



## **Scandcleft Project Trial 2 Comparison of speech outcome in one and two-stage palatal closure in 5-year-olds with UCLP**

Lundeborg, Inger ; Nyberg, Jill; Alaluusua, Suvi ; Willadsen, Elisabeth; Persson , Christina; Lohmander, Anette

*Published in:*  
Cleft Palate - Craniofacial Journal

*DOI:*  
[10.1177/1055665619888316](https://doi.org/10.1177/1055665619888316)

*Publication date:*  
2019

*Document version*  
Peer reviewed version

*Citation for published version (APA):*  
Lundeborg, I., Nyberg, J., Alaluusua, S., Willadsen, E., Persson, C., & Lohmander, A. (2019). Scandcleft Project Trial 2 Comparison of speech outcome in one and two-stage palatal closure in 5-year-olds with UCLP. *Cleft Palate - Craniofacial Journal*. <https://doi.org/10.1177/1055665619888316>

# Scandcleft Project Trial 2—Comparison of Speech Outcome in 1- and 2-Stage Palatal Closure in 5-Year-Olds With UCLP

Inger Lundeborg Hammarström, SLP, PhD<sup>1</sup>, Jill Nyberg, SLP, PhD<sup>2</sup>, Suvi Alaluusua, SLP, MA<sup>3</sup>, Jorma Rautio, MD<sup>3</sup>, Erik Neovius, MD, PhD<sup>2</sup>, Anders Berggren, MD<sup>4</sup>, Christina Persson, SLP, PhD<sup>5</sup>, Elisabeth Willadsen, SLP, PhD<sup>6</sup>, and Anette Lohmander, SLP, PhD<sup>7</sup>

## **ABSTRACT**

*Objective:* To investigate in-depth speech results in the Scandcleft Trial 2 with comparisons between surgical protocols and centers, and with benchmarks from peers without cleft palate.

*Design:* A prospective randomized clinical trial

*Setting:* Two Swedish and one Finnish Cleft Palate center.

*Participants:* The 112 participants were 5-year-old born with unilateral cleft lip and palate randomized to either lip repair and soft palate closure at 4 months, and hard palate closure at 12 months or lip repair at 3-4 months (Arm A) or a closure of both the soft and hard palate at 12 months (Arm C).

*Main outcome measures:* A composite measure dichotomized into velopharyngeal competency (VPC) or velopharyngeal incompetency (VPI), overall assessment of velopharyngeal function (VPC-Rate), Percentage of Consonants Correct (PCC-score), and consonant errors. In addition, number of speech therapy visits, average hearing thresholds, and secondary surgeries were documented to assess burden of treatment.

*Results:* Across the trial 53.5% demonstrated VPC and 46.5% VPI with no significant differences between arms or centers. In total, 27% reached age appropriate PCC-scores with no statistically significant difference between the arms. The Finnish center had significantly higher PCC-scores, the Swedish centers higher percentages of oral consonant errors. Number of speech therapy visits was significantly higher in the Finnish center.

*Conclusion:* At age 5, poor speech outcomes with some differences between participating centers were seen but could not be attributed to surgical protocol. As one center had very few participants, the results from that center should be interpreted with caution.

Trial registration: ISRCTN29932826

**Keywords:** randomized clinical trial, primary palatal repair, velopharyngeal competence, consonant proficiency, intercenter study, crosslinguistic

<sup>1</sup> Division of Speech and Language Pathology, Linköping University, Linköping, Sweden

<sup>2</sup> Stockholm Craniofacial Team, Karolinska University Hospital, Stockholm, Sweden

<sup>3</sup> Cleft Palate and Craniofacial Center, Helsinki University Central Hospital, Helsinki, Finland

<sup>4</sup> Department of Plastic Surgery, University Hospital, Linköping, Sweden

<sup>5</sup> Speech and Language Pathology Unit, Institute of Neuroscience and Physiology, University of Gothenburg, Sweden

<sup>6</sup> Department of Nordic Studies and Linguistics, University of Copenhagen, Denmark

<sup>7</sup> Division of Speech and Language Pathology, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden

Corresponding Author:

Inger Lundeborg Hammarström, IKE/Division of Speech Language Pathology, Linköping University, Campus US, S-581 83 Linköping, Sweden. Email: inger.lundeborghammarstrom@liu.se

## INTRODUCTION

Cleft lip and palate is one of the most common birth defects. The overall incidence of orofacial clefts is about 1 to 2.21 cases per 1000 live births, with a variability between countries and ethnic groups (Derijcke et al., 1996). In Finland the incidence of clefts is 2.56 cases per 1000 live births and abortions (Lithovius et al., 2014), and the incidence in Sweden is about 2/1000 births (Hagberg et al., 1998). Adequate facial growth and optimal speech development are important outcomes of palatal surgery in children born with cleft palate (CP) (Reddy et al., 2017). However, as a consequence of weak evidence regarding outcomes after primary palatal surgery there is no consensus regarding preferable protocol(s) (Peterson Falzone et al., 2010; Lohmander, 2011). Reddy et al. (2017) reviewed the results of comparisons of one- and two-stage surgical procedures to repair cleft palate reported in studies from 1960-2015. None of the 26 studies were prospective or included randomized comparison, which led the authors to conclude that there was no clear evidence of a relationship between type of surgical procedure and facial growth, speech and fistulas in patients born with unilateral cleft lip and palate. Recently, the outcome of three parallel multicenter randomized trials, the Scandcleft project, was published. The project was developed and executed by ten North European cleft teams divided into three trials investigating the impact of surgical protocol in children born with unilateral cleft lip and palate (UCLP) on outcome at 5 years of age (Semb et al., 2017). Participating centers in Trial 1 were two Danish and one Swedish, in Trial 2, two Swedish and one Finnish, and in Trial 3 two Norwegian and two British. The distribution of the centers into each trial was based on which local protocol they used. Furthermore, one surgical protocol served as a common method and was compared to the local protocol in each trial (Semb et al., 2017). The common surgical protocol was an early two-stage procedure with soft palate repair at 3–5 months of age and hard palate closure at 12 months of age (Rautio et al., 2017). In the present study,

more in-depth analyses were conducted to examine the influence of one- versus two-stage palatal repair on velopharyngeal closure and consonant proficiency, and the findings were compared to speech data obtained from children without cleft palate.

Prevalences of speech deviances between 30% and 50% are commonly reported and are in line with findings from a critical review of 34 articles describing speech results after different surgical protocols published from 1984-2010 (Lohmander, 2011). All results were based on assessments from audio recordings and showed that approximately 60-70% of the children had normal speech at age 5. The literature review was expanded by Britton et al. (2014) who also developed the following benchmarks: at least 50% of children at 5 years of age should have normal speech, 70% should be without evidence of structurally related problems of VPI or secondary surgery for VPI or fistula, and at least 50% should have no cleft speech errors requiring intervention. The study included a national British audit on 1100 children, which showed that 48% had speech within the normal range, 66% had no evidence of structurally related speech problems, and 60% had no serious cleft-related articulation errors (Britton et al., 2014). Thus, the mean outcome did not reach the benchmarks but there were differences between centers. However, no comparison of speech outcome after one- or two-stage surgical protocols was performed either in any of the included studies in the review, or in the national audit (Britton et al., 2014).

Speech outcome in consecutive series of children born with unilateral cleft lip and palate treated according to same surgical protocols are rare. We have found only one where the speech assessment had been performed by multiple listeners from recordings and reliability reported (Nyberg et al., 2014). The researchers evaluated speech outcome in 69 children born with UCLP treated according to the same one-stage protocol as the local protocol in Trial 2 (Nyberg et al., 2014). Perceptual findings suggested that 22% of the children had moderate to

severe velopharyngeal incompetence. Seventeen percent had already received pharyngeal flaps, and 30% had compensatory articulation at 5 years of age. In comparison to peers without cleft palate children with UCLP have been found to have significantly poorer speech at age 5 years (Lohmander and Persson, 2008; Klintö et al., 2016).

### *Cleft palate speech*

Speech errors related to a cleft palate condition are described as cleft palate speech and include both atypical resonance and nasal emission, and articulatory errors (Hutters and Brøndsted, 1987; Grunwell and Sell, 2001; Howard and Lohmander, 2011). Velopharyngeal insufficiency (VPI) compromises the normal balance between the oral and nasal cavity resulting in hypernasality, audible nasal escape, velopharyngeal friction sound and/or reduced pressure on high pressure consonants (Sweeney, 2011). These passive speech characteristics are also commonly described symptoms related to fistulae may occur after cleft palate repair when there is a breakdown of tissue during the healing process or are sometimes deliberately left at the surgery (Grunwell and Harding, 1998). The resulting speech errors are the unavoidable consequence of an oral-nasal communication when no effort is made by the speaker to lessen it, (Hutters and Brøndsted, 1987; Grunwell and Harding, 1998). Passive speech characteristics can also be related to other causes such as a too short velum after palatal closure (Chapman and Willadsen, 2011). Active speech errors, on the other hand are attributed to consonants being realized with a change in place and/or manner often in order to compensate for the lack of adequate intraoral air pressure for the pressure consonants (Hutters and Brøndsted, 1987; Grunwell and Harding, 1998).

### *Speech outcome measures*

One way of assessing articulatory proficiency is by using the measure Percentage of Consonants Correct (PCC). PCC is a quantitative measure originally proposed by Shriberg

and Kwiatkowsky (1982). They measured the proportion of correctly articulated consonants in phonetic transcriptions of conversational speech to assess the severity of involvement. Single-word material may also be used, provided that results are not related to the classification of severity of involvement based on spontaneous speech by Shriberg et al. (1997). Calculating the PCC in single word samples has been used to assess articulation skills in typically developed children without cleft palate (Dodd et al., 2003) and in children born with cleft palate (Lohmander and Persson, 2008; Scherer et al., 2008; Klintö et al., 2014). A restricted number of phonetically similar speech sounds enhances the validity of speech outcomes reported from different languages if the assessment is made primarily on speech sounds that are highly vulnerable to the cleft condition, that is, pressure consonants (Hutters and Henningsson, 2004). Also, embedding the target sound in a single-word context enhances reliability between raters (Klintö et al., 2011). Such a speech material will be short but sufficient for reporting speech outcome following surgical treatment (Lohmander et al., 2009). In a study of PCC-scores in typically developing 5-year-old Swedish speaking children without CP, using almost the same speech material as in the present study, a mean PCC-score of 96.3% was obtained (Lohmander et al., 2017b).

The reliability of perceptual judgements of atypical nasal resonance and nasal emission varies in different studies and hypernasality is considered the most difficult dimension to judge reliably (Watterson et al., 2007). In some the agreement is good (e.g. Sell et al., 2015), in others poor or moderate (e.g. Brunnegård et al., 2007; Nyberg et al., 2014). It is therefore important to strive to increase levels of agreement which can be achieved by focused training (John et al., 2006). Also, by the use of a method that allows the rater to stratify hypernasality in categories, reliability can be enhanced (Yamashita et al., 2018). An additional or alternative way of increasing inter-rater reliability is to use a composite velopharyngeal score

representing a sum of all variables associated with VPI. The use of such a score has been shown to correlate well with overall perceptual ratings of velopharyngeal function (Dotevall et al., 2002; Lohmander et al., 2009). An overall rating of velopharyngeal function (VPC-Rate) was recently shown to be a good predictor of a more complex composite score (VPC-Sum) (Lohmander et al., 2017c).

### *The Scandcleft study*

The speech assessment in the Scandcleft project was based on standardized methods developed within the project (Hutters and Henningsson, 2004; Lohmander et al., 2009; Lohmander et al., 2017a; Willadsen et al., 2017). The main results from the Scandcleft project reported recently, revealed no significant differences in VPC related to surgical protocol, that is, neither different staging nor timing or sequencing of palatal repair (Lohmander et al., 2017a). Differences in timing of hard palate repair after early soft palate closure at 3–5 months of age in the two-stage procedure, revealed a significantly higher occurrence of retracted oral consonant errors in the group with late hard palate repair (3 years) compared to early (1 year) (Willadsen et al., 2017).

**Table 1.** Patient numbers, participating centers, languages and surgical procedures in the three trials of the Scandcleft project

Trial no	Participating center	Language	Surgical procedure	Arm and no of patients
<b>1</b> (n=143/148)	1*, 2*, 3	Danish, Swedish	Lip and soft palate closure at 3-4 months, hard palate closure at 36 months	B, n=71
<b>2</b> (n=112/151)	4,5,6	Finnish, Swedish	Lip closure at 3-4 months, hard and soft palate closure at 12 months	C, n=55
<b>3</b> (136/149)	7,8,9,10	Norwegian, English	Lip closure at 3-4 months in combination with a single layer closure of the hard palate using a vomer flap and soft palate closure at 12 months	D, n=70
<b>Common</b>	To be carried out of all centres		Lip and soft palate closure at 3-4 months, hard palate closure at 12 months	A, n=195

\* All patients were operated in one center, but follow-ups were carried out at two centers. n total recorded/total

Early palatal closure of cleft palate has been advocated in the literature to enhance conditions for speech development (Kemp-Fincham et al., 1990). Whether the cleft palate should be closed in one or two stages is not clear, however. While Rohrich et al. (2000) recommended an early two-stage procedure with early soft palate closure at 6 months and hard palate repair at approximately 15 months of age, detailed speech outcome has not yet been reported. Thus, detailed speech results after a two-stage protocol with early soft palate repair and hard palate closure at 12 months compared with a one-stage protocol at around 12 months of age would be valuable. Finally, it is also important to compare the speech outcome with speech in peers without cleft palate.

### *Aim and research questions*

The aim of the present study was therefore to compare detailed speech outcome in 5-year-olds between the common surgical protocol in the Scandcleft project with lip repair and soft palate closure at 4 months, and hard palate closure at 12 months, and the local protocol used by all three centers in Trial 2 with lip repair at 3-4 months and a closure of both the soft and hard

palate at 12 months. Comparisons of outcome were made within and between the included centers.

The following research questions were addressed:

- Is there a difference in perceived characteristics associated with velopharyngeal function between the surgical protocols used and centers?
- Is there a difference in consonant proficiency measured by percentage of consonants correct (PCC) and in consonant errors between the surgical protocols used and centers?
- How does consonant proficiency of participants compare to benchmarks from peers without CP?
- Is there a relationship between perceived characteristics associated with velopharyngeal function or consonant proficiency and secondary pharyngeal surgery or between consonant proficiency and number of speech therapy visits?

## **METHODS**

### *Participants*

A total of 151 participants were recruited during 1997 - 2006 and randomized to the Scandleft Trial 2. Out of these 112 (35 girls and 77 boys) were included in this study. Three centers participated, two in Sweden (Center 5 and 6) and one in Finland (Center 4). Inclusion criteria were: being Caucasian, born with a non-syndromic unilateral cleft lip and palate (accepting a soft tissue bridge of  $\leq 5$  mm), living in Sweden or Finland since birth, one parent speaking the native language, Swedish or Finnish with the child, and Swedish or Finnish spoken as the mother tongue by the child. Also, the surgeons made a decision whether it was technically possible to perform the operation according to protocol. Exclusion criteria were

soft tissue bridge > 5mm, additional anomalies, or syndromes. For the current study of speech outcome, available audio and video recordings of good quality was an additional inclusionary criterion. This was not fulfilled for 39 participants. Distribution of participants between centers and arms is displayed in Table 2.

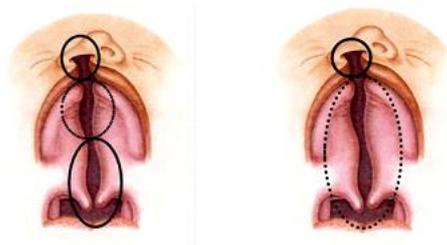
**Table 2.** Participating centers, languages and number of participants, and exclusions in each arm in Trial 2. Arm A (the common surgical method in the Scandleft project): lip repair and soft palate closure at 4 months, and hard palate closure at 12 months, Arm C (the local method) lip repair at 3-4 months and a closure of both the soft and hard palate at 12 months

Center no.	Language	Arm A no	Arm C no	Total no. of participants	Excluded due to missing/ low quality speechdata	Arm A good quality speech data	Arm C good quality speech data	Included no. of participants
4	Finnish	43	50	93	33	30	30	60
5	Swedish	12	9	21	4	9	8	17
6	Swedish	18	19	37	2	18	17	35
Total		73	78	151	39	57	55	112

### *Primary Surgery*

In Trial 2 the Scandleft common protocol, a two-stage surgical protocol, including lip and soft palate closure at 3-4 months and hard palate closure with a vomer flap at 12 months, Arm A (n=57) was compared with the local one-stage protocol (Arm C) with lip closure at 3-4 months, hard and soft palate closure at 12 months (n=55), which was the typical surgical protocol for the three participating centers (Figure 1). Five surgeons (two in Center 4, one in Center 5 and two in Center 6) performed all surgeries. Before starting the study, all participating surgeons attended two sessions involving actual surgery with the new protocol (Arm A). The surgeons all met annually for a refresher course based on clinical slides to emphasize adherence to the agreed protocol. A detailed description of the surgical procedures can be found in Rautio et al. (2017).

Ethical approval was obtained from the appropriate ethical committee in each country and the study is registered in the ISRCTN (Trial registration: ISRCTN29932826).



**Figure 1.** Surgical methods within the trial. Arm A, the Scandcleft common method, with lip and soft palate closure at 4 months and hard palate closure at 12 months; and Arm C, the local protocol, with lip closure at 4 months, hard and soft palate closure at 12 months.

### *Secondary pharyngeal surgery and palatal fistulae*

Out of the 10 participants who received secondary pharyngeal surgery (Figure 2), five participants had palatal re-repairs, all in Center 4, three in Arm A and two in Arm C and five participants had pharyngeal flap surgery, two in center four and three in Center 6, all in Arm A. The mean age for the operations was 4.27 years (3.15–4.97). A total of 11 participants had had fistula surgery, seven in Center 4 (six in Arm A), one in Center 5 (Arm A) and three in center 6, all in Arm C.

### *Hearing*

Pure tone audiometry was completed on both ears in 53 participants (46%) at the time of the 5 year assessment, Average hearing thresholds on four frequencies (500, 1000, 2000, and 4000 Hz) were calculated for the best ear and hearing loss was defined as mild (21-40 dBHL), moderate (41-60 dBHL) or severe (>60 dBHL). A total of 20 children (38%) had some degree of compromised hearing, of which 10 had binaural hearing loss. In total 10 out of the 53 participants had mild conductive hearing loss in the best ear and none had moderate or severe hearing loss. In the worst ear, 14 participants had mild hearing loss, six had moderate and

none had severe hearing loss. All the children with binaural hearing loss had mild problems (21-40 dBHL).

### *Speech therapy*

Information on the total number of *speech therapy visits* including all types, i.e. routine, review and therapy up to 5 years of age with the cleft site speech and language pathologist/therapist (SLP/T) and local SLP/T was collected. The total number by center and arm are presented in Table 3.

**Table 3.** Speech therapy visits (mean, standard deviation (SD), Median and Range) distributed by Center and Arm.

	Center 4		Center 5**		Center 6**	
	A (n=27)*	C (n=27)*	A (n=9)	C (n=8)	A (n=18)	C (n=17)
Mean (SD)	12.8 (0.5)	11.2 (0,5)	13.1 (2)	14.1 (6)	18.1 (6)	14.4 (0)
Median	5	5	8	14	18.5	15
Range	2-94	0-55	4-32	5-25	6-44	5-29

\*Data missing from 6 children in Center 4, \*\* The differences were significantly different according to Kruskal Wallis test ( $p=.000$ )A posthoc analysis(Mann Whitney U test) revealed a significantly higher number of speech therapy visits in center 5 ( $Z=-2.86$ ,  $p=.004$ ) and in center 6 ( $Z=-4.60$ ,  $p<.001$ ) compared to center 4 ( $p<.001$ ) but no difference between center 5 and 6.

## **Procedure**

### *Speech material*

A restricted single word test (a picture naming test) was developed for the two included languages and contained the most similar consonants in order to be comparable. Each target consonant was presented in word-initial strong stressed position in three different words. In addition, three words with /s/ in final position were included. Since Finnish does not include /f/ as a phoneme, and only the unvoiced stops /p t k/ in word-initial position, the Finnish single word test was constructed with six target words for the stops /p t k/, and only 30 words were included as compared to 33 words for Swedish. The first nine words in the test included high or semi-high vowels in similar phonetic context. Finally, retelling of the Bus story test

(Renfrew, 1969) was used.

### *Data collection and editing*

The 112 participants were audio and video recorded according to a method developed and standardized for data collection, including comparable speech material and elicitation (Lohmander et al., 2009; Lohmander et al., 2017a; Willadsen et al., 2017). The recordings were performed in a quiet room at the clinic at each one of the centers, with a super-VHS or a digital video camera or a digital video camera with external microphone of excellent quality (Sony ECM-MS 907). Simultaneously audio was recorded with a similar high quality microphone. The two microphones were placed approximately 40 cm from the child's mouth. The participant was asked to name a picture, and if unsuccessful, the SLP/T first asked if the picture represented X (the target word) or Y (a completely different word), if this also was unsuccessful, the SLP/T asked the child to repeat the target word. The procedure took approximately 40 minutes. All video recordings were later edited in Corel Video-Studio Pro X4 (Corel Corporation, Ottawa, DA). For the analysis and transcription of consonant proficiency and errors, the video recordings of the naming tests were edited so that each test word produced by the participant and repeated by the SLP/Ts was saved as a separate video clip (.mpeg format) to facilitate the assessment procedure. Duplications were made of 30% of the video clips for calculation of reliability.

For the assessments of hypernasality, the audio recordings of nine words with high vowels in each language were edited to a nine-word-string and saved in .wav format. For further details, see Lohmander et al. (2017a) and Willadsen et al. (2017). The audio recording of the retelling of the Bus story was saved as a .wav.file for each participant for assessment of overall velopharyngeal function (VPC-Rate) in connected speech.

### *Training*

In order to be able to evaluate the outcome measures in a reliable way, 30 SLP/Ts in the Scandleft speech group who had collected the material met every second year during the recruitment and data collection process (1999 – 2009) to discuss, and develop practice rules for transcription and rating of hypernasality (Lohmander et al., 2009). A final test for phonetic transcription, as well as for rating of hypernasality, was carried out in 2011 before the final data analysis of all the Scandleft material took place. The SLP/Ts with the highest intra-rater agreement (n=17; native speakers of British, Swedish, Danish, Norwegian, and Finnish) from the group of 30, performed the final phonetic transcriptions, ratings of hypernasality and the overall rating of velopharyngeal function.

### *Perceptual assessment*

The final data analysis took place after all participants had had their 5 year-check up at their cleft center where the data was collected. The assessments were performed during two sessions with four weeks in between. Each session lasted seven days.

### *Hypernasality*

The separate .mpeg files of each target word produced by each child were transcribed and the audio-recordings with the strings of the high vowel words were rated for hypernasality by three SLP/Ts: two speaking the same language as the child, but not from the same center, and one (Swedish) SLP/T with the highest intra-rater reliability rated all strings from both languages. The raters were not familiar with the children and they were also blinded to which arm the child was randomised. The Finnish-speaking children were evaluated by a SLP/T from the same center. However, this person had not been involved in the data collection of the 5-year-olds at the center. The files were presented from a laptop and the same type of high

quality headphones (Creative Aurvana Live, Creative Technology Ltd., Singapore) was used. The SLP/Ts were allowed to watch and listen to the clips as many times as they wished. For the analysis of target consonant realization narrow phonetic transcription was made using the International Phonetic Alphabet (IPA) and its extensions for disordered speech (ExtIPA). The hypernasality rating was performed from audio recordings in two steps. First the raters determined whether resonance was within normal limits and if not, the deviance was rated as either mild, moderate or severe. No written descriptions on the procedure were provided, but the SLP/Ts were all familiar with both procedure and terminology based on their previous consensus training.

#### *Overall assessment of velopharyngeal function, the VPC-Rate*

Velopharyngeal function was assessed by overall rating on a 3-point ordinal scale (VPC-Rate), where (0) was competent, (1) was marginal competent with evidence of minor problems suggesting borderline closure, or (2) incompetent with evidence of significant problems usually requiring surgical management (McWilliams et al., 1981; Karnell and Van Demark, 1986; Lohmander et al., 2017a). The VPC ratings were made from audio recordings at 5 years of age, but the assessment was undertaken two years after the main assessment in the trial. The ratings were performed independently by two listeners from each language, Swedish and Finnish who were not familiar with the children. The speech sample was an audio recording of the child's retelling of the Bus story in a Swedish and Finnish version.

#### *Phonetic transcription*

All assessments were performed individually and independently, and each child was assessed by two SLP/Ts. They could replay the files an unlimited number of times and with breaks when needed. They transcribed the target consonants in the single words from the video

recorded naming test on a laptop, wearing high quality headphones (Creative Aurvana Live, Creative Technology Ltd., Singapore). Narrow phonetic transcription was used according to the International Phonetic Alphabet with Extensions (ExtIPA, 2008).

## *Analysis*

### *Consonant proficiency*

All transcribed consonants that were judged as being correct were summed up and divided by the total number of occurring consonants to obtain the PCC-value. The *PCC-scores* were benchmarked against two PCC levels: > 90% corresponding to 0-3 errors, and < 79% corresponding to eight or more errors. The first level (>90%) is equivalent to -1SD, and the second (<79%) equivalent to -3SD in published reference data on similar restricted speech material (in Swedish) (Lohmander et al., 2017b). Since /s/-errors represented a high proportion of all consonant errors, the influence of the sound on articulatory proficiency was investigated by another calculation of PCC in which /s/-distortions and were counted as correct.

### *Consonant errors*

All transcribed consonant errors were categorized as either *developmental speech characteristics* (DSCs) or *cleft speech characteristics* (CSCs). The latter were then divided into active or passive CSCs. The active CSCs were subdivided into *non-oral* and *oral* CSCs. Non- oral CSCs included glottal and pharyngeal productions, active nasal fricatives, and nasal substitutions for plosives and fricatives. Active oral CSCs or retracted oral articulation refer to dental and alveolar consonants articulated at a retracted place within the oral cavity and realized as palatal, velar or uvular consonants. Passive errors, that is nasal emission, velopharyngeal friction, and weak pressure in unvoiced obstruents and nasalization of voiced

obstruents, were not rated as consonant errors if they otherwise were produced as the target sound with regards to place and manner of articulation. By this, the passive errors did not affect the PCC-scores that were calculated. They were instead registered to assist the evaluation of velopharyngeal incompetence and were included in the composite score used for this purpose (see above).

### *Velopharyngeal function*

A composite outcome measure, (VPC-Sum) was created based on inclusion of the ratings of hypernasality, the transcriptions of active non-oral consonant errors and the passive VPI symptoms (passive CSCs). The presence of an error type was divided into three categories based on the number of consonants affected; score 0 = less than three occurrences, score 1 = three to five occurrences and score 2 = six occurrences or more. Hypernasality got new scores based on the ratings; no presence = 0, mild degree = 1, moderate to severe degree = 2.

Thus, the new score (VPC-Sum) could be a maximum of 6 (maximum 2 each from the variables hypernasality, active non-oral errors and passive VPI-symptoms). For a more detailed description, see Lohmander et al. (2017a; 2017c). In this study the *VPC-Sum* is reported dichotomized as follows; 0-1 = *Velopharyngeal competence* and 2-6 =

*Velopharyngeal incompetence*.

The overall *VPC-Rate* measure is the calculated percentage of children falling into the scale value zero (0) on the overall rated velopharyngeal function from connected speech.

### ***Reliability***

The reliability calculation was performed on the speakers of Finnish or Swedish. As there was only one transcriber speaking Finnish, a Swedish speaking SLP/T also transcribed the Finnish material and the material was used for calculation of reliability. However, the results of the transcription of the Finnish material were based on the transcriptions from the Finnish speaking transcriber. However, the overall assessment of the perceived velopharyngeal function (VPC-Rate) was undertaken two years after the main assessment in the trial and performed independently by two Finnish speaking raters. One of them had previously also performed the transcriptions at the main assessment, whereas the second was on leave at that time.

Reliability of the transcription was assessed for each target consonant as percent exact agreement (Lim et al., 2012) between each two pairs of SLP/Ts. A random selection of 30 % of the material was re-transcribed four weeks after the first assessment to check for intra-rater reliability. The agreement was 67.9-100% for the Finnish material (mean 90%) and 71.9-100% (mean 94.2%) for the Swedish material. Inter-transcriber agreement was calculated from transcriptions on 100% of the material and ranged from 50-100% (mean 84.8%) for the Finnish material (one Finnish speaking and one Swedish speaking transcriber) where 17.2% of the transcriptions were below 70%. For the Swedish material the means were 86.1% (Center 5) and 83.9% (Center 6) with the range 71-100%. The data used in the analysis was the transcriptions from the transcriber with the highest reliability.

Inter-rater reliability for VPC-Rate, based on all material between two raters, was moderate (Weighted kappa score .52) for the Finnish and good (Weighted kappa score .73) for the Swedish (Fleiss and Cohen, 1973). Presented results are based on the ratings from the rater with the highest intra-rater reliability of hypernasality of the same children which was a weighted kappa score of 0.84 for the rater of the Finnish material and of 0.73 for the rater of the Swedish material (Lohmander et al., 2017a).

### ***Statistical analysis***

Descriptives were calculated and differences compared between the two arms and the centers in the trial. Due to skewed distribution and categorical data non parametric methods were used: Pearson Chi-square statistical analysis for comparisons between arms, Fisher's exact test for comparisons within and between centers, Kruskal-Wallis test and Mann-Whitney U-test for comparisons across centers. Univariate analyses between the outcome measures and the independent variables were performed with Spearman's rank correlation test. *P*-values less than .05 were considered statistically significant. All analyses were made using SPSS© Windows version 24 (SPSS INC. Chicago, IL).

## **RESULTS**

### ***Velopharyngeal function***

#### ***VPC-Sum***

In total 53.5% of all children (53% in Arm A and 54% in Arm C) were judged as having competent velopharyngeal function (VPC-Sum=0-1). Slight differences between centers and arms were seen but no difference was statistically significant.

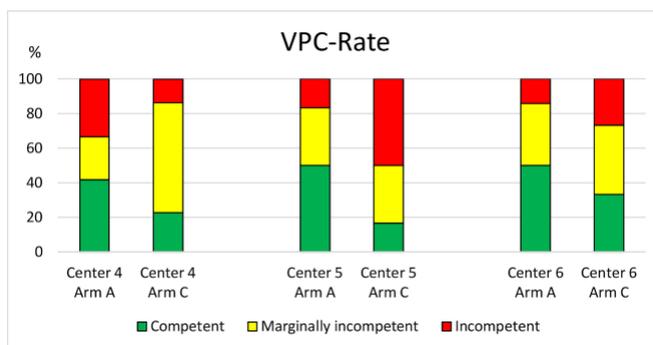
In total 46.5% of the children (47% in Arm A and 46% in Arm C) were judged as having incompetent velopharyngeal function (VPC-Sum=2-6). The differences between arms and centers were not statistically significant.

Six of ten children who had had secondary surgery was judged as having incompetent velopharyngeal function (VPC-Sum=2-6). Thus, another four children could be considered as having had primary VPI (50% of all participants). When comparing the number of children

judged as having primary VPI across centers and arms no statistically significant differences were seen.

### *VPC-Rate*

Thirty-six percent had competent velopharyngeal function as rated in connected speech, while 40% had marginally incompetent velopharyngeal function and 24% had incompetent velopharyngeal function. In Arm A 43% of the participants had a competent velopharyngeal function compared to 29% in Arm C. However, this difference was not statistically significant ( $Z=-0.19$ ,  $p=.85$ ). Center 6 had the highest prevalence of competent velopharyngeal function in both Arm A (54%) and in Arm C (41%) (Figure 2). There was no statistically significant difference across centers within Arm A and Arm C.



**Figure 2.** The proportions of the different overall ratings of velopharyngeal competence (VPC-rate) per centers and arms.

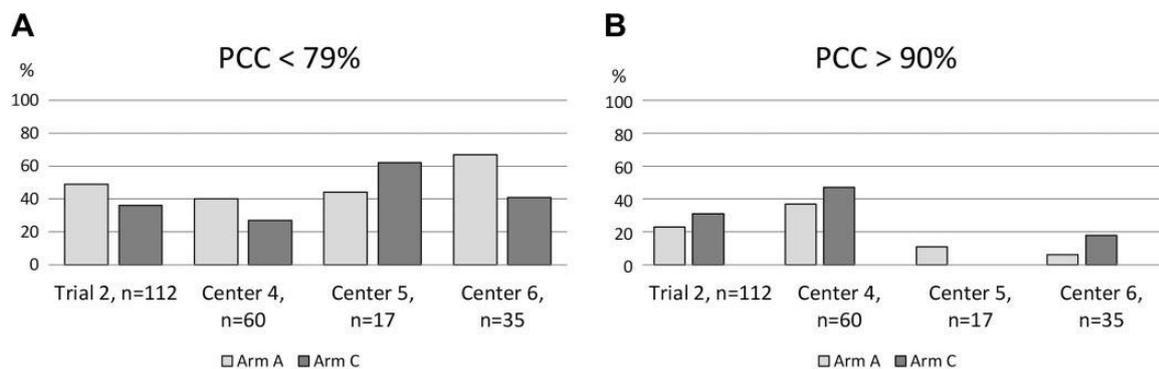
### *Consonant proficiency*

#### *Percentage of consonants correct (PCC) and benchmarks for PCC scores*

The mean PCC scores ranged between 65-79% with a mean of 70.7% (SD 22.5). The scores for PCC without s-distortions ranged from 71-87% with a mean of 78.3 (SD 22.5). No statistically significant difference between arms within centers was seen.

The percentage of children with a PCC-score below 79% in the whole group was 43% (49% in Arm A and 36% in Arm C). The difference between the arms was not statistically significant. The lowest percentage was found in Center 4 Arm C (Figure 3). There were no statistically significant differences between or within centers, however.

In total 27% of the children (23% in Arm A and 31% in Arm C) reached an age appropriate PCC level, that is > 90% (0-3 errors). The difference between the arms was not statistically significant. The best outcome was found in Center 4, Arm C (47 %) with a significantly higher percentage of children achieving a PCC-score > 90% than in Center 5 (Z = -2.399 p=.045) and 6 (Z -1.968 p=.049) (Figure 3).



**Figure 3.** Percentage of children with a PCC score <79% (Part A) and (Part B) >90% presented by centers and arms. PCC indicates percentage of consonants correct.

### ***Consonant errors***

#### *Developmental speech characteristics (DSCs)*

The proportion of participants having three or more DSCs was 28.3% in Center 4, 41.2% in Center 5 and 62.8% in Center 6. The difference between Center 4 and Center 6 was statistically significant (Z= -3.282, p<.001.)

### *Cleft speech characteristics (CSCs)*

The proportion of participants having three or more active CSCs was 26.7% in Center 4, 23.5% in Center 5 and 14.3% in Center 6. The differences between centers were not statistically significant ( $\chi^2(2) = 4.675, p=.097$ ) in the whole group or within Arm A ( $\chi^2(2) = 0.124, p=.94$ ). However, the differences between centers within Arm C were statistically significant between Center 4 and 5 with the higher prevalence in Center 5 ( $Z=-2.689, p=.007$ ).

### *Active non-oral CSCs and active oral CSCs*

The difference between arms for percentage of participants with three or more non-oral CSCs and oral retracted CSCs was already published and no statistically significant difference for percentage of participants with three or more non-oral CSCs and oral retracted CSCs was reported by Willadsen et al., 2017. However, Center 4 showed a statistically significant difference between arms regarding three or more non-oral errors in favour of the common Arm C (one child in Arm C compared to 11 children in Arm A) (Fishers exact test  $p=.001$ ) (Figure 4). No differences within arms were found in Center 5 or 6.

Regarding differences between centers, statistically significant differences were seen on oral CSCs (retracted oral errors), between Center 4 and 5 ( $Z= -3.006$ ) and Center 4 and 6 ( $Z = -2.956$ ), ( $p=.003$ ). No other differences regarding neither oral CSCs, nor non-oral CSCs were statistically significant.

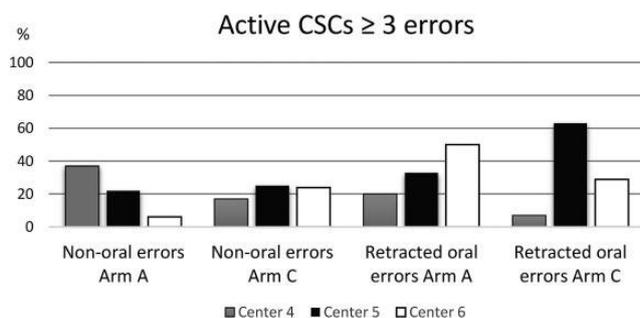


Figure 4. Greater than or equal to 3 active cleft speech characteristics per centers and arms.

### *Correlations with independent variables*

No significant relationship was found between history of secondary pharyngeal surgery and speech results in terms of VPC ( $r_s=.02$ ) and PCC ( $r_s= -.16$ ) were seen. A significant negative correlation was found between the number of speech therapy visits and PCC ( $r_s= -.49, p<.01$ ), that is, the children who had received a high number of visits had a low PCC.

## **DISCUSSION**

In the present study, the focus was on possible differences in speech outcome within the Scandleft Trial 2, i.e within and between centers. Although there were large differences in the number of children in the different three centers, the prevalence of perceived VPC and VPI from the dichotomization of the VPC-Sum were similar, with no statistically significant differences between arms within center or between centers were found. The center differences in VPC-Rate were interesting, however. For example, in all centers the highest percentages of children with VPC were found in the common Arm A. This was the common protocol with an early two-stage palatal procedure developed within the Scandleft project (Rautio et al., 2017) and therefore relatively novel to the surgeons. This protocol was also related to the highest total percentage of secondary pharyngeal surgeries, which means that the primary VPI rate was higher and might have been related to surgical experience as suggested by Rautio et al.

(2017). On the other hand, the total percentage secondary pharyngeal surgeries before age 5 years in the trial (9%) could be expected to increase throughout the years since the reported prevalence of VPI in the trial is high. The percentage of children falling into the scale value zero (0) on the overall rated velopharyngeal function from connected speech (VPC-Rate) was only 36%. This means that 64% were judged to have a velopharyngeal incompetency (24% VPI and 40% marginal VPI), which is far below the benchmark of 70% not having evidence of structurally related problems of VPI (Britten et al., 2014). Similar poor results (approximately 60% with VPI or marginal VPI) and similar percentage of secondary speech surgeries (12%) were reported earlier by Nyberg et al. (2014) in another cohort of children born with UCLP 1987–1997 before the Scandleft project was initiated. These children were the same age and had undergone the same one-stage protocol as the common one in the present study.

The impact of secondary speech surgeries is difficult to compare as the treatment might reflect a difference in attitude towards early secondary VPI surgeries, where teams' threshold for treatment of velopharyngeal insufficiency differ (Sommerlad, 2008). For example, Sell et al. (2017) reported a twice as high prevalence of secondary speech surgeries (25%) already at 5 years age in their large national study on 268 children whereas, the prevalence of passive speech symptoms was lower. In Nyberg et al. (2014) another 18% received secondary speech surgery after 5 years of age and the total percentage of secondary speech surgeries at 10 years of age was 30%. It is likely that the number of participants with secondary speech surgeries in the present study will increase by age.

The results on consonant proficiency were poorer than expected. Only approximately one third of the whole group reached age appropriate PCC and the best center scored as low as

47%, (the Finnish Center 4, Arm C). Nevertheless, it was significantly better than in the two Swedish centers. In Center 5 no child in Arm C had age appropriate PCC. The differences compared to the other two centers were statistically significant but should be interpreted with caution due to different sample sizes and also the number of exclusions. In Center 4, 93 children were recruited in total but high-quality recordings were only available for 60 children (64%). Center 5 recruited only 21 children and Center 6 recruited 37, of which high quality recordings were available for 17 (81%) in Center 5 and 35 (95%) in Center 6 (Semb et al., 2017). Due to the large number of exclusions, the group at Center 4 might not be representative of the population. Another explanatory factor, which was related to the differences in numbers of recruited children to the trial, could be that the surgeons perhaps had different levels of skill as a consequence of different numbers of performed cleft surgeries both throughout their careers and annually. Studies have shown clear advantages in speech outcomes for children born with CLP who were operated by surgeons who do large numbers of these procedures (Williams et al., 1999). It is possible that the surgeons in Center 4 which recruited 54% of all children, were the most experienced. Perhaps this could explain why consonant proficiency was significantly higher for the participants in Center 4 than in the two other centers and also of the significantly lower number of active CSCs. It might also explain the higher number of secondary pharyngeal surgeries in the common Arm A compared to the local Arm C. Thus, the highest numbers of secondary pharyngeal surgeries in the trial might be related to lower experience with the Arm A protocol.

The number of fistulae requiring surgery was 10% in the trial. The highest prevalence was observed in Center 4, Arm A and it is possibly also related to lower experience with that protocol. Interestingly, in Center 6 fistulae requiring surgery (8%) were only found in the local Arm C. Perhaps the higher fistulae prevalence could explain the unexpected high occurrence of oral retracted CSCs after one-stage palatal repair. However, Nyberg et al.

(2014) instead hypothesized that it was a phonological backing when they found an unexpected high prevalence of retracted oral articulation at the same age and after the same surgical protocol as the local Arm C. This might be an explanation for the high occurrence of retracted oral errors in Center 6 in the current study as well and possibly related to hearing loss related to otitis media with effusion.

Yet, another explanation of the differences between the results from Finnish Center 4 and the two Swedish centers could perhaps be attributed to language differences. Finnish has relative to Swedish a smaller phonemic inventory with 13 consonants and 8 vowels, whereas Swedish has 18 consonants and 9 vowels (Saaristo-Helin et al., 2011; McAllister, 1998). There are also differences in how co-occurring phonemes are realised. For example, contrary to Swedish, the plosives in Finnish are voiceless, unaspirated and the explosive burst is weak. Moreover, the most commonly used allophone of the /s/-sound is considered less “sharp” than the sibilant denoted by the IPA symbol [s]. Both the location of the greatest energy in the spectrum, and perceptually judged, this allophone is somewhere between IPA [s] and [S] (Suomi et al., 2008). Both plosives and the fricative /s/ sound are high pressure consonants reported to be vulnerable sounds to VPI (Hutters and Henningson, 2004) and therefore included in the Scandleft speech material. But perhaps the production of the Finnish variants of the sounds is not as challenging as the Swedish ones. This might make the comparison between speech results from the Finnish speaking children and Swedish speaking children difficult. Also, the combination of a smaller consonant inventory, and the less demanding articulation on both the plosives and the /s/-sound could be an explanation to why Finnish children have been reported to master all speech sounds at 2;6 years (Saaristo-Helin et al., 2011), whereas the completion of the phonetic inventory in Swedish children does not occur until after the age of 4 years (Nettelbladt, 2007). This might also be the explanation behind the much lower

proportions of Finnish participants having three or more developmental speech characteristics than the participants from the two Swedish centers.

Consonant proficiency as measured by PCC in this study is poor in comparison to scores from studies of Swedish children without cleft palate (Lohmander et al., 2017b). The speech material used in this study was restricted and only included a selection of singleton consonants in phonetically simple target words. This means that the PCC scores in the study could not be regarded as reflecting the participants true articulatory competence as the test did not include all the speech sounds and speech sound combinations of the two languages involved. It is possible that if non-restricted speech material had been used, the comparison to children without clefts would have been even more unfavorable. However, the restricted material was justified by the aim of the study (to compare speech outcome between two surgical protocols) and the inclusion of two different languages (Hutters and Henningsson, 2004). This means that the result only tells how well the participating children performed on the very carefully chosen consonants in the naming test. There are studies performed on 3-year-old children without cleft using the same or very similar target words (Willadsen, 2012; Klintö and Lohmander, 2017). A comparison of the results of the 5-year olds in the present study with the results from the 3-year old children without CLP is very disadvantageous since the 5-year olds scored below the 3-year old without CLP in the two above mentioned studies.

The poor PCC-scores in the present study are not fully explained by the presence of CSCs, and raise the question if there are other factors than the cleft condition and potential hearing loss involved, such as language or learning disabilities (Morris et al., 2003). The association between hearing loss and occurrence of backing which has been convincingly suggested by Shriberg et al. (2003), is an interesting factor which might have been significant for the high prevalence of retracted oral errors in the present study, particularly in Centers 5 and 6. Even

though few children showed a hearing loss at the assessment their hearing history was not known. The development of the consonant system could therefore have been influenced by hearing loss related to otitis media with effusion during their early childhood (Shriberg et al., 2003).

A correlation found between number of speech therapy visits and speech problems was found and could be interpreted in two ways. Either it's only natural that the children with the greatest needs receive the most visits or one could question the impact speech therapy has for children born with CLP. However, since the correlations were only low to moderate and the study protocol did not include type or intensity of therapy, this must be interpreted with caution. In future studies clearer descriptions and instructions on how to report speech therapy are needed.

### ***Limitations***

The findings in this study should be interpreted with caution since the number of participants was low in two centers, especially in Center 5. However, in Center 5 one surgeon performed all surgery and therefore the results from that center are considered consistent. Also, non-parametric tests were used to account for the different sample sizes in the participating centers. Another limiting factor was the high proportion of recruited children that were excluded in this study due to lack of sufficient recordings, especially in Center 4.

The reliability of the assessments was acceptable or good but could have been better. Even though most SLP/Ts stayed in the project throughout the years there were changes in staff, particularly in Center 4, which presumably have influenced the reliability since the lowest reliability was found for the Finnish material both regarding transcription and rating.

Although training of SLP/Ts was performed regularly, the effect of the training has not yet been evaluated.

A serious attempt to collect data of hearing was included in the study but unfortunately, the lack of data was substantial with no possibility to evaluate the impact on speech. Other factors that could have influenced speech outcome such as of cognition and language skill were not investigated. Finally, a measure of speech intelligibility or speech acceptability would have been valuable but no validated tools for cross-linguistic use were available at the time of the project.

## **CONCLUSION**

At age 5 years, poor speech outcome was found in Scandcleft Trial 2 with similar results for both arms and centers regarding perceived velopharyngeal competency (VPC) or velopharyngeal incompetency (VPI), and no significant differences regarding secondary pharyngeal surgeries or palatal fistulas. This indicates that palatal surgery in one or two stages did not influence studied speech outcome measures. In contrast, differences between the participating centers regarding the consonant proficiency outcome and the measures of consonant errors were substantial. These differences could be attributed to factors other than surgical protocol, such as the individual surgeon's skill or confidence with a new protocol. Finally, the efficacy of speech therapy should be investigated.

## REFERENCES

- Britton L, Albery L, Bowden, M, Harding-Bell A, Phippen G, Sell D. A cross –sectional cohort study of speech in five-year-olds with cleft palate. *Cleft Palate Craniofac J*. 2014;51;431-451.
- Brunnegård K, Lohmander A. A cross-sectional study of speech in 10-year-old children with cleft palate: results and issues of rater reliability. *Cleft Palate Craniofac J*. 2007;44:33-44.
- Chapman T, Willadsen, E. The development of speech in children with cleft palate. In: Howard S, Lohmander A, eds. *Cleft Palate Speech*. West-Sussex: Wiley-Blackwell; 2011: 23-40.
- Derijcke A, Eerens A, Carels C. The incidence of oral clefts: a review. *Br J Oral Maxillofac Surg*. 1996;34(6):488-494.
- Dodd B, Holm A, Hua Z, Crosbie S. Phonological development: a normative study of British-English-speaking children. *Clin Ling Phon*. 2003;17:617-643.
- Dotevall H, Lohmander A, Ejnell H, Bake B. Perceptual evaluation of speech and velopharyngeal function in children with and without cleft palate and the relationship to nasal airflow patterns. *Cleft Palate Craniofac J*. 2002;39:409-424.
- Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Edu Psychol Meas*. 1973;33:613-619.
- Grunwell P, Harding A. Notes and discussion active versus passive cleft-type speech characteristics. *Int J Lang Com Dis*. 1998;33:329-352.
- Grunwell P, Sell DA. Speech and cleft palate/velopharyngeal anomalies. In: Watson ACH, Sell DA, Grunwell P. *Management of Cleft Lip and Palate*. London: Whurr Publishers; 2001:68-86.
- Hagberg C, Larsson O, Milerad J. Incidence of cleft lip and palate and risks of additional malformations. *Cleft Palate Craniofac J*. 1998; 35: 546-549.

- Howard S, Lohmander A. Speech production and development. In: Howard S, Lohmander A, eds. *Cleft Palate Speech*. West-Sussex: Wiley-Blackwell; 2011:1-3.
- Hutters B, Brøndsted K. Strategies in cleft palate speech –with special reference to Danish *Cleft Palate J*. 1987; 24:126-136.
- Hutters B, Henningsson G. Speech outcome following treatment in cross-linguistic cleft palate studies: methodologic implications. *Cleft Palate Craniofac J*. 2004;41:544–549.
- IPA. The International Phonetic Alphabet [Internet]. International Phonetic Association; 2015 [accessed 2019 Feb 12]. Available from:  
[https://www.internationalphoneticassociation.org/sites/default/files/IPA\\_Doulos\\_2015.pdf](https://www.internationalphoneticassociation.org/sites/default/files/IPA_Doulos_2015.pdf)
- IPA. extIPA Symbols for Disordered Speech [Internet]. International Phonetic Association; 2008 [accessed 2019 Feb 12]. Available from:  
<https://www.internationalphoneticassociation.org/sites/default/files/extIPAChart2008.pdf>
- John A, Sell D, Sweeney T, Harding-Bell A, Williams A. The cleft audit protocol for speech – augmented: A validated and reliable measure for auditing cleft speech. *Cleft Palate Craniofac J*. 2006, 43 (3): 272-288.
- Karnell M, Van Demark D. Longitudinal speech performance in patients with cleft palate: Comparisons based on secondary management. *Cleft Palate J*. 1986;23:278–288.
- Kemp-Fincham SI, Kuehn DP, Trost-Cardamone JE. *Speech development and the timing of primary palatoplasty*. In: Bardach J, Morris HL, eds. *Multidisciplinary Management of Cleft Lip and Palate*. Philadelphia: WB Saunders; 1990:736–745.

- Klintö K, Salameh E K, Svensson H, Lohmander A. The impact of speech material on speech judgement in children with and without cleft palate. *Int J Commun Lang Disord.* 2011; 46(3): 348-360.
- Klintö K, Salameh EK, Olsson M, Flynn T, Svensson H, Lohmander A. Phonology in Swedish speaking 3-year-olds born with cleft lip and palate and the relationship with consonant production at 18 months. *Int J Commun Lang Disord.* 2014;49(2): 240-254.
- Klintö K, Salameh EK, Lohmander A. Phonology in Swedish speaking 5-year-olds born with unilateral cleft lip and palate and the relationship with consonant production at 3 years of age. *Int J Speech Lang Pathol.* 2016;18(2): 147-156.
- Klintö K, Lohmander A. Phonology in Swedish-speaking 3-year-olds born with unilateral cleft lip and palate treated with palatal closure in one or two stages. *J Plast Surg Hand Surg.* 2017;51:112-117.
- Lim S, Palethorpe N, Rodger S. Understanding the common interrater reliability measures. *Int J Ther Rehab.* 2012;19:488–496.
- Lithovius RH, Ylikontiola LP, Harila V, Sándor GK. A descriptive epidemiology study of cleft lip and palate in Northern Finland, *Acta Odontol Scand.* 2014; 72:5, 372-375.
- 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;45:32-41
- Lohmander A, Willadsen E, Persson C, Henningsson G, Bowden M, Hutter 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;46:347-362.
- Lohmander A. Surgical intervention and speech outcomes in cleft lip and palate. In: Howard S, Lohmander A, eds. *Cleft Palate Speech.* West-Sussex: Wiley-Blackwell; 2011:55-85.

- Lohmander A, Persson C, Willadsen E, Lundeborg I, Alaluusua S, Aukner R, Bau A, Boers M, Bowden M, Davies D, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 4. Speech outcomes in 5-year-olds—velopharyngeal competency and hypernasality. *J Plast Surg Hand Surg*. 2017a;51(1):27-37.
- Lohmander A, Lundeborg I, Persson C. SVANTE – The Swedish Articulation and Nasality Test – Normative data and a minimum standard set for cross-linguistic comparison. *Clin Ling Phon*. 2017b;31:137-154.
- Lohmander A, Hagberg E, Persson C, Willadsen E, Lundeborg I, Davies J, Havstam C, Boers M, Kisling-Møller M, Alaluusua S, et al. Validity of auditory perceptual assessment of velopharyngeal function and dysfunction: the VPC-Sum and the VPC-Rate. *Clin Ling Phon*. 2017c;31:589-597.
- McAllister R. *Talkommunikation* [Speech communication] Lund: Studentlitteratur; 1998: 193-200.
- McWilliams BJ, Glaser E R, Philips BJ, Lawrence C, Lavorato A S, Beery QC, Skolnick M L. A comparative study of four methods of evaluation velopharyngeal adequacy. *Plast Reconstr Surg*. 1981;68:1–10.
- Morris H, Ozanne A. Phonetic, phonological and language skills of children with a cleft palate. *Cleft Palate Craniofac J*. 2003; 40: 460-470.
- Nettelbladt U. Phonological development. In Nettelbladt U, Salameh, E-K. *Språkutveckling och språkstörning hos barn* [Language development and language impairment in children] Lund: Studentlitteratur; 2007: 57-94.
- Nyberg J, Peterson P, Lohmander A. Speech outcome at age 5 and 10 years in unilateral cleft lip and palate after one-stage palatal repair with minimal incision technique – A longitudinal perspective. *Int J Ped Otorhinolar*. 2014;78:1662-1670.
- Rautio J, Åbyholm F, Vindenes H, et al. Scandcleft randomised trials of primary surgery for

- unilateral cleft lip and palate: 2. Surgical results. *J Plast Surg Hand Surg*. 2017;51:13–19.
- Reddy RR, Reddy SG, Vaidhyanathan A, Bergé SJ, Kuijpers-Jagtman A-M. Maxillofacial growth and speech outcome after one-stage or two stage palatoplasty in unilateral cleft lip and palate. A systematic review. *J Cranio-Maxillo-Fac Surg*. 2017;45:995-1003.
- Renfrew C. *The Bus Story: A test of continuous speech*. North Place, Old Headington: Oxford; 1969.
- Rohrich RJ, Love EJ, Byrd HS, Johns DF. Optimal timing of cleft palate closure. *Plast Reconstr Surg*. 2000;106:413–421.
- Saaristo-Helin K, Kunnari S, Savinainen-Makkonen T. Phonological development in children learning Finnish: A review. *First Language*. 2011;31(3):342-363.
- Sell D, Mildinhal S, Albery L, Wills AK, Sandy JR, Ness AR. The cleft care UK study. Part 4: perceptual speech outcomes. *Orthod Craniofac Res*. 2015;18:36–46.
- Sell D, Southby L, Wren Y, Wills AK, Hall A, Mahmoud O, Waylen A, Sandy JR, Ness AR. Centre-level variation in speech outcome and interventions, and factors associated with poor speech outcomes in 5-year-old children with non-syndromic unilateral cleft lip and palate: The Cleft Care UK Study. Part 4. *Orthod Craniofac Res*. 2017;20:27-39.
- Semb G, Enemark H, Friede H, Paulin G, Lilja J, Rautio, J, Andersen M, Åbyholm F, Lohmander A, Worthington H, et al. Scandcleft randomized trials of primary surgery for unilateral cleft lip and palate – Planning and Management. *J Plast Surg Hand Surg*. 2017;51:2-13.
- Scherer N, Williams L, Proctor-Williams K. Early and later vocalization skills in children with and without cleft palate. *Int J Pediatr Otorhinolaryngol*. 2008;72: 827-840.
- Shriberg LD, Kwiatkowski J. Phonological disorders III: A procedure for assessing severity of involvement. *J Speech Hear Dis*. 1982;47:242-256.

- Shriberg LD, Austin D, Lewis BA, McSweeney JL, Wilson DL. The percentage of consonants correct (PCC) metric: extensions and reliability data. *J Speech Lang Hear Res.* 1997; 40(4):708-722.
- Shriberg LD, Kent RD, Karlsson HB, McSweeney JL, Nadler CJ, Brown RL. A diagnostic marker for speech delay associated with otitis media with effusion: backing of obstruents. *Clin Ling Phon.* 2003;17:529-547.
- Sommerlad BC. Cleft palate repair with minimal hard palate dissection and radical muscle reconstruction. In: Losee JE, Kirschner RE, editors. *Comprehensive Cleft Care*, Columbus, OH: McGraw-Hill, 2008. pp 413-429.
- Suomi K, Toivanen J, Ylitalo R. Finnish sound structure – Phonetics, phonology, phonotactics and prosody. *Studia Humaniora Ouluensia* 9, University of Oulu, Finland, 2008.
- Sweeney T. Nasality: Assessment and intervention. In: Howard S, Lohmander A, eds. *Cleft palate speech: Assessment and intervention*. Chichester: Wiley-Blackwell; 2011:199-220.
- Willadsen E. Influence of timing of hard palate repair in a two stage procedure on early speech development in Danish children with cleft palate. *Cleft Palate Craniofac J.* 2012; 49(5): 574-595.
- Willadsen E, Lohmander A, Persson C, Lundeborg I, Alaluusua S, Aukner R, Bau A, Boers M, Bowden M, Davies J, et al. Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 5. Speech outcomes in 5-year-olds—consonant proficiency and errors. *J Plast Surg Hand Surg.* 2017;51(1):38-51.
- Williams AC, Sandy JR, Thomas S, Sell D, Sterne JA. Influence of surgeons's experience on speech outcome in cleft lip and palate. *Lancet.* 1999;13:1697-1698.

Yamashita RP, Borg E, Granqvist S, Lohmander A. Reliability of hypernasality rating:  
Comparison of 3 different methods for perceptual assessment. *Cleft Palate Craniofac J.*  
2018;55(8):1060-1071.