least three years in the treatment of craniofacial anomalies judged individually and aurally 420 speech samples. These SLP reported having normal hearing, they had no contact with the subjects who had recorded their statements and had not received information about the objective of the study.

The five judges judged the presence (typical), the absence of occlusive /k/ and /g/ (omission), or even the presence of GS in the presented speech samples. The instructions given to the judgment of the samples followed those described for the establishment of the gold standard, with a difference: each experienced SLP heard with individual headphones, in a room reserved for this purpose. The judgments individually obtained were recorded in an answer sheet to be later compared to "gold standard" judgments.

A total of 2100 judgments (420 x 5 SLPs) were obtained by the 5 SLPs. From these, 1,050 trials were related to the speech of children with cleft lip and/or palate and 1,050 were related to speech from the control children. From the 1,050 trials performed for each group of children included in the study, 450 referred to single words (30 samples x 3 children x 5 judges), 450 the phrases vehicle (30 samples x 3 x 5 children judges) and 150 to phrases with recurrent sounds (10 samples x 3 children).

Data analysis

The percentage of agreement from the judgments was obtained by the samples of speech produced by children with cleft lip and/or palate and by speech samples produced by the control children. From the 1,050 judgments performed by each group of children (with cleft lip and/or palate or control), 450 judgments referred to the single words, 450 to vehicle phrases and 150 to short sentences with target sound recurrence. The percentage of the agreement for the judgments was then obtained, taking into account the type of consonant (/k/ or /g/) for the three types of speech stimuli in the study.

The Kappa index of Agreement was also used to measure the degree of agreement among judges regarding the studied stimuli (single words, phrase and sentence vehicle with target sound recurrence) and the consonants of interest (/k/ or /g/). The Kappa statistic is a measure used to verify compliance inter-judges that corrects the agreement reached by chance (distance the observations made from those expected, the effect of chance, indicating how legitimate the interpretations are). This conservative approach was also used, along with the percentage of agreement, in previous studies that aimed to obtain the inter-judges reliability in judgments of speech disorders presented by the population with CLP[13,19]. In this study, the Kappa coefficient was analyzed unifying the judgments from the five speech-language-therapysts, resulting in a single kappa value for each type of speech stimuli judged. Kappa values were interpreted according to the literature20 in which: 0.00 does not indicate agreement; 0.00 to 0.20 indicates poor agreement; 0.21 to 0.40 regular agreement; 0.41 to 0.60 moderate agreement; 0.61 to 0.80 substantial agreement and from 0.81 to 1.00 perfect agreement (or almost perfect). Confidence intervals were built with 95% statistical confidence and a significance level of 5% ($p < 0.05$) was adopted.

The Equality Test for Two Proportions, nonparametric, was used to compare the proportion of variable responses (types of stimulus and consonants) and their levels were statistically significant.

Results

As expected, there was 100% agreement (typical speech) for the judgments made for the speech samples produced by the control children. There was variability in the agreements obtained for the judgments made for the speech samples produced by children with cleft lip and/or palate, as seen in Tables 1 to 6.

Table 1 shows the percentage of agreement and the Kappa index obtained for the 1,050 performed judgments. It was obtained percentage of agreement of 69% (Kappa 0.76, rated substantial) for stimuli consisting of words and 79% (Kappa: 0.83, classified as almost perfect) for speech stimuli consisting of vehicle phrases. There was a higher percentage of agreement (98%) and Kappa coefficient index (0.99, classified as almost perfect) for stimuli consisting of phrases with target sound recurrence, suggesting influence of speech stimuli in the judgments made of speech production of children with CLP.

TABLE 1 - percentage of agreement and kappa index in different speech stimuli

Types of samples	Percentage	Kappa	p Value
SW	**69%**	**0,76**	**<0,001***
VP	**79%**	**0,83**	**<0,001***
SS	**98%**	**0,99**	**<0,001***

SW = single words; VP = vehicle-phrases and SS = sentences with target sound in recurrence

*Kappa Coefficient p<0.05

Table 2 shows the percentage of agreement for the speech stimuli of interest (single words, phrases and sentences vehicle with target sound recurrence), taking into account the type of response obtained (GS, omission or speech typical) . The results indicate that judges pointed out, in most cases, the "presence of GS" for all stimuli (SS = 98%; VP = 78% and SW = 56%) with greater compliance in sentences having the occlusive /k/ or /g/ recurrence.

TABLE 2 - percentage of agreement for each type of response in different speech stimuli

Samples	Agreement				Judgements
	GS	O	T	N (%)	
SW	253 (56%)	57 (13%)	0%	310 (69%)	450
VP	353 (78%)	4 (1%)	0%	357 (79%)	450
SS	147 (98%)	0%	0%	147 (98%)	450
Total					1050

GS = glottal stop; O = omission; T = typical speech; SW = single words; VP = vehicle phrases; and SS = sentences with target sound recurrence

Table 3 shows the percentage of agreement and the Kappa index obtained for the judgments in relation to the consonants /k/ and /g/, separately. The results show percentage of agreement with higher values for /k/ (77%; Kappa 0.86) than to /g/ (60%; Kappa 0.76) of isolated words. Likewise, they indicate the percentage of agreement with higher values for /k/ (87%; Kappa 0.92) than to /g/ (71%; Kappa 0.82) for vehicle phrases. In contrast, similar values for the two velar consonant (/k/ and /g/) were found when the stimulus consisting of sounds recurrence was judged.

TABLE 3 - percentage of agreement and kappa index for the speech stimuli in relation to the velar consonants

	/k/			/g/		
Samples	**Percentage**	**Kappa**	**Valor p**	**Percentage**	**Kappa**	**p value**
SW	77%	0,86	<0001*	<0001*	0,76	<0001*
VP	87%	0,92	<0001*	<0001*	0,82	<0001*
SS	96%	0,97	<0001*	<0001*	1,00	<0001*

SW = single words; VP = vehicle-phrases and SS = sentences with target sound in recurrence
*Kappa Coefficient p<0.05

Table 4 shows the percentage of agreement for the speech stimuli (single words, phrases and vehicle phrases with target sound recurrence), taking into account the type of response obtained (GS, omission or typical speech) for each velar consonant (/k/ and /g/). There was agreement about the presence of GS to the consonant /k/ in the three speech stimuli judged, with the highest percentage of agreement for the stimulus consisting of phrases with target sound recurrence followed by the vehicle phrases, and finally, the isolated words. This result differed from those obtained for the speech stimuli consisting of single words and vehicle phrases involving the consonant /g/, as on these samples judges agreed on the presence of GS or for the omission of the consonant, although the greater agreement has been for the presence of GS..

TABLE 4 - percentage of agreement for each type of response in the speech stimuli in relation to the velar consonants

	/k/ Agreement				/g/ Agreement			
Samples	**GS**	**O**	**T**	**Total**	**GS**	**O**	**T**	**Total**
SW	77%	0%	0%	225	35%	25%	0%	225
VP	87%	0%	0%	225	70%	2%	0%	225
SS	96%	0%	0%	75	100%	0%	0%	75

CA = compensatory articulation; O = omission; T = Typical; SW = single words; VP = vehicle phrases; e SS = sentences with target sound in recurrence

Table 5 shows the accuracy distribution for speech stimuli (single words, phrases and vehicle phrases with target sound recurrence) with consonants /k/ and /g/ and Table 6 shows the distribution of p values of correct answers to such stimuli. More specifically, Table 6 presents the p values of the distribution of correct sounds for both consonants, /g/ and /k/, taking into account: (a) isolated word x phrases with target sound recurrence; (b) vehicle phrases x phrases with target sound recurrence and (c) single words x phrase vehicle. The results indicated a statistically significant difference between stimuli analyzed by both investigated consonants (/k/ and /g/).

TABLE 5 - distribution of correct sounds for the speech stimuli consonant /k/ and /g/

Samples	**/g/**		**/k/**	
	Acerto	**%**	**Acerto**	**%**
SW	**135**	**60%**	**174**	**77%**
VP	**161**	**71%**	**196**	**87%**
SS	**75**	**100%**	**72**	**96%**

SW = single words; VP = vehicle-phrases and SS = sentences with target sound in recurrence

TABLE 6 - p values of the distribution of correct sounds for the speech sounds of consonants /k/ and /g/

	/g/		**/k/**	
	SS	**VP**	**SS**	**VP**
VP	**<0,001***	**-**	**0,031***	**-**
SW	**<0,001***	**0,010*-**	**<0,001***	**0,007***

*Test of Equality of Proportions

SW = single words; VP = vehicle-phrases and SS = sentences with target sound in recurrence

Discussion

The study aimed to establish agreement in auditory-perceptual judgments from five judges to judge speech samples involving the consonants /k/ or /g/, produced by three children with cleft palate or operated palate presenting GS and three children with typical speech (control group). The results showed 100% agreement in judgments obtained for the three speech stimuli (single word, vehicle phrase sentence with target sound recurrence) produced by the control children. These data suggest that in the presence of typical speech, speech-language-pathologysts were consistent in their responses.

On the other hand, when analyzing the judgments of speech samples produced by children with cleft lip and/or palate, it was found that substantial obtained agreement was almost perfect (Table 1), according to the Kappa index. These findings may result from the selected professional experience to participate and, therefore, agree with other information derived from previous studies that point to the importance of including experienced judges in studies aimed at identifying CA[2,6,8,13]. They may also have been influenced by the good quality of the collected data (recording, editing and storage), aspect which is considered essential in studies involving task of perception[5]. The inclusion of reference samples using the audio examples of each response option (typical speech, segment omission and presence of OS) in the CD-ROM for the judges to consult, before and during the judgments, may also have favored the agreement in the judgments. The literature reports that the training offered prior to the execution of speech judgments may have favored the agreement of the assessors in the proposed tasks, especially when it was used training programs structured for this end[11]. Although in the study it was not offered training prior to the judgments, the reference provided samples may have contributed to the findings obtained.

Particularly for the stimulus consisting of words it was obtained percentage of agreement classified as substantial and for the stimulus consisting of words inserted in vehicle phrases agreement was almost perfect (Table 1). In the study by Gooch et al.[8] were also used controlled words consisting of consonant + vowel + consonant (CVC) inserted in a sentence-vehicle. The results obtained by Gooch et al.[8] showed lower inter-judges agreement. Differences in agreement reported by Gooch et al.[8] and in the present study can be justified on methodological issues, and particularly in relation to the task given to the judges. While in the study of Gooch et al.[8] the task was to record the presence and the type of CA through phonetic transcription, in this study the task was written down in an elaborate sheet for this purpose, one of the three types of responses: presence of typical speech, omission segment or presence of CA. Still, for the stimulation of speech consisting of phrases with sound-target recurrence, the percentage of agreement showed higher values (98%, kappa 0.99) when compared to the other stimuli (vehicle phrases = 78% and single words = 56%) (Table 2). These data suggest that in addition to the speech-language-pathologists' experience, the recurrence of the velar consonant favored the agreement in the judged samples, as this recurrence gave opportunity to judges to hear at least four times the same target consonant in the same sample speech. Previous studies support the use of phrases consisting of a single consonant (recurrence), inserted at different positions in the word[15,19,21] to measure the altered speech aspects, including CA. Seeking to establish consistency and uniformity to document and describe the results of talks after treatment of the CLP and VPD, Henningsson et al.[1] (2008) proposed basic parameters for sample collection and speech analysis in subjects with CLP independently of the spoken language. In the proposal presented by the authors suggested the use of words and phrases to target sounds in recurrence, in order to favor the identification of speech disorders presented by the population with CLP and/or VPD.

Whereas the higher or lower recurrence of the target consonant may interfere with auditory-perceptual judgments, it was also decided to analyze the degree of concordance for the three types of speech stimuli also taking into account each of the velar consonants. The results showed a rate of agreement (and Kappa index) with higher values for /k/ than for /g/ for isolated words and phrases to the vehicle. However, similar values for the two velar consonant were found when the stimulus constituted by the target sounds recurrence was judged (Table 3). These results suggest that in the presence of the sentence with sound-target recurrence, the voicing of the consonant did not interfere in the judgments. The opportunity of judges to hear

at least four times, depending on the target in the same sentence seemed to favor the achievement of agreement in judgments made for both consonants. Comparisons with the literature involving the type of velar consonant (if voiced or voiceless) became difficult due to methodological differences between the studies. For example, previous research used words subsidiaries consisting of consonant + vowel + consonant (CVC), inserted in a vehicle sentence. However, Gooch et al.[8] did not aim to assess the influence of the consonant voicing in identifying the presence of CA and types of stimuli, making it difficult for comparisons between literature and the findings of this study.

In particular, when analyzing the three types of stimuli (typical speech, omission of segment or presence of GS) it was taken into account each of the velar consonants, and there was agreement on the presence of GS in consonant /k/ in all of the three types of judged speech stimuli. This result differed from that obtained for the stimuli consisting of single words and phrases involving the vehicle consonant /g/, suggesting that the specific constitution of each phonetic environment of the sample velar consonant may have adversely affected the outcome of the judgment of the consonant /g/, raising questions for the judges. For example, in the word "cuca" (bogery) there was recurrence of the sound /k/ (which may have eased the judgment) and in the word "gula" (gluttony) there was not this recurring pattern, although it has followed by recommendation of previous study(1) for the development the stimulus. More specifically, there was less agreement (and therefore more questions) when the stimulus of less extent (isolated words) involving the consonant /g/ was judged as shown in Table 4.

The results obtained by analyzing the distribution of correct sounds (and p values) for the speech samples of interest for each of the velar consonants indicated statistically significant differences between samples (Tables 5 and 6). Jointly, these results suggest that the formation of speech stimuli can influence the agreement of the judgments, regardless of whichever velar consonant was investigated. As discussed previously, higher agreements were obtained for the stimulus consisting of target sound recurrence, suggesting that this sample favors the identification of the type of change (in particular, the CA) by the judges. These findings confirm the literature recommendations on the importance of using speech samples phonetically controlled and standardized in order to allow comparison of results[1,5], especially in multicenter studies[6].

The use of standardized protocols for the Brazilian Portuguese that meet the international recommendations[1] may ease the identification of changes in speech production of children with CLP and therefore promote inter-judges agreement. In the US and Europe, the use of standardized speech samples (composed of words and phrases to target sounds in recurrence) has been recommended in the last decade[1,6,19,21] in order to ease the identification of CA by speech-language pathologists and enable the exchange of information between centers. This practice has been established at the national level, by initiatives from BrasilCleft task force aimed at standardization and systematization of collecting and analyzing primary post-intervention results of CLP, from the integration of the Brazilian services that work with craniofacial anomalies. Thus, this study sought to obtain information about factors intrinsic to the different types of speech sounds that can influence the identification of CA. Future studies are needed in order to verify that higher agreements are obtained for speech stimuli consisting of target sound recurrence when standardized speech samples (such as those being proposed by BrasilCleft task force) are produced by a population bigger in children with CA associated with CLP.

Conclusions

Data from the study showed higher agreements for speech stimuli involving target sounds in recurrence when compared to other stimuli (vehicle phrase and single words), regardless of whichever velar consonant (/k/ or /g/). For clinical and research purposes where the activity involves the auditory-perceptual evaluation, we suggest the use of speech stimuli with velar sounds recurrence, since they may favor the identification of speech disorders associated with CLP in particular, the CA.

References

1.Henningsson GE, Kuehn DP, Sell D, Sweeney T, Trost-Cardamone JE, Whitehill TL. Speech parameters group. universal parameters for reporting speech outcomes in individuals with cleft palate. Cleft Palate Craniofac J. 2008;45(1):2-17.

2.Lohmander A, Olsson M. Methodology for perceptual assessment of speech in patients with cleft palate: a critical review of the literature. Cleft Palate Craniofac J. 2004;41(1):64-70.
3.Nagarajan R, Savitha VH, Subramaniyan B. Communication disorders in individuals with cleft lip and palate: An overview. Indian J Plast Surg. 2009; 42(3):137-43
4.Peterson-Falzone SJ, Trost-Cardamone JE, Karnell MP, Hardin-Jones MA. The clinician's guide to treating cleft palate speech. St. Louis: Mosby; 2006.
5.Sell D. Issues in perceptual speech analysis in cleft palate and related disorders: a review. Int J Lang Commun Disord. 2005;40(2):103-21.
6.Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutters B. Methodology for speech assessment in the Scandcleft project--an international randomized clinical trial on palatal surgery: experiences from a pilot study. Cleft Palate Craniofac J. 2009; 46(4): 347-62.
7.Santelman L, Sussman J, Chapman K. Perception of middorsum palatal stops from the speech of three children with repaired cleft palate. Cleft Palate Craniofac J. 1999; 36(3): 233-42.
8.Gooch J, Hardin-Jones M, Chapman KL, Trost-Cardamone JE, Sussman J. Reliability of listener transcriptions of compensatory articulations. Cleft Palate Craniofac J. 2001; 38(1): 59-67.
9.Chapman KL, Hardin-Jones MA, Goldstein JA, Halter KA, Havlik RJ, Schulte J. Timing of Palatal Surgery and Speech Outcome. Cleft Palate Craniofac J. 2008; 45(3): 297-308.
10.Lohmander A, Persson C. A longitudinal study of speech production in Swedish children with unilateral cleft lip and palate and two-stage palatal repair. Cleft Palate Craniofac J. 2008; 45(1): 32-41.
11.Sell D, John A, Harding-Bell A, Sweeney T, Hegarty F, Freeman J. Cleft Audit Protocol for Speech (CAPS-A): a comprehensive training package for speech analysis. Int J Lang Commun Disord. 2009; 44(4): 529-48.
12.Nyberg J, Westberg LR, Neovius E, Larson O, Henningsson G. Speech results after one-stage palatoplasty with or without muscle reconstruction for isolated cleft palate. Cleft Palate Craniofac J. 2010; 47(1): 92-103.
13.Brunnegard K, Lohmander A. A cross-sectional study in 10-year-old children with cleft palate: results and issues of rater reliability. Cleft Palate Craniofac J. 2007; 44(1): 33-44.
14.Klinto K, Salameh EK, Svensson H, Lohmander A. The impact of speech material on speech judgement in children with and without cleft palate. Int J Lang Commun Disord. 2011; 46(3): 348-60.
15.Sell D, Harding A, Grunwell P. GOS.SP.ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). Int J Lang Commun Disord. 1999; 34(1):17-33.
16.Peterson-Falzone SJ, Hardin-Jones MA, Karnell MP. Communication disorders associated with cleft palate. Cleft Palate Speech. 3rd ed. St. Louis: Mosby, 2001:162-98.
17.Kummer AW. Velopharyngeal dysfunction and resonance disorders. In: ___ (ed). Cleft palate & craniofacial anomalies: effects on speech and resonance. San Diego:Singular. 2001:145-76.
18.Boersma P, Weenink D. Praat: doing phonetics by computer. Download da versão 5.3.53 em julho de 2013.
19.Prathanee B, Lorwatanapongsa P, Makarabhirom K, Suphawatjariyakul R, Manochiopinig S, Thinnaithorn R. Reliability of perceptual assessment for resonance disorders in patients with cleft palate. J Med Assoc Thai. 2012; 95(11): S73-9.
20.Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1): 159-74.
21.Sweeney T, Sell D. Relationship between perceptual ratings of nasality and nasometry in children/adolescents with cleft palate and/or velopharyngeal dysfunction. Int J Lang Commun Disord. 2008; 43(3):265-82.