Assessing speech at three years of age in the cleft palate population: a scoping review of assessment practices

Abstract

Background: There is no consensus in the UK regarding types of speech samples or parameters of speech which should be assessed at age-3 years in children with cleft palate+/- cleft lip (CP±L), despite cleft units routinely assessing speech at this age. Standardisation of assessment practices would facilitate comparisons of outcomes across UK Cleft Units; earlier identification of speech impairments - which could support more timely treatments; more reliable recording of therapy impacts and surgical interventions.

Aims: To explore assessment practices used to assess speech in three-year-old children with CP±L including speech parameters, methods of assessment, and nature of speech sample used.

Methods: A broad examination of literature was undertaken through the use of a scoping review conducted in accordance with Joanna Briggs Institute guidelines. Search terms were generated from a preliminary search and then used in the main search (Medline, CINAHL, Embase, AMED and PsycINFO).

Main Contribution: A combination of approaches (medical, linguistic, developmental and functional) is required to assess CP±L speech at age-3. A developmental approach is recommended at this age, considering the complexity of speech profiles at age-3, in which typically developing speech processes may occur alongside cleft speech characteristics. A combined measure for both nasal emission and turbulence, and an overall measure for velopharyngeal function for speech, show potential for assessment at this age. Categorical

1 ordinal scales are frequently used; the use of continuous scales has yet to be fully explored at

2 age 3.

3 Although single word assessments, including a subset of words developed for cross-

4 linguistic comparisons, are frequently used, more than one type of speech sample may be

5 needed to validly assess speech at this age. The lack of consensus regarding speech samples

6 highlights a need for further research into the types of speech samples 3-year olds can

7 complete; the impact of incomplete speech samples on outcome measures (particularly

8 relevant at this age when children may be less able to complete a full sample); the impact of

9 different speech samples on the validity of assessments and the reliability of listener

10 judgements.

11 Conclusions: Whilst a medical model and linguistic approaches are often central in

12 assessments of age-3 cleft speech, this review highlights the importance of developmental

13 and functional approaches to assessment. Cross-linguistic single word assessments show

14 potential, and would facilitate the comparison of UK speech outcomes with other countries.

15 Further research should explore the impact of different speech samples and rating scales on

16 assessment validity and listener reliability.

2

1
2
3
4   1   What this paper adds
5
6
7   2   **What is already known on this subject:**
8
9
10
11   3   Although speech is typically assessed at age-3 years in UK cleft units, assessment methods
12
13   4   vary. This prevents cross-unit outcome comparisons and is a barrier to measuring the impact
14
15   5   of therapy and surgical interventions. There is a need to broadly explore assessment practices
16
17
18   6   to guide the direction of assessments at age-3 years in the UK.
19
20
21   7   **What this study adds:**
22
23
24   8   This review highlights the importance of a developmental approach to assessment in the
25
26
27   9   assessment of children with CP±L at age-3 years. Whilst there is consensus as to the
28
29   10   parameters of assessment, there is no such consensus regarding speech samples or methods
30
31   11   of assessment.
32
33
34
35   12   **Clinical implications of the study:**
36
37
38   13   Children at age-3 are at a unique stage of speech development. Assessment procedures and
39
40
41   14   outcomes used with older age groups require adaptation to meet the needs of 3-year-olds,
42
43   15   both in terms of the speech sample used and the need to consider developmental and
44
45   16   functional outcomes. Whilst further research is needed to provide a foundation for decisions
46
47
48   17   regarding the selection of speech samples, and the impact of different speech samples and
49
50   18   methods of assessment on the reliability of listener judgements, cross-linguistic single word
51
52
53   19   assessments show potential and would facilitate international comparisons of outcomes.
54
55
56
57
58
59
60

1

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

1 Introduction

2 The speech outcome of individuals with a repaired cleft palate +/-cleft lip (CP±L) is a central

3 focus of all cleft teams, as speech outcomes are a primary measure of surgical success

4 (Grunwell and Sell 2001, pp. 68). To this end, a significant focus of cleft research has been on

5 developing speech assessment procedures and outcome measures. Perceptual speech

6 assessment with a basis in phonetic transcription is described as the 'gold standard' for the

7 assessment of speech in individuals with CP±L (Howard 2011, pp. 127). However, there are

8 significant challenges in designing assessment procedures and protocols which allow for

9 speech outcomes to be assessed over a lifespan, as particular speech samples and assessment

10 materials may be more appropriate at certain ages, and assessment objectives may differ

11 across age groups. For example, in children over 5, and adults speech assessment may focus

12 on the identification of persisting speech difficulties and the impact of anatomical changes

13 (e.g. adenoid atrophy, dentition and occlusion) on speech production (Sell and Pereira 2015).

14 However, for younger children, a focus of assessment may be to examine the impact of the

15 cleft on the process of normal speech sound development (phonology as well as articulation)

16 (Chapman and Willadsen 2011: 25).

17 Existing protocols such as the Cleft Audit Protocol for Speech-Augmented (CAPS-A)

18 (John et al. 2006, Sell et al. 2009) and the Americleft modification of the CAPS-A, CAPS-A-AM

19 (Chapman et al. 2016) have only been validated and tested for reliability on children aged 5

20 and above. This is understandable given the original intended use of the CAPS-A in the audit

21 of speech outcomes at age 5 years, an established age for assessing speech outcomes, hence

22 the focus of much attention in the last two decades. Only the Swedish Articulation and

23 Nasality Test (SVANTE) (Lohmander et al. 2005, Lohmander et al. 2009), a valid and reliable

1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1   assessment protocol available in Swedish and Norwegian, has been developed for use with

2   children under 5 years with CP±L (Lohmander et al. 2017a). The Universal Parameters (UPS)

3   (Henningsson et al. 2008), and the Pittsburgh Weighted Values for Speech Symptoms

4   Associated with Velopharyngeal Incompetence (PWSS) (McWilliams & Philips 1979) are

5   alternative assessment protocols; however, neither has been comprehensively validated

6   (Prathanee et al. 2011, Dudas et al. 2006).

7   Although UK cleft units routinely carry out speech assessments at age-3, unlike age-5,

8   there is no agreed assessment procedure or outcome measures. Agreed procedures and

9   outcome measures at age-3 would have several benefits. Firstly, valid and reliable outcome

10   measures could be used to identify children at risk of poor speech outcomes at age-5 and

11   encourage timely referral for intervention. Outcome measures would also facilitate the

12   comparison of outcomes across Cleft Units before 5-years, with the potential to identify

13   variability in outcomes and implement quality improvement initiatives in a similar way to

14   established practices in the UK at age-5 (Cleft Registry and Audit Network [CRANE] 2019). In

15   addition it could permit the collection of comparative longitudinal data, providing an

16   understanding of the impact of therapy and secondary surgery for speech through the

17   comparison of outcomes at age-3 and 5.

18   One well-recognised challenge in the assessment of speech in 3-year-olds is that at

19   this age children go through a period of significant developmental change, refining and

20   acquiring new skills related to their attention, language and speech (Dosman et al. 2012,

21   McLeod and Baker 2017:202). As such, 3-year-old children can be highly variable in how they

22   present in these areas despite being within normal developmental limits. Furthermore,

23   variability in presentation can also result from the known delay in expressive language and

2

1   speech (Cavalheiro et al. 2019) and an increased risk of neurodevelopmental disorders

2   (Tillman et al. 2018).   This variability presents a challenge when designing assessment

3   protocols, in particular the selection of speech samples and the parameters of speech which

4   can be assessed reliably. Given that children's speech sound systems are in a stage of

5   developmental change (irrespective of the cleft palate) it also raises the question as to the

6   focus of the assessment for children with cleft i.e. whether this focuses purely on those

7   parameters considered core to assessment of cleft speech and/or should include an

8   assessment of speech from a developmental perspective.

9   A broad scope of enquiry was required to examine the methods, parameters of

10 assessment (whether developmental or cleft specific) and speech samples utilised in the

11 assessment of speech in children with CP±L at age-3. This was achieved through the use of a

12 scoping review methodology (Khalil et al. 2016). This methodology was selected given that

13 the intention behind the work was to inform the future development of assessment

14 procedures at age-3 in the UK. Rather than answer an effectiveness question (Peters et al.

15 2015), evaluate study quality, or limit the review to a particular study design (Arksey and

16 O'Malley 2005) or country of origin, the scoping review methodology enabled inclusion of a

17 broad range of resource types, and the broad mapping of current practices, in order to clarify

18 the key concepts, identify gaps in the evidence base and make recommendations for future

19 practice in the UK. In doing so the review identified the extent to which there is consensus in

20 the parameters of assessment, procedures and methods; differences and similarities in

21 assessment at age-3 years compared to practice at age 5; and the types of speech samples

22 used (Munn et al. 2018).

3

Methods

Ethical Review

Approval for the study was gained in accordance with the Ethics and Governance procedures at Coventry University, study number P68435.

Design

To strengthen methodological rigour and to allow replication, the scoping review was undertaken using the Joanna Briggs Institute (JBI) guidance for scoping reviews (Peters et al. 2015) and followed the process outlined below.

Objective/Aim

To map the parameters of speech and types of speech samples used to assess speech in 3-year-old children with CP±L, with reference to the rating scales and methods used to assess them, and to consider key differences in speech assessments at age-3 and older age groups in order to inform the development of a UK assessment protocol.

Scoping Review Objectives

- Explore the parameters of speech typically assessed in 3-year-old children with CP±L, and consider if they are core to the assessment of individuals with CP±L.

- Explore the methods and rating scales used to assess the identified parameters of speech.

- Explore the types of speech samples are utilised in the assessment of 3-year-old children with CP±L.

4

1   •   Discuss how the parameters of speech assessment map onto different

2   theoretical approaches to assessment.

3   *Inclusion Criteria*

4   The inclusion criteria were developed with reference to the JBI methodology as

5   follows:

6   *Types of participants*

7   Children age-3 with CP±L. For inclusion in the review, literature had to specifically

8   include 3-year-olds with details provided as to how speech was assessed at this age (either

9   the parameters of speech, the measurement scales used and/or the speech sample). This

10   resulted in the inclusion of studies also assessing children at other ages i.e. longitudinal

11   studies; however, only data referencing assessment at age-3 was included in the review.

12   *Concept*

13   Literature and sources had to address the following concepts: speech assessments,

14   the types of speech samples and speech parameters assessed, including methods of

15   assessment. This allowed for the inclusion of sources which investigated assessment

16   processes as well as those reporting on speech outcomes.

17   *Context*

18   To broaden the examination, no preference was given to sources from any specific

19   country; however, resources were limited to English, or those with an available translation. It

20   was considered that the development of the Great Ormond Street Speech Assessment

5

1 (GOS.SP.ASS' 98) (Sell et al. 1999) marked the start of a new era in speech assessments and

2 outcome measures; this informed the timeframe, with studies reviewed from 1998-2018.

3 *Types of Sources*

4 Sources needed to be sufficiently detailed to enable the extraction of sufficient

5 information about assessment procedures, therefore conference abstracts were excluded.

6 *Search Strategy*

7 In accordance with Peters et al. (2015), an initial limited database search was

8 conducted using the EMBASE database to consolidate relevant search terms from the title,

9 abstract and keywords of identified studies and explore controlled vocabulary. Cleft specific

10 search terms as well as general terms such as 'speech assessment' were subsequently

11 included. The final list of search terms is shown in Table 1. Relevant Medical Subject Headings

12 (MeSH [vocabulary used to index publications]) were utilised in the database search.

13 <insert Table 1 here>

14 The electronic databases Medline, Cumulative Index of Nursing & Allied Health

15 Literature (CINAHL), Embase, AMED and PsycINFO were used. Citation tracking was used to

16 identify additional relevant sources.

17 *Resource Selection*

18 Figure 1 outlines the study selection processes utilising the flowchart advised in the

19 PRISMA Scoping Review Extension (PRISMA-ScR) flowchart (Tricco et al. 2016). Abstracts were

20 screened for relevance to the objectives of the review and the inclusion criteria, first by title,

21 then by abstract and if necessary, using the full text. With the agreement of the research team

6

1   and in an exception to the inclusion criteria, the normative data developed using the SVANTE

2   assessment (Lohmander et al.2017a) was included. Whilst the outcomes presented were not

3   related to children with CP±L, the assessment protocol which was primarily designed to assess

4   structurally based speech difficulties (i.e. cleft palate) and is suitable for children aged 3,

5   provides a detailed account of the assessment procedure and is highly relevant to the

6   objectives of this review. Consideration was also given to a large ongoing research project

7   taking place in the UK, The Cleft Collective Speech and Language Study

8   ([https://www.nbt.nhs.uk/bristol-speech-language-therapy-research-unit/bsltru-](https://www.nbt.nhs.uk/bristol-speech-language-therapy-research-unit/bsltru-research/cleft-speech-language-study)

9   [research/cleft-speech-language-study](https://www.nbt.nhs.uk/bristol-speech-language-therapy-research-unit/bsltru-research/cleft-speech-language-study)). As part of the development of the study, a national

10   survey had taken place regarding assessment practices in UK Cleft Centres at age-3. Given the

11   specific relevance of this survey to the aims of this review, with the agreement of the research

12   team, this was included and is referred to asWren 2013).

13   The final sources included in the review were discussed and confirmed by paired

14   members of the research team. All but one were research papers that appeared in peer-

15   reviewed journals.

16   <Insert Figure 1 here>

17   *Charting the results*

18   Studies meeting the inclusion criteria were read in full and charted according to the

19   country of origin, study aim, participant details, methodology and methods, details of the

20   speech assessment and speech sample, the parameters of speech assessed and the rating

21   scales utilised. The parameters of speech assessed across the CAPS-A, CAPS-A-AM, SVANTE,

7

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1   PWSS and UPS provided a framework for analysis defining those parameters core to the

2   assessment of speech in the CP±L population.

3

4   ## Results

5   35 sources were reviewed in the scoping review. These are listed in Appendix 1 and

6   detailed in Table 2.

7   <insert Table 2 here>

8   *Demographic Information*

9   The sources originated from 12 countries (see Figure 2). 34% originated from Sweden,

10   with Scandinavian countries producing 48% of the sources reviewed.

11   <Insert Figure 2 here>

12   Apart from one, almost all sources included were primary research and recruited

13   samples of children with CP±L. The sub-types are presented in Table 2. It is noteworthy that

14   many of the Scandinavian sources were part of/sub-studies relating to the Scandcleft Trial

15   (Lohmander et al.2017ab, Willadsen *et al.* 2017), a large multicentre cross-linguistic

16   randomised control trial, evaluating four different surgical protocols. The Scandcleft Trial only

17   recruited participants with Unilateral Cleft Lip and Palate (UCLP), contributing to the high

18   number of sources (n= 13/33 39.4%) which only included participants with UCLP. Two sources

19   did not record cleft type (Wren 2013 [reporting assessment practices]; Lohmander et al.

20   2017a [reporting on non-cleft children]) and were excluded in calculations regarding cleft

21   type.

8

1 *Which parameters of speech are typically assessed in 3-year-old children with CP±L,*

2 *and are they core to the assessment of individuals with CP±L?*

3 The parameters of speech assessed in each source are charted in Table 3. Each source

4 reported parameters which addressed their specific research aims. As such, not all sources

5 assessed all the parameters identified in this review. Therefore, where percentages are

6 presented, this refers only to those sources which directly assessed the parameter in

7 question.

8 <insert Table 3 here>

9 <u>Consonant Production</u>

10 Almost all sources assessed consonant production. Although phonetic transcription

11 underpinned the assessment of consonant production, a variety of methods were used,

12 shown in Figure 3.

13 <insert Figure 3 here>

14 The first principal method used to assess consonant production was to report

15 summary patterns by grouping errors according to their place of articulation, or the broader

16 categories of passive or active characteristics (see Hutters and Brøndsted [1987] for a

17 description of these processes). Such summary patterns can be viewed as a specific measure

18 of cleft speech characteristics (CSCs) and feature in the CAPS-A, CAPS-A-M, UPS, and SVANTE

19 protocols. Despite a strong trend across the studies to report on consonant articulation in

20 summary categories, different summary patterns were used. For example, Safaiean et al.

21 (2017) utilised the UPS summary patterns whilst Persson et al. (2006) identified

22 compensatory articulation using a 3-step process (yes/no scale --> categorising type of

9

1 compensatory articulation as retracted oral/pharyngeal/glottal/active nasal fricative/other -->

2 > recording frequency of error). Chapman et al. (2008) and Hutters et al. (2001) both recorded

3 the frequency with which CSCs/compensatory articulations occurred. The studies highlight

4 that at age-3 there is a consensus on the need to report CSCs, even within the context of a

5 developing sound system, albeit using different summary categories.

6   The second principal method was Percentage Consonants Correct (PCC) (Shriberg &

7 Kwiatkowski 1982). PCC was also used in its adjusted form (PCC-A) (Table 3) in which age-

8 appropriate speech distortions are classed as correct productions (Shriberg et al. 1997). For

9 example, Klintö et al. (2016, 2015, 2014a,b) utilised PCC-A, with simplifications of phonemes

10 used by more than 10% of the children scored as correct (Klintö et al. 2016 pp.151). Unlike

11 the summary patterns which specifically focus on CSCs, PCC is a broader measure of all the

12 speech errors including CSCs, developmental and other disordered speech patterns. PCC does

13 not differentiate between articulatory or phonologically based errors. This 'combined

14 approach' is recommended by Lohmander et al. (2017a) as articulation errors (arising from

15 structural or functional abnormalities) may be 'phonologised' (Harding and Grunwell 1995)

16 and become integrated into a child's sound system alongside or replacing developmental

17 phonological patterns. Summary patterns similarly provide phonological information, for

18 example, *backing* is an example of a common phonological process arising from retracted

19 articulation (Chapman 2003, Willadsen 2012). UK cleft units did not appear to use PCC in the

20 assessment of 3-year-olds (Wren 2013).

21   Other measures of consonant production included consonant/phoneme inventory

22 which was used in 85.7% of UK cleft units at age-3 (Wren 2013). This also appeared in several

23 other studies including the SVANTE (Lohmander et al.2017a, Klintö et al. 2014a, Morris &

10

1 Ozanne 2003, Chapman et al. 2008, Chapman 2004, Konst et al. 2003, Hattee et al. 2001). The

2 criteria for inclusion in the inventory varied across the studies. For example (Hattee et al.

3 2001) rated the presence of consonants, Morris & Ozanne (2003) required phonemes to be

4 heard twice to be included, and Lohmander et al. (2017a) reported correct productions in

5 more than 50% of the targets.

6 Several studies reported on both cleft and developmental phonological processes

7 (Willadsen et al. 2018, Chacon et al. 2017, Klintö et al. 2016, Klintö et al. 2014a, Klintö et al.

8 2014b, Willadsen 2012, Konst et al. 2003, Morris & Ozanne 2003 and Hutters et al. 2001). This

9 provides a broader perspective of an individual's phonological development from both a cleft

10 and developmental perspective. In contrast, whilst 78.57% of UK cleft units assess CSCs at

11 age-3, only 50% were reported to assess phonology (Wren 2013). Whilst there may be some

12 overlap between measures of CSCs and phonology, this indicates that a comprehensive

13 phonological assessment may not be considered a priority for assessment in UK cleft units at

14 age-3.

15 The sources reviewed highlight the importance, at age-3 years, of assessing

16 consonant production both from a cleft and developmental perspective. The methods used

17 in the studies to assess consonant production demonstrate that both PCC and summary

18 patterns can provide information about articulation and phonology.

19 <u>Resonance</u>

20 Both judgements of hypernasality and hyponasality are made in the CAPS-A, CAPS-A-

21 AM, UPS, PWSS and SVANTE outcome measures, indicating that these are two parameters

22 core to the assessment of speech in the CP±L population. Of those studies assessing

11

1 resonance, hypernasality was reported to be specifically assessed in 84.2% of the studies, as

2 shown in Table 3. Ordinal scales were most frequently used, although scales and descriptors

3 varied. For example, Lohmander et al. (2006) utilised a five-point scale with descriptors,

4 Chapman et al. (2008) used a four-point scale, and Pulkkinen et al. (2001) used a binary scale

5 rating the presence/absence of hypernasality. Whereas hypernasality was sometimes

6 assessed in the absence of hyponasality, hyponasality was always reported alongside

7 hypernasality (see Table 3). Wren (2013) reported that hypernasality and hyponasality are

8 assessed in 92.85% of UK cleft units at age-3.

9 Although, those studies assessing resonance predominately assessed hypernasality,

10 there was no consensus as to the type or length of scale used.

11 Nasal Airflow Errors

12 Nasal airflow errors (NAE) is a generic term to describe air escaping inappropriately

13 through the nose during the production of oral pressure consonants (Sweeney 2011: 200).

14 NAE featured in several of the studies reviewed and is therefore also considered core to the

15 assessment of cleft speech at this age. Wren (2013) separated out the parameters of nasal

16 emission and turbulence. In contrast, other studies only reported on audible nasal emission

17 (although this term was possibly used to refer to all NAE i.e. both nasal emission and nasal

18 turbulence) or used a single measure to encapsulate all types of nasal airflow, as per the CAPS-

19 A-AM and SVANTE protocols. As with measures of resonance, a variety of scales were used

20 e.g. Konst et al. (2003) measured the percentage of target consonants realised with nasal

21 escape, whereas Lohmander & Persson (2008) used a five-point scale to measure the

22 frequency with which nasal air leakage occurred.

1   Velopharyngeal Function

2   An overall measure of velopharyngeal function based on perceptual speech

3   assessment featured in some studies, as shown in Table 3. This parameter was assessed in

4   addition to those associated with velopharyngeal impairment i.e. hypernasality and NAE.

5   Methods of assessment varied significantly, with some studies utilising a composite score

6   (Swanson et al. 2017; Lohmander et al. 2006; Gunther et al. 1998), whilst others used rating

7   scales (Zanzi et al. 2002; Persson et al. 2006; El Ezzi et al. 2015; Lohmander et al. 2017a;

8   Larsson et al. 2017). Dayashankara et al. (2011) referred to the use of a clinical diagnosis and

9   Hamming et al. (2009) did not report the methods used. Wren (2013) did not record if an

10   overall measure of velopharyngeal function was used in UK cleft units at age-3.

11   Intelligibility

12   Whitehill et al. (2011) stated that a key objective of all cleft teams is for their patient's

13   speech to be understood and that intelligibility is a measure of how successfully cleft teams

14   have achieved this (Whitehill et al. 2011: 293). Whilst the SVANTE protocol assesses

15   intelligibility, there are well-reported challenges in both defining and measuring intelligibility

16   (see Whitehill [2002]). Despite these challenges, intelligibility was assessed in 11 studies,

17   using methods of orthographic transcription, or categories with descriptors (see Table 3).

18   From a clinical perspective, less than half of the UK cleft units (42.85%) reported that

19   intelligibility was assessed at age-3 (Wren 2013).

20   Voice

21   Only 20% of sources reported on the parameter of voice; this is summarised in Table

22   3. In contrast, voice was reported on in 78.57% of UK cleft units Wren (2013). One explanation

13

1 may be that UK SLTs routinely screen voice (as a parameter of assessment on the CAPS-A, and

2 in clinical assessment using the GOS.SP.ASS) and apply these or similar categories to the

3 assessment of voice at age-3. As such there may be a different culture in the assessment of

4 voice in the UK at age-3 in comparison to other countries.

5 *What methods and rating scales are used to assess the parameters of speech identified?*

6 *Methods of Assessment*

7   Phonetic transcription, as previously stated, underpinned measures of consonant

8 production. This was supplemented in the studies by Chacon et al. (2017), Chapman et al.

9 (2008), Chapman (2004) and Morris & Ozanne (2003) by computer software to aid the analysis

10 of articulation and phonology. Computer-based analysis was also used by Gugsch et al. (2008)

11 in the evaluation of the voice-specific measures of formants and fundamental frequency.

12 Orthographic transcription using naïve listeners was utilised in studies assessing intelligibility,

13 as shown in Table 3.

14   Categorical rating scales were commonly used (Frey et al. 2018, Chacon et al. 2017,

15 Lohmander et al.2017a, Swanson et al. 2017, El Ezzi et al. 2015, Klintö et al. 2014b,

16 Dayashankara et al. 2011, Hamming et al. 2009, Chapman et al. 2008, Lohmander & Persson

17 2008, Frederickson et al. 2006, Lohmander et al. 2006, Persson et al. 2006, Chapman 2004,

18 Konst et al. 2003, Zanzi et al. 2002, Gunther et al. 1998, Lohmander-Agerskov 1998,

19 Lohmander-Agerskov et al. 1998), however binary scales also featured (Larsson et al. 2017,

20 Hamming et al. 2009, Pulkkinen et al. 2001). Despite this, there was limited commonality in

21 the categorical scales across the studies with a range of scales, number of scalar points, and

22 descriptors used. As an alternative to categorical scales, Hodge & Gotzke (2007) utilised a

14

1   continuous scale to measure intelligibility, and Chapman et al. (2008) used Direct Magnitude

2   Estimation (DME) to measure articulation proficiency and hypernasality.

3   *What types of speech samples are utilised in the assessment of 3-year-old children with CP±L?*

4   The type of speech samples used is outlined in Table 2. Twenty percent of the studies

5   did not provide any information about this. Of the remaining studies, single-word naming was

6   most frequently used, in some instances in combination with other speech samples. A total

7   of eight sources (29%) used the picture naming test developed as part of the Scandcleft Trial,

8   referred to as the Restricted Word List (as shown in Table 2). This picture naming test has

9   been developed in seven different languages to facilitate cross-linguistic speech comparisons

10   (CLISPI n.d) through the assessment of phonetically similar units of speech across language

11   and context (Lohmander *et al.* 2009: 348).

12   Samples of spontaneous speech were also frequently used occurring in 37.14% of the

13   sources, sometimes in addition to single word naming. Wren (2013) reported that 64.28% of

14   UK cleft units utilised a spontaneous speech sample in the assessment of 3-year olds with

15   CP±L. Sentence repetition was used by Chacon et al. (2017); participants repeated six

16   sentences taken from the GOS.SP.ASS (Sell et al*.* 1999). The GOS.SP.ASS was frequently used

17   across UK Cleft units with 85.7% using this assessment at age-3 (Wren 2013). However, no

18   information is provided as to the type of speech sample used i.e. if it was used in its most

19   common form (using the accompanying sentences) or if short phrases or single words were

20   used.

15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

2   This scoping review aimed to map assessment practices at age-3 in the CP±L

3   population. This review found that the evaluation of palatal structure and function through

4   the assessment of resonance, NAE, overall measures of velopharyngeal function and CSCs was

5   a central feature of assessment in many of the sources reviewed, even at this young age.

6   Whilst this approach to assessment is in line with a medical model, it was used alongside a

7   linguistic approach to assessment which was underpinned by phonetic transcription.

8   Although many of the sources had the same assessment objective i.e. to assess palatal

9   function for speech, the assessment methods varied across the studies. The majority of

10   studies reported NAE in combination (e.g. both emission and turbulence in a single measure)

11   including the SVANTE, which is designed to measure speech outcomes at age-3. In contrast,

12   UK cleft units reported nasal emission and turbulence separately, this may be historical and

13   recorded as per the CAPS-A and the GOS.SP.ASS. There is a need for consensus in the UK as

14   to whether the objective at age-3 is to assess the type or overall presence of NAE. To guide

15   this decision making further research is needed to understand the implications of rating nasal

16   emission and turbulence separately or in combination, on listener reliability.

17   This scoping review indicates that an overall measure of velopharyngeal function was

18   assessed using varying methods. This included the use of validated and unvalidated rating

19   scales and clinical opinions. There is an emphasis on overall measures in the reporting of

20   speech outcomes in the CP±L population and this has been recommended by ICHOM (2017)

21   for use at older ages. The overall measure VPC-Rate (Lohmander et al. 2009) shows potential

22   for use with this age group; it has been shown to be reliable and efficient (Lohmander et al.

23   2017b) and has featured in a number of studies reporting outcomes associated with the

16

1   Scandcleft Trial, and used in a similar form in the SVANTE. In the UK, an overall score for

2   resonance and NAE is used to measure 5-year speech outcomes against National Standards

3   for Speech (Britton et al. 2014) and a velopharyngeal composite score using the CAPS-A has

4   also been validated (Pereira et al. 2013). Further research should compare the reliability of

5   hypernasality, NAE and overall velopharyngeal function at age-3, to support decisions

6   regarding the selection of outcome measures at this age

7   A linguistic approach underpinned by phonetic transcription allowed for both

8   phonetic and phonological analysis of consonant production and the subsequent

9   identification of cleft specific CSCs and calculations of PCC. The frequent reporting of

10   consonant production using cleft summary patterns highlights the importance of this

11   approach during the preschool years. Again, challenges in comparing speech outcomes across

12   the studies arise from the use of different criteria and summary patterns (Sell 2005) and the

13   need for consensus is relevant both at age-3 -years and for other age groups. PCC-A, using a

14   controlled speech sample allows for cross-linguistic comparisons and takes into consideration

15   sound distortions occurring as part of normal development, which is particularly appropriate

16   for assessment at age-3. PCC-A may, therefore, offer a partial solution although it has the

17   disadvantage of not capturing the qualitative nature of speech errors. In addition, to be

18   meaningful, PCC-A needs to be used alongside normative or comparison data (longitudinal

19   data or data to compare groups), which may account for why PCC-A has not been adopted in

20   the UK assessments at age-3. Perhaps a solution to the need for normative/comparison data

21   may be to compare PCC-A at ages 3 and 5-years to monitor progress, or to use assessments

22   with available normative data i.e. the DEAP (Dodd et al. 2002), although this has the

23   disadvantage of not being specifically designed to assess cleft speech.

17

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1   The assessment of speech from a developmental perspective has not been a core

2   feature in the reporting of cleft speech outcomes (which may account for why the assessment

3   of phonology only featured in 50% of UK cleft centres [Wren 2013]). However, this scoping

4   review highlights the use of both PCC-A and developmental phonological processes in speech

5   assessments at age-3 demonstrating the use of a combined linguistic and developmental

6   approach to consider delayed, typical and atypical speech production. The importance of a

7   developmental approach is evidenced by studies in this review, as both Chacon et al. (2017)

8   and Hutters et al. (2011) reported that children with CP±L present with more developmental

9   phonological processes at age-3 years than their non-cleft peers.

10   The assessment of intelligibility in the studies included in this review recognises the

11   need for a functional approach to speech assessment. Whilst not a cleft specific outcome

12   measure, a functional approach to assessment at age-3 is important given that many children

13   start attending pre-school education in the UK at this age and thus interact with a wider social

14   group. Most studies in the review used ordinal scales to assess intelligibility despite evidence

15   questioning their validity to measure this parameter (Whitehill et al. 2011). In the context of

16   the International Classification of Functioning, Disability and Health: Children and Youth

17   Version (ICF-CY; World Health Organization [WHO], 2007) McLeod et al. (2012: 649) report

18   that intelligibility is influenced by both production factors (Body Functions) and contextual

19   factors (Environmental Factors). Safaiean et al. (2018), Willadsen & Poulsen (2012) and Hodge

20   and Gotzke (2007) used orthographic transcription by naïve listeners. This method addresses

21   concerns regarding contextual factors and the validity of expert (cleft SLT) listeners rating

22   intelligibility but would be impractical to employ in clinical practice. Indeed, The Intelligibility

23   in Context Scale (ICS) (McLeod at al. 2012), which uses parent ratings of intelligibility, shows

18

1 potential for use with the cleft population. The ICS is validated and recommended by ICHO

2 (2017) for use at age 5 and 12 years but has yet to be validated for use at age-3; the challenges

3 and complexity of measuring intelligibility at this age persist.

4 In addition to intelligibility, the reliability and validity of using ordinal/categorical

5 scales to measure resonance and NAE has been challenged by growing evidence to suggest

6 that ratio or category-ratio scales may be more valid and reliable measures (Yamashita et al.

7 2018, Baylis et al. 2015). However, in the studies in this review, only Hodge & Gotzke (2007)

8 and Chapman et al. (2008) utilised non-categorical scales, using a continuous scale and DME

9 respectively, with good levels of reliability reported in both studies. The impact of different

10 rating scales on the validity and reliability of listener judgements has therefore yet to be fully

11 explored in 3-year-old children with CP±L and further research is required to determine if

12 findings at this age mirror that of other age groups.

13 Three types of speech samples were used in the studies: single word production,

14 spontaneous speech samples, and more infrequently, sentence repetition. The variety of

15 assessments used to assess single words is striking and well exemplified by Wren (2013) which

16 indicates there is no preferred single word assessment to evaluate speech at age-3 in the UK.

17 The use of supplementary assessment materials or unnamed picture naming assessments

18 suggests existing assessments may not be wholly adequate for the comprehensive

19 assessment of speech at age-3 in the CP±L population. The Restricted Word List (Lohmander

20 et al 2009), developed for cross-linguistic comparisons, would facilitate multi-centre research

21 and comparisons of outcomes internationally. However, given its design to allow cross-

22 linguistic comparisons it does not assess the full range of oral pressure consonants in English

23 and there is a need to expand upon this assessment to assess speech more comprehensively

19

1   (as per the SVANTE). Nonetheless, this review highlights its potential use in speech

2   assessments at age-3.

3   Speech samples may need to be used in combination to comprehensively assess speech

4   at age-3 when combining approaches to assessment i.e. using a medical model and linguistic,

5   developmental and functional approaches. Most common in the studies are single-word

6   naming samples, given that this is an age-appropriate task and there is evidence at age-5 that

7   single word naming samples enhance listener reliability in judgements of consonant

8   production (Klintö et al. 2011). However, evidence that there may be variability in 'speech

9   performance between single words and conversational speech' (Sweeney 2011: 206)

10   indicates that this sample may not be sufficient in isolation, particularly when measuring the

11   core parameters of hypernasality, NAE and consonant production. Although 37.14% of the

12   studies used a spontaneous speech sample, Klintö et al. (2014b), Persson et al. (2006), and

13   Lohmander and Persson (2008) favoured single words to assess resonance and NAE. Klintö et

14   al (2014b) used only single words reporting that a 'representative and standardised speech

15   sample with connected speech' (Klintö et al. 2014b pp.277) could not be achieved, thus

16   highlighting the challenges of spontaneous speech samples at this age. An alternative is

17   sentence or phrase repetition. Wren (2013) found that 85.7% of UK cleft units utilised the

18   GOS.SP.ASS and the accompanying speech sample (Sell et al. 1999) in addition to single word

19   assessments. The familiarity UK Cleft SLTs have with the GOS.SP.ASS may help them to

20   support 3-years olds in completing the GOS.SP.ASS sentences, sometimes as short phrases,

21   and may facilitate reliable listener judgements. The SVANTE also includes both single word

22   and sentence level speech samples.

23   <u>Limitations</u>

20

1   The parameters of speech identified in the study are those established as core to the

2   assessment of cleft speech and do not constitute novel information. The number of sources

3   (n=8/35 22.9%) included in this review which report on outcomes associated with the

4   Scandcleft Trial is a testament to the success of this research group in generating new

5   knowledge. However, it is important to recognise the impact this has on the overall results of

6   this scoping review, introducing an element of bias into the results given that the same

7   assessment principles and processes were used across this subset of studies. The inclusion of

8   sources in languages other than English would have provided a more diverse overview of

9   assessment practices at age-3 in other parts of the world and could potentially have informed

10   UK practice.

Conclusion

12   Whilst there are core parameters fundamental to the assessment of cleft speech, this

13   scoping review also highlights the unique assessment requirements of 3-year-old children,

14   particularly with reference to the selection of the speech sample and the importance of using

15   a developmental approach to assessment. The need to assess functional outcomes at this age

16   i.e. intelligibility, in addition to cleft specific outcomes, i.e. resonance and NAE, is highlighted.

17   Unlike age-5, speech outcomes at age-3 in the CP±L population have not been a

18   central focus in the UK or internationally; this is perhaps reflected in the variety of assessment

19   methods noted in this review. However, more recent studies relating to the Scandcleft study

20   highlight the potential of the Restricted Word List (Lohmander et al. 2009) for use in

21   assessments at age-3, particularly as this allows for cross-linguistic comparisons. This scoping

22   review highlights the need for a consolidated approach to assessing CP±L speech at age-3,

21

1 however, further research is needed to provide a foundation for assessments at this

2 important age in order to develop assessment protocols. This should consider the extent to

3 which 3-year-oldchildren with CP±L can complete different speech samples and the impact

4 this has on the parameters of speech which can be assessed and the validity and reliability of

5 listener judgements of those parameters using different measures.

22

Table 1. Search terms

Participant Type		Concept
Age	**Diagnosis**	
preschool	cleft palate	articulation
toddler	cleft lip and palate	cleft speech characteristics
kindergarten		cleft type characteristics
nursery		compensatory articulation
3 years old		velopharyngeal dysfunction (VPD)
aged 3 years		velopharyngeal insufficiency (VPI)
3 years of age		voice quality/dysphonia
		speech articulation
		speech analysis
		speech assessment
		speech and language assessment

Table 2. Articles included in the review recorded by author and date, country of origin, methods and methodology, and speech assessment

Author & Date	Country of Origin	Age of Participants	Cleft Type*	Methodology and Methods	Speech Assessment
Frey et al. (2018)	USA	15-36 mo	CPL	CPL recruited from a larger randomised experimental group design	Spontaneous speech sample
Willadsen et al. (2018)	Denmark	36 mo, mean 3.0 y, (Range 2.93-3.23 y)	UCLP NCCG	Randomised Control Trial (randomised by age at hard palate repair)	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
Chacon et al. (2017)	Australia	3 y-olds: 2;10-3;11 y 5 y-olds: 4;10-5;09 y	CPO UCLP BCLP	Cross-sectional observation study	GOS.SP.ASS x6 sentences DEAP Articulation subtest (Dodd et al. 2002), single words Informal list of 15 monosyllabic and 10 polysylabbic words
Larsson et al. (2017)	Sweden	Adopted group: 38 mo (Range 35-43 mo); Control group (also cleft) 37 mo (Range 34-42 mo)	UCLP	Cohort Study	SVANTE (Lohmander et al. 2005, 2015). Single word naming and spontaneous speech sample.
Lohmander et al. (2017a)	Sweden	Normative Data gained at 3, 5, 7, 10, 16, 19 y 3 y-olds: Mean age 2;11 (Range 2 y 8 mo-3 y 3 mo)	NCCG-Normative Data	Normative Referencing	SVANTE (Lohmander et al. 2005, 2015). Single word naming, sentence repetition, spontaneous speech sample.
Raud Westberg et al. (2017)	Sweden	Longitudinal assessment at age 1, 1 y 6 m and 3 y	UCLP	Randomised Control Trial	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
Safaiean et al. (2017)	Iran	3-5 y	CPL NCCG	Cohort Study	Persian Speech Intelligibility Test (Heydari et al. 2011) which involves picture naming
Swanson et al. (2017)	USA	Initial assessment at age 3 y	SMCP	Retrospective review	Pittsburgh Weighted Speech Scoring (McWilliams & Philips 1979) (no information provide as to what the speech sample was)
Klintö et al. (2016)	Sweden	Mean = 36 mo (Range 35-38 mo)	UCLP NCCG	Randomised Control Trial (some participants) Prospective longitudinal	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
El Ezzi et al. (2015)	Switzerland	3 y (no specific age range)	UCLP CP0	Retrospective cohort study	Not stated
Klintö et al. (2015)**same procedure as Klintö 2014a)	Sweden	Median: 36 mo (Range 35.5-37 mo)	UCLP NCCG	Prospective comparative study 5 y, and retrospective comparative study at 3 y	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
Klintö et al. (2014a)	Sweden	Median: 36 mo (Range 35.5-37 months)	UCLP NCCG	Randomised Control Trial (randomised by age at hard palate repair)	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009) Sample of spontaneous speech
Klintö et al. (2014b)	Sweden	Group 1: mean age 36 mo Group 2: mean age 35-39 mo	UCLP	Prospective comparison study	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009) Sample of spontaneous speech

Wren (2013)	UK	N/A Report on assessment practises of speech and language therapists	N/A	Survey of 14 UK Cleft Units (% of cleft units)	South Tyneside Assessment of Phonology (STAP) (Armstrong 1992): 21.42% Diagnostic Evaluation of Articulation & Phonology (Dodd *et al.* 2002): 28.57% PACS TOYS (Harding & Grunwell 1995): 28.57% Phonological Screening Assessment (PSA) (Stevens & Isles 2001): 14.28% Hart Screen Speech Assessment: 14.28% CLEAR Phonology Screening Assessment (CLEAR Resources 2006): 28.57% Great Ormond Street Speech Assessment (GOS.SP.ASS) (Sell et al. 1999): 85.71% Renfrew Action Picture Test (Renfrew 2011): 14.28% Preschool Language Scale Version 3 (PLS-3) (Boucher and Lewis 1997): 21.42% Spontaneous Speech: 64.28%
Willadsen (2012)	Denmark	3 y-old group: 36 mo	UCLP	RCT (age at HP repair)	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
Willadsen & Poulsen (2012)	Denmark	36 mo	UCLP NCCG	Cross-sectional study	Single word naming using the naming test developed in the Scandcleft study (Lohmander et al. 2009)
Dayashankara et al. (2011)	India	18-36 mo	UCLP BCLP CPO	Prospective cohort study	Not stated
Hamming et al. (2009)	USA	Younger age group: 3-4 y (up to 4.11)	UCLP BCLP CPO SMCP	Retrospective longitudinal	Not stated
Chapman et al. (2008)	USA	33-42 mo. Median = 39 mo	CPL	Prospective longitudinal	Goldman-Fristoe Test of Articulation (Goldman & Fristoe 1986), single word naming, spontaneous speech sample
Gugsch et al. (2008)	Germany	Mean age at baseline: 3 y11mo (Range 3.7-4.7)	UCLP BCLP CPO	Longitudinal prospective	Voice recordings in standardised form
Lohmander & Persson (2008)	Sweden	Longitudinal speech assessments at 18 mo, 3, 5, and 7 y	UCLP NCCG	Prospective Longitudinal study	Single word picture naming test
Hodge & Gotzke (2007)	Canada	Age range: 3 y 5 mo - 6 y 7 mo Mean 4 y 6 mo	UCLP BCLP CPO	Prospective between groups design	Single word repetition Spontaneous speech sample
Frederickson et al.(2006)	USA	Age range: 2 y 9 mo - 3 y 8 mo	UCLP NCCG	Prospective cross-sectional study	Goldman Fristoe Test of Articulation (Goldman & Fristoe 1986), single word naming
Lohmander et al. (2006)	Sweden	Longitudinal speech assessments at 3, 5, 7, 10 y	UCLP	Prospective longitudinal study (3-year results retrospectively collected)	Assessment not stated at age 3, data collected from records
Persson et al. (2006)	Sweden	Longitudinal speech assessments at 3, 5, 7, 10 y Age 3: Mean age: 3 y 0 mo Range 3 y 0 mo -3 y 3 mo	CPO NCCG	Longitudinal study (3-7yrs)	Single word naming Spontaneous speech if unable to complete naming
Chapman (2004)	USA	Longitudinal assessment, final assessment at 39 mo	UCLP BCLP	Multi-site longitudinal study	Spontaneous speech sample

Konst et al. (2003)	The Netherlands	Longitudinal follow up at age 2, 2.5 and 3 y	UCLP	Randomised Control Trial	Spontaneous speech sample
Morris & Ozanne (2003)	Australia	Assessment at age 2 y and 3 y	UCLP BCLP CPO	Comparative groups- longitudinal	Spontaneous speech sample. Goldman Fristoe Test of Articulation (Goldman & Fristoe 1986), single word naming
Zanzi et al. (2002)	Switzerland	Age 3.5 y	UCLP BCLP	Retrospective review	Spontaneous speech/Interview
Hattee et al. (2001)	UK	Longitudinal assessment at 9 mo, 18 mo and 3 y	UCLP CPO	Longitudinal study	South Tyneside Assessment of Phonology (STAP) (Armstrong & Ainsley 1988)
Hutters et al. (2001)	Denmark	Mean: 3 y 1 mo Range 2 y 11mo -3 y 3 mo	UCLP BCLP	Prospective cross-sectional study	Picture naming
Pulkkinen et al. (2001)	Finland	Longitudinal assessment at 3, 6, and 8 y	UCLP CPO	Longitudinal	Not stated
Gunther et al. (1998)	USA	Assessment at age 3 y	CPL	Retrospective review	Pittsburgh Weighted Speech Scale (McWilliams & Philips 1979) - no information on type of speech sample
Lohmander-Agerskov (1998)	Sweden	Longitudinal assessment at 9 mo, 18 mo, and 3, 5, 7, 10, 16 and 19 y	UCLP BCLP	Results combined from: Longitudinal study- retrospective and Cross-sectional study	Picture naming
Lohmander-Agerskov et al. (1998)	Sweden	Assessment at 3 and 5 y	UCLP BCLP CPO	Longitudinal study	Picture naming

*UCLP= Unilateral Cleft Lip and Palate; BCLP= Bilateral Cleft Lip and Palate; CPO= Cleft Palate Only; SMCP= Submucus cleft palate; CPL= Cleft Palate +/- Cleft Lip (used when specific type of cleft is not stated); NCCG= Non-Cleft Control Group

Additional References for Assessments:

ARMSTRONG, S. and AINLEY, M., 1992, South Tyneside assessment of phonology. Stass Publications.

BOUCHER, J. and LEWIS, V., 1997. Preschool language scale-3 (UK). London: The Psychological Corporation.

CLEAR RESOURCES., 2006, CLEAR Phonology Screening Assessment (Spilsby: Clear Resources).

GOLDMAN, R. and FRISTOE, M., 1986. Goldman Fristoe test of articulation. American Guidance Service.

GRUNWELL, P., and HARDING, A., 1995, PACS TOYS. Windsor, UK: NFER-Nelson.

HEYDARI, S., TORABI NEZHAD, F., AGHA RASOULI, Z., and HOSEYNI, F., 2011, Development of speech intelligibility measurement test for 3 to 5 years old normal children. Bimonthly Audiology-Tehran University of Medical Sciences, 20(1), 47-53.

STEVENS, N. and ISLES, D., 2011, Phonological Screening Assessment (PSA) (London: Taylor and Francis)

RENFREW, C.E., 2011. Action picture test. Speechmark.

Table 3. Summary of the parameters of speech assessed. Articles which utilised Percentage Consonants Correct categorised as assessing both articulation and phonology.

Author & Date	Parameters of assessment							
			Resonance					
	Articulation	Phonology	Hypernasality	Hyponasality	Nasal Airflow Errors (NAE)	Velopharyngeal Function	Intelligibility	Voice
Frey et al. (2018)							✓ Categorisation	
Willadsen et al. (2018)	✓	✓						
	PPC- Adjusted for age: age appropriate distortions were counted as correct, as were accompanying nasality or nasal airflow errors if the production did not cross a phoneme boundary							
Chacon et al. (2017)	✓	✓	✓					
	PCC PCC-R: in which sound distortions were also counted as correct. Percentage Vowels Correct.							
Larsson et al. (2017)	✓	✓			✓ Nasal air leakage	✓ Binary rating scale of yes/no to indicate competent velopharyngeal function.		
	PCC-Adjusted for age: misarticulations of /s, ɕ/ were scored as correct at age 3 with reference to typically developing 3-year olds. Audible nasal leakage or weak articulation was scored correctly. PCC by articulatory manner PCC by articulatory placement							
Lohmander et al. (2017a)	✓	✓	✓	✓	✓ Nasal air leakage	✓ A three-point scale to describe velopharyngeal function for speech	✓ A three-point ordinal scale	
		PCC						
Raud Westberg et al. (2017)	✓	✓						
	PCC-Adjusted for age							
Safaiean et al. (2017)	✓	✓	✓				✓ Orthographic transcription by naïve listeners	
		PCC						
Swanson et al. (2017)	✓		✓	✓	✓ Nasal emission/ nasal escape/nasal turbulence	✓ Composite score		✓ Hoarseness/breathiness, volume/tension
Klintö et al. (2016)	✓	✓						
	PCC-Adjusted: substitutions and simplifications of phonemes used							

Study								
	by more than 10% of children are scored correctly							
El Ezzi et al. (2015)	✓		✓	✓	✓ Audible nasal emission	Borel-Maisonny SIMPLIFIED score, which uses a four-point scale with descriptors.	✓ Not directly assessed but scalar point d) on Borel-Maisonny refers to 'no intelligibility'	✓ No information provided as to method
Klintö et al. (2015)*same procedure as Klintö 2014a)	✓	✓						
	PCC-Adjusted: substitutions and simplifications of phonemes used by more than 10% of children are scored correctly							
Klintö et al. (2014a)	✓	✓						
	PCC-Adjusted: substitutions and simplifications of phonemes used by more than 10% of children are scored correctly							
Klintö et al. (2014b)	✓	✓	✓	✓				
	PCC-Adjusted: substitutions and simplifications of phonemes used by more than 10% of children are scored correctly							
Wren (2013)	Consonant Inventory: 85.71% CSCs: 78.57%	50%	Resonance (hyper/hypo) 92.85%	Nasal emission: 85.71% Nasal Turbulence: 78.57%			42.85%	78.57%
Willadsen (2012)	✓	✓						
	PPC- Adjusted for age: age appropriate distortions were counted as correct, as were accompanying evidence of nasality or nasal airflow errors if the production did not cross a phoneme boundary							
Willadsen & Poulsen (2012)	✓	✓					✓ Orthographic transcription by naïve listeners	
	PPC- Adjusted for age: age appropriate distortions were counted as correct, as were accompanying evidence of nasality or nasal airflow errors if the production did not cross a phoneme boundary							
Dayashankara et al. (2011)	✓		✓			✓ Referred to as clinical diagnosis by surgeon and SLT	✓ Categorisation	
Hamming et al. (2009)			✓	✓		✓ Method not stated		✓ i-hoarseness and other dysphonias rated as inconsistent/consistent ii- Consistent dysphonias rated as mild/moderate/severe.
Chapman et	✓	✓	✓					

al. (2008)	Total Consonants Correct PCC by articulatory manner							
Gugsch et al. (2008)	✓ vowels							✓ Changes in the fundamental frequency of vowel production after surgery
Lohmander & Persson (2008)	✓	✓	✓		✓ Nasal air leakage			
	PCC PCC by articulatory manner PCC by articulatory placement							
Hodge & Gotzke (2007)							✓ Orthographic transcription by naïve listeners	
Frederickson et al (2006)	✓	✓	✓					
	PCC							
Lohmander et al. (2006)	✓		✓		✓ Nasal air leakage	✓ Composite measure of velopharyngeal insufficiency, formed by combining the two highest scores for hypernasality, nasal leakage and weak pressure consonants.	✓ Categorisation	
Persson et al. (2006)	✓		✓	✓	✓ Audible nasal emission	✓ 5 point scale to provide a global measure of velopharyngeal impairment.		
Chapman (2004)	✓		✓	✓				
Konst et al. (2003)	✓	✓			✓ Nasal escape			
Morris & Ozanne (2003)	✓	✓						
	PCC: with consonant distortions classified as correct. PCC by articulatory manner							
Zanzi et al. (2002)	✓				✓ Nasal emission	✓ Borel-Maisonny score, which uses a five-point scale with descriptors		
Hattee et al. (2001)	✓							
Hutters et al. (2001)	✓	✓						
Pulkkinen et al. (2001)	✓		✓		✓ Nasal emission			
Gunther et al. (1998)			✓		✓ Nasal escape	✓ Composite score		✓ Hoarseness rated on 4 point scale
Lohmander-Agerskov (1998)	✓		✓	✓	✓ Nasal escape-velopharyngeal friction sounds		✓ Categorisation (also included voice)	✓ Voice quality and intelligibility rated as normal, mildly distorted, severely distorted

Lohmander-Agerskov et al. (1998)	✓			✓	✓	Nasal escape-velopharyngeal friction sounds		✓ Categorisation	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1. Adapted PRISMA Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al. 2016)
Flow Diagram selection of sources

Identification

Records identified through
database searching
(n = 953)

Additional records identified
through other sources
(n = 3)

Screening

Records after duplicates removed
(n = 481)

Records screened by title
(n = 481)

Records excluded
(n = 268) e.g. genetic case studies, descriptions
of surgical procedures, dental studies

Eligibility

Records screened by abstract
(n = 259)

Records excluded, based upon abstract
(n = 192) e.g. language, phonological
awareness assessments, no reference to
age 3

Articles/Data assessed in full
for eligibility
(n = 57)

Full-text articles excluded, based
upon inclusion criteria
(n = 22) (e.g. those not reporting
assessment procedures)

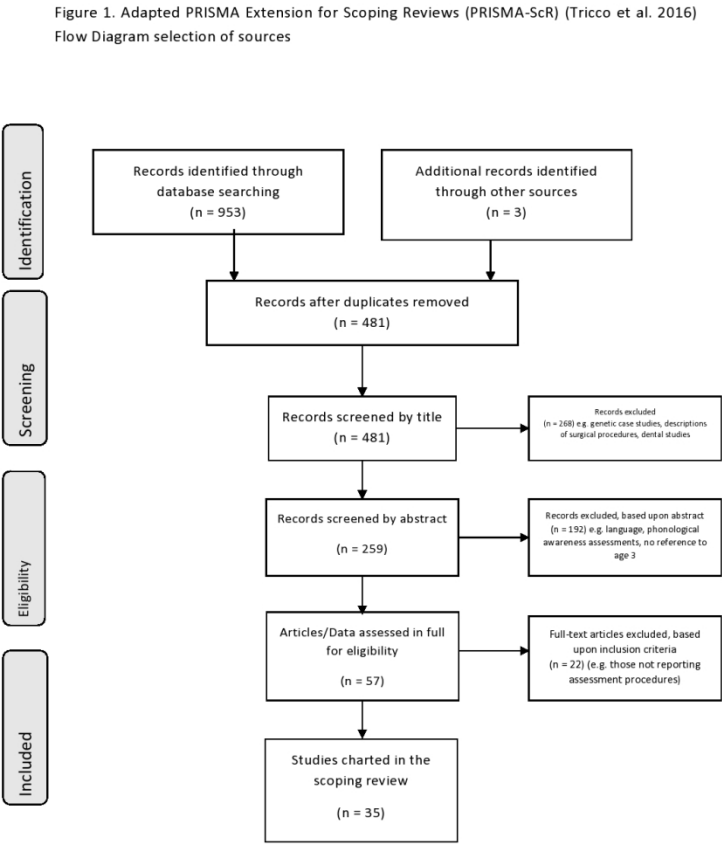**Included**

Studies charted in the
scoping review
(n = 35)

Figure 1. Adapted PRISMA Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al. 2016) Flow Diagram
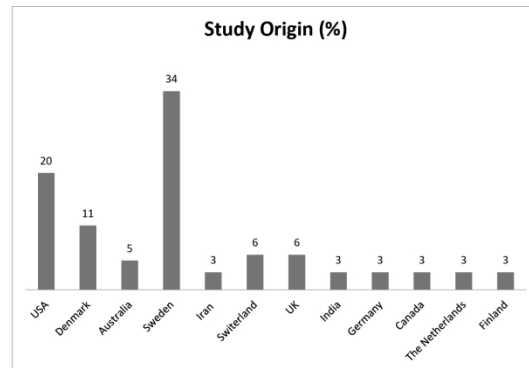selection of sources

437x618mm (72 x 72 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2. Study Origin

297x210mm (200 x 200 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
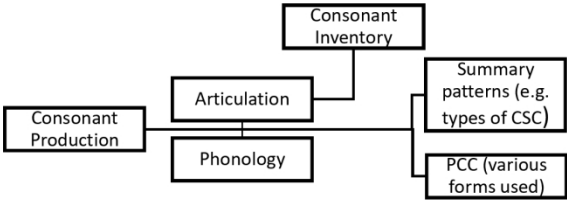46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3. Methods used to assess consonant production

210x297mm (200 x 200 DPI)

References

ARKSEY, H., & O'MALLEY, L., 2005, Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19-32.

BAYLIS, A., CHAPMAN, K. and WHITEHILL, T.L., 2015, Validity and reliability of visual analogue scaling for assessment of hypernasality and audible nasal emission in children with repaired cleft palate. The Cleft Palate-Craniofacial Journal, 52(6), 660-670.

BRITTON, L., ALBERY, L., BOWDEN, M., HARDING-BELL, A., PHIPPEN, G. and SELL, D., 2014, A cross-sectional cohort study of speech in five-year-olds with cleft palate±lip to support development of national audit standards: benchmarking speech standards in the United Kingdom. The Cleft Palate-Craniofacial Journal, 51(4), 431-451.

CAVALHEIRO, M.G., LAMÔNICA, D.A.C., DE VASCONSELLOS HAGE, SIMONE ROCHA and MAXIMINO, L.P., 2019, Child development skills and language in toddlers with cleft lip and palate. International Journal of Pediatric Otorhinolaryngology, 116, 18-21.

CHAPMAN, K.L., BAYLIS, A., TROST-CARDAMONE, J., CORDERO, K.N., DIXON, A., DOBBELSTEYN, C., THURMES, A., WILSON, K., HARDING-BELL, A. and SWEENEY, T., 2016, The Americleft Speech Project: a training and reliability study. The Cleft Palate-Craniofacial Journal, 53(1), 93-108.

CHAPMAN, K.L, and Willadsen, E., 2011, The Development of Speech in Children with Cleft Palate. In S. Howard S., and A. Lohmander (eds), Cleft Palate Speech: Assessment and Intervention (Chichester: Wiley-Blackwell), pp. 23-42.

CHAPMAN, K.L., HARDIN-JONES, M. and HALTER, K.A., 2003, The relationship between early speech and later speech and language performance for children with cleft lip and palate. Clinical linguistics & phonetics, 17(3), 173-197.

CLEFT PALATE INTERNATIONAL SPEECH ISSUES (CLISPI). n.d, Speech material to be collected at different ages. (available at: https://www.clispi.org/en/node/17) [accessed on 11 August 2018].

1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

CRANE., 2019, Annual Report on Cleft Lip and/or Palate 2019 [online] https://crane-database.org.uk/?!.iD=3l [accessed 01 October 2019].

DODD, B., ZHU, H., CROSBIE, S., HOLM, A. and OZANNE, A., 2002, Diagnostic evaluation of articulation and phonology (DEAP). Psychology Corporation.

DOSMAN, C. F., ANDREWS, D., and GOULDEN, K. J., 2012, Evidence-based milestone ages as a framework for developmental surveillance. Paediatrics & Child Health, 17(10), 561-568.

DUDAS, J. R., DELEYIANNIS, F. W., FORD, M. D., JIANG, S., and LOSEE, J. E., 2006. Diagnosis and treatment of velopharyngeal insufficiency: Clinical utility of speech evaluation and videofluoroscopy. Annals of Plastic Surgery, 56(5), 511-517.

GRUNWELL, P. and SELL, D., 2001, Speech and Cleft Palate/Velopharyngeal Anomalies. In A.C.H Watson., D.A. Sell., P. Grunwell (eds), Management of Cleft Lip and Palate (London: Whurr Publishers), pp. 68-86

HARDING, A. and GRUNWELL, P., 1995. Active versus passive cleft-type speech characteristics: implications for surgery and therapy. International Journal of Language & Communication Disorders, 30(S1), 242-247.

HENNINGSSON, G., KUEHN, D.P., SELL, D., SWEENEY, T., TROST-CARDAMONE, J.E. and WHITEHILL, T.L., 2008, Universal parameters for reporting speech outcomes in individuals with cleft palate. The Cleft Palate-Craniofacial Journal, 45(1), 1-17.

HOWARD, S., 2011, Phonetic Transcription for Speech Related to Cleft Palate. In S. Howard S., and A. Lohmander (eds), Cleft Palate Speech: Assessment and Intervention (Chichester: Wiley-Blackwell), pp. 127-144.

HUTTERS, B., BAU, A. and BRØNDSTED, K., 2001, A longitudinal group study of development in Danish children born with and without cleft lip and palate. International journal of language & communication disorders, 36(4), 447-470.

2

HUTTERS, B. and BRØNDSTED, K., 1987, Strategies in cleft palate speech--with special reference to Danish. The Cleft palate journal, 24(2), 126-136.

INTERNATIONAL CONSORTIUM FOR HEALTH OUTCOME MEASUREMENT (ICHOM)., 2017, Cleft Lip and Palate Data Collection version 3.0.6. [online] Available: http://www.ichom.org/medical-conditions/cleft-lip-palate/ [accessed 10 June 2018].

JOHN, A., SELL, D., SWEENEY, T., HARDING-BELL, A. and WILLIAMS, A., 2006, The cleft audit protocol for speech—augmented: A validated and reliable measure for auditing cleft speech. The Cleft palate-craniofacial journal, 43(3), 272-288.

KHALIL, H., PETERS, M., GODFREY, C. M., MCINERNEY, P., SOARES, C. B., & PARKER, D., 2016, An evidence-based approach to scoping reviews. Worldviews on Evidence-Based Nursing, 13(2), 118-123.

KLINTÖ, K., SALAMEH, E., SVENSSON, H. and LOHMANDER, A., 2011, The impact of speech material on speech judgement in children with and without cleft palate. International journal of language & communication disorders,46(3), 348-360.

LOHMANDER, A., LUNDEBORG, I. and PERSSON, C., 2017a, SVANTE–The Swedish Articulation and Nasality Test–Normative data and a minimum standard set for cross-linguistic comparison. Clinical linguistics & phonetics, 31(2), 137-154.

LOHMANDER, A., PERSSON, C., WILLADSEN, E., LUNDEBORG, I., ALALUUSUA, S., AUKNER, R., BAU, A., BOERS, M., BOWDEN, M. and DAVIES, J., 2017b, Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 4. Speech outcomes in 5-year-olds-velopharyngeal competency and hypernasality. Journal of Plastic Surgery and Hand Surgery, 51(1), 27-37.

LOHMANDER, A., WILLADSEN, E., PERSSON, C., HENNINGSSON, G., BOWDEN, M. and HUTTERS, B., 2009, Methodology for speech assessment in the Scandcleft project—an international

3

randomized clinical trial on palatal surgery: experiences from a pilot study. The Cleft Palate-Craniofacial Journal, 46(4), 347-362.

LOHMANDER, A., BORELL, E., HENNINGSSON, G., HAVSTAM, C., LUNDEBORG, I. & PERSSON, C., 2005. SVANTE Svenskt artikulations- och nasalitets-test, Manual [In Swedish]. (Lund: Studentlitteratur).

MCLEOD, S. and BAKER, E., 2017, Children's Speech: An Evidence-Based Approach to Assessment and Intervention (Boston: Pearson)

MCLEOD, S., HARRISON, L.J. and MCCORMACK, J., 2012, The intelligibility in context scale: Validity and reliability of a subjective rating measure. Journal of Speech, Language, and Hearing Research, 55(2), 648-656.

MCWILLIAMS, B. and PHILIPS, B.J., 1979, Velopharyngeal Incompetence audio seminars in speech pathology (Philadelphia: WB Saunders)

MUNN, Z., PETERS, M. D., STERN, C., TUFANARU, C., MCARTHUR, A., & AROMATARIS, E., 2018, Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. BMC Medical Research Methodology, 18(1), 143.

PETERS, M.D., GODFREY, C.M., KHALIL, H., MCINERNEY, P., PARKER, D. and SOARES, C.B., 2015, Guidance for conducting systematic scoping reviews. International journal of evidence-based healthcare, 13(3), 141-146.

PEREIRA, V. J., SELL, D., and TUOMAINEN, J., 2013, Effect of maxillary osteotomy on speech in cleft lip and palate: Perceptual outcomes of velopharyngeal function. International Journal of Language & Communication Disorders, 48(6), 640-650.

PRATHANEE, B., LORWATANAPONGSA, P., ANANTAPONG, D. and BUAKANOK, N., 2011, Thai speech parameters for patients with cleft palate in a universal reporting system. Asia Pacific Journal of Speech, Language and Hearing, 14(1), 31-49.

4

SAFAIEAN, A., JALILEVAND, N., EBRAHIMIPOUR, M., ASLESHIRIN, E. and HIRADFAR, M., 2017, Speech intelligibility after repair of cleft lip and palate. Medical Journal of The Islamic Republic of Iran (MJIRI), 31(1), 500-504.

SELL, D., and PEREIRA, V., 2015, Speech: perceptual and functional outcomes. In J. Losee., and RE, Kirschner (eds), Comprehensive Cleft Care. (London: McGraw Hill; editors), pp. 689-714

SELL, D., JOHN, A., HARDING-BELL, A., SWEENEY, T., HEGARTY, F. and FREEMAN, J., 2009, Cleft Audit Protocol for Speech (CAPS-A): a comprehensive training package for speech analysis. International Journal of Language & Communication Disorders, 44(4), 529-548.

SELL, D., 2005, Issues in perceptual speech analysis in cleft palate and related disorders: a review. International Journal of Language & Communication Disorders, 40(2), 103-121.

SELL, D. and GRUNWELL, P. 2001, Speech Assessment and Therapy. In A.C.H Watson., D.A. Sell., P. Grunwell (eds), Management of Cleft Lip and Palate (London: Whurr Publishers), pp. 227-257

SELL, D., HARDING, A. and GRUNWELL, P., 1999, GOS. SP. ASS.'98: an assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). International Journal of Language & Communication Disorders, 34(1), 17-33.

SHRIBERG, L.D., AUSTIN, D., LEWIS, B.A., MCSWEENY, J.L. and WILSON, D.L., 1997, The percentage of consonants correct (PCC) metric: Extensions and reliability data. Journal of Speech, Language, and Hearing Research, 40(4), 708-722.

SHRIBERG, L.D. and KWIATKOWSKI, J., 1982, Phonological disorders III: A procedure for assessing severity of involvement. Journal of speech and Hearing Disorders, 47(3), 256-270.

SWEENEY, T., 2011, Nasality- Assessment and Intervention. In S. Howard., and A. Lohmander (eds), Cleft Palate Speech: Assessment and Intervention (Chichester: Wiley-Blackwell), pp. 199-220.

5

TILLMAN, K. K., HAKELIUS, M., HÖIJER, J., RAMKLINT, M., EKSELIUS, L., NOWINSKI, D., and PAPADOPOULOS, F. C., 2018, Increased risk for neurodevelopmental disorders in children with orofacial clefts. Journal of the American Academy of Child and Adolescent Psychiatry, 57(11), 876.

TRICCO, A. C., LILLIE, E., ZARIN, W., O'BRIEN, K., COLQUHOUN, H., KASTNER, M., AND WILSON, K., 2016, A scoping review on the conduct and reporting of scoping reviews. BMC Medical Research Methodology, 16(1), 15.

WHITEHILL, T.L., 2002, Assessing intelligibility in speakers with cleft palate: a critical review of the literature. The Cleft palate-craniofacial journal, 39(1), 50-58.

WHITEHILL, T. L., GOTZKE, C. and Hodge, M., 2011, Speech Intelligibility. In S. Howard S., and A. Lohmander (eds), Cleft Palate Speech: Assessment and Intervention (Chichester: Wiley-Blackwell), pp. 293-304.

WHO., 2007, International Classification of Functioning, Disability and Health. Children and Youth Version [online]

WILLADSEN, E., LOHMANDER, A., PERSSON, C., LUNDEBORG, I., ALALUUSUA, S., AUKNER, R., BAU, A., BOERS, M., BOWDEN, M. and DAVIES, J., 2017, Scandcleft randomised trials of primary surgery for unilateral cleft lip and palate: 5. Speech outcomes in 5-year-olds-consonant proficiency and errors. Journal of plastic surgery and hand surgery, 51(1), 38-51.

WREN, Y., 2013, 'A survey of the assessment procedures at 24 and 36 months in UK Cleft Centres', The Cleft Collective Speech and Language Study. (Unpublished survey).

YAMASHITA, R.P., BORG, E., GRANQVIST, S. and LOHMANDER, A., 2018, Reliability of Hypernasality Rating: Comparison of 3 Different Methods for Perceptual Assessment. The Cleft Palate-Craniofacial Journal, 1-12.

6

Appendix 1:

<u>List of articles included in scoping review</u>

CHACON, A., PARKIN, M., BROOME, K. and PURCELL, A., 2017, Australian children with cleft palate achieve age-appropriate speech by 5 years-of-age. International Journal of Pediatric Otorhinolaryngology, 103, 93-102.

CHAPMAN, K.L., HARDIN-JONES, M.A., GOLDSTEIN, J.A., HALTER, K.A., HAVLIK, R.J. and SCHULTE, J., 2008, Timing of palatal surgery and speech outcome. The Cleft Palate-Craniofacial Journal, 45(3), 297-308.

CHAPMAN, K.L., 2004, Is presurgery and early postsurgery performance related to speech and language outcomes at 3 years of age for children with cleft palate? Clinical Linguistics & Phonetics, 18(4-5), 235-257.

DAYASHANKARA, R.J., SINGH, S., SUMA, G.N., GOKKULA, K., KUMAR, V. and SURSH, B.S., 2011, Comparative study of 2 palatoplasty techniques to assess speech and fistula in primary cleft palate patients. Journal of Dentistry for Children, 78(1), 13-18.

EL EZZI, O., HERZOG, G., BROOME, M., TRICHET-ZBINDEN, C., HOHLFELD, J., CHERPILLOD, J. and DE BUYS ROESSINGH, ANTHONY S., 2015, Grommets and speech at three and six years in children born with total cleft or cleft palate. International Journal of Pediatric Otorhinolaryngology, 79(12), 2243-2247.

FREDERICKSON, M.S., CHAPMAN, K.L. and HARDIN-JONES, M., 2006, Conversational skills of children with cleft lip and palate: a replication and extension. The Cleft Palate-Craniofacial Journal, 43(2), 179-188.

FREY, J.R., KAISER, A.P. and SCHERER, N.J., 2018, The influences of child intelligibility and rate on caregiver responses to toddlers with and without cleft palate. The Cleft Palate-Craniofacial Journal, 276–286.

GUNTHER, E., WISSER, J.R., COHEN, M.A. and BROWN, A.S., 1998, Palatoplasty: Furlow's double reversing Z-plasty versus intravelar veloplasty. The Cleft Palate-Craniofacial Journal, 35(6), 546-549.

GUGSCH, C., DANNHAUER, K. and FUCHS, M., 2008, Evaluation of the progress of therapy in patients with cleft lip, jaw and palate, using voice analysis- a pilot study. Journal of Orofacial Orthopedics, 69(4), 257-267.

HAMMING, K.K., FINKELSTEIN, M. and SIDMAN, J.D., 2009, Hoarseness in children with cleft palate. Otolaryngology—Head and Neck Surgery, 140(6), 902-906.

HATTEE, C., FARROW, K., HARLAND, K., SOMMERLAD, B. and WALSH, M., 2001, Are we ready to predict speech development from babble in cleft lip and palate children? International Journal of Language & Communication Disorders, 36(S1), 115-120.

HODGE, M. and GOTZKE, C.L., 2007, Preliminary results of an intelligibility measure for English-speaking children with cleft palate. The Cleft Palate-Craniofacial Journal, 44(2), 163-174.

HUTTERS, B., BAU, A. and BRØNDSTED, K., 2001, A longitudinal group study of speech development in Danish children born with and without cleft lip and palate. International Journal of Language & Communication Disorders, 36(4), 447-470.

KLINTÖ, K., SALAMEH, E. and LOHMANDER, A., 2016, 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. International Journal of Speech-Language Pathology, 18(2), 147-156.

KLINTÖ, K., SALAMEH, E. and LOHMANDER, A., 2015, Verbal competence in narrative retelling in 5 year olds with unilateral cleft lip and palate. International Journal of Language & Communication Disorders, 50(1), 119-128.

KLINTÖ, K., OLSSON, M., FLYNN, T., SVENSSON, H. and LOHMANDER, A., 2014a, Phonology in Swedish-speaking 3-year-olds born with cleft lip and palate and the relationship with consonant production at 18 months. International Journal of Language & Communication Disorders, 49(2), 240-254.

KLINTÖ, K., SVENSSON, H., ELANDER, A. and LOHMANDER, A., 2014b, Speech and phonology in Swedish-speaking 3-year-olds with unilateral complete cleft lip and

palate following different methods for primary palatal surgery. The Cleft Palate-Craniofacial Journal, 51(3), 274-282.

KONST, E.M., RIETVELD, T., PETERS, H.F. and PRAHL-ANDERSEN, B., 2003, Phonological development of toddlers with unilateral cleft lip and palate who were treated with and without infant orthopedics: a randomized clinical trial. The Cleft Palate-Craniofacial Journal, 40(1), 32-39.

LARSSON, A., SCHÖLIN, J., MARK, H., JÖNSSON, R. and PERSSON, C., 2017, Speech production in 3 year old internationally adopted children with unilateral cleft lip and palate. International Journal of Language & Communication Disorders, 52(5), 626-636.

LOHMANDER-AGERSKOV, A., 1998, Speech outcome after cleft palate surgery with the Gothenburg regimen including delayed hard palate closure. Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 32(1), 63-80.

LOHMANDER-AGERSKOV, A., SÖDERPALM, E., FRIEDE, H. and LILJA, J., 1998, A comparison of babbling and speech at pre-speech level, 3, and 5 years of age in children with cleft lip and palate treated with delayed hard palate closure. Folia Phoniatrica et Logopaedica, 50(6), 320-334.

LOHMANDER, A. and PERSSON, C., 2008, A longitudinal study of speech production in Swedish children with unilateral cleft lip and palate and two-stage palatal repair. The Cleft Palate-Craniofacial Journal, 45(1), 32-41.

LOHMANDER, A., LUNDEBORG, I. and PERSSON, C., 2017a, SVANTE–The Swedish Articulation and Nasality Test–Normative data and a minimum standard set for cross-linguistic comparison. Clinical linguistics & phonetics, 31(2), 137-154.

LOHMANDER, A., FRIEDE, H., ELANDER, A., PERSSON, C. and LILJA, J., 2006, Speech development in patients with unilateral cleft lip and palate treated with different delays in closure of the hard palate after early velar repair: a longitudinal perspective. Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 40(5), 267-274.

MORRIS, H. and OZANNE, A., 2003, Phonetic, phonological, and language skills of children with a cleft palate. The Cleft Palate-Craniofacial Journal, 40(5), 460-470.

PERSSON, C., LOHMANDER, A. and ELANDER, A., 2006, Speech in children with an isolated cleft palate: a longitudinal perspective. The Cleft Palate-Craniofacial Journal, 43(3), 295-309.

PULKKINEN, J., HAAPANEN, M., PAASO, M., LAITINEN, J. and RANTA, R., 2001, Velopharyngeal function from the age of three to eight years in cleft palate patients. Folia Phoniatrica et Logopaedica, 53(2), 93-98.

RAUD WESTBERG, L., HÖGLUND SANTAMARTA, L., KARLSSON, J., NYBERG, J., NEOVIUS, E. and LOHMANDER, A., 2017, Speech outcome in young children born with unilateral cleft lip and palate treated with one-or two-stage palatal repair

and the impact of early intervention. Logopedics Phoniatrics Vocology, DOI:
10.1080/14015439.2017.1390606

SAFAIEAN, A., JALILEVAND, N., EBRAHIMIPOUR, M., ASLESHIRIN, E. and
HIRADFAR, M., 2017, Speech intelligibility after repair of cleft lip and palate.
Medical Journal of The Islamic Republic of Iran (MJIRI), 31(1), 500-504.

SWANSON, J.W., MITCHELL, B.T., COHEN, M., SOLOT, C., JACKSON, O.,
LOW, D., BARTLETT, S.P. and TAYLOR, J.A., 2017, The effect of Furlow
palatoplasty timing on speech outcomes in submucous cleft palate. Annals of
Plastic
Surgery, 79(2), 156-161.

WILLADSEN, E., BOERS, M., SCHÖPS, A., KISLINGMØLLER, M., NIELSEN,
J.B., JØRGENSEN, L.D., ANDERSEN, M., BOLUND, S. and ANDERSEN, H.S.,
2018, Influence of timing of delayed hard palate closure on articulation skills in 3-
year-old Danish children with unilateral cleft lip and palate. International Journal of
Language & Communication Disorders, 53(1), 130-143.

WILLADSEN, E., 2012, Influence of timing of hard palate repair in a two-stage
procedure on early speech development in Danish children with cleft palate. The
Cleft Palate-Craniofacial Journal, 49(5), 574-595.

WILLADSEN, E. and POULSEN, M., 2012, A restricted test of single-word
intelligibility in 3-year-old children with and without cleft palate. The Cleft Palate-
Craniofacial Journal, 49(3), 6-16.

WREN, Y., 2013, A survey of the assessment procedures at 24 and 36 months in UK cleft centres, The Cleft Collective Speech and Language Study. (Unpublished survey).

ZANZI, M., CHERPILLOD, J. and HOHLFELD, J., 2002, Phonetic and ontological results after early palate closure in 18 consecutive children presenting with cleft lip and palate. International Journal of Pediatric Otorhinolaryngology, 66(2), 131-137.

Dear Dr McKean,

We would like to thank you and the reviewers for your thoughtful and detailed feedback on our submission. We have considered all of the feedback as detailed below and this has served to improve the quality of the manuscript.

In particular, the main focus of the changes has been to strengthen the rationale for the scoping review methodology and a significant re-write of the Discussion section to address theoretical approaches to assessment and relevance to clinical practice and future research.

Major changes to the original manuscript have been highlighted, and a detailed description of the changes is reported below.

Editor Comment	Response	Location
Why you have chosen to do a scoping review rather than a systematic review	The rationale for a scoping rather than systematic review has been strengthened in the manuscript. The aim of the review is to broadly map assessment practices, rather than focus on quality appraisal or effectiveness. This aim is to use this information to inform the design of a UK Assessment Protocol at age 3-years. The expansive inclusion criteria of a scoping review was also appropriate given the aim to provide a comprehensive overview of assessment practices. Had the scope of the review been limited to articles from a particular country, or type of assessment, not only would the usability of the review have been limited but potentially valuable information could have been missed. In addition, the aims of the review were not to aggregate findings but to address key issues through the charting of data which is in keeping with a scoping review methodology. Additional references have been introduced to support these justifications: Arksey, H., & O'Malley, L. (2005), Khalil et al. (2016), Munn et al. (2018), Peters et al. (2015) and Tricco et al. (2018). The lack of clarity between a scoping and systematic review methodology has been clarified by omitting comments on reliability, which is better suited to a systematic review and was not fully in line with the aims of this review. The aims of the scoping review have been more clearly stated. Of relevance is the following paragraph: "In order to examine the methods, parameters of assessment (whether developmental or cleft specific) and speech samples utilised in the assessment of speech in children with CP±L at age 3-years, a broad scope of enquiry was required. This was achieved through the use of a scoping review methodology (Khalil et al. 2016). This methodology was preferred given that the intention behind the work was to	*Abstract* Line 12-13 *What this paper adds* Line 5-6 **Main text** *Introduction* Page 3 Line 9-22 *Methods* Page 4 Line 10-11

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	inform the future development of assessment procedures at age 3-years in the UK. Rather than answer an effectiveness question (Peters et al. 2015), evaluate study quality, or limit the review to a particular study design (Arksey and O'Malley 2005) or country of origin, the scoping review methodology was used to broadly map current practices, in order to clarify the key concepts, identify gaps in the evidence base and make recommendations for future practice in the UK. In doing so the review will identify the extent to which there is consensus in the parameters of assessment, procedures and methods; differences and similarities in assessment at age 3 years compared to practice at age 5; the types of speech samples used (Munn et al. 2018)."	
Make clearer what this paper adds over and above simply providing a descriptive list of assessment approaches	The importance of a developmental approach to assessment at age 3 in the CP±L has been highlighted throughout the manuscript. This brings to the fore reasons why assessments at age 3 and those established for other age groups are different. This is particularly relevant in the CP±L population given that the focus of assessment at older age groups is often on cleft specific outcomes. In addition, the review highlights the need for multiple approaches to assessment to assess speech at this age. In the discussion and conclusion sections a greater emphasis has been placed on highlighting those areas in which assessment practices differ at age 3 in comparison to older age groups i.e. the type of speech sample used, the use of a developmental approach to assessment and the potential for the use of an overall measure of VP function. Finally, through a broad appraisal of the literature, the lack of consensus regarding speech samples and methods of assessment is revealed. This highlights potential research opportunities.	*Abstract* Main contribution Page 1 Line 16-21 Page 2 Line 1-2 Conclusion Page 2 Line 11-16 *What this paper adds* All **Main text** *Discussion* Page 16 Line 2-23 Page 17 Line 1-22 Page 18 Line 1-23 Page 19 Line 1-3 Page 20 Line 3-5 *Conclusion* Page 21 Line 12-23 Page 22

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

		Line 1-4
Theoretical motivations which might make researchers/clinicians choose some over others	The discussion section of the manuscript has been extensively re-written to highlight the different theoretical approaches to assessment, reasons as to why they have been used, and the need to use different theoretical approaches in conjunction to assess speech at age 3 years in this population. This includes the discussion of a medical model, linguistic, developmental and functional approaches to assessment.	**Main text** *Discussion* All
What you would recommend both for clinical practice and future research	Clinical Practice: The need for multiple theoretical approaches for assessment, used in combination to assess speech at age 3-years is highlighted. In particular, the need to take a developmental and functional perspective is discussed through the examination of the assessment methods and parameters of assessment. Key differences between assessments at older ages and age 3 are described.	*Abstract* Page 1 Line 16-21 Page 2 Line 1-2 Line 11-13 *What this paper adds* Line 8-9 **Main text** *Discussion* Page 17 Line 11-23 Page 18 Line 1-23 Page 19 Line 1-3 Page 20 Line 3-5 *Conclusion* Page 21 Line 15-20
	Research: In the amended manuscript clearer recommendations are made for future research these are: - The need for further research in functional outcomes i.e. intelligibility. The need for validation of the ICS at age 3 years is discussed. - The need for further research into the alternatives to categorical scales is highlighted. This is particularly relevant given evidence that "ratio or category-ratio scales may be more valid and reliable measures (Yamashita et al. 2018, Baylis et al. 2015)."	*Abstract* Page 2 Line 15-16 *What this paper adds* Line 16-19 **Main text** *Discussion* Page 16 Line 13-16 Line 21-23

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	- The need to research the impact of different speech samples in terms of completion rates and listener reliability is discussed. This is particularly relevant as age 3-years may not be able to produce the same complex samples used to assess outcomes in older age groups. The relationship between speech samples and reliability has not be investigated at age 3 the same way it has at age 5 (Klintö et al. 2011). - The potential for the use of the Restricted Word List to allow for cross-linguistic comparisons is discussed. However, the further development of the word list is required in order to complete a comprehensive assessment of disordered speech.	Page 17 Line 1-6 Page 18 Line 22-23 Page 19 Line 1-3 *Conclusion* Page 21 Line 21-23 Page 22 Line 1-4
Reviewer 1 Comment	**Response**	**Location**
Are the authors able to mention what motivated them to conduct a scoping review over other types of reviews (e.g. systematic review)?	Thank you for highlighting this point. We appreciate that we have not been sufficiently clear as to why a scoping review was selected as opposed to a systematic review. The justification and rationale for the scoping review has been clarified in the amended manuscript. The aim to review assessment practices at age 3-years is broad. The motivation of the review being not to aggregate findings or appraise quality but to understand the breadth and scope of assessment practices, in order to identify key concepts and core elements central to assessment at this age (which may or may not differ to assessments at older age groups) in order to identify gaps in the research, and if possible make recommendations regarding assessments protocols in the UK. With this in mind the methodology best suited to meet these objectives is a scoping rather than systematic review. Additional references have been included to support the rationale for using a scoping review methodology: Arksey, H., & O'Malley, L. (2005), Khalil et al. (2016), Munn et al. (2018), Peters et al. (2015) and Tricco et al. (2018). The potential for confusion as to the methodology may relate to the inclusion or reliability measures. As such comments about reliability have been edited from the manuscript as this potentially clouded the justification between scoping and systematic reviews.	*Abstract* Line 12-13 *What this paper adds* Line 5-6 **Main text** *Introduction* Page 3 Line 9-22 *Methods* Page 4 Line 10-11
Visually illustrated the search strategy using a flowchart	This reference has been updated to Tricco et al. (2018) who produced the PRISMA Extension for	*Figure 1*

that resembles the PRISMA guidelines for a systematic review (Moher et al., 2009).	Scoping Reviews. The flowchart is in line with the illustration used in this paper, and this reference is a better fit with the paper given the use of a scoping review.	
It could have been better if the authors restricted the inclusion criterion to studies that were carried out in countries where similar languages were spoken. In this way, the authors could have made equitable comparisons across studies and identified reliable assessment practices.	The justification for not limiting the study to a single country has been explained in the amended manuscript. This now reads: "To broaden the examination, no preference was given to literature from any specific country" An expansive inclusion criterion is in keeping with a scoping review methodology. A result of narrowing the review to certain countries i.e. to the UK only, would be to miss useful practices in other countries e.g. the use of a cross-linguistic assessment. Indeed, this review highlights the potential of cross-linguistic speech samples, to facilitate the comparison of outcomes between countries which would be lost if the remit of the review was narrowed. Indeed, given the rarity of CP±L (approx. 1 in 1000 births) there is a need to learn from research in other countries. Historically, the sharing of practices internationally has brought about significant changes to cleft care (e.g. the CSAG review in the UK in response to the Shaw et al. [1992] international study of cleft outcomes). A further justification for not limiting the review to a single country is that the papers only originated from 12 countries, with some countries only contributing a single paper. Selecting a single country would have significantly narrowed the breadth of the review. Whilst differences are explored, the inclusion of different countries also highlights similarities, indicating core parameters and methods of assessment which require further examination.	**Main text** *Methods* Page 5, Line 18-19
I wish the authors could have performed a quality appraisal of the studies that were included.	The amended manuscript clarifies why a scoping review rather than systematic review was selected. The Joanna Briggs Institute methodology was closely adhered to ensure the methodological rigour of the scoping review. However, a key difference between scoping and systematic reviews is quality appraisal and it was not the aim of this study to appraise the quality of the papers, but to broadly map assessment practices.	**Main text** *Introduction* Page 3 Line 9-22
I wonder why the authors did not compare the reliability of the assessment practices from all the included studies.	In the original manuscript, some of the confusion between scoping and systematic reviews may have been related to the reporting of reliability in the studies.	**Table 3** Recording of reliability has been removed

	The reporting of reliability does not fit well within the scoping review methodology nor the aims of this study. In light of this, and for clarity the reporting of reliability has been omitted in the amended manuscript.	
Finally, the authors should present their findings in a manner that would be applicable for clinical practice. This would ideally involve identifying assessment practices, speech samples, etc. that are more reliable over others and identifying speech samples that can be used for the 3-year olds from a developmental perspective of the speech and language skills.	Although reliability was not specifically examined the amended manuscript more specifically addresses the following implications for clinical practice: e.g. - The need for multiple theoretical approaches for assessment, used in combination to comprehensively assess speech. In particular, the need to use a developmental and functional perspective is highlighted through the examination of the assessment methods and parameters of assessment. - The potential of cross-linguistic assessments is discussed. - The potential use of an overall measure of velopharyngeal function is highlighted. However, key issues such as the lack of consensus as to the type of speech sample, and the limited use of sales other than ordinal, highlight the need for further research.	*Abstract* Page 1 Line 16-21 Page 2 Line 1-2 Line 11-13 *What this paper adds* Line 8-9 **Main text** *Discussion* Page 17 Line 11-23 Page 18 Line 1-23 Page 19 Line 1-3 Page 20 Line 3-5 *Conclusion* Page 21 Line 15-20
The findings do not seem to benefit a practising clinician or a researcher	As above the amended manuscript specifically addresses implications for clinical practice and research. In addition, the need for further research and the potential areas of this research are also highlighted through the broad examination of the literature through a scoping review. Specific recommendations for researchers are: - The need for validation of the ICS at age 3 years is discussed. - The limited use of "ratio or category-ratio scales, despite evidence to suggest this is a more valid measurement of resonance and NAE requires further investigation. Evidence has suggested that these types of scales can be reliably used with older age groups,	*Abstract* Line 15-16 *What this paper adds* Line 16-19 **Main text** *Discussion* Page 16 Line 13-16 Line 21-23 Page 17 Line 1-6 Page 18 Line 22-23

	however 3year olds have not been specifically investigated (Yamashita et al. 2018, Baylis et al. 2015). - The need to research the impact of different speech samples in terms of completion rates and listener reliability is discussed. This is particularly relevant as age 3-years may not be able to produce the same complex samples used to assess outcomes in older age groups. The relationship between speech samples and reliability has not been investigated at age 3 the same way it has at age 5 (Klintö et al. 2011). - The potential for the use of the Restricted Word List to allow for cross-linguistic comparisons is discussed. However, the further development of the word list is required in order to complete a comprehensive assessment of disordered speech.	Page 19 Line 1-3 *Conclusion* Page 21 Line 21-23 Page 22 Line 1-4
Reviewer 2 Comment	**Response**	**Location**
How was grey literature identified?	The Cleft Care UK Speech and Language Study is an ongoing longitudinal UK multi-centre study that is taking place in the cleft centre of the primary author. Information regarding the study is shared on an annual basis at the Craniofacial Society of Great Britain and Ireland's annual scientific conference. Given that the study is collecting data on speech outcomes at age 3-years, the primary author directly contacted the study team to identify if there was any relevant information connected to the study which could be considered for the review. This yielded the unpublished, yet highly relevant survey of assessment practices in the UK (Wren 2013). Indeed, the inclusion of this survey in the review may prevent the replication of this work by other researchers.	**Main text** *Method* Page 7 Line 6-12
Please explain what you mean by 'anatomical changes'	This has been clarified and now reads: "anatomical changes (e.g. adenoid atrophy, dentition and occlusion)"	**Main text** *Introduction* Page 1 Line 12-13
However, for younger children, a focus of assessment may be to examine the impact of the cleft on the process of normal speech development (articulation and phonology). This sentence seems to represent a very narrow view of speech development -	This has been changed and now reads: "the process of normal speech sound development"	**Main text** *Introduction* Page 1 Line 15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

what about e.g. syntax and morphology? If the interest is only in articulation and phonology, then 'the process of normal speech sound development' would seem more appropriate.		
It is unclear why information from grey literature did not need to be in full.	Clarification has been provided as to the inclusion criteria which applies to all the studies reviewed. This now reads: "Literature and sources had to address the following concepts: speech assessments, the types of speech samples and speech parameters assessed, including methods of assessment. This allowed for the inclusion of sources which investigated assessment processes as well as those reporting on speech outcomes."	**Main text** *Methods* Page 5 Line 13-16
It is not clear to me, what relevant Medical Subject Headings means	A definition has been provided. This now reads: "Relevant Medical Subject Headings (MeSH [vocabulary used to index publications]) were utilised in the database search."	**Main text** *Methods* Page 6 Line 12
Please explain/define the term 'consonant inventory'	This has been changed to consonant/phoneme inventory to reflect the different definitions used. Different examples of definitions of consonant/phoneme inventory used in the studies has been provided: "The criteria for inclusion in the inventory varied across the studies. For example (Hattee et al. 2001) rated the presence of consonants, Morris & Ozanne (2003) required phonemes to be heard twice to be included, and Lohmander et al. (2017a) reported correct productions in more than 50% of the targets."	**Main text** *Results* Page 11 Line 1-5
I suggest revise the very last part of the sentence to 'PCC<u> and</u> summary patterns',	This has been reworded to improve clarity and now reads: "The methods used in the studies to assess consonant production demonstrate that both PCC and summary patterns can provide information about articulation and phonology."	**Main text** *Results* Page 11 Line 15-18
P.20. Lines 48-51: While the category summary issue is true, PCC measures in cross-linguistic /cross-centre studies allow comparison of the overall articulation competence	This paragraph has been extensively re-written to reflect the pros and cons of both PCC-A and summary measures particularly in a cross-linguistic context. This section is particularly relevant: "Again, challenges in comparing speech outcomes across the studies arise from the use of different criteria and summary patterns (Sell 2005) and the need for consensus is relevant both at age 3 -years and for other age groups. PCC-A, using a controlled speech sample allows for cross-linguistic comparisons and takes into consideration sound distortions	**Main text** *Discussion* Page 17 Line 11-17

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

| | occurring as part of normal development, which is particularly appropriate for assessment at age 3-years. PCC-A may, therefore, offer a partial solution although it has the disadvantage of not capturing the qualitative nature of speech errors." | |