Dadansoddi ar gyfer Polisi
YMCHWIL GYMDEITHASOL Y LLYWODRAETH
GSR
GOVERNMENT SOCIAL RESEARCH
Analysis for Policy

Llywodraeth Cymru
Welsh Government

SOCIAL RESEARCH NUMBER:
55/2022
PUBLICATION DATE:
02/08/2022

Review of Early Language Screening Suitable for Children in Wales from Birth to 5 Years

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.

This document is also available in Welsh.

 Digital ISBN 978-1-80364-639-8

Title: Review of Early Language Screening Suitable for Children in Wales from Birth to 5 Years

Author(s): Sharon Baker[1], Sam Harding[1], Caitlin Holme[1], Rhonwen Lewis[2], Miriam Seifert[1], Yvonne Wren[1,2,3]

1 Bristol Speech and Language Therapy Research Unit, North Bristol NHS Trust

2 Cardiff Metropolitan University

3 University of Bristol

Full Research Report: Baker, S., Harding, S., Holme, C., Lewis, R., Seifert, M. and Wren, Y. (2022). *Review of Early Language Screening Suitable for Children in Wales from Birth to 5 Years.* Cardiff: Welsh Government, GSR report number 55/2022.
Available at: https://gov.wales/early-language-screening-review

For further information please contact:
Launa Anderson
Social Research and Information Division
Welsh Government
Cathays Park
Cardiff
CF10 3NQ

Email: SocialJusticeResearch@gov.wales

Contents

List of tables

List of figures

Glossary

Acronym/Key word	Definition
Acceptability	How a tool is likely to be received by those people administering it and those people who are being screened.
Auxillary verb	A verb that is used with another verb to show a tense e.g., I *have* played, I *will* play
Assessment bias	When one or more items on an assessment provide an advantage or disadvantage to specific groups of people e.g., asking questions related to a specific culture that some people would be unfamiliar with.
Bilingualism	The ability to produce two languages with at least a basic level of functional proficiency or use, regardless of the age at which the languages were learned.
Canonical babbling	Where consonant and vowels are produced in repetitive strings e.g., dadada, bububu.
Copula verb	A verb that is used to join an adjective or noun to another verb e.g., the grass *is* long, the girl *was* the captain
Diagnostic Power	How well a tool is able to accurately identify an individual with a given condition based on its sensitivity and specificity.
Developmental Language Disorder (DLD)	A persistent speech, language and communication need that cannot be explained by any particular cause. Also previously referred to as specific language impairment (SLI).
Dynamic Assessment	A way of assessing skills acquired alongside the potential for skills to be learnt.
Fine motor skills	Movements which require the coordination of small muscles e.g., pincer grip, transferring objects from one hand to another

Gross motor skills	Larger muscle movements involving the limbs and torso e.g., running, jumping, climbing
Impact	The result of using a screening tool e.g., accurate identification of SLCN allowing for early intervention.
Inter-rater reliability	The level of agreement between two reviewers.
Language Delay	A term used to describe a child whose receptive and expressive language skills fall well below expectations. This term is no longer in use following a consensus study (Bishop et al., 2016) but is present in some research prior to this date.
Mean Length of Utterance (MLU)	A measure of language production calculated by dividing the total number of utterances produced by the number of words or morphemes (the parts of words which mark plurals, tense etc) used in those utterances.
Morphology	The structure and form or words and morphemes (e.g., prefixes, suffixes).
Morphosyntactic structures	The parts of words that can change to affect meaning for example adding ‘s’ to a word for regular plurals and ‘ed’ to a word to indicate regular past tense.
Multilingualism	The ability to speak more than two languages with at least a basic level of functional proficiency or use, regardless of the age at which those languages were learned.
Normative data	Data from a reference population that characterises what is typical performance at a specific point or period of time.
Persistent Needs	A term used to describe the needs of children whose speech, language and communication difficulties persist even after intervention. Their problems might be restricted to speech, language and communication or they may have problems with these skills as part of

	another condition affecting their learning and/or development.
Personal social skills	Activities that children learn to complete to be able to look after themselves and function in society e.g., dressing, understanding others' emotions
Problem-solving skills	How children learn to be able to work out solutions to problems e.g., decision-making, lateral thinking.
Prudent Healthcare	A term introduced by Welsh Government with the following principles: achieve health and wellbeing with the public, patients and professionals as equal partners through co-production; care for those with the greatest health needs first, making the most effective use of skills and resources; do only what is needed, no more, no less and do no harm; reduced inappropriate variation using evidence based practices consistently and transparently.
Receptive Language	The understanding of spoken language which includes understanding of vocabulary and grammar.
Reliability	The consistency of a measure i.e., whether the same scores would be reproduced if the test was administered again. There are various types of reliability including test-retest reliability, inter-rater and intra-rater reliability.
Remediate	To correct something that is wrong or improve something considered negative.
Sensitivity	A measurement of how often a test correctly identifies when a person has the condition being tested for. Low sensitivity results in false negatives where people with the condition are incorrectly identified as not having the condition.
Specificity	A measurement of how often a test correctly identifies when a person does not have the condition being tested for. Low specificity can result in false positives where

	people without the condition are incorrectly identified as having the condition.
Speech, Language and Communication Needs (SLCN)	An umbrella term introduced in the Bercow (2008) report which describes all speech, language and communication needs that a child could be experiencing and is regardless of any underlying diagnosis or medical condition.
Standardisation	A process to ensure that assessments are conducted in the same way every time and are scored in a consistent manner.
Transient Needs	A term used to describe the needs of children whose speech, language and communication skills improve when additional support is provided.
Utility Equation	A method to critique assessment or screening tools. Includes measures of validity, reliability, acceptability, impact, cost and diagnostic power. Each of these are evaluated to determine whether a tool is scientifically robust enough to use.
Validity	The extent to which a test measures what it intended to measure. There are various types of validity including content validity, construct validity and face validity, which is the degree to which those using the tool consider that it measures the concept being investigated.

1. Introduction

Policy Background

1.1 This commission for a review of early language screening tools suitable for children aged between 0 and 4 years, 11 months across Wales is the first step in a process which is intended to lead to the identification of children with speech, language and communication needs. The contract for the work was awarded in June 2021 and work began in July of the same year. Once identified, these children will receive high quality support in the early years to develop their speech, language and communication skills on a scale according to their need (universal, population, targeted, specialist), as described in 'Talk with me: Speech, Language and Communication Delivery Plan' (Welsh Government, 2020) and 'Flying Start' (Welsh Government, 2017).

1.2 The Welsh Government is committed to ensuring that children from all backgrounds have the best start in life, particularly in relation to early identification of needs and provision of high quality intervention and support. It will also assist in ensuring that early years' provision builds strong key skills (such as speech, language and communication) from when children are very young. Moreover, it addresses the aims of the Healthy Child Wales Programme (Welsh Government, 2016) to assist children 'to meet all … developmental milestones enabling them to achieve school readiness' and 'to ensure early detection of …developmental…problems through an appropriate, universal screening programme' (p.5). The Welsh language strategy, Cymraeg 2050, aims to increase the number of people who can speak Welsh, including children, to one million by 2050. Therefore, it is important to consider how any speech, language and communication needs can be identified when children are using more than one language.

1.3 What this work is not trying to do is justify and defend the use of existing mechanisms for screening or to fit existing processes which have been developed for other populations to the context of early child development

surveillance in Wales. Nevertheless, there is a desire to build on existing work which has been carried out in a robust manner and to determine the degree to which currently available tools and systems are either fit for purpose as is or can be developed in such a way to fulfil the quality parameters identified as essential for accurate, valid and reliable screening.

1.4 Recently, work funded by Public Health England led to the development of the Early Language Identification Measure (Law et al., 2020) which has attempted to address the need for a tool for screening of early language measures and is now being used across England. This tool has been developed in response to the health and development needs of children in England and it is important to consider what evidence is available to support a screening programme in Wales. There is limited evidence available in relation to the Welsh context, where there is a higher prevalence of children being exposed to more than one language in varying measures both at home and in pre-school. Further information related to the Welsh context is provided in Annex F. The work reported in this review is intended to support decision making regarding the potential for adoption of existing tools for screening in Wales or the need to develop bespoke materials. Critical to the decision-making process is the capacity of existing tools to be used effectively in the Welsh context given differences relating to language use, which is unique to Wales, and the processes for surveillance in the Healthy Child Wales Programme Schedule for contacts (Welsh Government, 2016).

1.5 The work carried out and reported here sits in a framework of activity which includes raising public awareness of the importance of talking and ensuring parents are better informed regarding how to provide a language rich home learning environment and the provision of evidence-based interventions for children with identified needs, as outlined in the ‘Talk with Me, Speech, Language and Communication Delivery Plan’ (Welsh Government, 2020). The former may lead to an increase in concern from parents about their child’s speech, language and communication development. In order to either reassure parents that their child’s development is on track or to advise that their child would benefit from more direct action to address their speech,

language and communication needs, there is an associated need for a robust screening tool.

Overview of the report

1.6 This report addresses the objectives set out by Welsh Government in relation to early language screening. Each of the objectives are considered and detail is provided for the background/rationale/evidence base, methods and results.

Objective 1: Review current evidence regarding the advantages and disadvantages of early language screening in universal and targeted populations. Provide a set of recommendations: should screening take place at all? If so, what should be screened, by whom and at what age(s)?

1.7 Speech, language and communication needs is an umbrella term introduced in the Bercow report (2008) and describes the range of communication difficulties that children can encounter regardless of the underlying causes or features that they present with. These terms can be further broken down to describe the specific aspects that children may experience difficulties with in their development. This objective specifically reviews screening of 'early language'.

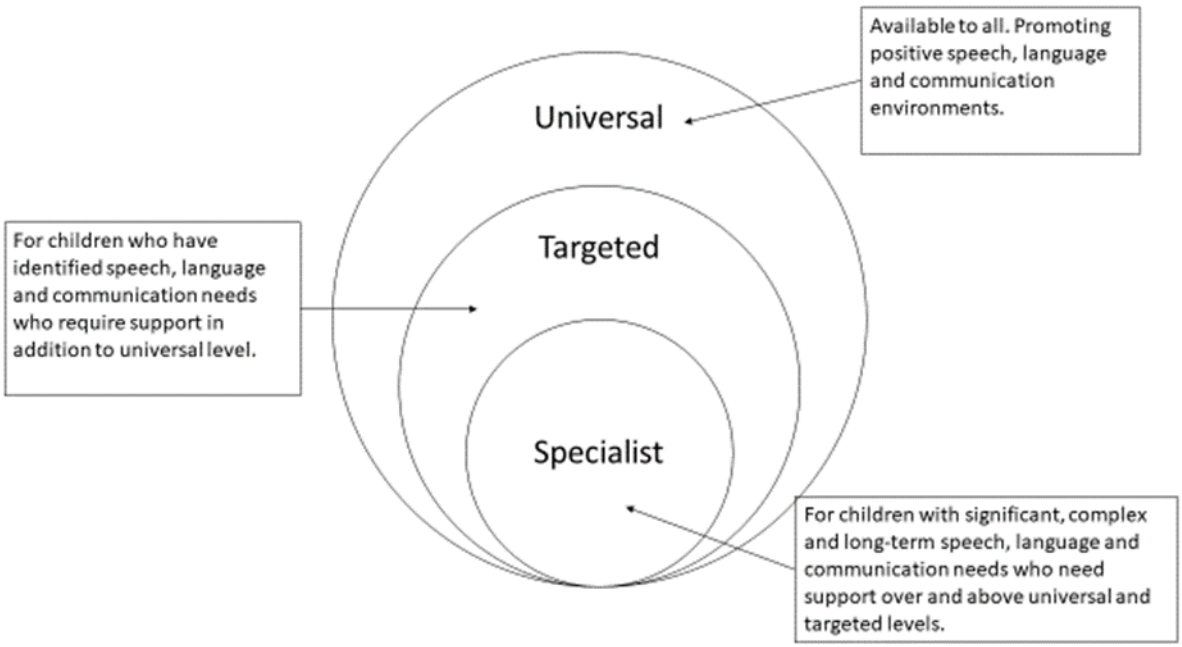
1.8 Language can be broken down into several components. Receptive language describes what a child understands, including an understanding of vocabulary (words) and grammatical forms (e.g., present tense versus past tense). Expressive language describes what is used, including vocabulary, grammatical structures and morphology (how words change, e.g., adding 's' to a word to make it a plural, for example cat vs cats). Historically, language delay is a term which was used to describe when a child's receptive and expressive skills were very late to develop in relation to expected norms (Law et al., 2013) and the term is used in this way throughout the review where it has previously been reported in the literature.

1.9 This objective also specifies that the review of screening should consider 'universal and targeted populations' as described by Better Communication

Trust (Lindsay, Dockrell, Law, & Roulstone, 2010). Universal interventions are provided to all and include training and awareness raising of speech, language and communication needs. Targeted interventions are those which are self-sustaining, for example language intervention groups delivered by early years practitioners to groups of children identified as being at risk of speech, language and communication needs. A third level, specialist interventions, are tailored to a child's specific speech, language and communication needs. Figure 1.1 provides an understanding of how these levels interact with each other. With regards to early language screening, the process for measurement and threshold for concern is consistent across universal and targeted populations, the difference between the two being marked by identified vulnerabilities in the targeted population.

1.10 This report provides an exploration of the advantages and disadvantages of screening to assist in determining whether screening should take place. It reports on the risk and protective factors for speech, language and communication development in the early years which provides information on what should be screened, when, and by whom.

Figure 1.1: Universal, targeted and specialist levels of intervention



Objective 2: Assess, critique and summarise the English and Welsh language screening/identification tools currently available to assess children's language ability and progress.

1.11 The report gives details of three scoping reviews conducted to meet this objective. These reviews provide an update to existing systematic reviews of early language screening in English, language screening in bilingual and multilingual populations and screening in the Welsh language. The aim of the first of these reviews was to identify tools which have been developed since the published systematic reviews and consider these for inclusion in a programme of screening. The latter two reviews were also carried out for the purpose of identifying tools but in addition, to determine recommendations for carrying out screening with bilingual and multilingual populations and children who are brought up in homes where Welsh is the primary language spoken.

1.12 Bilingual children are defined as those who are able to produce two languages and multilingual children as those who are able to produce more than two languages "with at least a basic level of functional proficiency or use, regardless of the age at which the languages were learned" (International Expert Panel on Multilingual Children's Speech, 2012).

Objective 3: Assess, critique and summarise what makes an effective screening/identification tool. Factors including validity, reliability, specificity and bi/multilingualism will need to be considered.

1.13 A number of the tools developed for identification of needs may not be fit for purpose. This could be because they are not assessing the 'right' speech and language variables, or that they are not valid and reliable, or that they are designed only for use by speech and language therapists. The existing literature relating to guidance on screening protocols is reported and features of effective screening tools, including details on validity, reliability, impact, acceptability, cost issues and diagnostic power are discussed. The tools identified from existing systematic reviews and from the three rapid

scoping reviews are subject to a critique to identify the most robust and suitable tools for consideration.

Objective 4: Provide a set of recommendations on the important components for an English and Welsh All Wales early language screening/identification tool.

1.14 The report considers the tools which have been critiqued as part of objective 3 against current practice for early language screening in Wales and a summary of the factors to consider at each age band is presented in graphical form. A series of recommendations are provided based on the work carried out.

Advantages and Disadvantages of Early Language Screening

1.15 Early language screening is defined by Berkman et al. (2015) as 'using standardized tools to detect the risk of a delay, which can be corroborated by a full-scale diagnostic evaluation' (p.14). Criteria for effective screening are detailed by the World Health Organisation (Wilson & Jungner, 1968). These criteria are used to determine whether a particular health concern should be part of a screening programme and can be used as a framework for considering the advantages and disadvantages of screening for speech, language and communication needs in children.

The condition should be an important health problem

1.16 The prevalence of primary language delay in children between the ages of 2 and 5 years is reported as ranging from 3 to 16 per cent (Law et al., 1998). This was determined following a systematic review of the literature in which the median reported prevalence of language delay was 7 per cent for all included studies. More recently, Norbury et al. (2016) reported prevalence of language difficulties in children aged 4 to 5 years as 9.92 per cent. This is based on a southern England population and may reflect an underestimate for some more socially deprived areas. Therefore, a large proportion of children are experiencing language difficulties in the early years. The impact

of this delayed development of language on children's later achievements has been well researched and documented.

1.17 Children who have speech and/or language delay have been reported to experience literacy difficulties at school-age, including difficulties in oracy, reading and with written language (Catts, 1997; McCormack et al., 2011; Wren et al., 2021). These children are also frequently reported to underachieve academically in comparison to those children who have not experienced a language delay (Berkman et al., 2015; Snowling et al., 2001; McKean et al., 2017). Literacy difficulties are known to continue into adulthood (Clegg et al., 2005).

1.18 Language difficulties can have a widespread impact upon social and emotional wellbeing (Norbury et al., 2016). The long-term impact of language difficulties has been documented as having an effect on behaviour. Behaviour difficulties have been identified as more likely to be present in children with language difficulties at age 4 to 6 years than peers with typically developing language (Willinger et al., 2003). It has been reported that language and behavioural difficulties can also impact on attendance at school, with truancy being reported as more likely in this group of children (Hinshaw, 1992). Moreover, up to 60 per cent of young offenders are reported as having speech, language and communication needs (Bryan, Freer & Furlong, 2007).

1.19 In a longitudinal study of children conducted by Beitchman et al. (2003), an association was reported between those children who had been identified with language difficulties at age 5 and those with a diagnosis of anxiety disorder at age 19. This was not evident for their peers who did not have any language difficulties in childhood. An earlier study of this cohort reported that those children without language difficulties had superior social adjustment at age 12 compared to their peers with diagnosed language difficulties (Beitchman et al., 1996).

1.20 Children identified as having language difficulties during their childhood are also at risk of difficulties with employment in adulthood. The longitudinal

study by Clegg et al. (2005) reported that many adults had experienced instability in employment with long periods of unemployment. Of those in employment most were in roles where they did not require any specific training or level of educational attainment.

1.21 Evidence is emerging on how young children's speech, language and communication development may have been impacted due to the public health restrictions introduced during thecoronavirus (COVID-19) pandemic, such as wearing of face masks and the closure of educational establishments. The children's communication charity, ICAN, published a report (2021) 'Speaking up for the Covid Generation' which discusses how children have missed opportunities to develop their communication skills due to the pandemic and have spent more time than usual on tablets/phones/computers, potentially affecting 1.5 million children. The report highlights teachers' concerns that for children where these public health restrictions have impacted upon a child's communication skills they may not be able to adequately catch up with their peers. It is not yet known what the long-term implications of this effect on their communication skills will be.

1.22 Identifying children who are at risk of speech, language and communication difficulties early in life will enable them to receive timely intervention thereby reducing the risk of possible wider effects, potentially providing long-term benefits into adulthood. It is important therefore that language difficulties are recognised as an important health problem.

The natural history of the condition should be known

1.23 Some longitudinal studies suggest that approximately 60 to 70 per cent of children with early language difficulties resolve spontaneously by the time they reach the age of 4 years (Dale et al., 2003; Paul et al., 1991; Rescorla et al., 2000). However, these studies did not look at clinical populations and the children were predominantly those with expressive language delays. Evidence from community population studies however, suggests that the trajectories of children with identified speech and language disorders are

relatively stable and that they do not simply grow out of a language delay (McKean et al., 2017; Norbury et al., 2016). Furthermore, evidence from the control arm of randomised controlled trials has shown that, whilst some children in waiting list conditions do make progress, few achieve levels concordant with their non-impaired peers and up to two thirds of children still need intervention after a twelve month waiting period (Broomfield & Dodd, 2011; Roulstone et al., 2003).

1.24 There is an argument that identification of children for intervention before any spontaneous improvement has taken place could constitute an ineffective use of resources. Arguably, this resource could be better used in focussing on a national population public health programme to promote wellbeing and prevention of language delay. However, it is known that for many children language difficulties persist and they require intervention to improve. Some of these children will continue to have difficulties in understanding and using language that follow them throughout childhood and into adulthood. An evidence-based approach to screening is important to enable identification of these children and to allow for monitoring and the provision of timely intervention.

The condition should have a recognisable pre-symptomatic stage

1.25 As discussed by Law et al. (1998), a developmental condition cannot be considered to have a pre-symptomatic stage; all children experience limited communication as part of their development. The systematic review by Law et al. (1998) suggests that the identification of a delay in language development should rather be considered pre-diagnostic.

1.26 There are many key milestones in children's language development during the first five years (Asmussen et al., 2018; Sheridan, Sharma & Cockerill, 2014). These provide guidelines for evaluating development, with acknowledgement that children will reach milestones at different rates. Some milestones, such as the child recognising their own name, are the same regardless of the language or languages the child is being exposed to. Others relate to the size of the child's productive vocabulary and in bilingual

or multilingual settings, consideration needs to be given to how this would be measured across all languages the child is exposed to. For the stages of linguistic development identified in Sheridan, Sharma and Cockerill (2014), the overall developmental sequence would be the same across monolingual and bilingual contexts, i.e., babbling will be followed by first words, word combinations and then the development of syntax over time. However, the specific lexical and grammatical milestones will be language-specific, depending on the presence and frequency of structures within the language or languages of the environment. Also, differences in amount of exposure due to bilingual or multilingualism can alter the developmental trajectory. These features of language and communication development can be used at the pre-diagnostic stage to determine whether or not a child is meeting the expected levels of development.

There should be an acceptable and effective form of treatment at the pre-symptomatic stage

1.27 Since it is not possible to identify a pre-symptomatic stage in a developmental condition, the focus of a review should instead be on whether there are acceptable and effective forms of intervention to reduce the likelihood of speech, language and communication needs occurring.

1.28 As described by Roulstone et al. (2012), there are a wealth of interventions used by speech and language therapists in the UK to promote development of children's speech, language and communication skills. Moreover, a systematic review by Law et al. (1998) reported that there is a wealth of evidence for the effectiveness of intervention for speech and language difficulties. Meta-analyses of studies included in this systematic review did not allow for direct recommendations of specific features of intervention, e.g., timing, frequency and duration of sessions. However, there is evidence that intervention for receptive and expressive language difficulties can be delivered as effectively indirectly, for example via a trained classroom assistant, as if it were delivered directly by a speech and language therapist (Levickis et al., 2014; Roberts & Keiser, 2011).

1.29 Consideration should also be given to whether the child is exposed to one or more languages. If they are experiencing language difficulties, it is likely that these difficulties are apparent in all languages being acquired by the child (Holm et al., 2013). For interventions involving bilingual or multilingual children, decisions regarding the language of delivery are made based on a number of factors. Where these interventions are delivered directly by the speech and language therapist, the Royal College of Speech and Language Therapists (RCSLT) guidelines state that therapy should be delivered in the home language (Pert & Bradley, 2018).

1.30 Where indirect interventions are being delivered, the language of the environment needs to be taken into account. Many early years settings adopt a Welsh-immersion strategy where all input given by the adults in the environment would be in Welsh. Within these settings, the children are likely to have mixed language abilities in Welsh, depending on patterns of language use within the home. Children are often supported in small groups containing children where Welsh is the primary language used at home and children who have had minimal exposure to Welsh (Duggan, Domsmukhambetova & Edwards, 2014). There is a lack of available evidence regarding the effectiveness of interventions for bilingual and multilingual children in this context.

1.31 Overall, there is evidence that intervention for speech, language and communication needs is effective. Therefore, if children were identified through screening there is the potential for intervention to be provided to remediate difficulties thereby reducing any impact that this may have on other social, emotional and behavioural factors.

There should be an agreed policy of whom to treat

1.32 Children can be described as having transient or persistent needs (ICAN, 2009). Historically, approximately 60 per cent of children with transient needs were described as being likely to catch up with their peers with some support (Dale et al., 2003; Paul et al., 1991). However, more recently other studies have argued that once clinical caseloads are considered, the

majority of children have stable trajectories and only a minority are likely to catch up with their peers without support. This means that for the majority of children, once identified with speech, language and communication needs, these needs are persistent (Broomfield & Dodd, 2011; McKean et al., 2017).

1.33 Some recent studies have begun to identify risk factors for persistent speech, language and communication needs. This will be discussed in detail later on in this report, but in summary, at present, there is not enough evidence to be conclusive about all of the potential risk factors (Bishop et al., 2016). Screening could therefore result in identification of children with speech, language and communication needs who may not require specialist intervention. Referral to specialist services such as speech and language therapy for these children will increase the burden of care for services which are already stretched.

1.34 It is possible that children identified as being at risk of speech, language and communication needs could receive intervention at a universal, population and then targeted level in the first instance while their skills are assessed dynamically to identify those with persistent needs. This would ensure those most in need of specialist intervention would be identified, thus providing prudent healthcare (Welsh Government, 2016), i.e., support from the right person, in the right place, at the right time.

1.35 Within this early screening process, children with speech, language and communication needs associated with other conditions such as autistic spectrum disorder are likely to be identified and signposted to appropriate support such as speech and language therapy and/or neurodevelopmental services. This early referral is advantageous in that it may aid diagnosis, timely access to appropriate services and support for children and their families.

There should be facilities for investigation and diagnosis

1.36 Screening tools are not able to distinguish between which children will make spontaneous improvements, which will have persistent needs and which have difficulties associated with underlying conditions such as autistic

spectrum disorder. Therefore, once children are identified as having speech, language and communication needs there should be an effective process for monitoring progress as part of a health surveillance programme. Onwards referral to services should be made where there is a serious impact on the child and family resulting in a requirement for additional support. Speech and Language Therapists have a wide range of diagnostic tools available to them for further investigation of speech, language and communication skills.

There should be a screening test available

1.37 As outlined in the systematic review by Law et al. (1998), there are concerns about the accuracy of screening tools for identification of language difficulties and the monitoring of these difficulties to identify whether a child's needs are transient or persistent. There are concerns regarding the reliability and validity of screening tools and practical concerns regarding how and by whom it should be administered (Dockrell & Marshall, 2015).

1.38 Referral to speech and language therapy services for children in Wales currently often follows a developmental review as part of the Healthy Child Wales Programme (Welsh Government, 2016). Developmental reviews carried out by early years practitioners or Health Visitors leads to the identification of differences between the accepted developmental milestones for that age and the behaviours exhibited by the child. Some tools, such as WellComm (Early Years) (GL Education, 2021), contain checklists to support this process and have been widely adopted for this purpose. In areas of Wales where Flying Start provision is available, families receive enhanced support from Health Visiting teams. Guidance has been provided in relation to screening for early identification of speech, language and communication needs in these Flying Start areas (Welsh Government, 2017).

1.39 Anecdotal evidence suggests that some tools have been informally translated into Welsh for use with Welsh-speaking or Welsh-English bilingual families. Translation of screening and assessment tools is problematic as language-specific aspects, such as word order and specific grammatical structures are not accounted for. Also, when evaluating a child's language

abilities based on what is expected within typical development, it is not appropriate to compare children with different linguistic and cultural experiences. It is in part for these reasons that this review has been carried out to ensure that best practice, based on current robust evidence, can be followed.

1.40 Early language screening can be hampered by the reporting of false positives and false negatives. The screening tool used to identify which children have speech, language and communication needs must be sensitive enough to identify all of those children who have difficulties with these skills. It also needs to have enough specificity to identify all of the children who do not exhibit any problems in this area. The challenge here is that there is a lack of appropriate standardised assessments that have sufficient sensitivity and specificity, particularly for use with multilingual populations (Bedore and Peña, 2008).

1.41 Where a screening tool has low sensitivity, there are likely to be false negatives, i.e., children who have language difficulties but are not identified by the screening tool. These children may struggle at school and will miss the opportunity of early intervention to remediate their difficulties. Similarly, insufficient specificity will lead to false positives where children are incorrectly identified as having language difficulties. This can cause unnecessary anxiety for parents and an additional burden for the family. This also puts increased pressure on speech and language therapy services.

Summary

1.42 Berkman et al.'s (2015) systematic review applied the WHO criteria for effective screening (Wilson & Jugner, 1968) to early screening of speech and language delays. At the end of the systematic review, they were unable to recommend early screening of speech and language due to the lack of availability of robust evidence. This was also the conclusion of the previous systematic review undertaken by Law et al. (1998) who reported that as there is little agreement as to who will have transient needs and who will

have persistent needs, it is not possible to recommend the introduction of a universal screening programme.

1.43 These findings were replicated in the Better Communication Research Programme report (Lindsay et al., 2010) where no single method of screening children's language and communication skills could be recommended. The report highlighted that it was unlikely that this could ever be developed, given the complexity of language development and how children's needs can alter. The recommendation was that a systemic approach should be adopted considering universal, targeted and specialist levels of involvement.

1.44 So, whilst there is a clear need, and there is strong evidence for the effectiveness of intervention, further consideration needs to be made as to whether screening can be enhanced by an alternative approach to a single one-off screen. If screening is to take place then knowledge of risk and protective factors for speech, language and communication development allows a better understanding of what should be screened and when.

Risk and protective factors for speech and language development

1.45 Potential risk and protective factors for speech, language and communication needs have been widely researched. The impact of these needs is far-reaching and can affect long-term outcomes in relation to education, social and emotional wellbeing and employment. Knowing what these risk and protective factors are is an important factor in prudent healthcare (Welsh Government, 2016). By identifying which children may require additional support and which children are likely to make spontaneous improvements in their language skills it is likely that children can receive the right level of support at the right time.

1.46 Evidence of the impact of public health restrictions during the COVID-19 pandemic on children's developing communication skills has been acknowledged and is beginning to emerge (Bergmann et al., 2022; Charney,

Camarata and Chern, 2021; Kartushina et al., in press). The importance of being aware of the risks and protective factors for speech, language and communication needs is crucial in being able to identify which children are most at risk.

1.47 This section summarises the evidence for risk and protective factors for speech, language and communication development and relates to objective 1. A broad overview is provided, based on the work of Asmussen et al. (2018). It includes consideration of physiological factors, family and environmental factors, bilingualism, and type of language difficulty.

Physiological factors

1.48 There are a number of physiological factors known to affect children's language development in the early years. These are factors which are present either antenatally or from birth. A summary of the risk and protective factors for these physiological factors is presented in Table 1.1.

Antenatal factors

1.49 Studies have explored the relationship between babies' exposure to substances when in the womb and their relationship with child development. Cohort studies have identified that once confounding factors such as maternal education and birthweight were considered, exposure to cigarette smoke did not have an effect on language development (Baghurst et al., 1992; Breslau et al., 2005; Falgreen et al., 2012). The study by Fergusson and Lloyd (1991) suggests that it is not exposure to cigarette smoke in the womb that poses a risk for cognitive development per se, but the relationship between smoking and other factors such as socio-economic status. Therefore, exposure to cigarette smoke in utero on its own should not be considered a risk factor for language development.

1.50 There is a reported association between high levels of alcohol in utero and children's cognition and behavioural difficulties (Cone-Wesson, 2005; O'Leary, 2004; Weinberg, 1997). However, there is a lack of studies where alcohol consumption in pregnancy and speech and language outcomes in

children have been measured (O'Keefe, Greene and Kearney, 2014). As a result it is not possible to draw any conclusions about children's exposure to alcohol in the womb and whether or not this is a risk factor for later speech, language and communication difficulties in the absence of cognitive difficulties. Similarly, a systematic review conducted by Frank et al. (2001) reported that there was no conclusive evidence of any association between cocaine use in pregnancy and children's later speech and language outcomes once other confounding factors had been considered.

1.51 It has been reported that prolonged use of Selective Serotonin Reuptake Inhibitors (SSRIs) used to treat depression can affect subsequent language development in children when used in pregnancy (O'Connor et al., 2016; Weikum et al., 2012). Avoidance of SSRIs in pregnancy is not likely to convey any benefit to children's language skills as symptoms of anxiety and depression in pregnancy have also been reported to have an association with language development in children (Skurtveit et al., 2014).

1.52 In summary, there is limited evidence in relation to exposure to substances in the womb and children's speech, language and communication outcomes. The evidence for any association between these two factors is weak once confounding factors are considered and therefore conclusions cannot be drawn.

Premature birth

1.53 Studies have confirmed that children who are born early are at risk of delayed language development (Foster-Cohen et al., 2010; Harrison & McLeod, 2010). In Harrison and McLeod's (2010) study, being born before 36 weeks gestation was associated with an increased chance of having language difficulties. Children born very premature, i.e., before 30 weeks gestation have been noted to have poorer language scores and this has been attributed to the development of the auditory cortex (Monson et al., 2018). Stipdonk, Franken and Dudink (2018) reported in their systematic review that atypical language development in children born before 37 weeks is likely to be associated with atypical brain development. Being born

prematurely can therefore be considered a risk factor for language development.

Biological sex

1.54 Many studies have reported girls having increased language skills compared to boys in the early years (Bornstein, Hahn & Haynes, 2004; Harrison & McLeod, 2010; Maccoby & Jacklin, 1974; Roy, Kersley & Law, 2005). This is manifest through an increased number of vocalisations when in the presence of others; increased use of gesture; quicker vocabulary acquisition; and being more likely to engage in narrative discourse in the early years (Eriksson et al., 2012; Haden & Ornstein, 2009; Lewis & Freedle, 1973; Ozçalişkan & Goldin-Meadow, 2010). Expressive vocabulary has been reported to be greater for girls growing up in bilingual homes compared to boys (Floccia et al., 2018). Girls were also reported to be less likely to have difficulties with comprehension of vocabulary at age 6 than boys (Ghassabian et al. 2014). These studies report small differences in these skills but are reliable nonetheless. The differences are also observed when the children are older with girls outperforming boys in language skills as reported by teacher ratings at ages 7, 9, 10 and 12 years (Hayiou-Thomas, Dale & Plomin, 2012). In contrast, a study by Rice et al. (2008) reported that boys were able to close the gap in their language abilities by age 7. Overall, the evidence suggests that boys are more likely to have early speech, language and communication needs than girls although the long term educational implications of this are not conclusive.

Hearing

1.55 The relationship between conductive hearing loss resulting from Otitis Media with Effusion (OME) and language development has been well researched. Roberts, Rosenfeld and Zeisel (2004) conducted a meta-analysis of all prospective and randomised controlled trials concerning OME and speech and language development. They reported that while there were some small effects of OME on receptive and expressive language development these may be overestimated because when confidence intervals were reviewed

they no longer reached statistical significance. Similarly, none of the included studies adjusted for confounding factors such as socio-economic status. Since this meta-analysis was published, a large cohort study by Harrison and McLeod (2010) reported that hearing loss was a risk factor for language difficulties. Using multi-factor analyses they identified that there was a greater odds of having either receptive and/or expressive language difficulties when there were ongoing hearing issues. However, this study was based on parent reported concerns and when children's language was formally assessed on a language measure, hearing no longer reached statistical significance as a risk factor. Whilst there is limited evidence available regarding the association between conductive hearing loss and language development there is evidence that children with sensori-neural hearing loss are at risk of later language difficulties (Ching et al., 2013), even when this loss is mild-to-moderate (Tomblin et al., 2015).

1.56 In summary, while there is evidence that a sensori-neural hearing loss can affect language development there is no robust evidence available to indicate that there is an association between children experiencing glue ear as a result of OME and language development. Therefore, conductive hearing loss cannot be considered as a risk factor for speech, language and communication needs.

Temperament

1.57 Some studies have reported an association between a child's temperament and their language development. Children who were considered to have a shy temperament were associated with reduced vocabularies at age 2 years (Prior et al., 2008). A study by Hauner et al. (2005) noted a negative association between children reported as having low sociability/approachability, negative mood and low task persistence with an increased severity of speech delay. Harrison and McLeod's (2010) cohort study reported reduced odds of receptive and expressive language difficulties for children described as having a more persistent temperament. Children with a more reactive temperament were more likely to have

receptive and expressive language difficulties as reported by their parents and also by testing on a formal measure of language. These characteristics of a child's temperament may therefore be indicative of their risk of speech, language and communication needs.

Summary

1.58 Table 1.1 summarises the risk and protective factors for all of the physiological factors discussed. Overall there is evidence that being born prematurely is a risk factor for later language difficulties and this may be as a result of how brain development has been affected. There is no evidence that exposure to substances in the womb is associated with language outcomes once other confounders are taken into consideration. There is strong evidence that boys are more likely to experience early speech, language and communication difficulties. Sensori-neural hearing loss has been demonstrated to be a risk factor for language development but there is limited evidence regarding the association between conductive hearing loss and speech, language and communication needs.

Table 1.1: Physiological factors associated with Risk and Protective factors for Speech, Language and Communication

Physiological Factors	Risk Factors	Protective Factors
Antenatal Factors	Premature birth	Being born full term
Biological Sex	Male	Female
Hearing	Sensorineural hearing loss	No hearing loss
Temperament	Shy children/low sociability. Reactive temperament	Sociable, more persistent temperament

Family and environmental factors

1.59 There are a number of factors which have been reported to affect children's language development which are related to their parents, family and/or environment. These include maternal age, parental level of education, household income and how parents interact and communicate with their

children (Collisson et al., 2016; Goisis, 2015; Hart & Risley, 1995; Harrison & McLeod, 2010; Rowe, Rawdenbush & Goldin-Meadow, 2012). It is important to note however that how parents communicate and interact with their child can have more influence on a child's speech, language and communication development than any demographic factors and this is becoming more widely recognised in the literature (Roulstone et al., 2011). It is also important to consider that there are many influences on a child's language development beyond their family. For example, the presence of other adults/children in a child's life and the range of places that they visit will influence the vocabulary that they are exposed to. This range of experiences and exposure to different events in a child's life has been reported to influence their language learning (Golinkoff et al., 2019).

Maternal age

1.60 In the study by Goisis (2015), children's vocabulary was assessed using the British Ability Scales and plotted alongside the age of mothers when they had their first child. Children's vocabulary scores increased with maternal age, until mothers were aged between 30 to 34 and then vocabulary scores reduced. This remained statistically significant when other confounders such as socio-economic status and maternal education levels were accounted for ($p>0.05$). A cohort study by Harrison and McLeod (2010) also identified older maternal age as being an important predictive factor in language development.

Maternal Mental Health

1.61 Maternal wellbeing has been reported as an important protective factor for children's language development (Harrison and McLeod, 2010). Children whose mothers were reported to have higher levels of psychological wellbeing were less likely to have receptive or expressive language difficulties reported by their parents and performed better on standardised language measures. A recent study by Clifford et al. (2021) also reported that maternal depression was associated with fewer words used by mothers; reduced conversational turns between mothers and their child; and reduced

number of child vocalisations. Therefore, although this evidence is limited it is an important consideration in relation to a child's risk and protective factors for speech, language and communication development.

Birth order and siblings

1.62 There is some evidence that children who are born first develop their vocabulary at a quicker rate than any subsequent siblings (Fenson et al., 1994; Pine, 1995; Zambrana et al., 2012). However, this rate of vocabulary acquisition slows down once children have learned approximately 50 words. Siblings are more likely to have a wider exposure to pronouns such as he, she, their, we etc (Berglund, Eriksson & Westerlund, 2005) which could be perceived as being beneficial for language development. Harrison and McLeod (2010) reported that children with older siblings were more likely to experience language difficulties than children with younger or no siblings. Therefore, whilst there is evidence that older siblings are associated with potential for experiencing language difficulties, there is no robust evidence regarding any long term effects of this.

1.63 A study of twins by Oliver, Dale and Plomin (2004) identified that while there may be some mild, genetic element to differences in language scores, vocabulary development is primarily affected by environmental influences, which they describe as exposure to new settings, people or items. This was also reported in the study of twins by Hayiou-Thomas, Dale and Plomin (2012) with the shared environment having the greatest effect on a child's language skills. Therefore, being a twin was not considered a risk factor for language development.

Family history of speech and language difficulties

1.64 A study by Collisson et al. (2016) reviewed a number of reported potential risk factors for language development. The most statistically significant relationship reported was that children were likely to experience a language delay where other children in the family had also previously had speech and language difficulties ($p=0.002$). The likelihood increased with the number of siblings who had experienced a delay in developing their language. Family

history of language difficulties, specifically Developmental Language Disorder (DLD) has been identified as a prognostic factor in many other studies (Choudhury and Benasich, 2003; Kalnak et al., 2012; Lahey and Edwards, 1995; Rice, Haney and Wexler, 1998; Reilly et al., 2010; Tomblin, 1989). There is strong evidence that history of speech, language and communication needs to be considered a predictive risk factor of persistent needs.

Family income/socio-economic status/parental education levels

1.65 One of the most cited studies of language development in relation to socio-economic status is that carried out by Hart and Risley (1995). They reported that by the age of 4 years old there was a 30 million word gap between children living in more affluent families in the United States and those in receipt of welfare. It is worth noting however, that in recent years, a number of critiques have been published of Hart and Risley's findings. Concerns have been raised about the representativeness of their small sample, and authors argue that findings may not capture a comprehensive picture of children's language environments due to methodological and theoretical biases (Kuchirko, 2019; Paugh & Riley, 2019; Sperry, Sperry & Miller, 2019). Other authors have attempted to replicate the study with Weisleder and Fernald (2017) reporting substantial variation in adult input across families from lower socioeconomic groups.

1.66 Rowe, Raudenbush and Goldin-Meadow (2012) reported that children had increased vocabularies in line with their parents' level of education. Those children whose parents had the highest level of qualifications had greater vocabularies, which continued to grow at an increased exponential rate as the child got older. There is some evidence that this gap in vocabulary size is not evident at age 12 months (Brushe et al., 2020) but becomes apparent by age 18 months (Fernald, Marchman & Weisleder, 2013).

1.67 Growing up in homes where parents have higher levels of education and/or income can be considered a protective factor. Associations between parental levels of education, level of income and also the quality of interactions

between parent and child have all been reported (Hoff, 2013; National Institute of Child Health & Human Development Early Childcare Research Network, 2003).

Quality of caregiver interactions

1.68 Many studies have reviewed the frequency and quantity of words used by parents with their children and how this affects their language development. A positive association has been reported where parents are responsive and sensitive to their child's interactions with children achieving their language milestones at an earlier age. This was found to continue to be of significance when the confounders of family income or other socio-economic factors were taken into consideration (Tamis-LeMonda, Bornstein & Baumwell, 2001). The amount and quality of infant directed speech has been found to be positively associated with the child's receptive and expressive vocabularies (Huttenlocher et al., 1991; Ma, Houston & Hirsh-Pasek, 2011). Hammer et al. (2001) reported that parents of children considered to have typically developing language had engaged more frequently in conversational activities whereas parents of children with DLD had engaged in school readiness activities such as teaching their children colours and the alphabet.

1.69 When considering responsiveness and quality of interactions, it is important to recognise that much of the research comes from studies of Western and middle-income populations, and therefore does not necessarily represent norms of interaction across the world (Kuchirko, 2019). While responsiveness has been observed in parents across cultures, researchers have documented variation in how and when parents respond, and the types of response that are usually given (Tamis-LeMonda, Kuchirko & Song, 2014).

1.70 Parents appear to be most responsive to their children when they are babbling (Ramirez-Esparza, Garcia-Sierra & Kuhl, 2014). This study also observed that it was not only the amount of words heard by the child but the quality of the interaction with the parent that predicted the child's receptive vocabulary at age 24 months. Other studies have reported similar outcomes

and the parent behaviours with the most protective factors for language development appear to be responsiveness to the child, parental use of gesture, using a wide variety of vocabulary, and use of one-word utterances (Lacroix, Pomerleau & Malcuit, 2002; Hurtado, Marchman & Fernald, 2008; Cartmill et al., 2013; Rowe, 2012; Swingley & Humphrey, 2018). A recent study by Shapiro et al. (2021) reported that fathers produced, on average, 51.9 per cent less child-directed speech than mothers. However, this varied by the age of the child with fathers more likely to use child-directed speech when their children were aged 2 years than when they were 6 months old. Fathers' repetition of their child's speech at age 2 years is associated with children using a wider vocabulary at age 4 years, even when mothers' repetition is controlled for (Conica, Nixon and Quigley, 2020).

1.71 Rowe (2008) brought socio-economic and quality of caregiver factors together to study children's language outcomes. This research reported that high levels of child-directed speech resulted in an increased vocabulary. Children whose parents had a high level of education and higher incomes were more likely to experience higher levels of child-directed speech. There was an association between those parents with a higher level of education having a greater understanding of child development and therefore engaging in more child-directed conversations resulting in improved language outcomes for the child. Therefore, parental engagement, responsiveness and increased exposure to vocabulary can be considered as protective factors for language development. Knowledge of child development is important and how to engage in child-directed speech activities can be considered a protective factor for language development regardless of socio-economic status.

Reading and books

1.72 Many studies have reported that the frequency of book-sharing between parents and children has a statistically significant association with a child's vocabulary and language development (Bus et al., 1995; Collisson et al., 2016; Hammer et al., 2001; Mol et al., 2008). A study by Farrant and Zubrik

(2013) reported that the number of books within the home can be used as a predictor for a child's vocabulary at age 5 since it is associated with the frequency of book reading at home. There is evidence that the interactions and discussions that take place through parents reading with their children are important and sharing books without text has been demonstrated to increase children's narrative skills (Leech & Rowe, 2014; Petrie et al., 2021).

Screentime

1.73 Early television watching (before the age of 2 years) has been associated with delayed language skills (Chonchaiya and Pruksananonda, 2008; Zimmerman, Christakis and Meltzoff, 2007) and decreased vocabulary (Pagani, Fitzpatrick and Barnett, 2013). However, if parents are engaged with their children while watching television this can allow for child-directed conversations and Zimmerman et al. (2009) report that this can remit the effects of children watching television alone. Anderson and Pempik (2005) argue that children will learn less from watching television than from real-life experiences. Therefore, there is some evidence that watching television can be detrimental for language development. Engaging children in conversations about what they are watching can reduce the impact of this but does not replace the opportunity for child-directed speech about real-life activities.

Childcare

1.74 Attending a childcare/pre-school setting has been associated with a positive effect on a child's language development (Collisson et al., 2016). There is also some evidence that children receiving care from a childminder or relative will have an increase in their vocabulary skills compared to children who do not have this range of caregivers (Morris, Melhuish & Gardiner, 2017).

1.75 Training early years practitioners about language development is important in this relationship between childcare providers and children's language development (Hoff, 2006). Also of importance is the age when the child receives this input. Starting childcare at an early age can be advantageous

for language development. Research has demonstrated that there are greater gains in vocabulary with earlier access to childcare and this is true for all children regardless of parental level of education or socio-economic status (Barnes & Melhuish, 2017; Becker, 2011; Sylva et al., 2011). Therefore, access to childcare when young can be considered a protective factor providing language enrichment.

Bilingualism/Multilingualism

1.76 Children growing up in bilingual or multilingual homes are reported to develop their language skills at a slower rate than those in monolingual homes initially. However, the evidence also suggests that by the age of 3 their language skills are within the typical range of development for all languages (Slobin, 1985). There is no evidence that children from bilingual or multilingual homes are more likely to experience an ongoing delay in the development of their language. They may use code-switching (where children use both languages within one sentence, however, this is not considered as a delay in their language development. Indeed, the exposure to differing vocabulary and grammatical structures can be viewed as an advantage to developing language skills (Genesee, 1988). Within an educational setting, there is some evidence that even when children enter an educational provision with little or no prior knowledge of the primary language used in that setting they are able to make gains in their acquisition of that language and to be as proficient as their peers by age 7 (Gathercole & Thomas, 2007a; McKean et al., 2015). Bilingualism or multilingualism should not be considered as a risk factor for language development.

Summary

1.77 Table 1.2 summarises the risk and protective factors in relation to family and environmental factors. The quality of interactions between parents and their child can influence language learning, regardless of socio-economic status. Increased adult responses to children, exposure to a wide range of vocabulary from a range of caregivers and exposure to different environments can be beneficial. Parents who understand child development

are more likely to engage in child-directed conversations during activities such as reading.

Table 1.2: Family and Environmental factors and associated Risk and Protective factors

Family and Environmental Factors	Risk Factors	Protective Factors
Maternal age	Mother aged <30 years old at birth of first child	Mothers aged >30 years old at birth of first child
Family history of SLCN	History of an immediate family member with Development Language Disorder	n/a
Socio-economic status	Low socio-economic status	n/a
Parents level of education	Low level of parental education	High level of parental education. Knowledge of child development.
Quality of interaction	Limited child-directed speech. Limited exposure to varied vocabulary	High levels of child-directed speech, exposure to wide vocabulary
Book sharing	Limited exposure to books or book-sharing activities	Frequent book sharing activities and access to books
Childcare	n/a	Exposure to childcare settings and other language role models

Type of speech, language or communication difficulty

1.78 Studies have focussed on many aspects of speech, language and communication development in order to identify whether there are particular skills that are indicative of later difficulties. I CAN (2009) described children as either having transient or persistent language needs. Transient needs are where language difficulties are identified but with support, children are able

to catch up with their peers. Persistent language needs are where children have ongoing language difficulties compared to their peers and/or expected norms and intervention is intending to reduce the impact of their needs. Some studies have highlighted particular features in children's development which are associated with an increased risk of language difficulties persisting into school age (Snowling et al., 2001; Stothard et al., 1998).

Gesture

1.79 Use of gesture has been reported as a positive predictor of later language performance (Rowe & Goldin-Meadow, 2009; Rowe, Raudenbush & Goldin-Meadow, 2012). A meta-analysis by Colonnesi et al. (2010) suggested that pointing is an early predictor of later language development.

Babbling

1.80 There is evidence that some speech errors can indicate later language difficulties. In particular, children who were delayed in the onset of canonical babbling, i.e., use of repeated consonant vowel sequences such as 'baba', 'dodo', were later identified as having smaller vocabularies at ages 18 months, 24 months and 30 months (Morgan & Wren, 2018). Morgan and Wren's systematic review identified that children with language difficulties were noted to have qualitative differences in their early babble when compared to peers who did not have any subsequent language difficulties. Delayed or absent canonical babbling can be viewed as a potential risk factor for persistent language difficulties.

Severity of difficulty

1.81 Both level of difficulty and breadth of need have been shown to be associated with later outcomes for language development. A study by Rescorla and Schwartz (1990) reported that a larger gap between expected levels of expressive language and levels attained at age 2 correlated with lower expressive language levels at age 3 to 4. With regards to breadth of need, the systematic review by Law et al. (1998) reported that children who experienced both receptive and expressive language difficulties were more

likely to have persisting language difficulties and achieve poorer educational outcomes than their peers. Therefore, language difficulties at an early age, particularly where there are receptive and expressive language difficulties, can be perceived as a risk factor for persistent language needs.

Vocabulary

1.82 Limited vocabulary in the early years has been identified as a potential risk factor for later language development. Dale et al. (2003) monitored children's language at age 3 and 4, having identified their vocabulary use at age 2. Of those children who had smaller vocabularies than expected at age 2, persistent language difficulties were evident in 44.1 per cent at age 3. At age 4, language difficulties were reported to be persistent in 40.2 per cent of the children originally identified as having smaller vocabularies. Similarly, while children may make progress with understanding of vocabulary, studies suggest that closing the gap with their peers is particularly challenging for this aspect of language development (Rice, 2013; Rowe, Raudenbush & Goldin-Meadow, 2012).

1.83 These findings for expressive vocabulary are in contrast to studies of receptive vocabulary, i.e., children's understanding of vocabulary, suggesting this is not a predictor of later language development. Evidence from Ghassabian et al. (2014) reported that 62 per cent of children who had difficulties with receptive vocabulary at age 6 years did not have any identified difficulties with vocabulary at age 18 months or 2 years 6 months.

Grammatical markers

1.84 The lack of use of grammatical markers can be an indicator of which children are more likely to have persistent language difficulties. Rice (2013) describes how children with DLD use fewer grammatical structures than their peers, even when they have similar mean lengths of utterance. They are less likely to use auxiliaries (e.g., he is, they are) regular past tense (e.g., walked) and irregular past tense (e.g., ran, spoke), and possessive 's' (e.g., Daddy's shoes). They acquire these skills at the same pace as their peers but are

delayed in their use. Reduced use of these morphosyntactic patterns can therefore be considered a risk factor for persistent language difficulties.

Summary

1.85 In summary, use of gesture (including pointing), a small gap between expected language level and actual language level, expressive delay only, wide vocabulary and use of morphosyntactic structures, can all be considered as protective factors in language development. In contrast, limited use of gesture, limited canonical babbling, a large gap between expected language level and actual language level, difficulties with both receptive and expressive language, limited use of morphosyntactic structures and a limited vocabulary can all be considered as risk factors for persistent language difficulties. These risk and protective factors are summarised in Table 1.3.

Table 1.3: Risk and Protective Factors in relation to Speech, Language and Communication skills

Developmental feature	Risk Factors	Protective Factors
Babble	Absent or limited babble	Using babble at appropriate developmental level
Gesture	Absent pointing and limited use of gestures	Use of pointing and gestures
Severity of Difficulty	Receptive and Expressive language difficulties. Large gap between expected level of language and level child is functioning at	Reaching developmental language milestones
Vocabulary	Limited expressive vocabulary	Wide vocabulary use
Grammatical markers	Lack of grammatical markers e.g., auxiliary verbs 'is/are', past tense '-ed' and plurals.	Using grammatical structures at an age appropriate level

Conclusion

1.86 As reported by Nelson et al. (2006), it is important to acknowledge that whilst there is a wealth of literature reporting potential risk factors, it is not possible to be fully conclusive about these due to the variability in results and the heterogeneity of included subjects in the design of the studies. Further to this, many of the studies did not control for confounding effects (Berkman et al., 2015). It is important to note that inclusion of data in Tables 1.1, 1.2 and 1.3 does not imply an equal level of risk or protection and that data on relative contributions of each factor is as yet unavailable.

1.87 A child's development of speech, language and communication is influenced by a wide range of factors, each of them potentially affecting the child in their own manner (McKean et al., 2015), i.e., each child will have their own complex, interrelated set of factors resulting in a unique set of circumstances for speech, language and communication development. Using knowledge of the risk and protective factors can help to ensure that every child can achieve their potential.

2. Methodology

2.1 This section provides detail on the methodology used for the scoping reviews undertaken to address objective 2. The formal identification and treatment of speech, language and communication needs currently falls to health services in the early years of life, and health boards across Wales have procedures for identification of these difficulties. Educational services and those responsible for providing Early Childhood Education and Care (ECEC) also have a considerable role.

2.2 The first stage before undertaking the reviews was to search for previous relevant systematic and meta-analyses. Three systematic reviews were identified which examined screening for speech and language delay in pre-school children (Berkman et al., 2015; Law et al., 1998; Nelson et al., 2006). In the earliest review, Law et al. (1998) suggested that primary care workers (e.g., health visitors, general practitioners, school nurses and nursery practitioners) should be involved in eliciting parental concerns and in making appropriate observations of children's communication behaviours. Nelson et al. (2006) considered the use of risk factors to guide the identification of who should be screened and found this was not supported by the evidence. They concluded that investigations to date were insufficient to determine optimal methods for screening, including the tools to use and the age at which to screen. Finally, Berkman et al. (2015) provided a formal update to the Nelson et al. (2006) review, and suggested that two parent-rated screening instruments, the MacArthur Bates Communicative Developmental Inventory (CDI) (Fenson et al., 1994) and the Language Development Survey (LDS; Rescorla & Alley, 2001), can accurately identify children who require further diagnostic evaluation from a speech and language therapist.

2.3 A limitation of these reviews is that they provide limited information about screening for bilingual or multilingual children and the specific considerations that should be taken into account for these populations. A further limitation is that while the reviews have incorporated studies taking place in the UK, no review has considered the Welsh context specifically. It was therefore

essential to review the literature on how screening may need to be adapted for a bilingual and Welsh-speaking context. The reviews carried out in this study took into account both bilingualism at the individual level, in terms of how each child's linguistic background may affect their performance on assessments, and the societal level, to reflect the varied linguistic landscape in Wales.

2.4 The three rapid scoping reviews undertaken and reported herein, aim to provide information needed to select appropriate screening tool(s) which can be used by people who are not speech and language therapists to screen for speech, language and communication needs in children under age 5. Each review has a specific aim, as outlined below.

Aims

English Language Screening Tools Review

2.5 The aim of this review was to build on and update the evidence base reported in previous systematic reviews (Berkman et al., 2015; Law et al., 1998; Nelson et al., 2006), and from this to identify additional tools that may be used for screening children aged 0 to 4 years 11 months.

Bilingual Language Screening Tools Review

2.6 The aim of this review was to identify existing tools that are available for use with preschool- aged bilingual/multilingual populations, and to provide recommendations regarding what factors to consider when screening children in a bilingual or multilingual context.

Welsh Language Screening Tools Review

2.7 The aim of this review was to identify existing tools that are available for use in Welsh, and to provide recommendations for what to consider when screening children who speak only Welsh or are Welsh-English bilingual.

Methods

2.8 Due to time constraints, a rapid scoping review methodology was chosen and was appropriate given the exploratory research objectives, which aimed to identify and synthesise available evidence and recommendations (Peters et al., 2020). Although rapid reviews have traditionally been used as a streamlined method of conducting a systematic review (Tricco et al., 2015), rapid scoping review methods are also becoming increasingly common (Rajmil, 2020; Tricco et al., 2017). To increase efficiency of the review process, measures recommended in Abrami et al. (2010) were employed, for example building upon existing reviews and limiting inclusion criteria to a specific time period.

Eligibility criteria

2.9 Inclusion and exclusion criteria for each of the reviews were based on previous systematic reviews (Berkman et al., 2015; Law et al., 1998; Nelson et al., 2006). Additional elements were considered to ensure that the eligibility criteria were relevant to the Welsh context. For example, an age range of birth to 5 years was selected to reflect the age at which children start compulsory schooling in Wales. In addition, the researchers considered tools designed for use in a primary health care or ECEC setting, to reflect that in Wales, screening is usually carried out by health visitors and early years practitioners.

2.10 Eligibility criteria varied slightly for each of the reviews to reflect the individual aims. For example, stringent criteria were applied for the English Language Screening Tools review, stating that each study must identify and validate a new assessment tool for screening speech and language. In the case of the Bilingual and Welsh Language reviews, the research team were interested in understanding more generally the benefits and challenges of adapting existing tools, and this flexibility was reflected in the inclusion criteria. Given the small number of studies identified for the Welsh Language review, the inclusion criteria were developed iteratively and more flexibility was given in the age range of children included in these studies.

Inclusion criteria for all three reviews

- Children's age range of 0 to 4 years, 11 months
- Sample size larger than 20
- Sample drawn from one of:
    - general population
    - clinical population of children with speech and language difficulties
    - mixed clinical/general population.
- Study design was description/validation of a test designed for identification of speech and/or language delay or difficulties
- Test was designed to be used in either:
- primary care health setting
- ECEC setting by non-specialist staff for early identification

Additional inclusion criteria for the Bilingual review

- Study design was translation/discussion of adaptation of an existing test for identification of speech and/or language delay or difficulties in bilingual and multilingual children.

Additional inclusion criteria for the Welsh Language review

- Initial aim was to review studies where the design was translation/discussion of adaptation of an existing test for identification of speech and/or language delay or difficulties in Welsh-speaking children. Due to the lack of studies including Welsh speaking children, inclusion criteria extended to include studies containing data relating to Welsh speech and/or language acquisition and age range extended 0 years to 6 years 10 months (with two papers containing children up to the age of 9 years 7 months and 11 years 2 months).

Exclusion criteria for all three reviews

- Study was related to children with following diagnoses: Attention Deficit Disorder/Attention Deficit and Hyperactivity Disorder; deafness/sensorineural

hearing loss; autism; psychiatric or emotional and behavioural difficulties; Down's Syndrome; cerebral palsy; dyslexia; other secondary speech and language delay (e.g., linked to learning difficulties, neurological conditions such as epilepsy, closed head injury); cleft lip/palate; children born preterm.

- Study design was a case report, systematic review, scoping review, discussion paper
- Test was designed for diagnosis by specialist, e.g., speech and language therapist

Additional exclusion criteria for the English Language Screening Tools review

- Families who are multilingual

Additional exclusion criteria for the Bilingual review

- Families who are monolingual

Additional exclusion criteria for the Welsh Language review

- No Welsh-speaking participants, no direct reference to speech or language acquisition, reports on language acquisition but in relation to perception or recognition only.

Search strategy

2.11 Core search terms were initially developed relating to the following areas:

- population of interest (preschool-aged children)
- language development and disorders
- screening.

2.12 Search terms were adapted from previous systematic reviews (Asmussen et al., 2018; Berkman et al., 2015; Law et al., 1998; Nelson et al., 2006; Roulstone et al., 2015) and were developed by members of the research team in consultation with a clinical librarian. As such terms which are no longer in use, e.g., mental handicap, were included as they would historically have been included in research papers.

2.13 Once agreement was reached, core search terms were piloted for the English Language Screening Tools review in the Ovid Medline database. Restrictions were applied to the search for the English Language Screening Tools review, including studies published only since 2015, which was the date of the most recent systematic review conducted. The Bilingualism and Welsh language review searches were conducted to include studies published from the year 2000, to reflect relevant research in this field within recent years.

2.14 For the Bilingualism review, additional search terms were developed relating to bilingualism and multilingualism and cultural diversity. Terms were adapted from existing systematic reviews on the topic of bilingual language development (Coalson et al., 2013; Hambly et al., 2013; Lund, Kohlmeier & Durán, 2017; Smith, Pacheco & Khorosheva, 2020; Surrain and Luk, 2017). These terms were then combined with the core search terms as shown in Table 2.1.

2.15 For the Welsh language review, search terms relating to Wales and the Welsh language were developed, although the number of terms was kept small as it was anticipated there would be few overall records. Terms were initially developed in English and searches were conducted in the English language initially. All terms were then translated into Welsh and the searches were run for a second time in the Welsh language.

2.16 Table 2.1 shows the distribution of topic search areas as they were combined across the four searches. For the full search strategy see Annex A.

Information sources

2.17 Searches were conducted across the following five databases: Ovid Medline, PsycInfo, CINAHL, ERIC and EMBASE. For the Welsh language review, issues of the Welsh language journal Gwerddon were also searched for relevant articles.

Source of evidence screening and selection

2.18 Results from the searches were collated, uploaded to EndNote and de-duplicated. At this point the two Welsh searches were combined to create one database of records for the Welsh review. After de-duplication, records were transferred into three separate databases in Rayyan, a computer programme designed for use in systematic reviews.

2.19 All screening at title and abstract level was conducted in Rayyan. Titles that were distinctly unrelated to the concept being studied were removed at this point. Abstracts were then reviewed against the inclusion criteria and additional records were excluded. Screening for each review was conducted by different members of the team, with 10 per cent of records being double-screened at abstract level to check for inter-rater reliability. The research team then met to discuss any differences and reached a consensus. Variation between research team members was minimal, allowing confidence for review of the remaining abstracts to be undertaken by one team member per review.

2.20 Full texts of selected studies were then sourced and screened. Where there was disagreement or uncertainty about inclusion, additional members of the research team were consulted and agreement was reached.

Table 2.1: Development of search terms

Search topic area	English Language Tools Review	Bilingualism review	Welsh Language review
Core search terms			
Population (preschool aged children)	X	X	X
Language development & disorders	X	X	X
Screening and identification	X	X	X
Additional search terms			
Validity and reliability	X		
Bi- and multilingualism		X	
Wales and Welsh language			X

Data extraction

Data extraction forms were developed based on the Joanna Briggs Institute guidance (Peters et al., 2020; Annex B).

3. Findings – Scoping Reviews

Systematic reviews are usually reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. This provides detailed information about each of the stages involved in the review process (PRISMA, 2020). The PRISMA diagrams below illustrate the process for each review undertaken including total numbers of references identified in each database, numbers excluded at each stage of screening and the reasons for exclusion. A brief summary of findings and recommendations from studies that were charted in each review are presented next. A comprehensive table including details of all included studies is presented in Annex C.

English Language Screening Tools Review

3.1 As displayed in Figure 3.1, a total of 2,432 unique records were identified through searches for the English Language Screening Tools review. Of these 2,090 records were excluded at title level and 264 at abstract level. Of the remaining 68 fulltext papers which were sourced and assessed for eligibility, data were charted for the 11 studies included in the final review. Table 3.1 details the 11 additional screening tools identified through the review.

Bilingual Screening review

3.2 As Figure 3.2 shows, of a total 1,282 unique records identified, 533 were excluded at title level and 652 at abstract level. The remaining 97 articles were then retrieved and screened at fulltext level. In addition, another relevant study (O'Toole et al., 2016) was identified through the reference lists of identified reports. At fulltext stage an additional 83 articles were excluded. The reasons are detailed in Figure 3.2 although this was most frequently due to the age range of the children exceeding the age range. Data were then extracted from the 16 articles that remained.

Figure 3.1: PRISMA diagram for English Language Screening Tools Review

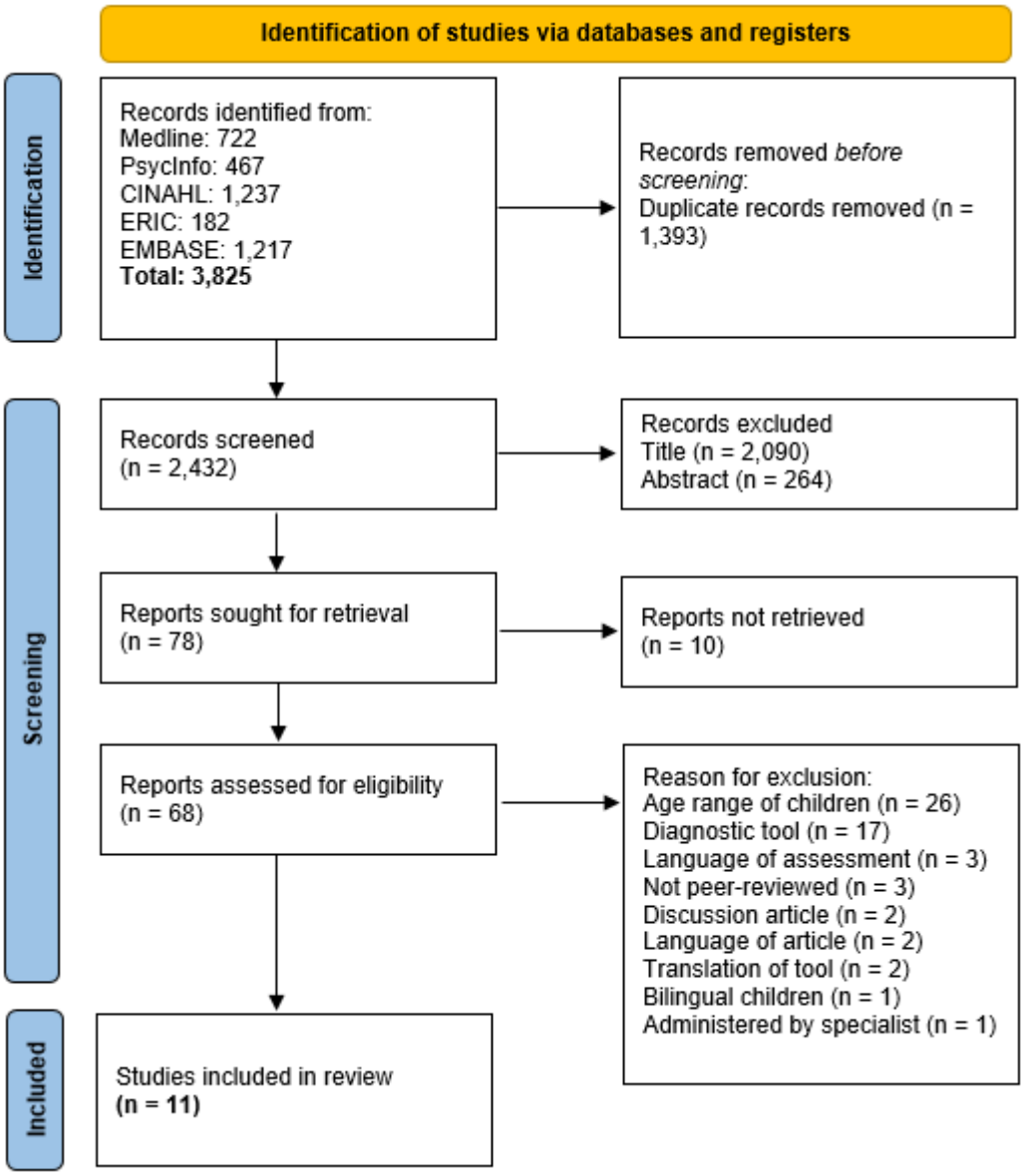
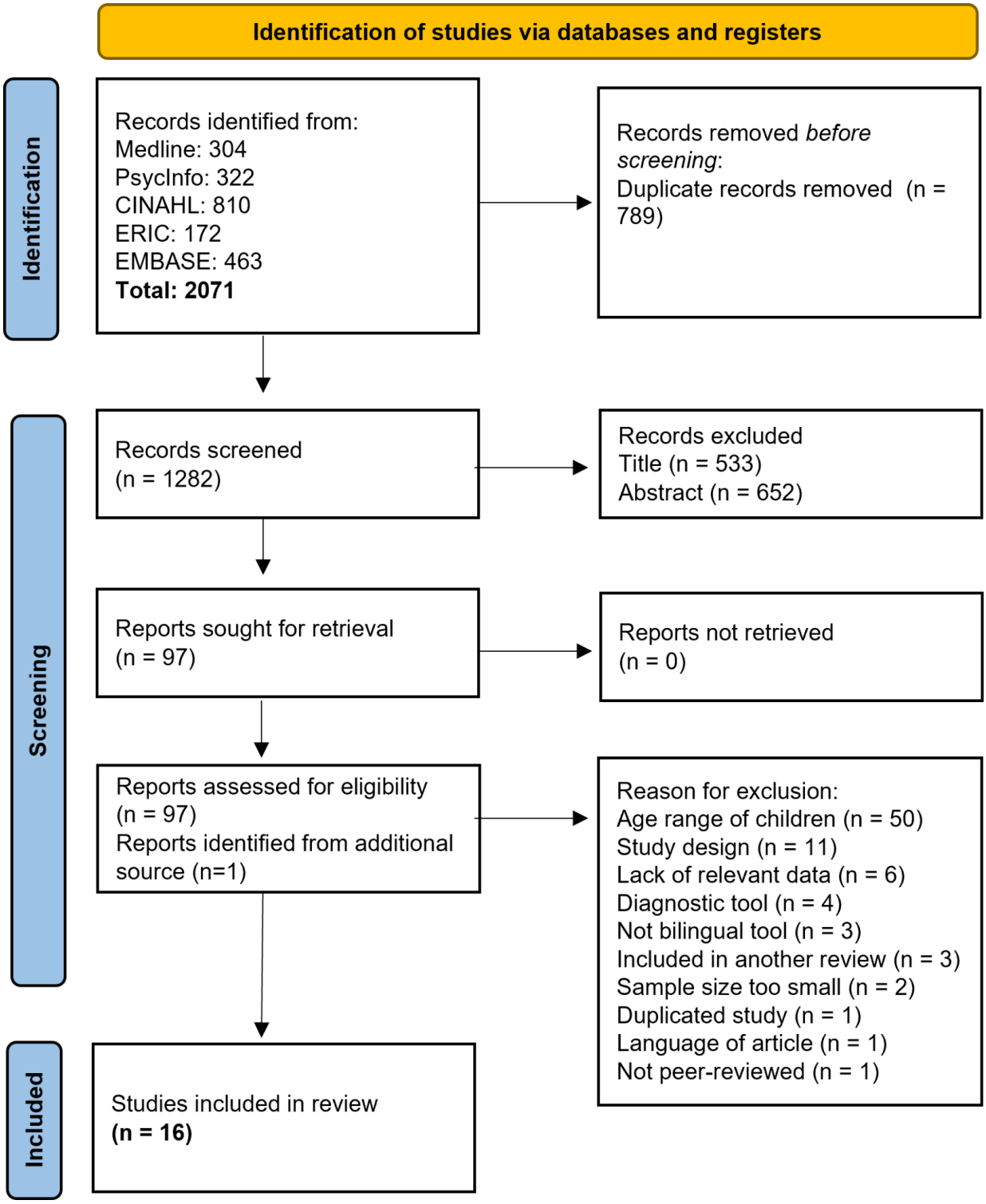


Table 3.1: Tools identified through the English Language Screening Tools Review

Tool identified	Aspect of speech/language assessed	Country of Origin	Reference
Toddler Communication Development Inventory (Italian version)	Expressive vocabulary	Italy	Chilosi et al. (2019)
Toddler Phonology Test	Phonology	Australia	Claessen et al. (2017)
Lena Developmental Snapshot	Receptive and expressive language	USA	Gilkerson et al. (2017)
Toddler Language and Motor Questionnaire	Language expression and comprehension	Iceland	Gudmundsson (2015)
FOCUS and Communication Function Classification System	Everyday communication function	Canada	Hidecker et al. (2017)
Parent Report of Children's Abilities – Revised (PARCA-R)	Vocabulary, sentence complexity, linguistic skills	UK	Johnson and Bountziouka (2020)
Dynamic Indicators of Vocabulary Skills (DIVS)	Vocabulary (assessed through fluency in picture naming and reverse definition)	USA	Marcotte et al. (2016)

Vocal Developmental Landmarks Interview (VDLI)	Vocal development	USA	Moeller et al. (2019)
Production of Infant Scale Evaluation (PRISE)	Vocal development	Iran	Oryadi-Zanjani (2018)
Sequenced Language Scale for Infants (SELSI)	Semantic-cognitive, phonological, syntactic abilities and pragmatic competence	Korea	Ruu and Sim (2019)
Early Years Foundation Stage Unique Child Communication Sheet (EYFS:UCCS)	Expressive and receptive language and attention	UK	Seager and Abbot-Smith (2017)

Figure 3.2: PRISMA diagram for bilingualism review



3.3 From the studies included in the Bilingualism Review, nine tools are represented in Table 3.2. Cattani et al. (2014) compared children's performance on three tools, and Thordardottir et al. (2006) across two tools. There was variability in whether studies administered tools in both of the child's languages or in the primary language of the home only. In most cases the administrator of the tool was a fluent speaker of the child's languages, most often a parent, caregiver or a trained member of the research staff. However, Cattani et al. (2014) compared children's performance on the CDI in English and their additional language with assessment administered in English only by monolingual speakers on the three tests of language development (British Picture Vocabulary Scale, Preschool Language Scale-4, and SETK-2[1]. In addition, Patterson, Rodríguez and Dale (2020) report on children's performance on dynamic assessment tasks which were administered to bilingual Spanish-English children by a graduate student with proficient but non-native ability in Spanish.

3.4 The most frequently studied language pairs were English-Spanish, in six studies (Core et al., 2013; DeAnda et al., 2016; Hardin, Scott-Little & Mereoiu, 2013; Mancilla-Martinez et al., 2016; Marchman & Martinez-Sussman, 2002; Patterson, Rodríguez & Dale, 2020; Vagh, Pan & Mancilla-Martinez, 2009) and English-French, in three studies (DeAnda et al., 2016; Legacy et al., 2017; Thordardottir et al., 2006). In addition, four studies included children learning a primary language (English or Swedish) and a range of additional languages to reflect variability in bilingual language backgrounds in these countries (Cattani et al., 2014; Floccia et al., 2018; Goh et al., 2017; Nayeb et al., 2021).

3.5 The tools used in one study (Floccia et al., 2018) were incorporated into the overall evaluation of screening tools, as this was the only study to include Welsh-English bilingual participants. Floccia et al. (2018) developed the UK Bilingual Toddlers Assessment Tool, an assessment which includes a combination of the CDI in British English, a CDI in the additional language,

[1] Sprachentwicklungstest für zweijährige Kinder-2

and a language exposure questionnaire. Normative data were gathered for 372 children representing 13 language pairs, including Welsh-English bilingual speakers. To assess children in Welsh, the Welsh-English Bilingual CDI was used (Mills, Gathercole & Ebanks, 2013), an adaptation which has been developed but does not yet have pilot data available.

Table 3.2: Tools identified through the Bilingualism review

Tool/adapted version of tool used	Aspect of speech/language assessed	References	Language(s) studied	Administered by	Was administrator a native speaker?
MacArthur Bates Communicative Development Inventory	Productive and receptive vocabulary, grammar (depending on version)	Armon-Lotem and Ohana (2017)	English, Hebrew	Usually parents, sometimes additional caregiver (e.g. teacher or another family member)	Yes
		Cattani et al. (2014)	English & CDIs available in: Arabic, French, Italian, Spanish, Catalan, German, Gaelic Irish, Dutch, Finnish, Polish, Slovak, Mandarin and Punjabi		
		Core et al. (2013)	English, Spanish		
		De Houwer, Bornstein and Putnick (2012)	French, Dutch		
		Legacy et al. (2017)	English, Canadian French		

		Mancilla-Martinez et al. (2016)	English, Spanish		
		Marchman and Martínez-Sussmann (2002)	English, Spanish		
		O'Toole et al. (2016)	English, Polish, Turkish, German, Hebrew, French, Portuguese & bilingual questionnaires for Maltese-English and Irish-English		
		Thordardottir et al. (2006)	English, French		
		Vagh, Pan and Mancilla-Martinez, (2009)	English, Spanish		
UK Bilingual Toddlers Assessment Tool (UKBTAT)	Productive and receptive language	Floccia et al. (2018); Plymouth University Baby Lab (2021)	English, Bengali; Cantonese; Dutch; French; German; Greek; Hindi/Urdu; Italian; Mandarin; Polish; Portuguese; Spanish; Welsh	Parents	Yes

Bayley Scales of Infant and Toddler Development –Third Edition (BSID-III)	Productive and receptive language	Goh et al. (2017)	English, Mandarin, Bahasa Melayu, Tamil	Trained examiners	Yes
Researcher developed screening tool	Productive and receptive vocabulary	Nayeb et al. (2021)	Swedish, Somali, Arabic, Kurmanji, Farsi, Sorani, Turkish	Child health nurses and bilingual pre-school staff/parents	Yes
Dynamic assessment tasks	Productive and receptive language, assessed via word-learning ability and understanding of abstract language	Patterson, Rodríguez and Dale (2020)	English, Spanish	Speech-language pathology graduate student	No
Peabody Picture	Receptive vocabulary	Thordardottir et al. (2006)	English, French	Research assistant	Yes

Vocabulary Test					
Reynell Developmental Language Scales	Receptive grammar	Thordardottir et al. (2006)	English, French	Research assistant	Yes
British Picture Vocabulary Scale III	Receptive vocabulary	Cattani et al. (2014)	English version used with bilingual children	Research assistant	No
Preschool Language Scale 4	Auditory component	Cattani et al. (2014)	English version used with bilingual children	Research assistant	No
Adapted English SETK-2 (Sprachentwicklungstest für zweijährige Kinder-2)	Object naming	Cattani et al. (2014)	English version used with bilingual children	Research assistant	No

3.6 Several studies also included questionnaires about children's exposure to each of their languages, therefore these have been listed separately in Table 3.3.

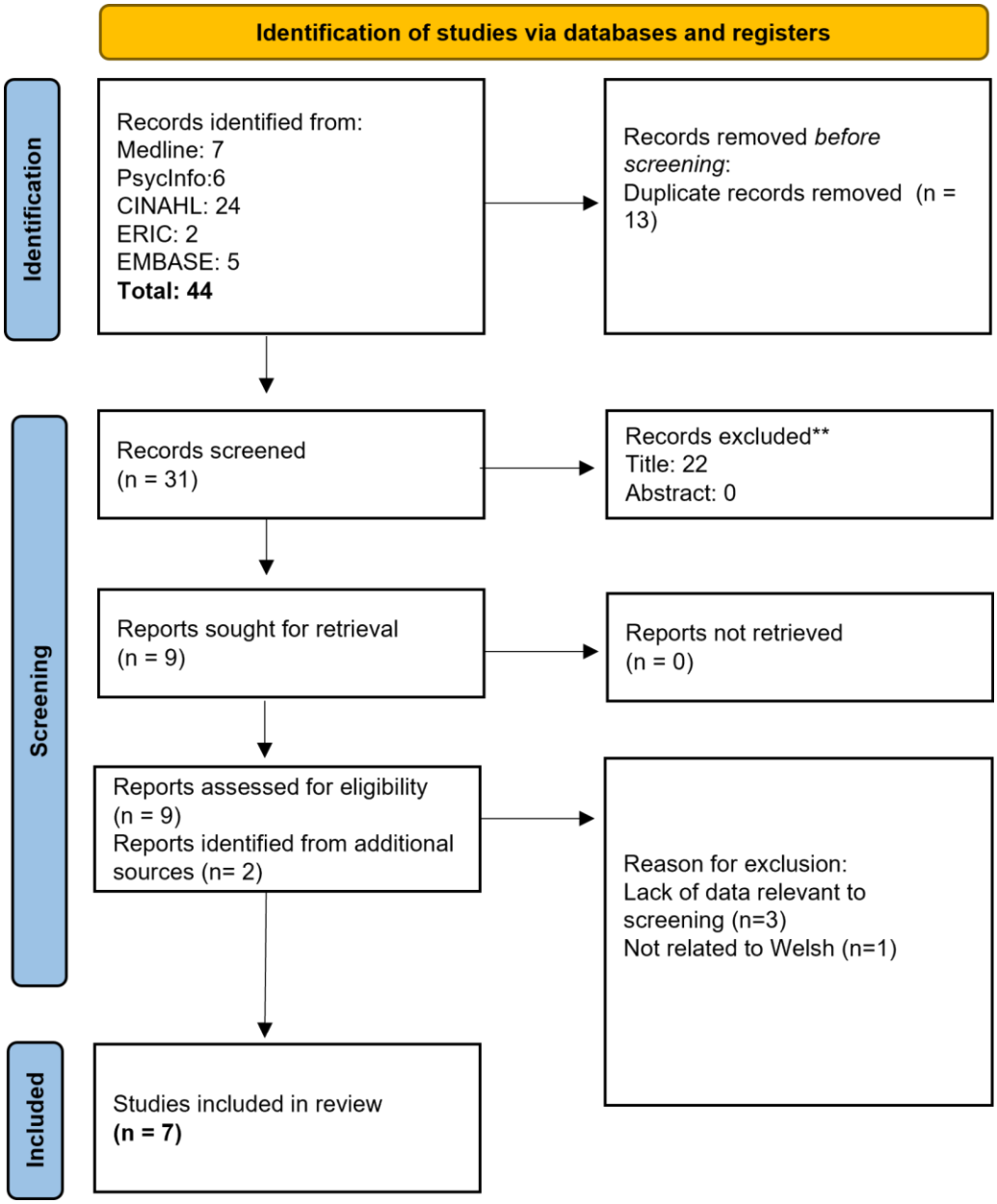
Table 3.3: Language exposure questionnaires used in studies in the bilingualism review

Language exposure questionnaire used	References
Questionnaire for Parents of Bilingual Children: Infants and Toddlers Version (PaBiQ-IT)	Armon-Lotem and Ohana (2017) O'Toole et al. (2016)
Plymouth Language Exposure Questionnaire	Cattani et al. (2014) Floccia et al. (2018)
The Language Exposure Assessment Tool (LEAT)	DeAnda et al. (2016) Legacy et al. (2017)
Family Bilingual Information & Observation (BIO) Questionnaire	Hardin, Scott-Little and Mereoiu (2013)

Welsh Language Review

3.7 For the Welsh language review, searches were first conducted in the English language, as displayed in Figure 3.3. This search resulted in 31 unique records, of which 22 were excluded at title level. In addition to the remaining nine studies that were sourced at fulltext level, two additional studies were identified through searches of reference lists (Borsley & Jones, 2001; Thomas et al., 2014). Four studies were excluded at fulltext level, leaving a total of seven studies identified through the English language searches.

Figure 3.3: PRISMA diagram for Welsh review conducted in English language



3.8 Following this, the search string was translated into Welsh and searches were conducted for a second time in the Welsh language. These searches identified zero records across all five databases. Then the archives of the Welsh-language journal Gwerddon was searched for articles of relevance to children's language development. This search returned a total of three relevant articles which were included in the fulltext Welsh review. Two of these were subsequently excluded as they reported on the use of Welsh within a specific context but did not contain data related to screening or acquisition, leaving a final total of eight studies. None of the studies contained reference to tools used for screening so were not taken further for review.

Discussion

3.9 Statistical results of included studies can be found in Annex C. This section presents a summary of results from the studies identified in the three reviews, before concluding with recommendations arising from the reviews.

English Language Screening Tools Review

3.10 Unlike the three reviews previously undertaken by Berkman et al. (2015), Law et al. (1998) and Nelson et al. (2006), the current work was not intended as a systematic review of research into the value of screening and intervention for speech and language delays in children, or to evaluate the evidence on screening and treating children for speech and language delays or disorders. The aim was to build on and update the identification of potential screening tools appropriate for use with the target population.

3.11 To that end the retention of 11 articles not represented in the three previous systematic reviews, allowed identification of an additional 11 potential screening tools. These were then added to the tools previously listed in the three systematic reviews and other large scale research, to allow for evaluation of their utility for use in the Welsh context.

Bilingualism and Welsh Language Reviews

3.12 The aim of the Bilingualism and Welsh Language reviews was to identify tools that may be appropriate for use in the Welsh context. Although none were identified as a result of the Welsh Language review, two additional tools were incorporated into the screening tool database from the Bilingualism Review. Of the 16 studies identified, only two tools were appropriate for inclusion given their availability for use with Welsh speakers: the UK Bilingual Toddlers Assessment Tool (UKBTAT) (Floccia et al., 2018), and the Welsh-English adaptation of the CDI (Mills, Gathercole & Ebanks, 2013), which was used to gather pilot data for the UKBTAT.

3.13 As well as identifying tools, the Bilingualism and Welsh language reviews also aimed to consider which factors are important in adapting screening practices for a bilingual and Welsh-speaking context. Key findings from studies included in the Bilingualism and Welsh language reviews are therefore summarised in themes below.

Language background and exposure

3.14 Children's language development is inextricably linked to the social and cultural environment in which they grow up (De LamoBeddau1979! White & Jin, 2011). Language exposure has been found to correlate with the size of dual language learner's vocabulary in each of their languages (DeAnda et al., 2016). Therefore, many studies in this review emphasised the importance of considering bilingual children's linguistic backgrounds, and the contexts and settings in which they use each of their languages.

3.15 Armon-Lotem and Ohana's (2017) study of Hebrew-English speaking bilingual children found that frequency of exposure to each language was a major factor contributing to maintenance of the first language and successful acquisition of a second language. Similarly, Floccia et al.'s (2018) findings show that the amount of exposure to each language is a robust predictor of vocabulary. Exposure may be particularly relevant when it is not feasible to assess bilingual children in each of their languages. For example, Cattani et al. (2014) investigated the amount of exposure that is necessary for bilingual

children to achieve similar results to monolingual children when assessed in English only. They found that when children received 60 per cent exposure to English in the home, they could achieve scores on English assessments that were comparable to monolingual norms. If children had less than 60 per cent exposure to English then the authors state that they should be assessed in their additional language in order to fully capture their linguistic abilities. However, it is important to note that this study included only simultaneous bilingual children i.e. children who acquired two languages from birth; frequency of exposure may have a different effect on vocabulary development when children begin to learn English at an older age.

3.16 Given the importance of exposure for bilingual language development, two studies in this review set out to create tools to provide a reliable picture of children's language backgrounds. DeAnda et al. (2016) developed the Language Exposure Assessment Tool (LEAT), validated with children exposed to French and English or Spanish and English. Hardin et al.'s (2013) Family Bilingual Information and Observation (BIO) Questionnaire was developed for use with parents of Latino heritage in the USA. Other tools used by authors in these studies to gather information about language exposure are detailed in Table 7.1. These tools aim to provide detailed and systematic information about children's language backgrounds, including information like communicative partners, language exposure during activities within and outside the home, and achievement of milestones in each of their languages. As Hardin et al. (2013, p. 503) comment, "Looking at a child's language development without an understanding of the child's language history and exposure to language presents an incomplete picture of the child's language development".

3.17 In the Welsh context, studies have shown that differences in patterns of language exposure can affect the acquisition of certain structures. Thomas et al. (2014) investigated the production of plurals in Welsh in children aged between 7 and 11 years and in another study, examined the ability of 4 to 6 year old children to mark gender in nouns in Welsh (Thomas, 2007). Both studies provide support for the notion that the specific patterns of input within

a child's life should be taken into account when determining the milestones for language development.

Standardisation

3.18 There is a risk of assessment bias when using standardised assessments for bilingual children, which may lead to over- or under-identification of language difficulties (Roseberry-McKibbin, 2021). De Lamo White and Jin (2011) report that this may be due to content bias, where speakers perform poorly on a test due to unfamiliarity with the culturally bound content of test items; or linguistic bias, where the assessor is unfamiliar with normative linguistic variation in the speaker's culture and therefore perceives a linguistic difference as a feature of impairment.

3.19 The study by Goh et al. (2017) set out to investigate potential test bias when tools standardised in monolingual environments are used in multilingual settings. They used differential item functioning to assess the Bayley Scales of Infant and Toddler Development (BSID-III; Bayley, 2006), and found possible bias either favouring bilingual children or putting them at a disadvantage in five of 16 test items.

3.20 Even when individual tools that are appropriate to each of the child's languages are used, there is the potential for issues with bias and standardisation in scoring. For example, Cattani et al. (2014) and O'Toole et al. (2016) both used adaptations of the CDI to assess children's development across a range of language pairs. They found that the total number of words tested across adaptations of the tool was vastly different. Although scores in each language can be adjusted to account for this, O'Toole et al. (2016) found a significant correlation between the number of words on the checklist and the number of words used by children ($p<.001$), suggesting that the availability of words provided to parents may influence how they report on their child's vocabulary development. Finally, Core et al. (2013) note that when a tool is translated or adapted, there is potential for overlap in the lexical items that appear in each tool used to assess the child's languages. If the tool is a direct translation, the child's maximum

score will be the same as the number of items they can produce in just one of their languages. This means that the composition of the assessment tools can influence the degree of overlap in the results, and therefore the extent to which it truly represents the bilingual child's linguistic abilities.

Screening in both languages

3.21 Studies in the review emphasised the importance of assessing bilingual children's language development in each of their languages in order to get a full picture of their abilities. They also highlight the inadequacy of monolingual normative data as a comparison point for bilingual language development. For example, Thordardottir et al. (2006) compared the vocabulary development of French-English bilingual children and monolingual children in each respective language. They found that bilingual children scored significantly lower on their English scores than monolingual English children across measures of vocabulary and syntax ($p=0.000$). The bilingual children's scores in French were also lower than monolingual French scores, although this difference did not reach significance. These results illustrate the issues that may arise in interpreting scores when bilingual children are assessed in just one of their languages.

3.22 The findings of Armon-Lotem and Ohana (2017) in turn highlight the benefit of normative data specifically for bilingual children. They assessed the language development of bilingual English-Hebrew children who were relatively homogeneous in language background and demographics. The authors identified two children within their sample who were at risk of a language disorder, as they scored more than one standard deviation below the mean in vocabulary production in both languages when compared with the bilingual normative data. This finding demonstrates the importance of data which is normed on bilingual populations when assessing children who are acquiring more than one language. For the Welsh context, a useful example of how normative data may be standardised is provided by the Prawf Geirfa Cymraeg, a receptive vocabulary test for 7-11 year olds (Gathercole & Thomas 2007b, 2009). Their data reflect different norms for

the varied linguistic contexts in Wales: homes where only English is spoken, homes where only Welsh is spoken, and homes where both English and Welsh are spoken.

3.23 Further support for screening in both languages is provided by Nayeb et al. (2021). They directly evaluated the validity of screening approaches for bilingual children, using four models: 1) screening in mother tongue (defined by the authors as the child's first language) only, 2) screening in Swedish only, 3) screening in both languages and combining scores, 4) screening in Swedish and parent report of mother tongue development. Children were aged 29-33 months, had varied home languages and were all acquiring Swedish in the community as a second language. Results showed that screening in just one of the child's languages led to low specificity, resulting in a high number of children being incorrectly identified as having language difficulties. The most effective method was found to be model three, direct screening in both of the child's languages, which resulted in sensitivity of 0.83 and specificity of 0.86.

3.24 The specific language combinations that children are exposed to has also been shown to be important. Floccia et al. (2018) found that the linguistic distance between English and the additional language of bilingual children living in the UK was a predictive factor in their vocabulary outcomes. For example, children had more advanced productive vocabularies in their additional language if the sound system or phonology of that language was similar to English, while receptive vocabulary was higher if the second language had the same word order typology as English. Similarly, O'Toole et al. (2016) argue that it is important to consider factors specific to the developmental trajectories of languages that may contribute to differential language acquisition. An example given is that two languages represented in their sample were Maltese and Polish, both of which have complex inflectional systems, which may lead to smaller vocabularies in the early stages of language development.

3.25 Finally, O'Toole et al. (2016) also emphasised the need to consider additional 'bilingual factors' which are specific to children learning more than one language. They define these factors as variables such as language exposure and contextual opportunities to use each language, but also wider societal factors, such as the prestige value attached to certain languages and perceived societal value. For example, Armon-Lotem and Ohana (2017) found that in a society where two languages (in this case Hebrew and English) are widely valued, spoken and supported at a governmental level, children are more likely to achieve relatively balanced bilingualism. This is highly relevant to Wales, where Welsh is an official language and should therefore not be treated less favourably than English. This may lead to differences in language profiles of children being brought up in communities where both Welsh and English are frequently spoken compared with children from families who have moved into Wales speaking minority languages (i.e. a language spoken by a minority of people within Wales). It is worth noting, however, that the only study in this review that assessed Welsh language development found no statistically significant difference in the vocabulary sizes of Welsh-English bilingual children when compared to children in the UK acquiring minority languages alongside English (Floccia et al., 2018).

Scoring

3.26 How to score screening assessments across two languages is also important to consider with this population. There is debate about whether bilingual vocabulary assessments should be counted according to 'total vocabulary' (TV) or 'total conceptual vocabulary' (TCV). TV is the sum of all words that are produced by a bilingual child in both of their languages, while TCV counts each concept a child produces rather than each individual word, so that words that are translation equivalents are only counted once e.g. the Welsh word 'ci' and the English word 'dog' would be counted twice as TV and once as TCV (De Houwer, Bornstein & Putnick, 2012). Calculating TV scores only might suggest that bilingual children know a larger number of lexical meanings than they actually do; however, TCV assumes that one lexical meaning has an equivalent translation in two languages, when in fact

one word may represent a range of concepts in different languages (Core et al., 2013).

3.27 Across the studies reviewed here, researchers found that bilingual children had statistically significantly larger total vocabulary scores than their total conceptual vocabularies (Core et al., 2013; Thordardottir et al., 2006; O'Toole et al., 2016). In fact, De Houwer, Bornstein and Putnick (2012) found that bilingual children understood more words than monolingual children at 13 months; however, when translation equivalents were removed from the analysis, this advantage disappeared. The mean proportion of translation equivalents in bilingual children's vocabulary at age 16 months as 48.53 per cent in Legacy et al.'s study (2017) and the number of equivalents increased with age. These findings suggest that interpreting bilingual children's production and comprehension of concepts rather than words could provide more accurate screening results.

3.28 However, it is incorrect to assume that comprehension of concepts can be compared with monolingual normative data. Core et al. (2013) found that when bilingual children's TCV scores were compared with monolingual norms, more bilingual children fell into the 'at-risk' range than when TV was used. Similarly, Thordardottir et al. (2006) found that the TCV scores of bilingual French-English speaking children failed to reach the expected range when compared to monolingual data.

3.29 An alternative approach to scoring screening tests for bilingual children could be the use of dynamic assessment. Dynamic assessment approaches reduce the risk of cultural and linguistic bias as they tap into children's learning potential, rather than their pre-existing knowledge (Lazewnik et al., 2018; Wright, Karem et al., 2019). However, given the typical test-teach-retest format, dynamic assessment may be inappropriate to the time constraints of a screening procedure (De Lamo White & Jin, 2011). A graduated prompting approach was used by Patterson, Dale and Rodríguez (2020) which involved demonstration, feedback and modelling, with a mean time of administration of four to six minutes per task. Dynamic language

tasks were administered to children, including asking them to identify similarities between two items by function (e.g., a plastic spoon and fork), and teaching children novel adjectives then asking them to identify which object could be labelled with that adjective. The study found that results from two of the tasks were positively and statistically significantly correlated with children's scores on a direct language assessment ($p<0.05$). They therefore proposed that a shorter adapted dynamic assessment task may be appropriate as part of a non-biased screening assessment.

Listening to parents

3.30 The majority of studies included in this review used a measure that involved asking parents for their report of their child's vocabulary development. Mancilla-Martinez et al. (2016) argue that parent report is a valuable tool for assessment of bilingual children, as it is cost-effective and avoids task and context bias. Gathering information about language development from parents or carers is particularly important when the assessor does not have proficiency in all of the child's languages. Research suggests that the parents of bilingual children with language impairment are more likely than monolingual parents to feel that professionals did not acknowledge or listen to their concerns about their child's speech and language (Crutchley, 2000).

3.31 Studies in this review provided evidence for the validity of parent report for assessing bilingual children's language development. Two studies using versions of the CDI in Spanish and English found significant correlations between children's productive vocabulary as reported by parents and when directly assessed or observed (Mancilla-Martinez et al., 2016; Marchman & Martinez-Sussman, 2002). There were however stronger associations between parent report and productive vocabulary than receptive vocabulary, suggesting that parents may be better able to report on the words their child actually uses than those they only understand (Mancilla-Martinez et al., 2016). Legacy et al. (2017) also found that parents tended to over-report their child's vocabulary comprehension, when scores were compared to direct assessment of comprehension. In addition, Nayeb et al. (2021)

assessed sensitivity and specificity of a model combining direct assessment in Swedish with parent report of language development in the mother tongue; results showed high specificity but low sensitivity, suggesting that using this method alone may lead to under-identification of children with speech, language and communication needs.

3.32 Given that parents may have different levels of proficiency in the languages their child uses, in some cases it may be necessary for different informants to report on the child's development in each of their languages. To assess the validity of this approach, Marchman and Martinez-Sussman (2002) compared the results of CDI forms completed by just one individual, and those completed by multiple caregivers. They found that correlations between reported and observed vocabulary remained consistent when completed by single or multiple reporters. Vagh, Pan and Mancilla-Martinez (2009) assessed the validity of reports on English-Spanish bilingual children's vocabulary development when completed by parents, teachers, and a composite of both. They found a statistically significant association between children's scores on direct assessments and parent reports, but a non-significant association with teacher reports. There was insufficient evidence in their study to support assessment via composite parent-teacher reports, as parent report alone provided a better fit with predictive statistical models. This was the case even where parents were non-native speakers of English and had varying levels of exposure to their child's daily use of English.

3.33 As well as asking parents to report on vocabulary, studies also requested information from parents about their own concerns around their child's speech and language. Armon-Lotem and Ohana (2017) found that combining vocabulary data from the CDI with parental report of concern was a reliable method of identifying language delay in bilingual children. In addition, O'Toole et al. (2016) compared the incidence of risk factors in the children with results at the 10th and 90th percentile. They found that in the group of children who were in the 10th percentile and therefore at risk of language impairment, there was a high rate of parental concern. Asking

about parental concern may therefore provide reliable information to help identify children in need of further support. In comparison, the incidence of other potential risk factors like socio-economic status and family history of speech and language difficulties varied across each group, providing inconclusive data.

Conclusion

3.34 In this chapter the results of the three rapid scoping reviews conducted are presented. Comprehensive literature searches were conducted with the aim of identifying tools appropriate for screening preschool-aged children living in Wales and providing recommendations of factors to consider in screening for identification of speech, language and communication needs for children across Wales.

3.35 Through the English Language Screening Tools review, the researchers built on the evidence base of previous systematic reviews and identified an additional 11 relevant studies published since the most recent review in 2015. The tools identified were incorporated into the screening database and their evaluation is discussed in the following chapters.

3.36 Findings from the Bilingualism and Welsh Language Reviews enabled identification of two new tools which may be appropriate for use within a Welsh context. The results of studies included in these two reviews were summarised and analysed thematically to produce recommendations detailed in chapter 7.

4. Findings - Protocols for screening of early language development

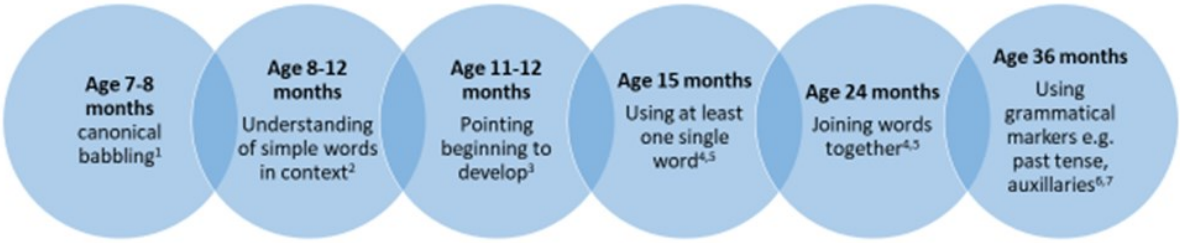
4.1 This section reviews the process of screening and considers what research tells us is important for accurate identification of speech, language and communication needs in young children. The following discussion includes a review of what specific speech and/or language skills should be screened, when screening for these skills should take place to have the most impact, and who is best placed to screen children's developing language.

What should be screened?

4.2 With knowledge of the risk and protective factors for language development, it is possible to identify particular skills in a child's development that should be screened for identification of potential ongoing concerns. As discussed in the introduction, this includes use of babble, use of gesture (including pointing), understanding of language, understanding and use of vocabulary, and use of grammatical markers.

4.3 There is evidence of when children typically develop these skills that can be used to guide when they should be screened. Whilst this evidence is available it is important to remember that this should be used for guidance only and should not be used as an absolute. Children will have natural fluctuations in development of all skills. A summary of which speech, language and communication skills should be screened at different ages is depicted in Figure 4.1.

Figure 4.1: Summary of milestones in early language development



Key: [1]Morgan & Wren (2018); [2]Sheridan, Sharma & Cockeril (2014); [3]Colonessi et al. (2010); [4]Dale et al. (1989); [5]Mayor & Plunkett (2010); [6]Hadley & Holt (2006); [7]Hadley & Short (2005).

Babble development

4.4 The systematic review by Morgan and Wren (2018) reported that canonical babbling is present in infants at around age 7 to 8 months and is not found in infants older than age 10 months including those who were born preterm, with bilingual backgrounds or from households with a low socio-economic status. The confidence and fluidity of babble increases with age with sounds such as ‘b, d, m, n and w’ being the most frequently occurring consonants reported close to the onset of babbling. In particular, sounds made using the lips or the front of the tongue are first to appear with sounds using the back of the tongue and fricatives such as ‘f’ and ‘s’ occurring later in development.

4.5 As discussed previously, there is evidence that lack of babble can be indicative of children who will have persisting language difficulties (Morgan and Wren, 2018). Therefore, identifying the presence of babbling in babies aged 7 to 8 months can be a useful indicator for screening for later language difficulties.

Comprehension of language

4.6 Children’s comprehension (understanding) of simple words in context has been described by Sheridan, Sharma and Cockerill (2014) as starting to develop around the ages of 8 to 12 months. Fernald, Perfors and Marchman (2006) reported that children who demonstrated quicker and more accurate language comprehension skills at age 25 months developed accelerated

expressive vocabulary between the ages of 2 to 3 years. Receptive language has also been shown to have an association with later language scores. Children with poorer scores at age 18 to 20 months are more likely to meet the criteria for Developmental Language Disorder at age 4 (Bishop et al., 2012). Therefore, screening for receptive language from ages 15 months to 25 months can be used for identification of children with potential for later expressive language and/or vocabulary difficulties.

Use of gesture

4.7 Pointing has been reported as developing, on average, at age 11 to 12 months (Colonessi, et al., 2010). The meta-analysis by Colonessi et al. (2010) draws together all of the research and found a strong effect for pointing in relation to later language development, indicating that this form of communication supports language development. While the relationship between pointing and language development was evident from age 12 months, the strongest evidence was from when children were aged 15 to 20 months. Therefore, screening whether children are pointing at this age can be used as a potential clinical marker for language difficulties.

Vocabulary

4.8 It has been reported that delayed or limited vocabulary is one potential risk factor for later language difficulties (Dale, 2003; Rice, 2013; Rowe, Raudenbush & Goldin-Meadow, 2012). However, children's development of vocabulary is complex and extremely variable. Bates et al. (1994) reported on the vocabulary of 1,803 infants between the ages of 8 months and 2 years 6 months using the MacArthur Bates CDI. This tool has been reported as having reliable correlation with direct observation of children's vocabulary (Dale et al., 1989). The study reported that between the ages of 8 months and 11 months the majority of children were not using any words. After this age, the number of words used increased greatly as did the variability in children's vocabulary use, with the range of words used varying most at age 1 years 4 months to 2 years. The median and range of reported number of words used at varying ages in this study are detailed in Table 4.1.

Table 4.1: Measure of vocabulary use in children aged 8 months to 2 years 6 months as reported by Dale et al. (1989)

Age range	Median number of words used	Range of number of words used
0 years 8 months – 0 years 11 months	0	0 – 3
1 year 0 months	6	0 – 52
1 year 4 months	40	0 – 347
1 year 8 months	170	3 – 544
2 years 0 months	311	not reported
2 years 6 months	574	208 – 675

4.9 Estimates of expressive vocabulary sizes have also been reported by Mayor and Plunkett (2010), using data from 20 years of research using the MacArthur Bates CDI. They report that at age 16 months children will be using a mean of 73.8 words and at 2 years of age this has increased to a mean of 599 words. By age 2 years 6 months the mean number of words used is reported as being 1,313.6.

4.10 These two studies demonstrate the complexity and variance in the development of vocabulary, which makes it extremely difficult to recommend a definitive age for screening of this skill, and to identify at which point a delay in vocabulary should warrant any further investigation or intervention. Arguably, a child using no words expressively at age 1 year 8 months would indicate some level of clinical concern and it would be appropriate to monitor their vocabulary development for a period of time.

4.11 Studies of children with DLD indicate that a gap in vocabulary size is consistent over time. Children with DLD will use fewer words at a young age

and attain vocabulary at the same rate as their peers but the gap will never close (Rice, 2012; 2013). In the same studies, mean length of utterance (MLU) is reported to be an unreliable measure of persistent language difficulties, as children with DLD can achieve the same MLU as their peers but with a delay of approximately two years. Monitoring the vocabulary acquisition of children initially identified as being delayed may provide some indication of whether they are catching up with their peers or maintaining a gap, which could be indicative of ongoing language difficulties.

4.12 Late emergence of word combinations has been reported to be a reliable indicator of children who would later be diagnosed with DLD (Rudolph & Leonard, 2016). This was more reliable than delayed vocabulary indicating that this communicative skill may be of benefit in screening for children who are likely to have persistent language needs. The American Academy of Pediatrics' guidelines state that children should be beginning to use some words expressively by age 15 months and they should be combining words by age 24 months (Hagan, Shaw & Duncan, 2008). If these communicative skills are not present at this age, parents and professionals should seek further assessment and advice.

4.13 Monitoring of children's vocabulary development is important for identification of children with persistent language needs. This is an area with much complexity and a one-off screen of this skill may not give enough information to be conclusive. Screening for these markers could be beneficial in accurate identification of children with persistent language difficulties.

Grammatical markers

4.14 The use of grammatical markers in children's expressive language has been described as developing later in children with persistent language difficulties compared to their typically developing peers (Rice, 2013). Hadley and Holt (2006) measured the use of third person singular present tense, past tense, plus copula and auxiliary verbs (e.g., is, are, was). They reported that the use of all of these grammatical markers developed and grew between the

ages of 24 to 36 months and Hadley and Short (2005) reported that in typically developing children, tense morphemes are established by the age of 36 months. Therefore, by the age of 3 years children should be expected to be using a range of grammatical markers in their expressive language. Screening should include examination of children's sentences for the presence or absence of these grammatical markers for identification of children with persistent language needs.

When should screening take place?

4.15 As discussed previously, children may vary in when their speech, language and communication skills develop. Many children may not achieve particular language milestones within the average timeframe but then may make good progress and within a few months have caught up with their peers. In contrast, a child may achieve early milestones such as babbling but then not achieve later skills when expected (Reilly et al., 2010). A child's development will be influenced by biological and genetic factors, their range of experiences and interactions, and how others communicate and interact with them (McKean et al., 2016). For these reasons, monitoring of skills may be more beneficial than a one-off screen for identification of potential speech, language and communication needs.

4.16 Currently in Wales, the Healthy Child Wales programme (Welsh Government, 2016) recommends surveillance of a child's communication skills, by Health Visiting teams at age 15 months and 27 months at a universal level. The aim of this programme is to ensure all children are supported in reaching their developmental milestones and achieve school readiness. At age 15 months, there are some communication developmental milestones that children could be expected to have achieved as detailed above.

4.17 There is evidence from a population-based study, that children who were identified as having a language delay at age 18 months, went on to spontaneously develop their language skills without any intervention (Wake et al., 2011). The conclusion therefore was that this age was too young to

identify children with persistent language difficulties. In this study, children who had no expressive vocabulary at age 18 months were randomly assigned to either an intervention programme or a control group of watch and wait. At age 2 years, both groups of children had similar expressive vocabulary skills with no statistically significant differences between the groups. Expressive vocabulary was the only risk factor considered and it may be that at this age, this alone is insufficient to make a decision about which children are at risk of ongoing delay and that other early communication milestones also need to be taken into consideration.

4.18 A study by McKean et al. (2016) reported that it is possible to identify children who are at risk of persistent language difficulties at age 12 months from parental report of early communication skills. When this information was coupled with knowledge of other risk factors, such as family characteristics and communicative behaviours of the child's parents, the reliability of diagnostic accuracy increased. This was as reliable as parent report alone at age 2 years. Moreover, monitoring children's speech, language and communication needs from age 12 months allows for the provision of early advice and guidance to parents.

4.19 While it may not always be possible to screen for language difficulties at age 12 months, there is some evidence that intervention for language delay is effective for children whose needs have been identified by age 2. Buschmann et al. (2009) provided parent-led intervention to 2-year-old children who were identified as having language delay. Those in the intervention group achieved significantly improved language outcomes at age 3 compared to the control group.

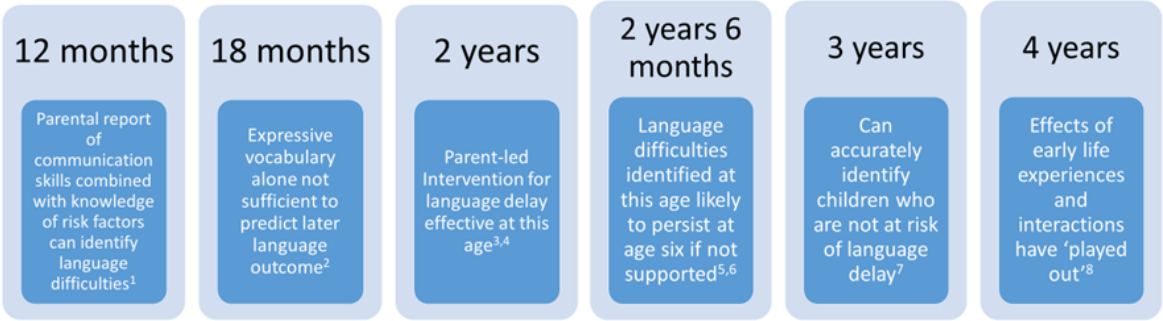
4.20 Screening at age 2 has also been shown to be predictive of children at risk for persistent difficulties. In a study by Miniscalco, Westerlund and Lohmander (2005), children who were identified as having language delay at age 2 years 6 months continued to present with language difficulties at 6 years old. They reported that this was statistically significant in comparison to a control group of children who had developmentally normal language

skills at age 2 years 6 months. Ghassabian et al. (2014) identified that when children were identified with delayed language comprehension at age 2 years 6 months they were more likely to continue to present with a delay in comprehension of language at 6 years old than those who were reported to have delayed language comprehension at age 18 months. Therefore, there is some evidence that screening around age 2 years 6 months is reliable and can have an impact on outcomes for children. The Healthy Child Wales programme of surveillance with screening at age 27 months is in line with this recommendation.

4.21 By 3 years of age, the reliability with which we can predict that children with no identified language needs are not at risk of persistent language difficulties increases. However, it is still not possible to accurately predict all of the children who are at risk (Law et al., 2012). The reliability of using potential risk factors alone as predictors for persistent language difficulties under the age of 4 years is still debatable. This is due to the wide variation in children's development and our knowledge that some children show signs of language difficulties that are not evident at a young age (Asmussen et al., 2018).

4.22 However, there is an argument that waiting until 4 years of age to provide any intervention for those potentially at risk of persistent speech, language and communication needs may be too late and could impact on education. A longitudinal study of children's language development highlights that by 4 years of age, the impact of early life experiences and interactions have peaked. It is still possible to improve language outcomes, for example by increasing the frequency of how often children are read to, but the effects of this lessen after the age of 4 years (McKean et al., 2017). Therefore, identification of children before the age of 4 years is extremely important in relation to receiving timely intervention. Moreover, the provision of preventative approaches has been shown to have benefits (McKean, 2016) and early identification promotes and allows for this evidence-based, effective, early intervention. A summary of the evidence relating to screening for early speech, language and communication needs is displayed in Figure 4.2.

Figure 4.2: A summary of key knowledge relevant to screening for early language difficulties



Key: [1] McKean et al. (2016); [2] Wake et al. (2011); [3] Buschmann et al. (2009); [4]Girolametto et al. (1995) ; [5] Miniscalco, Westerlund & Lohmander (2005); [6]Ghassabian et al. (2014); [7] Law et al. (2012); [8] McKean et al. (2017)

Who should carry out screening?

4.23 There is evidence that parents are able to reliably report on their child's language and communication skills and that this information may be as sensitive as formal screening tests (Law et al., 1998). The study by Laing et al. (2002) indicated that formal screening of children's language and communication needs by a Health Visitor had the same level of sensitivity as that of parents reporting their child's level of communicative functioning.

4.24 Parent reporting of language at age 2 years onwards using the MacArthur Bates CDI has been reported as being a reliable measure when other language measures have been used as a reference tool (Feldman et al., 2000; Feldman et al., 2005). However, at age 12 months to 1 years 7 months there is evidence that there is overreporting of children's language skills in families with a lower socio-economic status (Feldman et al., 2000; Reese & Read, 2000).

4.25 In their study on the efficacy of the Language Development Survey screening measure, Klee et al. (1998) demonstrated that parental report of expressive vocabulary correlated with formal expressive language measures meaning that parents can accurately identify the number of words that their child is using. For children who were not identified as having language delay from this screening tool there was an over-reporting of the number of words

used when compared with formal language assessment. However, there is evidence from formal assessment of speech and language skills, that performance can be affected by many factors including a child's personality, their previous experience of testing, attention skills and cultural differences (Camilleri & Law, 2007). Children may not always perform to the best of their abilities in structured settings or in the presence of unfamiliar adults.

4.26 As reported by Dockrell (2001), parents are most reliable in their reporting of their child's communication skills if the information being reported is current and not relying on memory of when they achieved a particular skill. Therefore, the communication skill they are reporting on should be emergent. Dale et al. (1989) also reported that parents can accurately report on their child's communication skills but rather than asking them to think of any words that their child is using it is more reliable to present them with a vocabulary list and ask them to identify which words are being used.

4.27 Parents may not seek support for their children even when there is parental concern that the child may have speech, language and communication needs (Skeat et al., 2010). A qualitative study by Glogowska and Campbell (2004) reported that parents often compare their child's development to how their siblings developed or how other children they know are presenting. Parents reported that they felt that they should give their child time to develop spontaneously and watch how they progress (Glogowska and Campbell, 2010). In Skeat et al.'s study (2010), only 40 per cent of children who were noted to have low vocabulary scores at age 2 presented for help by the age of 3. Therefore, guidance from a professional who has experience of speech, language and communication development may aid parents in the screening process. A collaborative approach is recommended where parents should report on their child's level of functioning, alongside support from professionals with knowledge of speech, language and communication difficulties, to determine whether any support is required for the child and/or family.

4.28 For many screening tools, a health or educational professional, for example a Health Care Support Worker or early language practitioner, is required to complete the tool. This can either be through prompting and encouraging or direct observation of a skill. It is important that time is taken to ensure the child is relaxed and performing to the best of their ability or an underestimation of their performance may occur. On some occasions it may be necessary to abandon the screen and complete it on another day, for example if the child is tired or unwell. The person responsible for screening needs some level of understanding of the tool and knowledge of how to interact with children to be able to administer it sensitively and accurately.

4.29 In their report on the development of ELIM as a screening tool for use at age 2 years 6 months in England, Law et al. (2020) found that the highest levels of sensitivity and specificity were achieved through a combination of a list of words produced as reported by parents, combined with professional observation and report. In qualitative interviews, parents discussed the importance of their own role in reporting on their child's development, while at the same time they valued the opportunity for a professional to observe their child and respond to any concerns. This report found that for a brief screening measure to be effective, it must be underpinned by a collaborative relationship and shared decision-making between parent and practitioner.

Approaches to screening

4.30 It has been suggested that screening may be enhanced by several factors, which will be discussed in more detail below. There is currently no evidence regarding which of these approaches is likely to be most successful in the identification of children who will have persistent speech, language and communication needs. It has been recommended that further research is undertaken in the early years in developing and monitoring pathways for surveillance to provide the robust level of evidence required to provide evidence-based practice (McKean et al., 2017). Similarly, Rudolph & Leonard (2016) conclude that prospective research must focus on identification of the features of language development in toddlers who are

later diagnosed with DLD to aid understanding of who is at risk. The COVID-19 pandemic has resulted in services being delivered in alternative ways and consideration can be made as to whether these approaches could be delivered in a virtual or non-face-to-face setting.

Using known risk factors for language difficulties

4.31 One approach, suggested by Law et al. (2012) is to use knowledge of how a child's language is developing against expected norms and combining this with the risk factors known to impact on long term language outcomes. This is also advocated by Levickis et al. (2014) who suggest that using a language screen alongside known risk factors such as a measure of parent/child interactions may be beneficial in identifying which children would benefit from speech and language therapy intervention at a targeted level. They suggest a play session would be sufficient to identify whether parents are using responsive communicative behaviours. However, there is a question over how acceptable this would be to parents. In Law et al.'s (2020) development of the ELIM, the pilot tool incorporated a section requesting information on familial and social risk variables. This section was removed from later iterations of the tool as these variables were found to have the lowest levels of sensitivity and specificity in predicting risk for speech, language and communication needs. Moreover, concerns were expressed by both parents and practitioners that some of the questions about risk factors were intrusive in the context of a screening questionnaire. These findings therefore highlight the importance of considering acceptability for all stakeholders when deciding what should be screened and by whom.

Use of ongoing monitoring rather than a single screen

4.32 A one-off screen could be insufficient whereas ongoing monitoring of skills may be more beneficial and lead to an increase in the early identification of children who are at risk of persistent language difficulties (Law et al., 2012). This is mirrored in general developmental assessment where increased accuracy in identification of cases may be related to repetition of the test after an agreed time interval and use of an alternative screen (Bellman,

Byrne & Sege, 2013). Klee et al. (1998) discuss re-screening children where there are concerns about language development 3 to 6 months later. They suggest that this would improve the predictive validity of a screening tool and would highlight children who would be more likely to benefit from intervention. Ongoing monitoring could take place virtually as well as in person. Some children may be more confident in a virtual setting while others may be more inhibited. If only a single screen is conducted it may not be possible to fully understand the child's communicative potential.

Use of dynamic assessment

4.33 Dynamic assessment of children's language skills, in particular their word learning, has been suggested as a method of identifying which children would benefit from additional support with their language (Camilleri & Law, 2014). This approach aims to assess potential for learning rather than a static assessment of skills by prompting, encouraging or cueing within the assessment to determine whether this supports the child's responses (Hasson & Botting, 2010). The person administering this type of screen would need skills in the assessment of children's speech, language and communication needs in order to know how to undertake these supportive mechanisms. This type of assessment has been reported as being fair for children from different language and cultural backgrounds and in being reliable in predicting children who have persistent language needs, regardless of whether they are monolingual or bilingual (Hasson et al., 2013). However, this research was conducted with speech and language therapists conducting the assessments and so consideration would need to be made regarding who could conduct this type of assessment and interpret a child's potential for learning.

4.34 Related to this, it is worth noting initial results from the study by Patterson, Dale and Rodríguez (2020) identified through the Bilingual Review. They suggested that the use of a graduated prompting approach to dynamic assessment enabled a reduced length of assessment and may be a useful procedure for screening in healthcare or early years settings. This approach

may be difficult to deliver in a non face-to-face setting. The practitioner working with the child would need to be able to closely monitor the child's responses and adapt accordingly which would be difficult if working in a virtual manner.

Using a staged approach

4.35 Another suggestion proposed by Law et al. (1998) is to work on the basis of risk and take a staged approach. This would include children who are known to be at risk due to being born prematurely or coming from a low socio-economic background. Law et al. (1998) also describe children who have had adverse childhood experiences within this group. Whilst not all children who are at risk of persistent language needs would fall within this group they would have a greater potential for language difficulties than other children and hence this could represent an efficient use of resources.

4.36 Currently in Wales, children living in Flying Start areas receive this staged approach in relation to speech, language and communication (Welsh Government, 2017). At a universal level, awareness is raised about the importance of speech, language and communication to all families. At the targeted level, children identified as at risk of speech, language and communication needs are offered early, targeted intervention. Finally, children living in Flying Start areas can be supported by specialist services aiming to reduce the negative impact of any identified speech, language and communication need.

4.37 Another example of a staged approach is given by Law et al. (2020) in their development of an intervention co-designed with parents and practitioners. The authors found that developmental reviews provided a unique framework for reaching out to all families without stigma, as these appointments are widely valued and accepted by families. Initial findings from the intervention development show that use of a staged approach focussing only on areas of disadvantage may risk stigmatising and alienating the most affected, while missing the opportunity to reduce effects for all who are negatively impacted to a greater or lesser extent. For delivery to be equitable, the model must be

not just proportionate but tailored, considering the specific barriers and enablers for each individual family.

Focus on prevention rather than screening

4.38 Finally, shifting resources to prevention of language difficulties rather than screening for them has been suggested as a method of improving outcomes for children (Law et al., 2013). A staged approach using universal, targeted and specialised tiers may be of benefit to all children who may potentially develop difficulties with speech, language and communication. At the primary prevention phase (universal level) the aim would be to prevent and reduce the number of children experiencing speech, language and communication needs by implementing a health promotion approach. At the secondary phase (targeted level) children who are at increased risk of speech, language and communication needs due to the presence of some of the risk factors previously discussed would be targeted with the aim of reducing or slowing some of those effects. At the third level (specialised), children who have a clear, identified speech, language and communication need would receive intervention with the aim to minimise any of the negative effects on quality of life. This approach has been adopted by Welsh Government in Flying Start regions across Wales (Welsh Government, 2017).

5. Findings - Identification and critique of existing English and Welsh Language screening tools

5.1 The results presented in this section relate to objective 2 and 3 of the specification. Three systematic reviews of English language screening have previously been undertaken (Berkman et al., 2015; Law et al., 1998; Nelson et al., 2006), all of which report on the screening tools present in literature and proposed for use with preschool children. The tools included in these reviews were extracted and added to a list of tools identified in the Child Talk programme of work (Roulstone et al., 2015). These four sources were then supplemented by tools identified in the first of the rapid scoping reviews reported above (English Language Screening Tools Review). The collation of tools led to a list of 107 English language tools identified that could be potentially used for screening children's language skills.

5.2 The purpose of a screening tool is to accurately and reliably identify children who might benefit from intervention to address speech, language and communication and is not intended to be diagnostic in relation to what specific aspects may require support. In order to be effective as a screen, tools which purport to carry out screening must fulfil certain key criteria. These criteria include measures of validity, reliability, acceptability, impact, cost and diagnostic power and each of these factors need to be reviewed alongside each other to fully evaluate the tools. Applying these measures can be described as using a utility equation (Amin, Seng and Ang, 2006). Further details about this utility equation and the thresholds for measurements of each of the criteria can be found in Annex E.

5.3 Prior to applying the utility equation, 64 of the 107 tools identified were excluded. A summary of the reasons for exclusion is provided in Table 5.1.

5.4 A total of 29 were excluded given their status as diagnostic tools used specifically by Speech and Language Therapists to determine a child's level of communication and to identify particular areas of concern. A further ten were excluded for being assessment tools designed to measure a specific

aspect of language, e.g., Test of Reception of Grammar specifically measures understanding of grammatical constructs.

5.5 Of the remaining 68 tools, three were identified as intervention tools rather than screening tools (Derbyshire Language Scheme, Nuffield Early Language Intervention and Broad Target Recast). Four of the suggested tools rely on the screener's own knowledge to make a decision about the child without any guidance (clinical judgement, health/education professional judgement, parental judgement and Healthy Child Wales Developmental Framework). Three tools were for use with children aged 5 years and above and were specific to literacy (Dynamic Indicators of Basic Early Literacy Skills, Phonological Awareness Literacy Screening and Phonological Knowledge Profile). Two of the tools were considered as monitoring tools only as no guidance is available regarding at what point concerns should be raised about a child's functioning (Early Literacy and Developmental Growth Indicators and Foundation Phase profile).

5.6 Four tools included insufficient information to include them in the review (Parent Language Checklist, Speech and Language Screening Questionnaire, Teaching Talking and Parent Questionnaire with comprehension item). While details were available online regarding Teaching Talking, there was no technical report available and although the manufacturer was approached to provide further detail none was received.

5.7 Finally, nine tools were excluded for other reasons. Information regarding VroegTijdige Onderkenning Ontwikkelingsstoornissen Language Screening Instrument and Sequenced Language Scales for Infants was not available in English and details of the Production of Infant Scale Evaluation was only available in conference proceedings. Focus on the Outcomes of Communication Under Six was designed as an outcome measure, not as a screening tool and similarly Communication Function and Classification System is a tool to describe a child's level of communicative functioning. Early Years Foundation Stage Unique Child Communication Sheet appears to be an informal tool for monitoring a child's progress in England. Law et al.

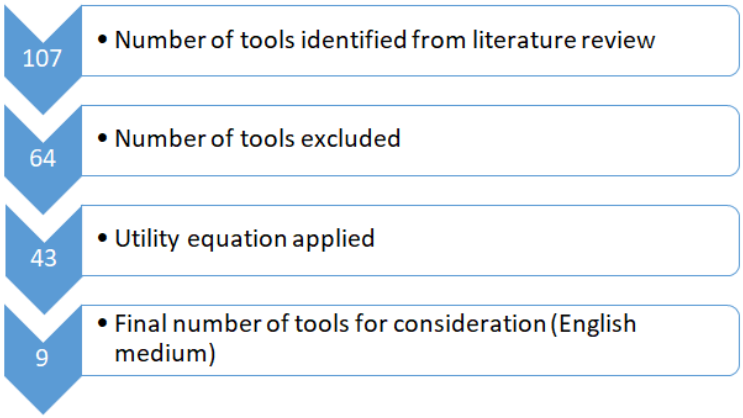
(1998) discuss that generally a screening tool is considered to be appropriate if it can be completed in 30 minutes or less. Therefore, Bayley Scales of Infant and Toddler Development was excluded as the tool takes 60 to 90 minutes to complete in total, which is longer than would be expected for a screening tool. Similarly, the long version of the MacArthur Bates Communicative Developmental Inventories was excluded, and the short form only considered. Finally, Mary Sheridan's milestones were excluded as while they do provide guidance on what a child should be achieving at a particular stage they are written as a reference point to guide judgement and not as a screening tool.

Table 5.1: Summary of the reasons for exclusion of screening tools

Reason for Exclusion	**Number Excluded**
Diagnostic test for Speech and Language Therapy use	29
Specific aspect of speech/language/communication assessed (usually by Speech and Language Therapist)	10
Intervention tool	3
Relies on users' theoretical knowledge to make a judgement	4
Above age threshold	3
Monitoring tool	2
Unable to locate specific details about tool	4
Other	9
Total number excluded	64

5.8 In total, 43 tools remained to be critiqued and the utility equation was applied to each of these to determine their value as a screening tool for early language skills. Tools were excluded at this stage where at least two of the following existed: where tools did not measure the specific aspects of communication identified in the 'what should be screened and when' section; where one or more measurements of reliability and/or validity were below the recommended threshold for scientific rigour; where tools had been designed and used for research purposes but had not been used in clinical settings. Alongside this, tools were only included if they were currently able to be obtained, either through purchase or available with no cost. The application of the utility equation to the excluded tools is detailed in Annex E. This resulted in nine English language tools left for inclusion. A summary of the stages involved in inclusion and exclusion of tools is detailed in Figure 5.1, leading to the final selection and reporting of nine screening tools.

Figure 5.1: Summary of the number of tools at each stage of inclusion/exclusion



Bilingual and Welsh Language screening tools

5.9 Only two tools were identified from the Bilingual and Welsh reviews as potentially useful in screening children who speak Welsh. These were the MacArthur Bates CDI and the UK Bilingual Toddlers Online Assessment Tool (UKBTAT). The MacArthur Bates was considered alongside the English version of the tool and the UKBTAT was considered separately. Given the

small number of tools available from these reviews, the utility equation was not used as a criterion for exclusion but rather for the purpose of describing the tool. The UKBTAT is available in other languages as well as Welsh.

Summary

5.10 Once the utility equation was applied just nine tools were left for final inclusion. There were two tools suitable for use with children from bilingual homes, one of which was also included in the final list of tools available in English. Therefore, ten tools were considered in total and full details of these tools are provided in the following chapter.

6. Conclusions: Identifying which tools to use

6.1 From the initial list of 107 English medium screening tools, a total of nine tools were identified as being the most robust from a scientific and clinical perspective. This was supplemented by a tenth tool, which was identified from the separate review of tools designed for use with bilingual or multilingual children. Each of the tools is described in detail in this chapter in chronological order according to the age at which they are first recommended for use. Specific details of the utility equation considerations are discussed in order to determine the usefulness of the tool.

Screening tools for consideration at age 15 months

6.2 As part of the Healthy Child Wales programme, practitioners from health visiting teams routinely meet with parents and their child at age 15 months. As this is embedded in the workforce activity, it is anticipated that this has a high level of acceptability. There is also evidence that it is possible to screen for many of the aspects of speech, language and communication skills highlighted previously as being indicative for identification of persistent needs at age 15 months. There is limited evidence for the efficacy of screening alone at this age per se, however early identification at a universal level at this age allows for preventative measures to be taken and for further monitoring of skills for those most in need. This is in line with the recommendations of how screening should take place.

6.3 The majority of the tools identified by the review for age 15 months relied exclusively on parent report. Parents are often able to reliably report on their child's language and communication skills (Dale et al., 1989; Dockrell, 2001; Klee et al., 1998; Law et al., 1998; McKean et al., 2016) and so a screening tool based on parent report could be beneficial in highlighting children where there are potential concerns and who may benefit from monitoring of their speech, language and communication skills.

6.4 Completion of a screen by parents would also be cost-effective, requiring minimal additional time from health or early years practitioners who could

focus on discussing the outcome of the screen with the family. However, at age 15 months there is some evidence that there can be overreporting of children's communicative abilities of parents with lower socio-economic status (Feldman et al., 2000; Feldman et al., 2005; Reese & Read, 2000). Parental reporting alone at age 15 months should be interpreted with caution and Law et al. (2020) recommend a collaborative approach with health professionals, as described in their use of the ELIM tool at age 2 years. Alongside the parent reported measures, there were also tools identified which were suitable to be administered by a health or education professional. These tools could be adapted for use to allow for parent reporting of some of the measures.

6.5 The following tools were identified in the review and are suitable for use at age 15 months. They are all reported to be quick to administer, i.e. up to 30 minutes maximum.

Ages and Stages Questionnaire (ASQ-3)

6.6 This tool screens for all aspects of language development at the appropriate development level. It also screens for other developmental milestones including gross motor, fine motor, problem-solving and personal-social. It is widely used and has been validated in large, diverse, standardisation samples. It is suitable for use from age 1 month to 5 years 6 months.

6.7 There are differing reported levels of sensitivity and specificity depending on which assessment is used as the reference tool. When the Preschool Language Scale was used, sensitivity was reported as 50 per cent for comprehension items and 59 per cent for expressive items, below the recommended threshold of 72 per cent (Law et al., 2000), but specificity was higher at 79 per cent for comprehension and 83 per cent for expression (Frisk et al., 2009). When Bayley Scales of Infant Development were used, sensitivity was reported as 75 per cent and specificity as 81 per cent (Frisk et al., 2009). On their website, the authors report the sensitivity as between 75 per cent and 100 per cent with an overall agreement for all of the ages

measured being 86 per cent and the specificity as between 70 per cent and 100 per cent with an overall agreement of 85 per cent.

6.8 The ASQ-3 has a high reliability level of 92 per cent (Frisk et al., 2009) and a positive predictive value of 47 per cent (Schonaut et al., 2013). The authors report a high level of validity of 0.82 to 0.88, test re-test reliability of 0.91 and inter-rater reliability of 0.92.

6.9 The tool has been translated into other languages and although Welsh is not one of these, the authors invite expressions of interest to those who wish to consider translations into other languages. ASQ-3 can be completed on paper forms or there is an option to set up the tool so it can be completed by parents electronically on a mobile, tablet or computer. ASQ-3 is reported to take 10 to 15 minutes to complete. It requires a professional to score the tool and determine if there are areas of concern and this is reported to take approximately 3 to 5 minutes.

6.10 The starter kit is available as a commercial tool and this includes printable PDF questionnaires (this is an American tool hence the price provided in American dollars). Prices for the online version of the ASQ are available on request.

Brigance Preschool Screen

6.11 This tool is widely used in the USA and Canada. There are two separate tools depending on age and so this information refers to the preschool screen, suitable for children aged 0 to 35 months. It is reported to take around 10 to 15 minutes per child to complete the screen and alongside language, other developmental milestones are screened including physical development, academic skills/cognitive development, self-help and social-emotional skills.

6.12 Glascoe (2002) reports the tool as having sensitivity of 76 to 77 per cent and specificity of 85 to 86 per cent using a combination of Bayley Scales and the Infant Behavior Record, the Receptive-Expressive Emergent Language Test, the Preschool Language Scales, Rosetti Infant Toddler Language Scale, the

Vineland Adaptive Behavioral Scale and the Alberta Infant Motor Scales. However, Frisk et al. (2009) report lower sensitivity and specificity when comparing the tool to the Preschool Language Scale alone. Using this as a reference assessment, the tool achieved a sensitivity of 55 per cent for comprehension and 90 per cent for expression and a specificity of 69.6 per cent for comprehension and 75 per cent for expression. The scores for comprehension do not reach the threshold of 72 per cent and 88 per cent as suggested by Law et al. (2000) and only sensitivity for expression achieves this level. Glascoe (2002) reports an excellent reliability level of 0.98-0.99. It is important to note that the study by Glascoe (2002) used an American population and the study by Frisk et al. (2009) used a Canadian population. Caution is required around how replicable these results would be for a UK population, given the finding that children from the UK have been reported to have lower scores for both comprehension and expressive vocabulary than American children of the same age (Hamilton, Plunkett & Schafer, 2000).

6.13 The tool involves asking parents questions about skills a child has mastered as well as direct observations that are made by the professional administering the screen. There is an online system available where data can be entered to produce a report highlighting at what level the child is functioning and whether this falls above or below the ‘cut-off’ for concern.

6.14 The complete screening kit including technical manual is available as a commercial tool. Additional data sheets can be purchased in packs of 60 for data collection. The online management system for scoring and managing data is available to purchase and is charged per child for a one year licence.

LENA Developmental Snapshot

6.15 This tool screens for all aspects of language development from age 2 months to 36 months of age and provides standardised data for children in the United States.

6.16 The authors of the tool detail its reliability in the technical manual which is freely available online. The tool has been developed with scientific rigour however sensitivity and specificity are not reported. The tool has been

correlated with a number of speech and language therapy assessments and the average reported correlation is 0.93 (range 0.81-0.97). This indicates a high level of validity. It is also reported to have a high level of test retest reliability where participants completed the tool on a monthly basis and correlations with previous test scores were reported to be between 0.93-0.98.

6.17 At present the tool is available in English and Spanish. Parents are sent a hyperlink to complete the tool online. A summary of their responses is provided by the tool highlighting whether the child is functioning at a developmentally appropriate level. They also suggest the tool can be used to monitor a child's development. There is an initial set-up fee and then an annual fee. There is also a monthly charge for the first 50 screens completed and an additional fee for each additional 50 screens.

MacArthur Bates Communicative Developmental Inventory

6.18 This tool is used to screen children's language skills by parental reporting of their child's understanding and use of vocabulary and is recommended for use with children aged 8 to 30 months of age. There are both long and short versions of the form - the long version taking around 20 to 40 minutes to complete while the short form takes around 10 minutes. It has been standardised on children in the UK (Hamilton, Plunkett & Schafer, 2000), and a UK version of the form is available to download for free. This study highlighted that American children have higher scores for both comprehension and expressive vocabulary than children of the same age in the UK. The tool has been adapted for use in many languages including Welsh. There are a number of free resources describing how the tool should be administered and a free scoring programme which is available by contacting the authors via their website.

6.19 The short form has been reported as having high reliability of 0.97 to 0.99 and a good level of concurrent validity reported as being 0.74 to 0.93 (Fenson et al., 2000). Skarakis-Doyle, Campbell and Dempsey (2009) report the tool's sensitivity in accurately identifying children with language delay as

88.9 per cent and specificity of 98 per cent. However, it is worth noting that the children in this study were aged 30 to 45 months and so at age 15 months it may not be as reliable. Participants in this study were from Canada and therefore may not be replicable in a UK population.

Schedule of Growing Skills (SOGs) II

6.20 Within Wales, the Schedule of Growing Skills-II (SOGs II) has been used by Health Visitors for a number of years and has been included in the final list of tools exclusively due to its high level of external validity. The tool measures a number of developmental milestones including hearing and language skills, speech and language skills and interactive social skills. It is suitable from birth to 5 years and is reported to be easy and quick to administer.

6.21 The SOGs II was standardised in 1996 on a cohort of 360 children with a concurrent validation study using 11 clinical case studies (GL Assessment, 1996). It is reported as having a high level of internal consistency of 0.91. Williams et al. (2013) reviewed the tool in a small study of children in Wales where 24 out of 35 children were Welsh speakers. The sensitivity of the tool was reported as 0 to 25 per cent, while the specificity was 92 to 100 per cent. This low level of sensitivity resulted in children who were identified as having language difficulties not being identified. The authors proposed using an alternate cutoff for calculating the developmental quotient which increased the sensitivity of the tool to 0.68 to 0.73 and the specificity to 0.80 to 0.91 (Williams et al., 2013).

6.22 The SOGs II is likely to already be available to many Health Visitors meaning no or low cost for it to be used. However, difficulty was encountered in locating data regarding the psychometric properties of the test; and so it is not possible to assess how robust the tool is for accurate identification of children with language difficulties.

WellComm (Early Years)

6.23 Many Health Visitors and early years practitioners in Wales will be familiar with the WellComm (Early Years) screening tool. This tool is suitable for

children aged 6 months to 6 years and is reported to be quick to administer. Once the tool has been completed, results are colour coded to indicate whether further monitoring or assessment is required. There are also recommended activities for intervention to go alongside the tool where this is highlighted as being required.

6.24 The technical information in the manual states that it has a sensitivity of 100 per cent where a child has both comprehension and expressive language difficulties and 88 per cent when the child has just one of these (with the Reynell Developmental Language Scales used as the reference test). The specificity of the tool is reported as 58 per cent which will result in false positives, i.e., children being identified as having difficulties with language when they don't. This could result in children receiving intervention when it is not required and/or over referral to speech and language therapy services. However, these scores are based on assessment and screening of children aged 3 to 6 years and there are no validation measures provided for younger children.

6.25 WellComm (Early Years) is sold by GL Education who can provide a quote online for the tool. There is also an online reporting tool that can be used to keep track of children's scores and progress.

Summary of tools suitable for use at age 15 months

6.26 The three parent-reported screening tools (ASQ-3, LENA Developmental Snapshot and MacArthur Bates CDI) are considered to be the most robust based on our review of 107 tools and the outcome of the utility equation. The MacArthur CDI only measures vocabulary and it has been reported that this alone may not be sufficient to screen for language delay in this age group (Wake et al., 2011). However, it has been included because an adapted Welsh-English bilingual version is available, and it could be used as a tool to monitor vocabulary development over time. It was previously identified as one of only two tools considered to be fit for purpose by Berkman et al. (2015) in their systematic review.

6.27 The other three screening tools described here (Brigance Preschool Screen, SOGs II and WellComm (Early Years) are all designed to be completed by a health/education professional. None of these tools achieve all of the recommended levels of the utility equation and so cannot be recommended over the others.

Screening tools for consideration at age 27 months

6.28 The Healthy Child Wales programme recommends contact with families of children aged 27 months for child surveillance purposes and tools for use at this age are reported in this section. The reliability of screening measures to accurately identify children with persistent speech, language and communication needs increases from this time point onwards.

6.29 At this age, research indicates that parents may still be the most appropriate person to report on their child's level of communication skills. The parent-reported tools described for use at age 15 months could all also be used at this age. The other tools described for use at age 15 months which are to be used by professionals can also still be used at age 27 months. The tools discussed in this section could not be used at an age 15 months screen but are suitable for use with children aged 27 months.

Early Language Identification Measure (ELIM)

6.30 The Early Language Identification Measure (ELIM) has recently been produced through extensive consultation with stakeholders including Health Visitors and parents in England and as a result has a high level of acceptability (Law et al., 2020). It has been designed for use at age 2 years 6 months. There are three steps to the measure: Assessment (word list and observation), Conversation (Further Exploration and Signposting), and Intervention (Offering Tailored Support). The tool is reported to be quick to administer and fosters a supportive discussion between the health professional and the parent regarding the child's language development.

6.31 A large-scale study was undertaken to determine the psychometrics of the tool. The Preschool Language Scale (PLS -5) was used as the reference tool

and the sensitivity is reported to be 0.94 and the specificity 0.65. This tool is still in its infancy in terms of its use in England and further developments have been recommended by the authors to enhance its usability for parents and Health Visitors.

UK Bilingual Assessment Tool (UKBTAT)

6.32 Also for consideration at this age is the UKBTAT which, alongside the MacArthur Bates CDI, is suitable for screening children who are bilingual or multilingual. This tool has been standardised to use with children aged 24 months to 35 months and is designed to be administered by both parents and professionals. It has been standardised in Welsh as well as many other languages, i.e., Bengali, Chinese, Cantonese, Chinese Mandarin, Dutch, European Portuguese, French, German, Greek, Hindi, Urdu, Italian and Polish. The tool can also be used with children who speak other languages though a standardised score is not available for these. It can be used with children who are bilingual or multilingual.

6.33 The tool uses the MacArthur Bates CDI as its basis and invites parents to complete a checklist of words their child understands and uses. A list of words in their child's other language(s) is also completed. Floccia et al. (2018) identified that the strongest predictor of a child's comprehension and use of vocabulary was related to the consistency and proportion of English spoken by parents. Therefore, a practitioner completes a measure of language exposure with the parents and the scores are entered into a computer system for analysis. A percentile score is provided for the child's performance based on the responses to each of the three measures. These percentiles were demonstrated to have good validity with no over- or underprediction of scores reported using predicted and observed data (Floccia et al., 2018).

6.34 The tool is available to clinicians free of charge and they need to register on the website for access. This tool could also be considered as a possibility for use with children who speak languages other than English.

Summary of tools suitable for use at age 27 months

6.35 The tools designed to be administered by professionals should be considered alongside the parent-reported tools. There is no one definitive tool with a track record of robust research that can be recommended for this age group. Figure 6.1 depicts all of the evidence regarding what should be screened, the known risk factors for persistent language difficulties, what tools are suitable for use at either 15 or 27 months and suggested action points. Figure 6.2 depicts tools that can be used between ages 30-36 months as the reliability of screening increases at this age.

Figure 6.1: Summary of evidence and appropriate screening tools (age 15-30 months)

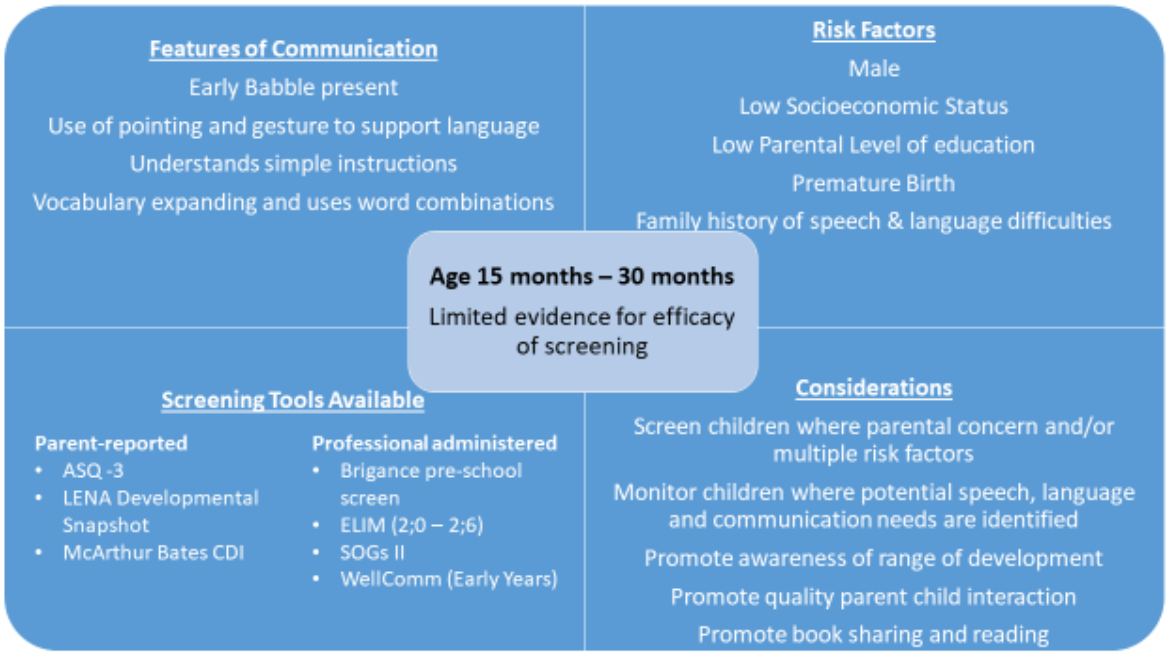
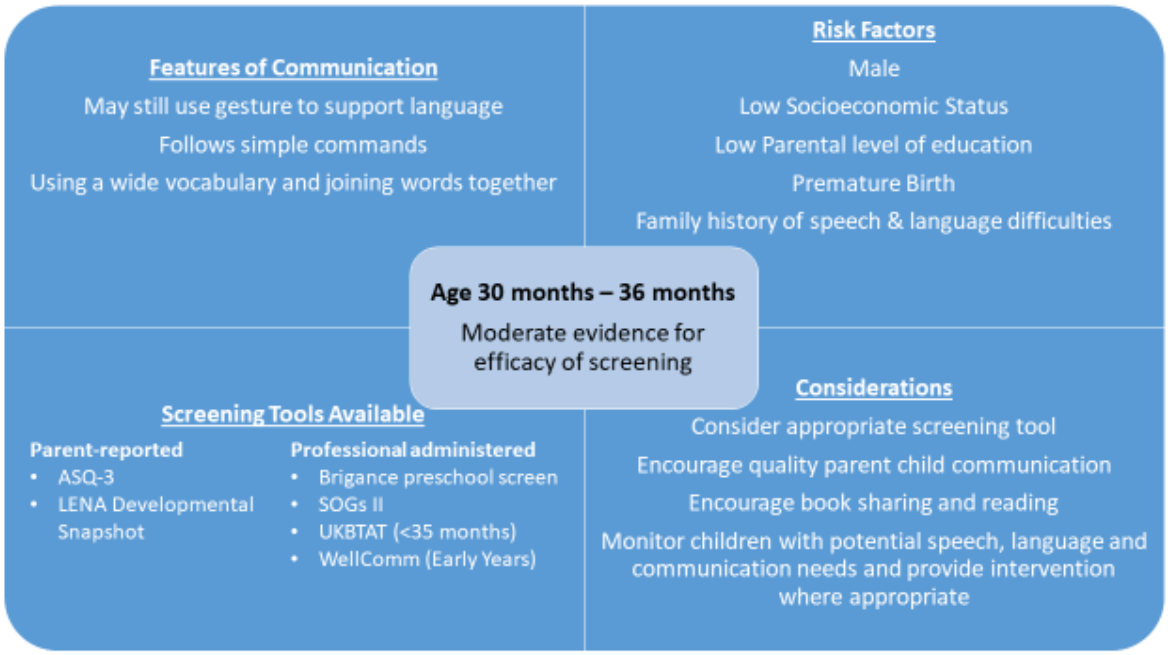


Figure 6.2: Summary of evidence and appropriate screening tools (age 30-36 months)



Screening tools for consideration at age 3 years and above

6.36 At age 3 to 4 years, there is increasing evidence of the reliability of screening for language difficulties. Some of the tools already discussed are still suitable for use with this age (i.e., ASQ-3, Brigance Preschool Screen, SOGs II and WellComm (Early Years)). There are two additional screening tools for consideration that have not previously been discussed.

Nuffield Early Language Intervention (NELI) Language Screen

6.37 This screening tool has been developed alongside the NELI programme. It is reported to take around 10 minutes to complete and can be used to monitor children's language development following intervention with the NELI programme. The tool is an app which early years practitioners use to screen for language difficulties. It is suitable for use with children aged 3 years 6 months to 8 years.

6.38 In a large-scale study it is reported to have excellent reliability of 0.95 (West et al., 2021). There are no reported measures of sensitivity or specificity or other information regarding the psychometric properties of the tool. The app

can be downloaded from the website for a free trial with ongoing costs per year depending on the size of the setting.

Language Link

6.39 This tool is also suitable for practitioners working in education to use as a screening tool and is frequently used across Wales. It is designed for use with children aged 4 to 7 years and is reported to be quick and easy to administer. There is a programme of intervention to go alongside the tool for those children who are identified as being at risk of language difficulties.

6.40 The technical report with the tool reports that it has good internal consistency of 0.90 and test re-test reliability of 0.89. The validity of the tool is reported to be between 0.59 to 0.76 with the lower end of this scale falling outside of the acceptable levels of validity. Language Link has a start-up fee per school with an ongoing subscription cost annually per school.

Summary of tools suitable for use at age 3 years and above

6.41 There are a number of tools that are suitable for this age range, and with the exception of the ASQ-3, they are for use by either health or education professionals. The NELI Language Screen is the only tool where the reported measures reach the threshold of acceptability but they only report on reliability and not sensitivity or specificity. Each of the tools has its own relative strengths and weaknesses and so it is not possible to be conclusive in recommending one specific tool for use at this age range. Evidence for what should be screened, known risk factors, appropriate tools that can be used and suggested action points are summarised in Figure 6.3 for age 3 to 4 years and Figure 6.4 for age 4 to 5 years.

Figure 6.3: Summary of evidence and appropriate screening tools (age 3-4 years)

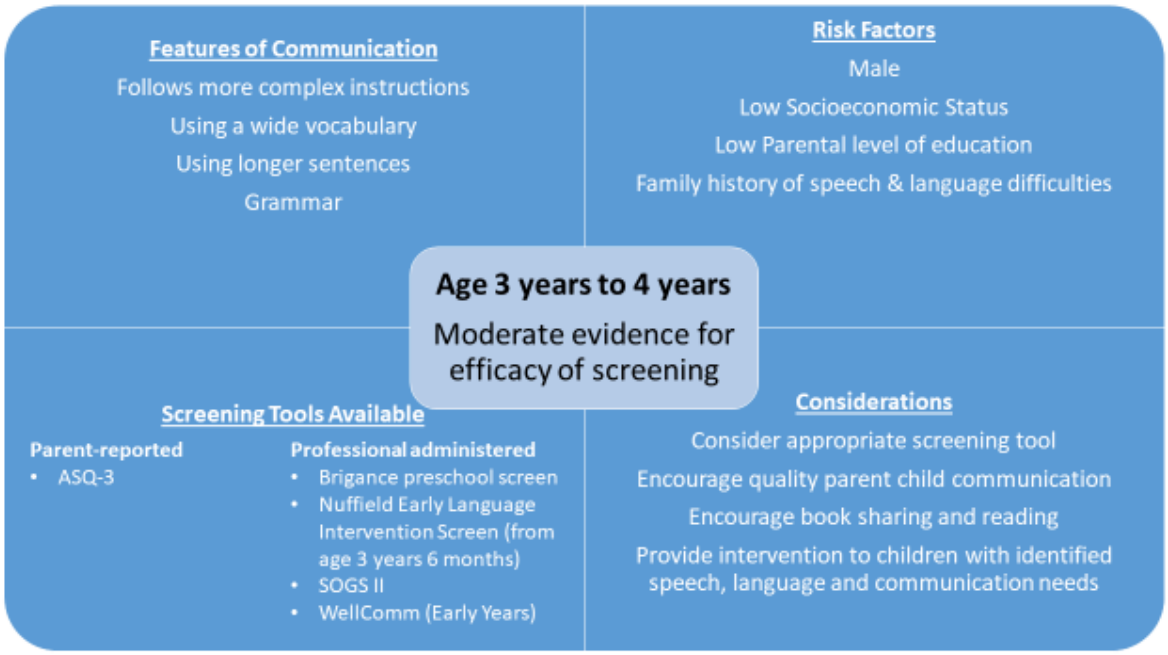
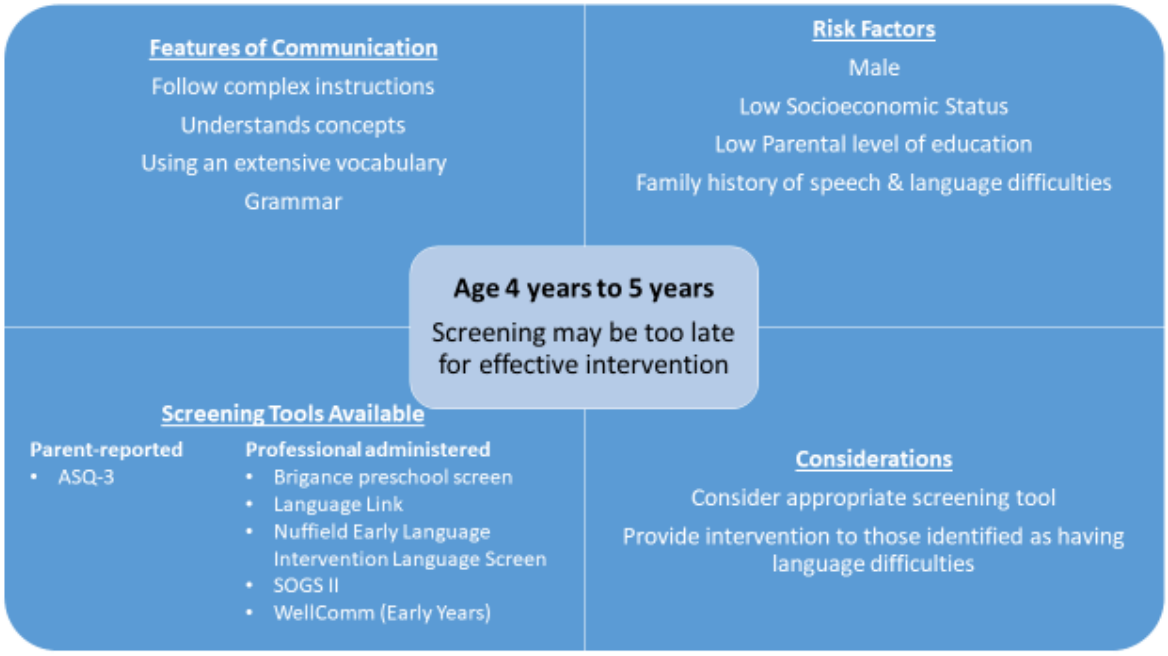


Figure 6.4: Summary of evidence and appropriate screening tools (age 4-5 years)



Summary of Tools

6.42 Following a rigorous, scientific review there is no one screening tool which can be recommended for use with children aged under 5 years in Wales. Each of the tools considered in the final list of those included has its relative merits but none alone reach all thresholds for reliability, validity, sensitivity and specificity, measure all aspects of communication as determined from the review of the literature, or are suitable for all ages. The utility equations for the final nine tools are summarised in Table 6.1. Measures which do not meet recommended thresholds and tools which are only valid for specific age ranges are highlighted in the table.

Table 6.1: Final tools for inclusion

Name of tool	**Impact**	**Reliability**	**Validity**	**Acceptability**	**Feasibility**	**Cost** (all correct at time of publication)
Ages & Stages Questionnaire	Assesses all relevant features of communication at age appropriate time points	Sensitivity: 48*-75% Specificity: 79-83%	Not reported	Used across the UK in many settings. Suitable for ages 1 – 66 months.	Reported to take 10-15 minutes for parents to complete and 2-3 minutes for professional to score. Parent reported.	$295 (starter kit) / £236
Brigance Preschool Screen	Assesses all relevant features of communication at age appropriate time points	Sensitivity: 55.6*-90.9% Specificity: 69.6*-86% Internal consistency 0.94-0.98 Test re-test 0.92-0.99 Interrater reliability 0.93-0.95	Not reported	Used in USA and Canada	Two separate tools age 0 - 35 months and 3-5 years. Reported to take 10-15 minutes per child to administer. Administered by	$309 for 0-35 months kit - £247 $279 for 3-5 yrs kit - £223

					profession al.	
Early Language Identification Measure	Assesses all relevant features of communic ation at age appropriate time points	Sensitivity: 0.94 Specificity: 0.65*	Not reported	Stakehold er consultatio n (parents and Health Visitors) in design	Designed for use at age 2;0-2;6** Estimate of length of time to administer not provided as varies depending on family. For use by health visitors alongside parents.	Available
Language Link	Assesses all relevant features of communic ation at age appropriate time points	Internal consistency: 0.898 Test re-test: 0.89	0.585*-0.759	For age 4-7 years** Widely used in schools across south Wales. Suggested activities to target language skills and monitor developme nt.	Reported to be easy to administer taking up to 25 minutes per child. School staff administer .	£425 per school start up fee and then £275 per school annually

NELI Language Screen	Assesses all relevant features of communication at age appropriate time points	0.95 Sensitivity and Specificity not reported*	Not reported	Newly developed to go alongside Nuffield Early Language Intervention For age 3;6 -8 years**	Easy to administer using an app. Reported to take 10 minutes per child to administer. For school staff to administer.	Annual subscription cost for schools £275-£395 per year
Lena Developmental Snapshot	Assesses all relevant features of communication at age appropriate time points	Test re-test: 0.95-0.98	0.93	For age 2 months-36 months** Can be completed online via an app (internet access required)	Parent reported. Reported to take up to 15 minutes to administer. Goes alongside a parental intervention measure.	$299 start up fee £240, $99 annual fee - £79 Monthly subscription $90 for up to 51 participants then $49 for each additional 50 £72 and £39 respectively.

MacArthur-Bates Communicative Development Inventory (short form)	Assesses vocabulary	0.97-0.99	Concurrent validity: 0.74-0.93	Lots of research and used for over 30 years worldwide. For ages 8 - 30 months**	Reported to take approximately 20 minutes to administer. Parent reported. Welsh adaptation available.	Free to download
Schedule of Growing Skills III	Assesses all relevant features of communication at age appropriate time points	Internal consistency: 0.91 Sensitivity: 0-25% Specificity: 92-100% (where 24/35 children were Welsh speakers). No other measures of sensitivity/sensitivity reported*	Not reported	Widely used across Wales as developmental screen. Suitable for children aged 0- 5 years	Reported to be able to assess, score and provide feedback within an hour. Professionals administer.	Available to purchase. Cost unknown
UK Bilingual Toddler Assessment Tool	Assesses vocabulary using CDI plus Language Exposure Measure	Not reported (other than CDI data) Norms developed for Welsh	Not reported	Suitable for children aged 2 years**	Time to administer not reported. Parent reported Can be completed online Available in many languages	Available free for clinicians

					including Welsh	
WellComm (Early Years)	Assesses all relevant features of communication at age appropriate time points	Sensitivity: 88-100% Specificity: 58.5%*	Not reported	Widely used in Wales	Reported to take 20 minutew to administer. Administered by Early Years Practitioners. Provides activities to target and monitor development.	Available to purchase Cost unknown

*measurement does not meet minimum recommended threshold i.e. Reliability (internal consistency) = 0.6-0.7, Construct Validity = 0.8, Sensitivity = 0.8, Specificity = 0.8

**specified age range for use does not span from age 15 months to 4 years 11 months

7. Recommendations

7.1 The work carried out and reported in this review has identified a number of recommendations with regards to screening children's early language development in Wales. These are addressed in this chapter under the headings of the key objectives for this work.

Should screening take place?

7.2 Until recently, a number of reviews concluded that the evidence base did not support universal screening of early language development to identify children at risk of persistent speech, language and communication needs. It was argued this was due to lack of robust research to determine who will have transient and who will have persistent needs. However, this conclusion appeared to be based on a presumption of a single point screening protocol whereas recent evidence has highlighted the changing nature of language development. Language concerns in infancy and early childhood appear to resolve without intervention for some children, while difficulties emerge in others beyond the very early stages of language development. This has led to the suggestion that screening should take place at multiple times and not as a single event, with an acknowledgement that the outputs from screening appear to be more dependable from age 2 years and 6 months onwards.

7.3 With regards to what should be screened, a variety of risk and protective factors have been identified which could be useful in providing information regarding the likelihood of ongoing problems with speech, language and communication. These include physiological, family and environmental, and communication factors, though not all are of equal value and not all can be easily measured.

7.4 ***Recommendation 1:*** A systemic or preventative approach to screening is more beneficial than a single screen for language difficulties and should be adopted in preference to a single screen at one point in time.

7.5 ***Recommendation 2:*** Key developmental communication milestones can be used in the screening of early language development.

7.6 ***Recommendation 3:*** Known risk factors should be combined with monitoring of language skills and/or dynamic assessment to enhance screening. This would require active child participation, in a modification of an assessment that is both fluid and responsive to the child's ability.

7.7 ***Recommendation 4:*** Parents are able to provide reliable information about their child's current level of communicative functioning from age 2 years. Parents should be involved in discussions about their child's development and any concerns they have about it.This will culminate in shared decision making regarding any further action required.

Summary of English and Welsh language screening/identification tools

7.8 Of the many tools identified in the reviews, just ten have been included for consideration to be used in a programme for screening early language development for children in Wales.

7.9 Table 7.1 provides a summary of the ten tools. Information on whether or not they provide a screen for all aspects of language development is provided together with information on which age range each tool is suitable for. Consideration has been given as to whether the tool could be used without the requirement for face-to-face contact. A literature review was not conducted as there is likely to be limited published data but a practical approach using knowledge of the tool and of working in a virtual way has been applied. Finally, information is included on whether the data available for the screening meets recommended thresholds for robustness.

Table 7.1: Summary of recommended tools

Name of Tool	Screens for all aspects of language development*	Suitable for age 15-26 months	Suitable for age 27 months - 4 years 11 months	Available data meets recommended thresholds**	Available in Welsh	Suitable virtual use	Cost
Ages and Stages Questionnaire (ASQ)	✓	✓	✓	✗	✗	✓	£200-£300
Brigance Preschool screen	✓	✓	✓	✗	✗	✗	£400-£500
Early Language Identification Measure (ELIM)	✓	Suitable for age 24 - 30 months only	Suitable for age 24-30 months only	✗	✗	✓	£0
Language Link	✓	✗	Suitable from age 4 years	✗	✗	✗	£425 initial fee plus £200-300 per annum
Nuffield Early Language Intervention (NELI) Language Screen	✓	✗	Suitable from age 3 years 6 months	✓	✗	✗	£300-400 per annum
Language Environmental Analysis (LENA)	✓	✓	Suitable up to age 36	✓	✗	✓	£200-£300

Developmental Snapshot			months only				start up fee £0-£100 per annum £100-£200 per 100 participants
MacArthur Bates Communication Development Inventory (CDI)	✗	✓	Suitable up to age 30 months only	✓	✓	✓	£0
Schedule of Growing Skills (SOGs)	✓	✓	✓	✗	✗	✓	Unknown
UK Bilingual Toddlers Assessment Tool (UKBTAT)	✗	Age 2 years only	Age 2 years only	✓	✓	✓	£0
WellComm (Early Years)	✓	✓	✓	✗	✗	✗	Unknown

** Screens for every age specific aspect of language considered to be important for identifying risk of persistent language needs*

***Recommended thresholds (as described previously): Reliability (internal consistency) = 0.6-0.7, Construct Validity = 0.8, Sensitivity = 0.8, Specificity = 0.8*

7.10 ***Recommendation 5:*** Only those tools where data are available to make a judgement regarding their robustness should be used.

7.11 ***Recommendation 6:*** Where these data are not available, it is possible to make a plan to acquire the evidence needed to determine suitability. This is vital given the recognition that many tools available or in use do not have the required data but have face validity amongst the workforce.

7.12 ***Recommendation 7:*** While no one tool is currently suitable for use across the age range, consideration should be given to whether there is value in developing such a tool. This would assist in the confidence of practitioners in using the tool and the reliability with which it is administered.

What makes an effective screening/identification tool?

7.13 An effective screening tool is one which is valid, reliable, has impact, is acceptable, is cost effective and has diagnostic power. Validity, reliability and diagnostic power can be measured quantitatively with robust guidance regarding minimal figures. Acceptability is equally important. If the tool selected is unacceptable to either families or health and education professionals, then its use in practice may be reduced and the anticipated gains from a programme of screening will not be realised. The tool or tools must also be cost effective; this does not mean they must be 'cheap'; rather that the costs of implementation are outweighed by the benefits realised. If identification of SLCN is earlier, and intervention leads to fewer children with persistent problems, there are likely to be gains in school attainment, wellbeing and later social and economic independence. In this way, the screening programmes can be seen to have impact.

7.14 ***Recommendation 8:*** Screening tools should be evaluated based on their utility which includes validity, reliability, impact, acceptability, cost and diagnostic power. For reliability, a measure of 0.6-0.7 is considered acceptable with scores of 0.8 and above being good. Sensitivity and specificity scores of 0.8 have been recommended as a measure of a robust tool (Law et al., 1998).

Recommendations on the important components for a Bilingual English and Welsh All Wales early language screening/identification tool

7.15 The reviews have identified a number of important considerations to be included in a plan to develop a programme for screening early language development for children in Wales. These include the following general principles relating to the screening of children being brought up in bilingual and multilingual homes.

7.16 ***Recommendation 9:*** Information about children's language background and exposure to each of their languages should be collected as this is crucial to the screening process for bilingual children.

7.17 ***Recommendation 10:*** Bilingual and multilingual children should be assessed in all of their languages in order to obtain an accurate picture of their language development.

7.18 ***Recommendation 11:*** The use of standardised assessments with bilingual and multilingual children should be considered with care, and interpretation of results must take into account the potential for cultural and linguistic bias.

7.19 ***Recommendation 12:*** Normative data for monolingual children should not be used as a comparison for bilingual children's language development.

7.20 ***Recommendation 13:*** Measurements of vocabulary in bilingual and multilingual children should not be used as they may fail to accurately represent their competence in either language, and are likely to be inappropriate when compared to monolingual normative data.

7.21 ***Recommendation 14:*** Parents should be consulted about their child's speech and language development (in all languages where applicable) and any concerns they may have should be discussed.

8. Summary

8.1 This review of early language screening suitable for children in Wales has built on existing work and carried out novel reviews to determine the relative merits of early language screening in universal and targeted populations. Risk and protective factors for early language development have been identified and reported, highlighting their multifactorial and complex nature and the varied impact on transient and persistent language needs.

8.2 The research team's understanding of how language develops together with these risk and protective factors provide a useful basis for exploring screening with regards to informing what should be screened when, by whom and how. The variation which exists in typical development provides a challenge to this but nevertheless, the evidence shows that the results of screening from age 2 years and 6 months are more dependable and that parents are a good source of information throughout the screening process. There is also increasing recognition that a one-off screen may miss important indicators as children's speech, language and communication needs change over time, with some needs not becoming apparent until later in childhood.

8.3 Following reviews of early language screening tools for use with children growing up in monolingual English, bilingual and multilingual homes (including Welsh), tools have been identified for consideration of their potential usefulness in the Welsh context. An exploration of what constitutes an effective screening tool facilitated an analysis of the identified tools to determine which fulfil the criteria required to be considered a robust tool. From the initial reviews which identified 107 tools available for use with children growing up in English speaking homes, 16 for those in bilingual homes and another eight for those in Welsh speaking homes, just ten tools fulfilled the criteria for inclusion. Of these ten, no one tool can be considered sufficient to fulfil all criteria (screens for all aspects of language development; suitable for ages 15 months to 4 years, 11 months; meets recommended

thresholds for validity, reliability and diagnostic power), nevertheless each has its merits

8.4 The recommendations detail more than just the tools available for use. They summarise key principles which need to be considered in the screening of early language skills of children in Wales. Whilst there is much which is universal in this process and the learning from research carried out in England and elsewhere can be used, there are nevertheless differences for children in Wales in terms of the language environment and also the Healthy Child Wales Programme. This means a novel approach, which incorporates all that is useful from existing work, is needed to ensure that screening of early language skills for children in Wales is fit for purpose in identifying those children who would benefit from additional support and intervention when they need it. Ultimately, a robust programme for screening early language which is feasible and cost-effective to deliver and is acceptable to the key stakeholders involved is vital. This can lead to an increase in the number of children entering school with the speech, language and communication skills they need to succeed, emotionally, socially and educationally.

References

Abrami, P.C., Borokhovski, E., Bernard, R.M., Wade, C.A., Tamim, R., Persson, T., Bethel, E.C., Hanz, K. and Surkes, M.A. (2010) 'Issues in conducting and disseminating brief reviews of evidence', Evidence and Policy, 6(3), pp. 371–389. doi: 10.1332/174426410X524866.

Albanese, M.A., Mejicano, G., Mullan, P., Kokotailo, P. and Gruppen, L. (2008) 'Defining characteristics of educational competencies', Medical Education, 42(3), pp. 248-255. doi: 10.1111/j.1365-2923.2007.02996.x.

Amin, Z., Seng, C.Y. and Eng, K.H. (2006) 'Practical Guide to Medical Student Assessment'. Singapore: World Scientific Publishing Co. Pte.Ltd

Anderson, D.R. and Pempek, T. A. (2005) 'Television and very young children', American behavioral scientist, 48(5), pp. 505-522. doi: 10.1177/0002764204271506.

Armon-Lotem, S. and Ohana, O. (2017) 'A CDI study of bilingual English-Hebrew children–frequency of exposure as a major source of variation', International Journal of Bilingual Education and Bilingualism, 20(2), pp. 201–217. doi: 10.1080/13670050.2016.1179257.

Asmussen, K., Law, J., Charlton, J., Acquah, D., Brims, L., Pote, I. and McBride, T. (2018) 'Key Competencies in early cognitive development: Things, People, Numbers and Words. Early Intervention Foundation'. (Accessed: 14 September 2021)

Baghurst, P.A., Tong, S.L., Woodward, A. and McMichael, A.J. (1992) 'Effects of maternal smoking upon neuropsychological development in early childhood: Importance of taking account of social and environmental factors'. Paediatric and Perinatal Epidemiology, 6(4), pp. 403-15. doi: 10.1111/j.1365-3016.1992.tb00784.x.

Barnes, J. and Melhuish, E.C. (2017) 'Amount and timing of group-based childcare from birth and cognitive development at 51 months: A UK study'. International Journal of Behavioral Development, 41(3), pp. 360–370. doi: 10.1177/0165025416635756.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J.S., Reilly, J. and Hartung, J. (1994) 'Developmental and stylistic variation in the composition of early vocabulary'. Journal of child language. 21(1), pp. 85-123. doi: 10.1017/s0305000900008680.

Bayley, N. (2006) Bayley Scales of Infant and Toddler Development– Third Edition. San Antonio, TX: Harcourt Assessment.

Becker, B. (2011). 'Social disparities in children's vocabulary in early childhood. Does pre-school education help to close the gap?' The British Journal of Sociology, 62(1), 69–88. doi: 10.1111/j.1468-4446.2010.01345.x.

Bedore, L.M. and Peña, E.D. (2008) 'Assessment of bilingual children for identification of language impairment: Current findings and implications for practice', International Journal of Bilingual Education and Bilingualism, 11(1), pp. 1–29. doi: 10.2167/beb392.0.

Beitchman, J.H., Wilson, B., Brownlie, E.B., Walters, H., Inglis, A. and Lancee, W. (1996) 'Long-Term Consistency in Speech/Language Profiles: II. Behavioural, Emotional, and Social Outcomes', Journal of the American Academy of Child and Adolescent Psychiatry, 35(6), pp. 815-825. doi: 10.1097/00004583-199606000-00022.

Beitchman, J.H., Wilson, E.B., Johnson, C.J., Atkinson, L., Young, A. and Adlaf, E. (2003) 'Fourteen-year follow-up of speech/language impaired and control children: psychiatric outcome', Journal of the American Academy of Child and Adolescent Psychiatry, 40(1), pp. 75–82. doi: 10.1097/00004583-200101000-00019.

Bellman, M., Byrne, O. and Sege, R. (2013) 'Developmental assessment of children'. BMJ, 346. doi: https://doi.org/10.1136/bmj.e8687

Bellman, M., Lingham, S., Aukett, A. (1996) Schedule of Growing Skills II. London: GL Assessment. https://support.glassessment.co.uk/media/2406/sgs-ii-ref-manual (Accessed 26 June 2022)

Bercow, J. (2008) The Bercow Report: A review of services for Children and Young People with Speech, Language and Communication needs. (Accessed 3 February 2022).

Berglund, E., Eriksson, M. and Westerlund, M. (2005) 'Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children', Scandinavian Journal of Psychology, 46(6), 485–491. doi: 10.1111/j.1467-9450.2005.00480.x.

Bergmann, C., Dimitrova, N., Alaslani, K., Almohammadi, A., Alroqi, H., Aussems, S., Barokova, M., Davies, C., Gonzalez-Gomez, N., Gibson, S.P. and Havron, N. (2022) Young children's screen time during the first COVID-19 lockdown in 12 countries. Scientific reports, 12(1), pp.1-15. (accessed 7th June 2022).

Berkman, N.D., Wallace, I., Watson, L., Coyne-Beasley, T., Cullen, K., Wood, C. and Lohr, K.N. (2015) 'Screening for Speech and Language Delays and Disorders in Children Age 5 years or younger: A systematic review', U.S. Preventive Services Task Force Agency for Healthcare Research and Quality (US) Rockville (MD) Rockville (MD): Report No.: 13-05197-EF-1. (Accessed: 30 December 2021).

Binks, H. and Thomas, E. (2019) 'Long-term outcomes for bilinguals in minority language contexts: Welsh-English teenagers' performance on measures of grammatical gender and plural morphology in Welsh.' Applied Psycholinguistics, 40 (4), pp. 1019-1049.

Bishop, D.V., Holt, G., Line, E., McDonald, D., McDonald, S. and Watt, H. (2012) 'Parental phonological memory contributes to prediction of outcome of late talkers from 20 months to 4 years: a longitudinal study of precursors of specific language impairment', Journal of Neurodevelopmental Disorders, 4(1), pp. 1-12. doi: 10.1186/1866-1955-4-3.

Bishop, D.V.M., Snowling, M.J., Thompson, P.A. and Greenhalgh, T., CATALISE consortium. (2016) 'CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children', PLoS ONE, 11(7). doi: 10.1371/journal.pone.0158753.

Bornstein, M., Hahn, C.S. and Haynes, O. (2004) 'Specific and general language performance across early childhood: Stability and gender considerations', First Language, 24(3), pp. 267–304. doi: 10.1177/0142723704045681.

Borsley R.D. and Jones, B.M. (2001) 'The development of finiteness in early Welsh', Journal of Celtic Language Learning (Accessed: 30 December 2021).

Breslau, N., Paneth, N., Lucia, V.C. and Paneth-Pollak, R. (2005) 'Maternal smoking during pregnancy and offspring IQ', International Journal of Epidemiology, 34(5), pp. 1047–1053. doi: 10.1093/ije/dyi163.

Brigance, A. H., & Glascoe, F. P. (2002). Brigance infant and toddler. Billerica, MA: Curriculum Associates. (Accessed 27th October 2021).

Broomfield, J. and Dodd, B. (2011) Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial. Int J Lang Commun Disord, 46(6), pp. 628-640. doi: 10.1111/j.1460-6984.2011.00039.x.

Brushe, M.E., Lynch, J.W., Reilly, S., Melhuish, E. and Brinkman, S.A. (2020) 'How many words are Australian children hearing in the first year of life?', BMC Pediatrics, 20, 52. doi: 10.1186/s12887-020-1946-0.

Bryan, K., Freer, J. and Furlong, C. (2007) 'Language and communication difficulties in juvenile offenders', International Journal of Language and Communication Disorders, 42(5), pp. 505-520. doi: 10.1080/13682820601053977.

Bus, A.G., Van Ijzendoorn, M.H. and Pellegrini, A.D. (1995) 'Joint book reading makes for success in learning to read: A metaanalysis on intergenerational transmission of literacy', Review of Educational Research, 65(1), pp. 1–21. doi: 10.2307/1170476.

Buschmann, A., Jooss, B., Rupp, A., Feldhusen, F., Pietz, J. and Philippi, H. (2009) 'Parent based language intervention for 2-year-old children with specific expressive language delay: a randomised controlled trial. Archives of disease in childhood 94:2 110-116.

Camilleri, B. and Law, J. (2007) ‘Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary’, Advances in Speech Language Pathology, 9(4), pp. 312-322. doi: 10.1080/14417040701624474.

Camilleri, B. and Law, J. (2014) ‘Dynamic assessment of word learning skills of pre-school children with primary language impairment’, International Journal of Speech-Language Pathology, 16(5), pp. 507-516. doi: 10.3109/17549507.2013.847497.

Carraccio, C., Wolfsthal, S.D., Englander, R., Ferentz, K. and Martin, C. (2002) ‘Shifting paradigms: from Flexner to Competencies’, Academic Medicine, 77(5), pp. 361-367. doi: 10.1097/00001888-200205000-00003.

Cartmill, E.A., Armstrong, B.F., Gleitman, L.R., Goldin-Meadow, S., Medina, T.N. and Trueswell, J. C. (2013) ‘Quality of early parent input predicts child vocabulary 3 years later’, Proceedings of the National Academy of Sciences, 110(28), pp. 11278–11283. doi: 10.1073/pnas.1309518110.

Cattani, A., Abbot-Smith, K., Farag,R., Krott, A., Arreckx, F., Dennis, I. and Floccia, C. (2014) ‘How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests?’, International Journal of Language and Communication Disorders, 49(6), pp. 649–671. doi: 10.1111/1460-6984.12082.

Catts, H.W. (1997) ‘The Early Identification of Language-based reading difficulties’, Language, Speech and Hearing Services in Schools, 28(1), pp. 86-89. doi: 10.1044/0161-1461.2801.86.

Charney, S.A., Camarata, S.M. and Chern, A. (2021) Potential impact of the COVID-19 pandemic on communication and language skills in children. Otolaryngology–Head and Neck Surgery, 165(1), pp.1-2.

Chilosi, A.M., Pfanner, L., Pecini, C., Salvadorini, R., Casalini, C., Brizzolara, D. and Cipriani, P. (2019) ‘Which linguistic measures distinguish transient from persistent language problems in Late Talkers from 2 to 4 years? A study on Italian speaking children’. Res Dev Disabil, 89, pp. 59-68. doi: 10.1016/j.ridd.2019.03.005.

Ching, T.Y.C., Dillon, H., Marnane, V., Hou, S., Day, J., Seeto, M., Crowe, K., Street, L., Thomson, J., VanBuynder, P., Zhang, V., Wong, A., Burns, L., Flynn, C., Cupples, L., Cowan, R.S.C., Leigh, G., Sjahalam-King, J. and Yeh, A. (2013) ‘Outcomes of early- and late-identified children at 3 years of age: Findings from a prospective population-based study’, Ear and Hearing, 34(5), pp. 535–552. doi: 10.1179/1467010014Z.000000000168.

Chonchaiya, W. and Pruksananonda, C. (2008) Television viewing associates with delayed language development. Acta Paediatrica, 97(7), pp. 977-982. doi: 10.1111/j.1651-2227.2008.00831.x.

Chondrogianni, V. and John, N. (2018) ‘Tense and plural formation in Welsh-English bilingual children with and without language impairment’, Int J Lang Commun Disord, 53(3) pp. 495-514. doi: 10.1111/1460-6984.12363.

Chondrogianni, V. and Kwon, H. (2019) ‘The development of English tense and agreement morphology in Welsh–English bilingual children with and without specific language impairment’. Applied Psycholinguistics, 40(4), pp. 821-852. doi:10.1017/S0142716418000772.

Choudhury, N. and Benasich, A.A. (2003) A family aggregation study. Journal of Speech, Language and Hearing Research, 46(2), pp. 261-272. doi: 10.1044/1092-4388(2003/021).

Claessen, M., Beattie, T., Roberts, R., Leitao, S., Whitworth, A. and Dodd, B. (2016) ‘Is two too early? Assessing toddlers’ phonology’, Speech, Language and Hearing, 20(2), pp. 91-101. doi: 10.1080/2050571X.2016.1222723.

Clegg, J., Hollis, C., Mawhood, L. and Rutter, M. (2005) ‘Developmental language disorder—a follow-up in later adult life: Cognitive, language and psychosocial outcomes’, Journal of Child Psychology and Psychiatry, 46, pp. 128–149.

Coalson, G. A., Peña, E. D. and Byrd, C. T. (2013) ‘Description of multilingual participants who stutter’, Journal of Fluency Disorders, 38(2), pp. 141–156. doi: 10.1016/j.jfludis.2013.02.008.

Collisson, B.A., Graham, S.A, Preston, J.L., Rose, M.S., McDonald, S. and Tough, S. (2016). Risk and protective factors for late talking: An epidemiologic investigation, The Journal of Pediatrics, 172, pp. 168-174. doi: 10.1016/j.jpeds.2016.02.020.

Colonnesi, C., Stams, G.J., Koster, I. and Noom, M.J. (2010) ‘The Relation Between Pointing and Language Development: A Meta-Analysis’, Developmental review, 30(4), pp. 352–366. doi:10.1016/j.dr.2010.10.001.

Cone-Wesson, B. (2005) Prenatal alcohol and cocaine exposure: influences on cognition, speech, language, and hearing. Journal of communication disorders, 38(4), 279-302. doi: 10.1016/j.jcomdis.2005.02.004.

Conica, M., Nixon, E., & Quigley, J. (2020). Fathers’ but not mothers’ repetition of children’s utterances at age two is associated with child vocabulary at age four. Journal of experimental child psychology, 191, 104738. doi: 10.1016/j.jecp.2019.104738.

Core, C., Hoff, E., Rumiche, R. and Señor, M. (2013) ‘Total and conceptual vocabulary in Spanish-English bilinguals from 22 to 30 months: Implications for assessment’, Journal of Speech, Language, and Hearing Research, 56(5), pp. 1637–1649. doi: 10.1044/1092-4388(2013/11-0044).

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. Journal of applied psychology, 78(1), 98.

Council on Linkages between Academia and Public Health Practice. The Competencies Project. http://www.trainingfinder.org/competencies/background.htm.

Cronbach, L.J., Gleser, G.C. and Nanda, H.R.N. (1972) 'The dependability of behavioural measurements: Theory of generalizability for scores and profiles', New York: John Wiley.

Cronbach, L.J., Nageswari, R. and Gleser, G.C. (1963) 'Theory of Generalizability: A libration of reliability theory', The British Journal of Statistical Psychology, 16, pp. 137-163.

Crutchley, A. (2000) 'Bilingual children in language units: Does having "well-informed" parents make a difference?', International Journal of Language and Communication Disorders, 35(1), pp. 65–81. doi: 10.1080/136828200247250.

Dale, P.S., Bates, E., Reznick, J.S. and Morisset, C. (1989) 'The validity of a parent report instrument of child language at twenty months', Journal of child language, 16(2), pp. 239-249. doi: 10.1017/s0305000900010394.

Dale, P.S., Price, T.S., Bishop, D.V.M. and Plomin, R. (2003) 'Outcomes of early language delay, 1: Predicting persistent and transient language difficulties at 3 and 4 years', Journal of Speech, Language, and Hearing Research, 46, pp. 544–560. doi: 10.1044/1092-4388(2003/044).

De Houwer, A., Bornstein, M. H. and Putnick, D. L. (2012) 'A bilingual-monolingual comparison of young children's vocabulary size: Evidence from comprehension and production', Applied Psycholinguistics, 35(6), pp. 1189–1211. doi: 10.1017/S0142716412000744.

De Lamo White, C. and Jin, L. (2011) 'Evaluation of speech and language assessment approaches with bilingual children', International Journal of Language and Communication Disorders, 46(6), pp. 613–627. doi: 10.1111/j.1460-6984.2011.00049.x.

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P. and Friend, M. (2016) 'The Language Exposure Assessment Tool: Quantifying Language Exposure in Infants and Children', Journal of Speech, Language, and Hearing Research, 59(6), pp. 1346–1356. doi: 10.1044/2016_JSLHR-L-15-0234.

Dockrell, J. E. (2001) 'Assessing language skills in preschool children', Child Psychology and Psychiatry Review, 6(2), pp. 74-85. doi: 10.1111/1475-3588.00325.

Dockrell, J.E. and Marshall, C.R. (2015) 'Measurement Issues: Assessing Language Skills in Young Children', Child and Adolescent Mental Health, 20(2), pp. 116-125. doi: 10.1111/camh.12072.

Downing, S. M. and Yudkowsky, R. (2009) 'Assessment in Health Professionals Education'. London: Routledge.

Early Intervention Foundation. (2018) Key competencies in early cognitive development: Things, people, numbers and words. Early Intervention Foundation. (Accessed: 14 June 2021).

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovic-Umek, L., Gayraud, F., Kovacevic, M. and Gallego, C. (2012) 'Differences between girls and boys in emerging language skills: evidence from 10 language communities', British Journal of Developmental Psychology, 30(2), pp. 326–343. doi: 10.1111/j.2044-835X.2011.02042.x.

Falgreen, E.H.L., Kesmodel, U.S., Wimberley, T., Underbjerg, M., Kilburn, T.R. and Mortensen, E.L. (2012) 'Effects of tobacco smoking in pregnancy on offspring intelligence at the age of 5', Journal of Pregnancy, doi: 10.1155/2012/945196.

Farrant, B.M. and Zubrick, S.R. (2013) 'Parent–child book reading across early childhood and child vocabulary in the early school years: Findings from the Longitudinal Study of Australian Children First Language', 33(3), pp. 280–293. doi: 10.1177/0142723713487617.

Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Kurs-Lasky, M., Janosky, J.E. and Paradise, J.L. (2000) Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. Child development, 71(2), pp. 310-322. doi: 10.1111/1467-8624.00146.

Feldman, H.M., Dale, P.S., Campbell, T.F., Colborn, D.K., Kurs-Lasky, M., Rockette, H.E. and Paradise, J.L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. Child development, 76(4), pp. 856-868. doi: 10.1111/j.1467-8624.2005.00882.x.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., Pethick, S.J., Tomasello, M., Mervis, C.B. and Stiles, J. (1994) 'Variability in early communicative development', Monographs of the society for research in child development, 59(5), i–185. doi: 10.2307/1166093.

Fenson, L., Pethick, S., Renda, C., Cox, J., Dale, P. S. and Reznick, J. S. (2000) 'Short-form versions of the MacArthur Communicative Developmental Inventories', Applied Psycholinguistics, 21, pp. 95–116.

Fergusson, D. M. and Lloyd, M. (1991) 'Smoking during pregnancy and its effects on child cognitive ability from the ages 8 to 12 years', Paediatric and Perinatal Epidemiology, 5(2), pp. 189–200. doi: 10.1111/j.1365-3016.1991.tb00700.x.

Fernald, A., Perfors, A. and Marchman, V.A. (2006) 'Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year', Developmental psychology, 42(1), pp. 98–116. doi: 10.1037/0012-1649.42.1.98.

Fernald, A., Marchman, V.A. and Weisleder, A. (2013) 'SES differences in language processing skill and vocabulary are evident at 18 months', Developmental Science, 16(2), pp. 234-248. doi: 10.1111/desc.12019.

Floccia, C., Sambrook, T.D., Luche, C.D., Kwok, R., Goslin, J., White, Cattani, L.A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J. and Plunkett, K. (2018) 'Vocabulary of 2-Year-Olds Learning English and an Additional Language: Norms and Effects of Linguistic Distance', Monographs of the Society for Research in Child Development, 83(1), pp. 1-6. doi: 10.1111/mono.v83.1.

Floccia, C., Sambrook, T.D., Delle Luche, C., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J. and Plunkett, K. (2018) 'III: Analysis and results for study1: Estimating the effect of linguistic distance on vocabulary development', Monographs Society Res Child, 83(1), pp. 43-60. doi: 10.1111/mono.12350.

Foster-Cohen, S.H., Friesen, M.D., Champion, P.R. and Woodward, L.J. (2010) 'High prevalence/low severity language delay in preschool children born very preterm', Journal of Developmental and Behavioral Pediatrics, 31(8), pp. 658-67. doi: 10.1097/DBP.0b013e3181e5ab7e.

Frank, D.A., Augustyn, M., Knight, W.G., Pell, T. and Zuckerman, B. (2001) Growth, development, and behavior in early childhood following prenatal cocaine exposure: a systematic review. Jama, 285(12), pp. 1613-1625. doi: 10.1001/jama.285.12.1613.

Frank, J.R. (2005) 'The CanMEDS 2005 Physicians Competency Framework. Better Standards. Better Physicians; Better Care', Royal College of Physicians and Surgeons of Canada, Ottawa, Canada. (Accessed: 30 December 2021).

Frisk, V., Montgomery, L., Boychyn, E., Young, R., Van Ryn, E., McLachlan, D. and Neufeld, J. (2009) 'Why Screening Canadian pre-schoolers for language delays is more difficult than it should be', Infants and Young Children' 22(4), pp. 290-308. doi: 10.1097/IYC.0b013e3181bc4db6.

Gathercole, V.C.M. and Thomas, E.M. (2007a) 'Language transmission in bilingual families in Wales', Cardiff: Welsh Language Board (Accessed: 30 December 2021).

Gathercole, V.C.M. and Thomas, E.M. (2007b) 'Prawf Geirfa Cymraeg, Fersiwn 7-11', (Welsh Vocabulary Test, Version 7-11). (Accessed: 30 December 2021).

Gathercole, V.C.M. and Thomas, E.M. (2009) 'Bilingual first language development: Dominant language takeover, threatened minority take-up', Bilingualism: Language and Cognition, 12(1), pp. 213-37. doi: 10.1017/S1366728909004015.

Genesee, F. and Bourhis, R. Y. (1988) 'Evaluative reactions to language choice strategies: The role of sociostructural factors', Language and Communication, 8(3-4), pp. 229–250. doi: 10.1016/0271-5309(88)90020-1.

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V.W., Verhulst, F.C. and Tiemeier, H. (2014) 'Early lexical development and risk of verbal and nonverbal cognitive delay at school age', Acta Paediatrica, 103(1), pp. 70-80. doi: 10.1111/apa.12449.

Gilkerson, J., Richards, J.A., Greenwood, C.R. and Montgomery, J.K. (2017) 'Language assessment in a snap: Monitoring progress up to 36 months', Child Language Teaching and Therapy, 33(2), pp. 99-115. doi:10.1177/0265659016660599.

Girolametto, L., Pearce, P.S. and Weitzman, E. (1995) 'The Effects of Focused Stimulation for Promoting Vocabulary in Young Children with Delays: A Pilot Study', Journal of Children's Communication Development, 17(2), pp. 39-49. doi: 10.1177/152574019501700205.

GL Education. (2021) WellComm: A Speech and Language Toolkit for the Early Years. (Accessed: 27 October 2021)

Glascoe, F.P. (2002) 'The Brigance Infant and Toddler Screen: Standardization and Validation', Journal of developmental and behavioural paediatrics, 23(3), pp. 145-150. doi: 10.1097/00004703-200206000-00003.

Glogowska, M. and Campbell, R. (2004) Parental views of surveillance for early speech and language difficulties. Children and Society, 18(4), pp. 266-277. doi: 10.1002/chi.806

Goh, S.K.Y., Tham, E.K.H., Magiati, I., Sim, L., Sanmugam, S., Qiu, A., Daniel, M.L., Broekman, B.F.P. and Rifkin-Graboi, A. (2017) 'Analysis of Item-Level Bias in the Bayley-III Language Subscales: The Validity and Utility of Standardized Language Assessment in a Multilingual Setting', Journal of Speech, Language, and Hearing Research, 60, pp. 2663–2671. doi: 10.1044/2017_JSLHR-L-16-0196.

Goisis, A. (2015) 'How are children of older mothers doing?: Evidence from the U.K', Biodemography and Social Biology, 61(3), pp. 231-251. doi: 10.1080/19485565.2014.1001887.

Golinkoff, R.M., Hoff, E., Rowe, M.L., Tamis-LeMonda, C.S. and Hirsh-Pasek, K. (2019) 'Language Matters: Denying the Existence of the 30-Million-Word Gap Has Serious Consequences', Child Development, 90(3), pp. 985-992. doi: 10.1111/cdev.13128.

Gudmundsson, E. (2015) 'The Toddler Language and Motor Questionnaire: A mother-report measure of language and motor development', Res Dev Disabil, 45-46, pp. 21-31. doi: 10.1016/j.ridd.2015.07.007.

Guiberson, M. and Atkins, J. (2012) 'Speech-language pathologists' preparation, practices, and perspectives on serving culturally and linguistically diverse children', Communication Disorders Quarterly, 33(3), pp. 169–180. doi: 10.1177/1525740110384132.

Haden, C.A. and Ornstein, P. A. (2009) 'Research on talking about the past: The past, present, and future', Journal of Cognition and Development, 10(3), pp. 135–142. doi: 10.1080/15248370903155718.

Hadley P.A. and Holt J.K. (2006) 'Individual differences in the onset of tense marking: a growth-curve analysis', Journal of speech, language, and hearing research, 49(5), pp. 984-1000. doi: 10.1044/1092-4388(2006/071).

Hadley, P. and Short, H. (2005) 'The onset of tense marking in children at-risk for specific language impairment', Journal of Speech, Language, and Hearing Research, 48, pp. 1344–1362. doi: 10.1044/1092-4388(2005/094).

Hagan, J., Shaw, J.S. and Duncan, P. M. (2008) Bright futures: Guidelines for health supervision of infants, children, and adolescents (3rd ed.) Elk Grove Village, IL: American Academy of Pediatrics

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miękisz, A., Kacprzak, A., Banasik, N. and Foryś-Nogala, M. (2017) 'How does L1 and L2 exposure impact l1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom', Frontiers in Psychology, 8(SEP). doi: 10.3389/fpsyg.2017.01444.

Hambly, H., Wren, Y., McLeod, S. and Roulstone, S. (2013) 'The influence of bilingualism on speech production: A systematic review', International Journal of Language and Communication Disorders, 48(1), pp. 1-24. doi: 10.1111/j.1460-6984.2012.00178.x.

Hamilton, A., Plunkett, K. and Schafer, G. (2000) 'Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than the USA', Journal of Child Language, 27, pp. 689-705. doi: 10.1017/s0305000900004414.

Hammer, C.S., Tomblin, J.B., Zhang, X. and Weiss, A.L. (2001) 'Relationship between parenting behaviours and specific language impairment in children', International Journal of Language and Communication Disorders, 36(2), pp. 185-205. doi: 10.1080/13682820010019919.

Hardin, B. J., Scott-Little, C. and Mereoiu, M. (2013) 'Developing the BIO questionnaire: A bilingual parent report tool for prekindergarten english learners of latino heritage', Journal

of Research in Childhood Education, 27(4), pp. 485–509. doi: 10.1080/02568543.2013.824940.

Harrison, L.J. and McLeod, S. (2010) 'Risk and protective factors associated with speech and language impairment in a nationally representative sample of 4- to 5-Year-old children', Journal of Speech Language and Hearing Research, 53(2), pp. 508–529. doi: 10.1044/1092-4388(2009/08-0086).

Hart, B. and Risley, T.R. (1995) Meaningful differences in the everyday experience of young American children. Baltimore, MD: Paul H Brookes Publishing

Hasson, N. and Botting, N. (2010) 'Dynamic assessment of children with language impairments: A pilot study', Child Language Teaching and Therapy, 26(3), pp. 249- 272. doi: 10.1177/0265659009349982.

Hasson, N., Camilleri, B., Jones, C., Smith, J. and Dodd, B. (2013) 'Discriminating disorder from difference using dynamic assessment with bilingual children', Child Language Teaching and Therapy, 29(1), pp. 57-75. doi: 10.1177/0265659012459526.

Hauner, K.K., Shriberg, L.D., Kwiatkowski, J. and Allen, C.T. (2005) A subtype of speech delay associated with developmental psychosocial involvement. Journal of speech, language and hearing research, 48(3), pp. 635-650. doi: 10.1044/1092-4388(2005/044).

Hayiou-Thomas, M.E., Dale, P.S. and Plomin, R. (2012) 'The etiology of variation in language skills changes with development: A longitudinal twin study of language from 2 to 12 years', Developmental Science, 15(2), pp. 233–249. doi: 10.1111/j.1467-7687.2011.01119.x.

Hickey, T.M., Lewis, G. and Baker, C. (2014) 'How deep is your immersion? Policy and practice in Welsh-medium preschools with children from different language backgrounds', International Journal of Bilingual Education and Bilingualism, 17(2), pp. 215-234. doi: 10.1080/13670050.2013.866629.

Hidecker, M.J., Cunningham, B.J., Thomas-Stonell, N., Oddson, B. and Rosenbaum, P. (2017) 'Validity of the Communication Function Classification System for use with preschool children with communication disorders', Dev Med Child Neurol, 59(5), pp. 526-530. doi: 10.1111/dmcn.13373.

Hinshaw, S.P. (1992) 'Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms', Psychological Bulletin, 111(1), pp. 127–55. doi: 10.1037/0033-2909.111.1.127.

Hoff, E. (2006) 'How social contexts support and shape language development', Developmental Review, 26(1), pp. 55-88. doi: 10.1016/j.dr.2005.11.002.

Hoff, E. (2013) 'Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps', Developmental psychology, 49(1), pp. 4-14. doi: 10.1037/a0027238.

Holm, A., Stow, C. and Dodd, B. (2013) Bilingual Children with Phonological Disorders: Identification and Intervention, in B. Dodd (ed.) Differential Diagnosis and Treatment of Children with Speech Disorder, 2nd edn. London: Whurr, pp. 275-88.

Hurtado, N., Marchman, V.A. and Fernald, A. (2008) 'Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children', Developmental Science, 11(6), pp. F31-9. doi: 10.1111/j.1467-7687.2008.00768.x.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. and Lyons, T. (1991) 'Early vocabulary growth: Relation to language input and gender', Developmental Psychology, 27(2), pp. 236–248. doi: 10.1037/0012-1649.27.2.236.

I CAN. (2009) The Cost of the Nation of Children's Poor Communication. (Accessed 30 December 2021).

I CAN. (2021) Speaking up for the Covid Generation, (Accessed 8 October 2021).

International Expert Panel on Multilingual Children's Speech. (2012) Multilingual children with speech sound disorders: Position paper. Accessible at: https://www.csu.edu.au/research/multilingual-speech/position-paper. (Accessed 1 January 2022)

Irwin, L.G., Siddiqi, A. and Hertzman, G. (2007) Early child development: a powerful equalizer. Vancouver, BC: Human Early Learning Partnership (HELP).

Johnson, S. and Bountziouka, V. for the PARCA-R Study Group. (2020) 'Using the PARCA-R to assess children's cognitive and language development at two years of age', Infant, 16(4), pp. 159-63.

Joint Centre for Education in Medicine. (1997) 'The Good Assessment Guide: A Practical Guide to Assessment and Appraisal for Higher Specialist Training." Joint Centre for Education in Medicine, London, UK, pp. 34-39.

Kalnak, N., Peyrard-Janvid, M., Sahlén, B. and Forssberg, H. (2012) Family history interview of a broad phenotype in specific language impairment and matched controls. Genes, Brain and Behavior, 11(8), pp. 921-927. doi: 10.1111/j.1601-183X.2012.00841.x.

Kartushina, N., Mani, N., Aktan-Erciyes, A.S.L.I., Alaslani, K., Aldrich, N.J., Almohammadi, A., Alroqi, H., Anderson, L., Andonova, E., Aussems, S. and Babineau, M. (2022) COVID-19 first lockdown as a unique window into language acquisition: What you do (with your child) matters. In press

Klee, T., Carson, D., Gavin, W., Hall, L., Kent, A. and Reece, S. (1998) 'Concurrent and predictive validity of an early language screening program', Journal of Speech, Language and Hearing, 41(3), pp. 627-641. doi: 10.1044/jslhr.4103.627.

Kohnert, K. and Medina, A. (2009) 'Bilingual children and communication disorders: A 30-Year research retrospective', Seminars in Speech and Language, 30(4), pp. 219–233. doi: 10.1055/s-0029-1241721.

Kuchirko, Y. (2019) 'On differences and deficits: A critique of the theoretical and methodological underpinnings of the word gap', Journal of Early Childhood Literacy, 19(4), pp. 533–562. doi: 10.1177/1468798417747029.

Lacroix, V., Pomerleau, A. and Malcuit, G. (2002) 'Properties of adult and adolescent mothers' speech, children's verbal performance and cognitive development in different socioeconomic groups: a longitudinal study', First Language, 22(2), pp. 173–196. doi: 10.1177/014272370202206503.

Lahey, M. and Edwards, J. (1995) Specific language impairment: Preliminary investigation of factors associated with family history and with patterns of language performance. Journal of Speech, Language, and Hearing Research, 38(3), pp. 643-657. doi: 10.1044/jshr.3803.643.

Laing, G. J., Law, J., Levin, A. and Logan, S. (2002) 'Evaluation of a structured test and a parent led method for screening for speech and language problems: prospective population based study', British Medical Journal, 16;325(7373):1152. doi: 10.1136/bmj.325.7373.1152.

Law, J., Boyle, J., Harris, F., Harkness, A. and Nye, C. (1998) 'Screening for speech and language delay: A systematic review of the literature', Health Technology Assessment, 2(9), pp. 1-184. doi: 10.3310/hta2090.

Law, J., Boyle, J., Harris, F., Harkness, A. and Nye, C. (2000) 'The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature', Developmental Medicine and Child Neurology, 42(3), pp. 190-200. doi: 10.1017/s0012162200000335.

Law, J., Charlton, J. and Asmussen, K. (2017) 'Early Intervention Foundation report: Language as a child wellbeing indicator'. (Accessed: 14 June 2021)

Law, J., Charlton, J., Dockrell, J., Gascoigne, C., McKean, C. and Theakston, A. (2017) 'Early Language Development: Needs, provision and intervention for preschool children from socio-economically disadvantage backgrounds. A report for Education Endowment Foundation'. (Accessed: 31 December 2021)

Law, J., Charlton, J., McKean, C., Watson, R., Roulstone, S., Holme, C., Gilroy, V., Wilson, P. and Rush, R. (2020) 'Identifying and Supporting Children's Language Needs Public Health England', (Accessed 27 October 2021)

Law, J., Reilly, S. and Snow, P.C. (2013) 'Child speech, language and communication need re-examined in a public health context: a new direction for the speech and language therapy profession', International Journal of Language and Communication Disorders, 48(5), pp. 486-496. doi: 10.1111/1460-6984.12027.

Law, J., Rush, R., Anandan, C., Cox, M. and Wood, R. (2012) 'Predicting language change between 3 and 5 years and its implications for early identification', Pediatrics, 130(1), pp. e132-e137. doi: 10.1542/peds.2011-1673.

Lazewnik, R., Creaghead, N.A., Smith, A.B., Prendeville, J.A., Raisor-Becker, L. and Silbert, N. (2019) 'Identifiers of language impairment for Spanish-English dual language learners', Language, Speech, and Hearing Services in Schools, 50(1), pp. 126–137. doi: 10.1044/2018_LSHSS-17-0046.

Leech, K. A. and Rowe, M.L. (2014) 'A Comparison of Preschool Children's Discussions with Parents During Picture Book and Chapter Book Reading', First Language, 34(3), pp. 205–226. doi: 10.1177/0142723714534220.

Legacy, J. Reider, J., Crivello, C., Kuzyk, O., Friend, M., Zesiger, P. and Poulin-Dubois, D. (2017) 'Dog or Chien? Translation Equivalents in the Receptive and Expressive Vocabularies of Young French-English Bilinguals Jacqueline', Journal of Child Language, 44(4), pp. 881–904. doi: 10.1017/S0305000916000295.

LENA. (2021) 'LENA Developmental Snapshot Developmental Snapshot', (Accessed 27 October 2021).

Levickis, P., Reilly, S., Girolametto, L., Ukoumunne, O.C. and Wake, M. (2014) 'Maternal Behaviors Promoting Language Acquisition in Slow-to-Talk Toddlers', Journal of Developmental and Behavioral Pediatrics, 35(4), pp. 274-281. doi: 10.1097/DBP.0000000000000056.

Lewis, M. and Freedle, R. (1973) 'Mother–infant dyad: The cradle of meaning'. In Communication and affect: Language and thought pp. 127–155. (Accessed: 31 December 2021).

Lindsay, G., Dockrell, J., Law, J. and Roulstone, S. (2010) 'The Better Communication Research Programme: Improving provision for children and young people with speech, language and communication needs', (Accessed 8 October 2021)

Linn, R. L. and Miller, D. (2005) 'Reliability and other desired characteristics', in Measurement and Assessment in Teaching, 9th edn, New Jersey: Pearson Educational International, pp. 104-133.

Lund, E.M., Kohlmeier, T.L. and Durán, L.K. (2017) 'Comparative Language Development in Bilingual and Monolingual Children With Autism Spectrum Disorder: A Systematic Review', Journal of Early Intervention, 39(2), pp. 106–124. doi: 10.1177/1053815117690871.

Ma, W., Golinkoff, R. M., Houston, D. M. and Hirsh-Pasek, K. (2011) 'Word learning in infant- and adult-directed speech', Language Learning and Development, 7(3), pp. 185–201.

MacArthur Bates CDI. (2015) 'MacArthur Bates Communicative Developmental Inventories (MB-CDIs) MacArthur-Bates Communicative Development Inventories', (Accessed 27 October 2021).

Maccoby, E.E. and Jacklin, C.N. (1974) The psychology of sex differences Stanford, CA: Stanford University Press.

Maggi, S., Irwin, L.J., Siddiqi, A. and Hertzman, C. (2010) 'The social determinants of early child development: An overview', Journal of Paediatrics and Child Health, 46, pp. 627-35. doi: 10.1080/15475441.2011.579839.

Mancilla-Martinez, J., Gámez, P.B., Vagh, S.B. and Lesaux, N.K. (2016) 'Parent Reports of Young Spanish–English Bilingual Children's Productive Vocabulary: A Development and Validation Study', Language, Speech and Hearing Services in Schools, 47(1), pp. 1–15. doi: 10.1044/2015_LSHSS-15-0013.

Marchman, V. A. and Martínez-Sussmann, C. (2002) 'Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish', Journal of Speech, Language, and Hearing Research, 45(5), pp. 983–997. doi: 10.1044/1092-4388(2002/080).

Marcotte, A.M., Clemens, N.H., Parker, C. and Whitcomb, S.A. (2016) 'Examining the Classification Accuracy of a Vocabulary Screening Measure With Preschool Children', Assessment for Effective Intervention, 41(4), pp. 230-242. doi:10.1177/1534508416632236.

Marmot, M., Allen, J., Goldblatt, P., Boyce, T., McNeish, D., Grady, M. and Geddes, I. (2010) 'Fair society, healthy lives: The Marmot Review', (Accessed 14 June 2021).

Mayor, J. and Plunkett, K. (2011) 'A statistical estimate of infant and toddler vocabulary size from CDI analysis', Developmental Science, 14(4), pp. 769-785. doi: 10.1111/j.1467-7687.2010.01024.x.

Mayr, R., Howells, G. and Lewis, R. (2015) 'Asymmetries in phonological development: the case of word-final cluster acquisition in Welsh-English bilingual children', J Child Lang, 42(1), pp. 146-79. doi: 10.1017/S0305000913000603.

McCormack, J., Harrison, L.J., McLeod, S. and McAllister, L. (2011) 'A Nationally Representative Study of the Association between Communication Impairment at 4-5 years old and Children's Life Activities at 7-9 years old', Journal of Speech, Language and Hearing Research, 54(5), pp. 1328-1348. doi: 10.1044/1092-4388(2011/10-0155).

McHugh M. L. (2012) Interrater reliability: the kappa statistic. Biochemia medica, 22(3), 276–282.

McKean, C., Law, J., Mensah, F., Cini, E., Eadie, P., Frazer, K. and Reilly, S. (2016) 'Predicting meaningful differences in school-entry language skills from child and family factors measured at 12 months of age', International Journal of Early Childhood, 48(3), pp. 329-351. doi: 10.1007/s13158-016-0174-0.

McKean, C., Mensah, F.K., Eadie, P., Bavin, E.L., Bretherton, L., Cini, E. and Reilly, S. (2015) 'Levers for language growth: Characteristics and predictors of language trajectories between 4 and 7 years', PLoS ONE, 10(8). doi: 10.1371/journal.pone.0134251.

McKean, C., Reilly, S., Bavin, E.L., Bretherton, L., Cini, E., Conway, L., Cook, F., Eadie, P., Prior, M., Wake, W. and Mensah, F. (2017) 'Language Outcomes at 7 years: Early Predictors and Co-Occurring Difficulties', Pediatrics, 139:3. doi: 10.1542/peds.2016-1684.

Messick, S. (1989) 'Validity', in Educational Measurement, 3rd edn, R. L. Linn, ed., American Council on Education and Macmillan., New York, pp. 13-104.

Miller, G.E. (1990) 'The assessment of clinical skills/competence/performance.', Academic Medicine, 65(9), pp. 63-67. doi: 10.1097/00001888-199009000-00045.

Mills, D., Gathercole, V. and Ebanks, N. (2013) 'The Bangor Welsh Communicative Development Inventory: Words and Gestures'. Bangor University. Requested copy from first author via personal communication 7 October 2021.

Miniscalco, C., Westerlund, M. and Lohmander, A. (2005) 'Language skills at age 6 years in Swedish children screened for language delay at 2½ years of age', Acta Paediatrica, 94(12), pp. 1798-1806. doi: 10.1111/j.1651-2227.2005.tb01856.x.

Moeller, M.P., Thomas, A.E., Oleson, J. and Ambrose, S.E. (2019) 'Validation of a Parent Report Tool for Monitoring Early Vocal Stages in Infants', J Speech Lang Hear Res, 62(7), pp. 2245-2257. doi: 10.1044/2019_JSLHR-S-18-0485.

Mol, S.E., Bus, A.G., De Jong, M.T. and Smeets, D.J. (2008) 'Added value of dialogic parent–child book readings: A meta-analysis', Early Education and Development, 19(1), pp. 7-26. doi: 10.1080/10409280701838603.

Monson, B.B., Eaton-Rosen, Z., Kapur, K., Liebenthal, E., Brownell, A., Smyser, C.D., Rogers, C.E., Inder, T.E., Warfield, S.K. and Neil, J.J. (2018) 'Differential Rates of Perinatal Maturation of Human Primary and Nonprimary Auditory Cortex', eNeuro, 5:1. doi: 10.1523/ENEURO.0380-17.2017.

Morgan, L. and Wren, Y.E. (2018) 'A systematic review of the literature on early vocalizations and babbling patterns in young children', Communication Disorders Quarterly, 40(1), pp. 3-14. doi: /10.1177/1525740118760215.

Morris, S.P., Melhuish, E. and Gardiner, J. (2017) 'Study of early education and development (SEED): Impact study on early education use and child outcomes up to age three', (Accessed 31 December 2021).

Munro, S., Ball, MJ., Müller, N., Duckworth, M. and Lyddy, F. (2005) Phonological acquisition in Welsh – English bilingual children, Journal of Multilingual Communication Disorders, 3(1), PP. 24-49, doi: 10.1080/14769670410001683467.

National Institute of Child Health and Human Development Early Childcare Research Network. (2003) 'Social functioning in first grade: Associations with earlier home and childcare predictors and with current classroom experiences', Child Development, 74(6), pp. 1639–1662. doi: 10.1046/j.1467-8624.2003.00629.x.

Nayeb, L., Wallby, T., Westerlund, M., Salameh, E.K. and Sarkadi, A. (2015) 'Child healthcare nurses believe that bilingual children show slower language development, simplify screening procedures and delay referrals', Acta Paediatrica, 104(2), pp. 198–205. doi: 10.1111/apa.12834.

Nayeb, L., Lagerberg, D., Sarkadi, A., Salameh, E. and Eriksson, M. (2021) 'Identifying language disorder in bilingual children aged 2.5 years requires screening in both languages', Acta Paediatrica, 110(1), pp. 265–272. doi: 10.1111/apa.15343.

Nelson, H.D., Nygren, P., Walker, M. and Panoscha, R. (2006) 'Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force', Pediatricsl 117(2), pp. 298-319. doi: 10.1542/peds.2005-1467.

Norbury, C.F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G. and Pickles, A. (2016) The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. Journal of Child Psychology and Psychiatry, 57(11), pp. 1247-1257. doi: 10.1111/jcpp.12573.

Nunnally, J. C. and Bernstein, I. H. (1994) Psychometric theory, 3rd edn, McGraw Hill, New York.

O'Connor, E., Rossom, R.C., Henninger, M., Groom, H. C. and Burda, B.U. (2016) 'Primary care screening for and treatment of depression in pregnant and postpartum women: evidence report and systematic review for the US Preventive Services Task Force', Jama, 315(4), pp. 388–406. doi: 10.1001/jama.2015.18948.

O'Keeffe, L.M. Greene, R.A. and Kearney, P.M. (2014) The effect of moderate gestational alcohol consumption during pregnancy on speech and language outcomes in children: a systematic review. Systematic reviews, 3(1), pp. 1-11. doi: 10.1186/2046-4053-3-1.

O'Leary, C.M. (2004) Fetal alcohol syndrome: Diagnosis, epidemiology, and developmental outcomes. Journal of Paediatric Child Health, 40(1-2), pp. 2–7. doi: 10.1111/j.1440-1754.2004.00280.x.

O'Toole, C., Gatt, D., Hickey, T.M., Miękisz, A., Haman, E., Armon-Lotem, S., Rinker, T., Ohana, O., dos Santos, C. and Kern, S. (2016) 'Parent report of early lexical production in bilingual children: a cross-linguistic CDI comparison', International Journal of Bilingual Education and Bilingualism, 20(2), pp. 124–145. doi: 10.1080/13670050.2016.1179258.

Oliver, B., Dale, P. and Plomin, R. (2004) 'Verbal and nonverbal predictors of early language problems: An analysis of twins in early childhood back to infancy', Journal of Child Language, 31(3), pp. 609-631. doi: 10.1017/S0305000904006221.

Office for National Statistics. (2011). 2011 Census. (Accessed 7 March 2022).

Oryadi-Zanjani, M.M. (2018) 'Production of Infant Scale Evaluation (PRISE) in Persian normal hearing children: A validation study', Int J Pediatr Otorhinolaryngol, 113, pp. 76-81. doi: 10.1016/j.ijporl.2018.07.036.

OxEd and Assessment. (2021) Language Screen OxEd and Assessment. (Accessed 27 October 2021)

Ozçalişkan, S. and Goldin-Meadow, S. (2010) 'Sex differences in language first appear in gesture', Developmental Science, 13, pp. 752–760. doi: 10.1111/j.1467-7687.2009.00933.x

Pagani, L.S., Fitzpatrick, C. and Barnett, T.A. (2013) Early childhood television viewing and kindergarten entry readiness. Pediatric research, 74(3), pp. 350-355. doi: 10.1038/pr.2013.105.

Patterson, J.L., Rodríguez, B.L. and Dale, P.S. (2020) 'Dynamic assessment language tasks and the prediction of performance on year-end language skills in preschool dual language learners', American Journal of Speech-Language Pathology, 29(3), pp. 1226–1240. doi: 10.1044/2019_AJSLP-19-00120.

Paugh, A.L. and Riley, K.C. (2019) 'Poverty and Children's Language in Anthropolitical Perspective', Annual Review of Anthropology, 48, pp. 297–315. doi: 10.1146/annurev-Anthro-102218-011224.

Paul, R., Looney, S.S. and Dahm, P.S. (1991) 'Communication and socialization skills at ages 2 and 3 in "late-talking" young children', Journal of Speech, Language, and Hearing Research, 34(4), pp. 858-865. doi: 10.1044/jshr.3404.858.

Pert, S. and Bradley, B. (2018) 'Clinical Guidelines for Speech and Language Therapists: Bilingualism Bilingualism – Guidance', Royal College of Speech and Language Therapists. (Accessed 31 December 2021).

Peters, M.D.J., Marnie. C., Tricco, A.C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C.M. and Khalil, H. (2020) 'Updated methodological guidance for the conduct of scoping reviews', JBI Evidence Synthesis, 18(10), pp. 2119–2126. doi: 10.11124/JBIES-20-00167.

Pine, J. M. (1995) 'Variation in vocabulary development as a function of birth-order', Child Development, 66, pp. 272–281. doi: 10.2307/1131205

Plymouth University BabyLab. (2021) 'UK Bilingual Toddlers Assessment Tool UKBTAT', (Accessed 29 October 2021)

Prior, M., Bavin, E. L., Cini, E., Reilly, S., Bretherton, L., Wake, M. and Eadie, P. (2008) Influences on communicative development at 24 months of age: Child temperament, behaviour problems, and maternal factors. Infant Behavior and Development, 31(2), pp. 270-279. doi: 10.1016/j.infbeh.2007.11.001.

PRISMA. (2020) www.prisma-statement.org (Accessed 4 March 2022).

Public Health England. (2020) Identifying and supporting children's early language needs. Identifying and Supporting Children's Early Language Needs: summary report' (Accessed 14 June 2021).

Rajmil, L. (2020) 'Role of children in the transmission of the COVID-19 pandemic: A rapid scoping review', BMJ Paediatrics Open, 4(1), pp. 1–6. doi: 10.1136/bmjpo-2020-000722.

Ramírez-Esparza, N., García-Sierra, A. and Kuhl, P.K. (2014). 'Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development', Developmental Science, 17(6), pp. 880–891. doi: 10.1111/desc.12172.

Reese, E. and Read, S. (2000) Predictive validity of the New Zealand MacArthur communicative development inventory: Words and sentences. Journal of Child Language, 27(2), pp. 255-266. doi: 10.1017/s0305000900004098.

Reilly, S., Wake, M., Ukoumunne, O.C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P. and Bretherton, L., 2010. Predicting language outcomes at 4 years of age: findings from Early Language in Victoria Study. Pediatrics, 126(6), pp.e1530-e1537. doi: 10.1542/peds.2010-0254.

Rescorla, L. and Alley, A. (2001) Validation of the language development survey (LDS): A parent report tool for identifying language delay in toddlers. Journal of Speech, Language, and Hearing Research, 44(2), pp. 434-445. doi: 10.1044/1092-4388(2001/035).

Rescorla, L., Dahlsgaard, K. and Roberts, J. (2000) 'Late-talking toddlers: MLU and IPSyn Outcomes at 3;0 and 4;0', Journal of Child Language, 27(3):643-64. doi: 10.1017/s0305000900004232.

Rescorla, L. and Schwartz, E. (1990) 'Outcome of toddlers with specific expressive language delay', Applied Psycholinguistics, 11(4), pp. 393–407. doi: 10.1017/S0142716400009644.

Rethans, J.J., Norcini, J.J., Baron-Maldonado, M., Blackmore, D., Jolly, B.C., LaDuca, T., Lew, S., Page, G. G. and Southgate, L.H. (2002) 'The relationship between competence and performance: implications for assessing practice performance', Medical Education, 36(10), pp. 901-909. doi: 10.1046/j.1365-2923.2002.01316.x.

Rice, M. L. (2012) 'Toward epigenetic and gene regulation models of specific language impairment: looking for links among growth, genes, and impairments', Journal of neurodevelopmental disorders, 4(1), pp. 1-14. doi: 10.1186/1866-1955-4-27.

Rice, M. L. (2013) 'Language growth and genetics of specific language impairment', International journal of speech-language pathology, 15(3), pp. 223-233. doi: 10.3109/17549507.2013.783113.

Rice, M. L., Haney, K. R. and Wexler, K. (1998) Family histories of children with SLI who show extended optional infinitives. Journal of Speech, Language, and Hearing Research, 41(2), pp. 419-432. doi: 10.1044/jslhr.4102.419.

Robert, E. (2009) 'Accommodating "new" speakers? An attitudinal investigation of L2 speakers of Welsh in south-east Wales', International Journal of the Sociology of Language, 195, pp. 93–115. doi: 10.1515/IJSL.2009.007.

Roberts, M.Y. and Kaiser, A.P. (2011) 'The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis', American Journal of Speech-Language Pathology, 20(3), pp. 180-199. doi: 10.1044/1058-0360(2011/10-0055).

Roberts, J.E., Rosenfeld, R.M. and Zeisel, S.A. (2004) 'Otitis Media and Speech and Language: A Meta-Analysis of Prospective Studies', Pediatrics, 113(3), pp. 238-248. doi: 10.1542/peds.113.3.e238.

Roseberry-McKibbin, C. (2021) 'Utilizing Comprehensive Preassessment Procedures for Differentiating Language Difference From Language Impairment in English Learners', Communication Disorders Quarterly, 42(2), pp. 93–99. doi: 10.1177/1525740119890314.

Roulstone, S., Law, J., Rush, R., Clegg, J. and Peters, T. (2011) Investigating the role of language in children's early educational outcomes, (Accessed 31 December 2021).

Roulstone, S.E., Marshall, J.E., Powell, G.G., Goldbart, J., Wren, Y.E., Coad, J., Daykin, N., Powell, J.E., Lascelles, L., Hollingworth, W., Emond, A., Peters, T.J., Pollock, J.I., Fernandes, C., Moultrie, J., Harding, S.A., Morgan, L., Hambly, H.F., Parker, N.K. and Coad, R.A. (2015) 'Evidence-based intervention for preschool children with primary speech and language impairments: Child Talk – an exploratory mixed-methods study', Programme Grants for Applied Research, 3(5), pp. 1–408. doi: 10.3310/pgfar03050.

Roulstone, S., Peters, T.J., Glogowska, M. and Enderby, P. (2003) A 12-month follow-up of preschool children investigating the natural history of speech and language delay. Child Care Health Dev, 29(4), pp.245-55. doi: 10.1046/j.1365-2214.2003.00339.x.

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S. and Lindsay, G. (2012) Exploring interventions for children and young people with speech, language and communication needs: A study of practice. (Accessed 16 September 2021).

Rowe, M. L. (2012) 'A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development', Child Development, 83(5), pp. 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x.

Rowe, M.L. and Goldin-Meadow, S. (2009) 'Early gesture selectively predicts later language learning', Developmental Science, 12(1), pp. 182-187. doi: 10.1111/j.1467-7687.2008.00764.x.

Rowe, M.L., Raudenbush, S.W. and Goldin-Meadow, S. (2012) 'The Pace of Vocabulary Growth Helps Predict Later Vocabulary Skill', Child Development, 83(2), pp. 508-525. doi: 10.1111/j.1467-8624.2011.01710.x.

Roy, P., Kersley, H. and Law, J. (2005) 'The Sure Start Language Measure Standardisation Study', (Accessed 31 December 2021).

Rudolph, J. M. and Leonard, L. B. (2016) 'Early language milestones and specific language impairment', Journal of Early Intervention, 38(1), pp. 41-58. doi: 10.1177/1053815116633861.

Ryu, S.H. and Sim, Y-J. (2019) 'The validity and reliability of DDST II and bayley iii in children with language development delay', Neurology Asia, 24(4), pp. 355-361. (Accessed 31 December 2021).

Salameh, E.K., Nettelbladt, U., Håkansson, G. and Gullberg, B. (2002) 'Language impairment in Swedish bilingual children: A comparison between bilingual and monolingual children in Malmö', Acta Paediatrica, 91(2), pp. 229–234. doi: 10.1080/080352502317285261.

Saxton, M. (2010) Child language: Acquisition and development. London: Sage Publishing.

Schonhaut, L., Armijo, I., Schönstedt, M., Alvarez, J. and Cordero, M. (2013) 'Validity of the Ages and Stages Questionnaires in Term and Preterm Infants', Pediatrics, 131, pp. 1468-1474. doi: 10.1542/peds.2012-3313.

Seager, E. and Abbot-Smith, K. (2017) 'Can early years professionals determine which preschoolers have comprehension delays? A comparison of two screening tools', Child Language Teaching and Therapy, 33(1), pp.67-79. doi:10.1177/0265659016650977.

Selleck, C. (2018). 'We're not fully Welsh': Hierarchies of belonging and 'new' speakers of Welsh. In Smith-Christmas, C., Ó Murchadha, N.P., Hornsby, M. and Moriaty, M. (eds.), New Speakers of Minority Languages: Linguistic Ideologies and Practices (pp. 45-65). London et al: Palgrave MacMillan.

Shapiro, N.T., Hippe, D.S. and Ramírez, N.F. (2021) How Chatty Are Daddies? An Exploratory Study of Infants' Language Environments. Journal of Speech, Language, and Hearing Research, 64(8), pp. 3242-3252. doi: 10.1044/2021_JSLHR-20-00727.

Sharp, K.M. and Gathercole, V.C.M. (2013) 'Can a novel word repetition task be a language-neutral assessment tool? Evidence from Welsh–English bilingual children', Child Language Teaching and Therapy, 29(1), pp. 77-89. doi:10.1177/0265659012465208.

Sheridan, M., Sharma, A. and Cockerill, H. (2014) From birth to five years: children's developmental progress (4th ed.) Routledge. (Accessed 14 September 2021).

Shriberg, L.D., Friel-Patti, S., Flipsen Jr, P. and Brown, R.L. (2000) 'Otitis Media, Fluctuant Hearing Loss, and Speech-Language Outcomes: A Preliminary Structural Equation Model', Journal of Speech, Language and Hearing Research, 43, pp. 100-120. doi: 10.1044/jslhr.4301.100.

Shrout, P.E. and Fleiss, J.L. (1979) 'Intraclass correlations: Uses in assessing rater reliability', Psychological Bulletin, 86, pp. 420-428.

Siu, A.L. (2015) 'Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force recommendation statement', Pediatrics, 136, pp. e474. doi: 10.1542/peds.2015-1711 PMID: 26152670.

Skarakis-Doyle, E., Campbell, W. and Dempsey, L. (2009) 'Identification of children with language impairment: Investigating the classification accuracy of the MacArthur–Bates

Communicative Development Inventories, Level III'. 18(3), pp. 277-88. doi: 10.1044/1058-0360(2009/08-0035).

Skeat, J., Eadie, P., Ukomunne, O., and Reilly, S. (2010) Predictors of parents seeking help or advice about children's communication development in the early years. Child: care, health and development, 36(6), pp. 878-887. doi: 10.1111/j.1365-2214.2010.01093.x.

Skurtveit, S., Selmer, R., Roth, C., Hernandez-Diaz, S. and Handal, M. (2014) 'Prenatal exposure to antidepressants and language competence at age three: results from a large population-based pregnancy cohort in Norway', BJOG, 121, pp. 1621–1632. doi: 10.1111/1471-0528.12821.

Slobin, D.I. (1985) The cross-linguistic study of language acquisition, vols 1 and 2. Lawrence Erlbaum Associates.

Smith, R. (2018) 'Can we improve on Wilson and Jungner's principles of screening for disease?' CMAJ, 9, pp. E414-5. doi: 10.1503/cmaj.180330.

Smith, B.E., Pacheco, M.B. and Khorosheva, M. (2021) 'Emergent Bilingual Students and Digital Multimodal Composition: A Systematic Review of Research in Secondary Classrooms', Reading Research Quarterly, 56(1), pp. 33–52. doi: 10.1002/rrq.298.

Snowling, M.J., Adams, J.W., Bishop, D.V.M. and Stothard, S.E. (2001) 'Educational attainments of school leavers with a preschool history of speech-language impairments', International Journal of Language and Communication Disorders, 36(2), pp. 173-183. doi: 10.1080/13682820010019892.

Spearman, C. (1904) 'The proof and measurement of association between two things', American Journal of Psychology, 15, pp. 72-101.

SpeechLink MultiMedia Limited (2020) Language Link Infant Language Link. Accessible at: https://speechandlanguage.info/infant (Accessed 27 October 2021).

Sperry, D. E., Sperry, L. L. and Miller, P. J. (2019) 'Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds', Child Development, 90(4), pp. 1303–1318. doi: 10.1111/cdev.13072.

Squires, J. and Bricker, D. (2009) Ages and Stages Questionnaires®, Third Edition (ASQ®-3): A Parent-Completed Child Monitoring System. Baltimore: Paul H. Brookes Publishing Co., Inc.

StatsWales (2019) Percentage of adults who speak Welsh by local authority. National Survey for Wales. (Accessed 4 February 2022).

Stipdonk, L.W., Franken, M.C.J. and Dudink, J. (2018) Language outcome related to brain structures in school-aged preterm children: A systematic review. PloS one, 13(6), e0196607. doi: 10.1371/journal.pone.0196607.

Stothard, S.E., Snowling, M.J., Bishop, D.V., Chipchase, B.B. and Kaplan, C.A. (1998) 'Language-impaired preschoolers: A follow-up into adolescence', Journal of Speech, Language, and Hearing Research, 41(2), pp. 407-418. doi: 10.1044/jslhr.4102.407.

Surrain, S. and Luk, G. (2019) 'Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005-2015', Bilingualism, 22(2), pp. 401–415. doi: 10.1017/S1366728917000682.

Swingley, D. and Humphrey, C. (2018) 'Quantitative linguistic predictors of infants' learning of specific english words', Child development, 89(4), pp. 1247–1267. doi: 10.1111/cdev.12731.

Sylva, K., Stein, A., Leach, P., Barnes, J., Malmberg, L. E. and FCCC-team. (2011) 'Effects of early child-care on cognition, language, and task-related behaviours at 18 months: An English study', British Journal of Developmental Psychology, 29(1), pp. 18–45. doi: 10.1348/026151010X533229.

Tamis-LeMonda, C.S., Bornstein, M.H. and Baumwell, L. (2001) 'Maternal responsiveness and children's achievement of language milestones', Child Development, 72(3), pp. 748–767. doi: 10.1111/1467-8624.00313.

Tamis-LeMonda, C.S., Kuchirko, Y. and Song, L. (2014) 'Why Is Infant Language Learning Facilitated by Parental Responsiveness?', Current Directions in Psychological Science, 23(2), pp. 121–126. doi: 10.1177/0963721414522813.

Thomas, E. (2007) 'Natur prosesau caffael iaith gan blant: Marcio cenedl enwau yn y Gymraeg', Gwerddon, 1, pp. 53-81. (Accessed 6 June 2021)

Thomas, E.M. and Mueller Gathercole, V.C. (2007) Children's productive command of grammatical gender and mutation in Welsh: An alternative to rule-based learning. First Language, 27(3), 251-278.

Thomas, E., Williams, N., Jones, L., Davies, S. and Binks, H. (2014) 'Acquiring complex structures under minority language conditions: Bilingual acquisition of plural morphology in Welsh', Bilingualism: Language and Cognition, 17(3), pp. 478-494. doi:10.1017/S1366728913000497.

Thordardottir, E., Rothenberg, A., Rivard, M-E. and Naves, R. (2006) 'Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages?', Journal of Multilingual Communication Disorders, 4(1), pp. 1–21. doi: 10.1080/14769670500215647.

Tomblin, J.B. (1989) Familial concentration of developmental language impairment. Journal of Speech and Hearing Disorders, 54(2), pp. 287-295. doi: 10.1044/jshd.5402.287.

Tomblin, J.B., Harrison, M., Ambrose, S.E., Walker, E.A., Oleson, J.J. and Moeller, M.P. (2015) 'Language outcomes in young children with mild to severe hearing loss', Ear and Hearing, 2(36), pp. 76S-91S. doi: 10.1097/AUD.0000000000000219.

Tricco, A.C., Antony, J., Zarin, W., Strifler, L., Ghassemi, M., Ivory, J., Perrier, L., Hutton, B., Moher, D. and Straus, S.E. (2015) 'A scoping review of rapid review methods', BMC Medicine, 13(1). doi: 10.1186/s12916-015-0465-6.

Tricco, A.C., Thomas, S.M., Antony, J., Rios, P., Robson, R., Pattani, R., Ghassemi, M., Sullivan, S., Selvaratnam, I., Tannenbaum, C. and Straus, S.E. (2017) 'Strategies to prevent or reduce gender bias in peer review of research grants: A rapid scoping review', PLoS ONE, 12(1), pp. 1–12. doi: 10.1371/journal.pone.0169718.

Ukoumunne, O.C., Wake, M., Carlin, J., Bavin, E.L., Lum, J., Skeat, J., ... and Reilly, S. (2012) 'Profiles of language development in pre-school children: A longitudinal latent class analysis of data from the Early Language in Victoria Study', Child: care, health and development, 38(3), pp. 341-349. doi: 10.1111/j.1365-2214.2011.01234.x.

Unicef. (2012) Inequities in early childhood development: What the data say: Evidence from the Multiple Indicator Cluster Surveys. New York: UNICEF. (Accessed 31 December 2021).

Vagh, S.B., Pan, B.A. and Mancilla-Martinez, J. (2009) 'Measuring Growth in Bilingual and Monolingual Children ' s English Productive Vocabulary Development : The Utility of Combining Parent and Teacher Report Children ' s Measuring Growth in Bilingual and Monolingual The Utility of Combining English Productiv', Child Development, 80(5), pp. 1545–1563. doi: 10.1111/j.1467-8624.2009.01350.x.

Van der Vleuten, C. (2000) 'Validity of final examinations in undergraduate medical training', BMJ, 321(7270), pp. 1217-1219. doi: 10.1136/bmj.321.7270.1217.

Van der Vleuten, C. P. (1996) 'The assessment of professional competence: developments, research and practical implications', Adv Health Sci Educ, 1, pp. 41-67. doi: 10.1007/BF00596229.

Wake, M., Tobin., Girolametto,L., Ukoumunne, O.C ., Gold, L., Levikis, P., Sheehan, J., Goldfeld, S. and Reilly, S. (2011) 'Outcomes of population based language promotion for slow to talk toddlers at 2 and 3 years: Let's Learn Language cluster randomised controlled trial', British Medical Journal, 343(7821), pp. 460 – 470. doi: 10.1136/bmj.d4741.

Weikum, W.M., Oberlander, T.F., Hensch, T.K. and Werker, J.F. (2012) 'Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech

perception', Proceedings of the National Academy of Sciences, 109(Supplement 2), pp. 17221–17227. doi: 0.1073/pnas.1121263109.

Weinberg, N.Z (1997). Cognitive and behavioral deficits associated with parental alcohol use. Journal of the American Academy of Child & Adolescent Psychiatry, 36(9), pp. 1177–1186. doi: 10.1097/00004583-199709000-00009.

Weisleder, A. and Fernald, A. (2017) 'Talking to children matters: Early language experience strengthens processing and builds vocabulary', 24(11), pp. 2143–2152. doi: 10.1177/0956797613488145.Talking.

Welsh Government (2014) 'A review of research evidence on the effectiveness of different approaches to promoting early speech and language development'. Authors: Duggan, B., Dosmukhambetova, D. and Edwards, V. (Accessed 14 September 2021).

Welsh Government. (2016) 'Prudent Healthcare: Securing Health and Wellbeing for Future Generations Prudent healthcare', (Accessed 31 December 2021)

Welsh Government. (2016) An overview of the Healthy Child Wales Programme. (Accessed 14 June 2021)

Welsh Government (2016) Healthy Child Wales Programme. (Accessed 30 December 2021)

Welsh Government. (2017) Flying Start Annex - Guidance on Speech, Language and Communication. Flying Start speech, language and communication: guidance. (Accessed 20 January 2022).

Welsh Government. (2020) Siarad Gyda Fi/Talk with me: Speech, Language and Communication (SLC) Delivery Plan. (Accessed 14 June 2021).

West, G., Snowling, M.J., Lervag, A., Buchanan-Worster, E., Duta, M., Hall, A., McLachlan, H. and Hulme, C. (2021) 'Early language screening and intervention can be delivered successfully at scale: Evidence from a cluster randomized control trial', Journal of Child Psychology and Psychiatry, 62(12), pp. 1425-1434. doi: 10.1111/jcpp.13415.

Williams, C.J. and McLeod, S. (2012) 'Speech-language pathologists' assessment and intervention practices with multilingual children', International Journal of Speech-Language Pathology, 14(3), pp. 292–305. doi: 10.3109/17549507.2011.636071.

Williams, M. and Cooper, S. (2021) 'Adult new speakers of Welsh: Accent, pronunciation and language experience in South Wales', Languages, 6(2), 86, doi: 10.3390/languages6020086.

Williams, M.E., Hutchings, J., Bywater, T., Daley, D. and Whitaker, C.J. (2013) 'Schedule of Growing Skills II: Pilot Study of an Alternate Scoring Method', Psychology, 4(3), pp. 143-152. doi: 10.4236/PSYCH.2013.43021.

Willinger, U., Brunner, E., Diendorfer-Radner, G., Sams, J., Sirsch, U. and Eisenwort, B. (2003) 'Behaviour in children with language development disorders', Canadian Journal of Psychiatry, 48, pp. 607–614. doi: 10.1177/070674370304800907.

Wilson, J.M.G. and Jungner, G. (1968) Principles and practice of screening for disease. Geneva: World Health Organisation.

Wren, Y., Hambly, H. and Roulstone, S. (2013) 'A review of the impact of bilingualism on the development of phonemic awareness skills in children with typical speech development', Child Language Teaching and Therapy, 29(1), pp. 11-25. doi: 10.1177/0265659012464880.

Wren, Y., Miller, L. L., Peters, T. J., Emond, A. and Roulstone, S. (2016) 'Prevalence and predictors of persistent speech sound disorder at eight years old: Findings from a population cohort study', Journal of Speech, Language, and Hearing Research, 59(4), pp. 647-673. doi: 10.1044/2015_JSLHR-S-14-0282.

Wren, Y., Pagnamenta, E., Peters, T. J., Emond, A., Northstone, K., Miller, L. L. and Roulstone, S. (2021) Educational outcomes associated with persistent speech disorder. International Journal of Language & Communication Disorders, 56(2), pp. 299-312. doi: 10.1111/1460-6984.12599.

Wright Karem, R., Washington, K.N., Crowe, K., Jenkins, A., Leon, M., Kokotek, L., Raisor-Becker, L. and Westby, C. (2019) 'Current Methods of Evaluating the Language Abilities of Multilingual Preschoolers: A Scoping Review using the International Classification of Functioning, Disability and Health - Children and Youth Version', Language Speech Hear Serv Sch, 50, pp. 434–451. doi: 10.1044/2019_LSHSS-18-0128.

Wylie, K., McAllister, L., Davidson, B., Marshall, J. and Law, J. (2014) 'Adopting public health approaches to communication disability: Challenges for the education of speech-language pathologists', Folia Phoniatrica et Logopaedica, 66(4-5), pp. 164-175. doi: 10.1159/000365752.

Zambrana, I. M., Ystrom, E. and Pons, F. (2012) 'Impact of gender, maternal education, and birth order on the development of language comprehension: a longitudinal study from 18 to 36 months of age', Journal of Developmental and Behavioral Pediatrics, 33(2), pp. 146–155. doi: 10.1097/DBP.0b013e31823d4f83.

Zimmerman, F.J., Christakis, D.A. and Meltzoff, A.N. (2007) Associations between media viewing and language development in children under age 2 years. The Journal of pediatrics, 151(4), pp. 364-368. doi: 10.1016/j.jpeds.2007.04.071.

Zimmerman, F.J., Gilkerson, J., Richards, J.A., Christakis, D.A., Xu, D., Gray, S. and Yapanel, U. (2009) Teaching by listening: The importance of adult-child conversations to language development. Pediatrics, 124(1), pp. 342-349. doi: 10.1542/peds.2008-2267.

Annex A: Search strategies for literature review

English Language Screening Tools

Medline

1 Child, preschool/

2 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,kf.

3 1 or 2

4 language development/

5 communication disorders/

6 language disorders/

7 language development disorders/

8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

10 speech disorders/

11 voice disorders/

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp mass screening/

14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kf.

15 screen*.ti,ab,kf.#

16 diagnos*.ti,ab,kf.

17 detect*.ti,ab,kf.

18 "surveillance".ti,ab,kf.

19 exp "diagnostic techniques and procedures"/

20 13 or 14 or 15 or 16 or 17 or 18 or 19

21 exp evaluation studies/

22 "predictive value of tests".ti,ab,kf.

23 "sensitivity and specificity".ti,ab,kf.

24 (predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity).ti,ab,kf.

25 (acceptability or satisfaction).ti,ab,kf.

26 21 or 22 or 23 or 24 or 25

27 3 and 12 and 20 and 26

28 limit 27 to yr="2015 -Current"

PsycInfo

1 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,id.

2 language development/

3 language disorders/

4 language delay/

5 communication disorders/

6 speech disorders/

7 speech development/

8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.

9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.

10 exp screening/

11 exp screening tests/

12 exp diagnosis/

13 exp evaluation/

14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,id.

15 screen*.ti,ab,id.

16 diagnos*.ti,ab,id.

17 detect*.ti,ab,id.

18 "surveillance".ti,ab,id.

19 predictive validity/

20 test validity/

21 test reliability/

22 test standardization/

23 "predictive value of tests".ti,ab,id.

24 "sensitivity and specificity".ti,ab,id.

25 (predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity).ti,ab,id.

26 (acceptability or satisfaction).ti,ab,id.

27 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9

28 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18

29 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26

30 1 and 27 and 28 and 29

31 limit 30 to yr="2015 -Current"

CINAHL

Limit Published Date: 20150101-20211231

S28 S1 AND S9 AND S17 AND S26

S27 S1 AND S9 AND S17 AND S26

S26 S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25

S25 TI ((acceptability or satisfaction)) OR AB ((acceptability or satisfaction)) OR MW ((acceptability or satisfaction))

S24 TI ((predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity)) OR AB ((predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity)) OR MW ((predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity))

S23 TI ("sensitivity and specificity") OR AB ("sensitivity and specificity") OR MW ("sensitivity and specificity")

S22 TI "predictive value of tests" OR AB "predictive value of tests" OR MW "predictive value of tests"

S21 MH “reliability and validity”

S20 MH "Instrument validation"

S19 MH "Predictive validity"

S18 MH "validation studies"

S17 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16

S16 TI "surveillance" OR AB "surveillance" OR MW "surveillance"

S15 TI detect* OR AB detect* OR MW detect*

S14 TI diagnos* OR AB diagnos* OR MW diagnos*

S13 TI screen* OR AB screen* OR MW screen*

S12 TI ((parent* n3 (checklist* or survey* or questionnaire*))) OR AB ((parent* n3 (checklist* or survey* or questionnaire*))) OR MW ((parent* n3 (checklist* or survey* or questionnaire*)))

S11 MH "diagnosis+"

S10 MH ("health screening+")

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR AB ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR MW ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))))

S7 TI (((language n3 (impair* or delay* or disorder* or difficult*)))) OR AB (((language n3 (impair* or delay* or disorder* or difficult*)))) OR MW (((language n3 (impair* or delay* or disorder* or difficult*))))

S6 MH "voice disorders"

S5 MH "speech disorders"

S4 MH "communicative disorders"

S3 MH "language disorders"

S2 MH "language development"

S1 TI ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR AB ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR MW ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")))

ERIC

Limit Date Published: 20150101-20211231

S27 S1 AND S9 AND S18 AND S25

S26 S1 AND S9 AND S18 AND S25

S25 S19 OR S20 OR S21 OR S22 OR S23 OR S24

S24 TI ((acceptability or satisfaction)) OR AB ((acceptability or satisfaction)) OR KW ((acceptability or satisfaction))

S23 TI ((predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity)) OR AB ((predictive or concurrent* or valid* or reliab* or

standardi* or sensitivity or specificity)) OR KW ((predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity))

S22 TI ("sensitivity and specificity") OR AB ("sensitivity and specificity") OR KW ("sensitivity and specificity")

S21 SU “reliability”

S20 SU "predictive validity"

S19 SU "validity"

S18 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17

S17 TI "surveillance" OR AB "surveillance" OR KW "surveillance"

S16 TI detect* OR AB detect* OR KW detect*

S15 TI diagnos* OR AB diagnos* OR KW diagnos*

S14 TI screen* OR AB screen* OR KW screen*

S13 TI (parent* n3 (checklist* or survey* or questionnaire*)) OR AB (parent* n3 (checklist* or survey* or questionnaire*)) OR KW (parent* n3 (checklist* or survey* or questionnaire*))

S12 SU "clinical diagnosis"

S11 SU "identification"

S10 SU "screening tests"

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR AB (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR KW (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))

S7 TI ((language n3 (impair* or delay* or disorder* or difficult*))) OR AB ((language n3 (impair* or delay* or disorder* or difficult*))) OR KW ((language n3 (impair* or delay* or disorder* or difficult*)))

S6 SU "voice disorders"

S5 SU "speech impairments"

S4 SU "communication disorders"

S3 SU "language impairments"

S2 SU "language acquisition"

S1 TI ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")) OR AB ((child* or baby or babies or infant* or toddler* or

preschool* or "pre-school*")) OR KW ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))

Embase

1 preschool child/ 5

2 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,kw.

3 1 or 2

4 language development/

5 communication disorder/

6 language disability/

7 developmental language disorder/

8 speech disorder/

9 voice disorder/

10 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

11 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp screening/

14 exp screening test/

15 exp language test/

16 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kw.

17 screen*.ti,ab,kw.

18 diagnos*.ti,ab,kw.

19 detect*.ti,ab,kw.

20 "surveillance".ti,ab,kw.

21 exp diagnosis/

22 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23 exp evaluation study/

24 predictive validity/

25 validity/

26 reliability/

27 "predictive value of tests".ti,ab,kw.

28 "sensitivity and specificity".ti,ab,kw.

29 (predictive or concurrent* or valid* or reliab* or standardi* or sensitivity or specificity).ti,ab,kw.

30 (acceptability or satisfaction).ti,ab,kw.

31 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30

32 3 and 12 and 22 and 31

33 limit 32 to yr="2015 -Current"

Bilingualism Review

Medline

1 Child, preschool/

2 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,kf.

3 1 or 2

4 language development/

5 communication disorders/

6 language disorders/

7 language development disorders/

8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

10 speech disorders/

11 voice disorders/

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp mass screening/

14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kf.

15 screen*.ti,ab,kf.

16 diagnos*.ti,ab,kf.

17 detect*.ti,ab,kf.

18 "surveillance".ti,ab,kf.

19 exp "diagnostic techniques and procedures"/

20 13 or 14 or 15 or 16 or 17 or 18 or 19

21 bilingu*.ti,ab,kf.
22 multilingu*.ti,ab,kf.
23 multi-lingu*.ti,ab,kf.
24 "dual language".ti,ab,kf.
25 ("ESL" or "EAL" or "English as an additional language").ti,ab,kf.
26 "first language".ti,ab,kf.
27 "second language".ti,ab,kf.
28 "linguistically diverse".ti,ab,kf.
29 translingu*.ti,ab,kf.
30 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29
31 3 and 12 and 20 and 30
32 limit 31 to yr="2000 -Current"

PsycInfo

1 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,id.
2 language development/
3 language disorders/
4 language delay/
5 communication disorders/
6 speech disorders/
7 speech development/
8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.
9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.
10 exp screening/
11 exp screening tests/
12 exp diagnosis/
13 exp evaluation/
14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,id.
15 screen*.ti,ab,id.
16 diagnos*.ti,ab,id.
17 detect*.ti,ab,id.

18 "surveillance".ti,ab,id.

19 multilingualism/

20 bilingualism/

21 english as second language/

22 bilingu*.ti,ab,id.

23 multilingu*.ti,ab,id.

24 multi-lingu*.ti,ab,id.

25 "dual language".ti,ab,id.

26 ("ESL" or "EAL" or "English as an additional language").ti,ab,id.

27 "first language".ti,ab,id.

28 "second language".ti,ab,id.

29 "linguistically diverse".ti,ab,id.

30 translingu*.ti,ab,id.

31 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30

32 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9

33 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18

34 1 and 31 and 32 and 33

35 limit 34 to yr="2000 -Current"

CINAHL

Limit Published Date: 20000101-20211231

S31 S1 AND S9 AND S17 AND S29

S30 S1 AND S9 AND S17 AND S29

S29 S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S2

S28 TI translingu* OR AB translingu* OR MW translingu*

S27 TI "linguistically diverse" OR AB "linguistically diverse" OR MW "linguistically diverse”

S26 TI "second language" OR AB "second language" OR MW "second language"
3

S25 TI "first language" OR AB "first language" OR MW "first language"

S24 TI (("ESL" or "EAL" or "English as an additional language")) OR AB (("ESL" or "EAL" or "English as an additional language")) OR MW (("ESL" or "EAL" or "English as an additional language"))

S23 TI "dual language" OR AB "dual language" OR MW "dual language"

S22 TI multi-lingu* OR AB multi-lingu* OR MW multi-lingu*

S21 TI multilingu* OR AB multilingu* OR MW multilingu*

S20 TI bilingu* OR AB bilingu* OR MW bilingu*

S19 MH "multilingualism"

S18 MH "english as a second language"

S17 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16

S16 TI "surveillance" OR AB "surveillance" OR MW "surveillance"

S15 TI detect* OR AB detect* OR MW detect*

S14 TI diagnos* OR AB diagnos* OR MW diagnos*

S13 TI screen* OR AB screen* OR MW screen*

S12 TI ((parent* n3 (checklist* or survey* or questionnaire*))) OR AB ((parent* n3 (checklist* or survey* or questionnaire*))) OR MW ((parent* n3 (checklist* or survey* or questionnaire*)))

S11 MH "diagnosis+"

S10 MH ("health screening+")

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR AB ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR MW ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))))

S7 TI (((language n3 (impair* or delay* or disorder* or difficult*)))) OR AB (((language n3 (impair* or delay* or disorder* or difficult*)))) OR MW (((language n3 (impair* or delay* or disorder* or difficult*))))

S6 MH "voice disorders"

S5 MH "speech disorders"

S4 MH "communicative disorders"

S3 MH "language disorders"

S2 MH "language development"

S1 TI ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR AB ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR MW ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))

ERIC

Limit Date Published: 20000101-20211231

S33 S1 AND S9 AND S18 AND S31

S32 S1 AND S9 AND S18 AND S31

S31 S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30

S30 TI translingu* OR AB translingu* OR KW translingu*

S29 TI "linguistically diverse" OR AB "linguistically diverse" OR KW "linguistically diverse"

S28 TI "second language" OR AB "second language" OR KW "second language"

S27 TI "first language" OR AB "first language" OR KW "first language"

S26 TI (("ESL" or "EAL" or "English as an additional language")) OR AB (("ESL" or "EAL" or "English as an additional language")) OR KW (("ESL" or "EAL" or "English as an additional language"))

S25 TI "dual language" OR AB "dual language" OR KW "dual language"

S24 TI multi-lingu* OR AB multi-lingu* OR KW multi-lingu*

S23 TI multilingu* OR AB multilingu* OR KW multilingu*

S22 TI bilingu* OR AB bilingu* OR KW bilingu*

S21 SU "English (Second language)”

S20 SU "bilingualism"

S19 SU "multilingualism"

S18 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17

S17 TI "surveillance" OR AB "surveillance" OR KW "surveillance"

S16 TI detect* OR AB detect* OR KW detect*

S15 TI diagnos* OR AB diagnos* OR KW diagnos*

S14 TI screen* OR AB screen* OR KW screen*

S13 TI (parent* n3 (checklist* or survey* or questionnaire*)) OR AB (parent* n3 (checklist* or survey* or questionnaire*)) OR KW (parent* n3 (checklist* or survey* or questionnaire*))

S12 SU "clinical diagnosis"

S11 SU "identification"

S10 SU "screening tests"

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR AB (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR KW (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))

S7 TI ((language n3 (impair* or delay* or disorder* or difficult*))) OR AB ((language n3 (impair* or delay* or disorder* or difficult*))) OR KW ((language n3 (impair* or delay* or disorder* or difficult*)))

S6 SU "voice disorders"

S5 SU "speech impairments"

S4 SU "communication disorders"

S3 SU "language impairments"

S2 SU "language acquisition"

S1 TI ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")) OR AB ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")) OR KW ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))

Embase

1 preschool child/

2 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,kw.

3 1 or 2

4 language development/

5 communication disorder/

6 language disability/

7 developmental language disorder/

8 speech disorder/

9 voice disorder/

10 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

11 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp screening/

14 exp screening test/

15 exp language test/

16 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kw.

17 screen*.ti,ab,kw.

18 diagnos*.ti,ab,kw.

19 detect*.ti,ab,kw.

20 "surveillance".ti,ab,kw.

21 exp diagnosis/

22 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23 multilingualism/

24 bilingualism/

25 english as a second language/

26 bilingu*.ti,ab,kw.

27 multilingu*.ti,ab,kw.

28 multi-lingu*.ti,ab,kw.

29 "dual language".ti,ab,kw.

30 ("ESL" or "EAL" or "English as an additional language").ti,ab,kw.

31 "first language".ti,ab,kw.

32 "second language".ti,ab,kw.

33 "linguistically diverse".ti,ab,kw.

34 translingu*.ti,ab,kw.

35 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34

36 3 and 12 and 22 and 35

37 limit 36 to yr="2000 -Current"

Welsh review (conducted in English language)

Medline

1 Child, preschool/

2 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,kf.

3 1 or 2

4 language development/

5 communication disorders/

6 language disorders/

7 language development disorders/

8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kf.

10 speech disorders/

11 voice disorders/

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp mass screening/

14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kf.

15 screen*.ti,ab,kf.

16 diagnos*.ti,ab,kf

17 detect*.ti,ab,kf.

18 "surveillance".ti,ab,kf.

19 exp "diagnostic techniques and procedures"/

20 13 or 14 or 15 or 16 or 17 or 18 or 19

21 exp Wales/ not New South Wales.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

22 "Welsh".ti,ab,kf.

23 "Cymru".ti,ab,kf.

24 "Cymraeg".ti,ab,kf.

25 21 or 22 or 23 or 24

26 3 and 12 and 20 and 25

27 limit 26 to yr="2000 -Current"

PsycInfo

1 (child* or baby or babies or infant* or toddler* or preschool* or "pre-school*").ti,ab,id.

2 language development/

3 language disorders/

4 language delay/

5 communication disorders/

6 speech disorders/

7 speech development/

8 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.

9 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,id.

10 exp screening/

11 exp screening tests/

12 exp diagnosis/

13 exp evaluation/

14 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,id.

15 screen*.ti,ab,id.

16 diagnos*.ti,ab,id.

17 detect*.ti,ab,id.

18 "surveillance".ti,ab,id.

19 ("wales" not "new south wales").ti,ab,id.

20 "welsh".ti,ab,id.

21 "cymru".ti,ab,id.

22 "cymraeg".ti,ab,id.

23 19 or 20 or 21 or 22

24 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9

25 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18

26 1 and 23 and 24 and 25

27 limit 26 to yr="2000 -Current"

CINAHL

Limit Published Date: 20000101-20211231

S24 S1 AND S9 AND S17 AND S22

S23 S1 AND S9 AND S17 AND S22

S22 S18 OR S19 OR S20 OR S21

S21 TI "cymraeg" OR AB "cymraeg" OR MW "cymraeg"

S20 TI "cymru" OR AB "cymru" OR MW "cymru"

S19 TI "welsh" OR AB "welsh" OR MW "welsh"

S18 MH "wales+" NOT MH "new south wales"

S17 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16

S16 TI "surveillance" OR AB "surveillance" OR MW "surveillance"

S15 TI detect* OR AB detect* OR MW detect*

S14 TI diagnos* OR AB diagnos* OR MW diagnos*

S13 TI screen* OR AB screen* OR MW screen*

S12 TI ((parent* n3 (checklist* or survey* or questionnaire*))) OR AB ((parent* n3 (checklist* or survey* or questionnaire*))) OR MW ((parent* n3 (checklist* or survey* or questionnaire*)))

S11 MH "diagnosis+"

S10 MH ("health screening+")

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR AB ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))) OR MW ((((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))))

S7 TI (((language n3 (impair* or delay* or disorder* or difficult*)))) OR AB (((language n3 (impair* or delay* or disorder* or difficult*)))) OR MW (((language n3 (impair* or delay* or disorder* or difficult*))))

S6 MH "voice disorders"

S5 MH "speech disorders"

S4 MH "communicative disorders"

S3 MH "language disorders"

S2 MH "language development"

S1 TI ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR AB ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))) OR MW ((((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")))

ERIC

Limit Date Published: 20000101-20181231

S27 S1 AND S9 AND S18 AND S26

S26 S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25

S25 KW "wales" NOT KW "new south wales"

S24 AB "wales" NOT AB "new south wales"

S23 TI "wales" NOT TI "new south wales"

S22 TI "cymru" OR AB "cymru" OR KW "cymru"

S21 TI "cymraeg" OR AB "cymraeg" OR KW "cymraeg"

S20 SU "Welsh"

S19 TI "Welsh" OR AB "Welsh" OR KW "Welsh"

S18 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17

S17 TI "surveillance" OR AB "surveillance" OR KW "surveillance"

S16 TI detect* OR AB detect* OR KW detect*

S15 TI diagnos* OR AB diagnos* OR KW diagnos*

S14 TI screen* OR AB screen* OR KW screen*

S13 TI (parent* n3 (checklist* or survey* or questionnaire*)) OR AB (parent* n3 (checklist* or survey* or questionnaire*)) OR KW (parent* n3 (checklist* or survey* or questionnaire*))

S12 SU "clinical diagnosis"

S11 SU "identification"

S10 SU "screening tests"

S9 S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 TI (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR AB (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*))) OR KW (((speech* or phon*) n3 (impair* or delay* or disorder* or difficult*)))

S7 TI ((language n3 (impair* or delay* or disorder* or difficult*))) OR AB ((language n3 (impair* or delay* or disorder* or difficult*))) OR KW ((language n3 (impair* or delay* or disorder* or difficult*)))

S6 SU "voice disorders"

S5 SU "speech impairments"

S4 SU "communication disorders"

S3 SU "language impairments"

S2 SU "language acquisition"

S1 TI ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")) OR AB ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*")) OR KW ((child* or baby or babies or infant* or toddler* or preschool* or "pre-school*"))

S28 S1 AND S9 AND S18 AND S26

Embase

1 preschool child/

2 (child* or baby or babes or infant* or toddler* or preschool* or "pre-school*").ti,ab,kw.

3 1 or 2

4 language development/

5 communication disorder/

6 language disability/

7 developmental language disorder/

8 speech disorder/

9 voice disorder/

10 (language adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

11 ((speech* or phon*) adj3 (impair* or delay* or disorder* or difficult*)).ti,ab,kw.

12 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11

13 exp screening/

14 exp screening test/

15 exp language test/

16 (parent* adj3 (checklist* or survey* or questionnaire*)).ti,ab,kw.

17 screen*.ti,ab,kw.

18 diagnos*.ti,ab,kw.

19 detect*.ti,ab,kw.

20 "surveillance".ti,ab,kw.

21 exp diagnosis/

22 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23 exp wales/ not "new south wales".mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]

24 "welsh".ti,ab,kw.

25 "cymru".ti,ab,kw.

26 "cymraeg".ti,ab,kw.

27 23 or 24 or 25 or 26

28 3 and 12 and 22 and 27

29 limit 28 to yr="2000 -Current"

Welsh review (conducted in Welsh)

Medline

1 "plentyn".ti,ab,kf.

2 cyn-ysgol.ti,ab,kf.

3 "babi".ti,ab,kf.

4 "baban".ti,ab,kf.

5 "babanod".ti,ab,kf.

6 1 or 2 or 3 or 4 or 5

7 datblygiad iaith.ti,ab,kf.

8 (Anhwylder or anhwylderau cyfathrebu).ti,ab,kf.

9 anhwylderau iaith.ti,ab,kf.

10 anhwylderau datblygiad iaith.ti,ab,kf.

11 (iaith adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,kf.

12 ((lleferydd or ffonoleg or seineg or ffoneteg) adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,kf.
13 sgrinio.ti,ab,kf.
14 ((Rhiant or rhieni) adj3 (Rhestr wirioor or arolwg or holiadur)).ti,ab,kf.
15 Deiagnosis.ti,ab,kf.
16 Canfod.ti,ab,kf.
17 Gwyliadwriaeth.ti,ab,kf.
18 Technegau a prosesau diagnostig.ti,ab,kf.
19 Astudiaethau gwerthuso.ti,ab,kf.
20 7 or 8 or 9 or 10 or 11 or 12
21 13 or 14 or 15 or 16 or 17 or 18 or 19
22 6 and 20 and 21

PsycInfo

1 "plentyn".ti,ab,id.
2 cyn-ysgol.ti,ab,id.
3 "babi".ti,ab,id.
4 "baban".ti,ab,id.
5 "babanod".ti,ab,id.
6 1 or 2 or 3 or 4 or 5
7 datblygiad iaith.ti,ab,id.
8 (Anhwylder or anhwylderau cyfathrebu).ti,ab,id.
9 anhwylderau iaith.ti,ab,id.
10 anhwylderau datblygiad iaith.ti,ab,id.
11 (iaith adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,id.
12 ((lleferydd or ffonoleg or seineg or ffoneteg) adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,id.
13 sgrinio.ti,ab,id.
14 ((Rhiant or rhieni) adj3 (Rhestr wirioor or arolwg or holiadur)).ti,ab,id.
15 Deiagnosis.ti,ab,id.
16 Canfod.ti,ab,id.
17 Gwyliadwriaeth.ti,ab,id.

18 Technegau a prosesau diagnostig.ti,ab,id.

19 Astudiaethau gwerthuso.ti,ab,id.

20 7 or 8 or 9 or 10 or 11 or 12

21 13 or 14 or 15 or 16 or 17 or 18 or 19

22 6 and 20 and 21

CINAHL

S22 S6 AND S13 AND S21

S21 S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20

S20 TI Astudiaethau gwerthuso OR AB Astudiaethau gwerthuso OR MW Astudiaethau gwerthuso

S19 TI Technegau a prosesau diagnostig OR AB Technegau a prosesau diagnostig OR MW Technegau a prosesau diagnostig

S18 TI Gwyliadwriaeth OR AB Gwyliadwriaeth OR MW Gwyliadwriaeth

S17 TI Canfod OR AB Canfod OR MW Canfod

S16 TI Deiagnosis OR AB Deiagnosis OR MW Deiagnosis

S15 TI (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur))) OR AB (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur))) OR MW (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur)))

S14 TI sgrinio OR AB sgrinio OR MW sgrinio

S13 S7 OR S8 OR S9 OR S10 OR S11 OR S12

S12 TI (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR AB (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR MW (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau)))

S11 TI ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR AB ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR MW ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau)))

S10 TI anhwylderau datblygiad iaith OR AB anhwylderau datblygiad iaith OR MW anhwylderau datblygiad iaith

S9 TI anhwylderau iaith OR AB anhwylderau iaith OR MW anhwylderau iaith

S8 TI ((Anhwylder or anhwylderau cyfathrebu)) OR AB ((Anhwylder or anhwylderau cyfathrebu)) OR MW ((Anhwylder or anhwylderau cyfathrebu))

S7 TI datblygiad iaith OR AB datblygiad iaith OR MW datblygiad iaith

S6 S1 OR S2 OR S3 OR S4 OR S5

S5 TI "babanod" OR AB "babanod" OR MW "babanod"

S4 TI "baban" OR AB "baban" OR MW "baban"

S3 TI "babi" OR AB "babi" OR MW "babi"

S2 TI cyn-ysgol OR AB cyn-ysgol OR MW cyn-ysgol

S1 TI "plentyn" OR AB "plentyn" OR MW "plentyn"

ERIC

S22 S6 AND S13 AND S21

S21 S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20

S20 TI Astudiaethau gwerthuso OR AB Astudiaethau gwerthuso OR KW Astudiaethau gwerthuso

S19 TI Technegau a prosesau diagnostig OR AB Technegau a prosesau diagnostig OR KW Technegau a prosesau diagnostig

S18 TI Gwyliadwriaeth OR AB Gwyliadwriaeth OR KW Gwyliadwriaeth 0

S17 TI Canfod OR AB Canfod OR KW Canfod

S16 TI Deiagnosis OR AB Deiagnosis OR KW Deiagnosis

S15 TI (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur))) OR AB (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur))) OR KW (((Rhiant or rhieni) n3 (Rhestr wirioor or arolwg or holiadur)))

S14 TI sgrinio OR AB sgrinio OR KW sgrinio

S13 S7 OR S8 OR S9 OR S10 OR S11 OR S12

S12 TI (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR AB (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))) OR KW (((lleferydd or ffonoleg or seineg or ffoneteg) n3 (Nam or oedi or Anhwylder or anhawster or anawsterau)))

S11 TI ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau)) OR AB ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau)) OR KW ((iaith n3 (Nam or oedi or Anhwylder or anhawster or anawsterau))

S10 TI anhwylderau datblygiad iaith OR AB anhwylderau datblygiad iaith OR KW anhwylderau datblygiad iaith

S9 TI anhwylderau iaith OR AB anhwylderau iaith OR KW anhwylderau iaith

S8 TI ((Anhwylder or anhwylderau cyfathrebu)) OR AB ((Anhwylder or anhwylderau cyfathrebu)) OR KW ((Anhwylder or anhwylderau cyfathrebu))

S7 TI datblygiad iaith OR AB datblygiad iaith OR KW datblygiad iaith

S6 S1 OR S2 OR S3 OR S4 OR S5

S5 TI "babanod" OR AB "babanod" OR KW "babanod"

S4 TI "baban" OR AB "baban" OR KW "baban"

S3 TI "babi" OR AB "babi" OR KW "babi"

S2 TI cyn-ysgol OR AB cyn-ysgol OR KW cyn-ysgol

S1 TI "plentyn" OR AB "plentyn" OR KW "plentyn"

Embase

1 "plentyn".ti,ab,kw.

2 cyn-ysgol.ti,ab,kw.

3 "babi".ti,ab,kw.

4 "baban".ti,ab,kw.

5 "babanod".ti,ab,kw.

6 1 or 2 or 3 or 4 or 5

7 datblygiad iaith.ti,ab,kw.

8 (Anhwylder or anhwylderau cyfathrebu).ti,ab,kw.

9 anhwylderau iaith.ti,ab,kw.

10 anhwylderau datblygiad iaith.ti,ab,kw.

11 (iaith adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,kw.

12 ((lleferydd or ffonoleg or seineg or ffoneteg) adj3 (Nam or oedi or Anhwylder or anhawster or anawsterau)).ti,ab,kw.

13 7 or 8 or 9 or 10 or 11 or 12

14 sgrinio.ti,ab,kw.

15 ((Rhiant or rhieni) adj3 (Rhestr wirioor or arolwg or holiadur)).ti,ab,kw.

16 Deiagnosis.ti,ab,kw.

17 Canfod.ti,ab,kw.

18 Gwyliadwriaeth.ti,ab,kw.

19 Technegau a prosesau diagnostig.ti,ab,kw.

20 Astudiaethau gwerthuso.ti,ab,kw.

21 14 or 15 or 16 or 17 or 18 or 19 or 20

22 6 and 13 and 21 0

Annex B: Data charting table for reviews

Data Charting Evidence source details and characteristics	
Citation details • Authors • Title • Year • Journal	
Study aim	
Study population • Number of participants • Child age range • Home language(s) • Type of bilingualism (Welsh/bilingualism review only) • Child gender • Demographic information (e.g. SES, ethnicity)	
Screening details • Main screener/tool used • Reference tools used • Aspect(s) of speech & language assessed • Type of screening tool (e.g. parent report, observation, direct assessment) • Who completed the screening tool? • Was the person completing the tool a native speaker (Welsh/bilingualism review only)	

Results • Study methodology • Statistical outcomes (sensitivity, specificity, validity, reliability) • Summary of language outcomes • Free-text – what does this tell us about screening?	-

Annex C: Full table of studies included in reviews

English language screening tools: Table of demographic information in studies

Reference			Study design		Study population				
Authors & Year	**Title**	**Journal**	**Aim**	**Methodology**	**No. participants**	**Child age range**	**Home language(s)**	**Child gender**	**Demographic Information**
Chilosi et al. (2019)	Which linguistic measures distinguish transient from persistent language problems in Late Talkers from 2 to 4 years? A study on Italian speaking children	Research in Developmental Disabilities	To identify the early language characteristics of LTs (late talkers) whose outcome was either a transient delay or a Developmental Language Disorder (DLD). The aim of our prospective study was to provide some diagnostic guidelines based on LTs early linguistic profiles and developmental language trajectories.	Cohort	50	18-34 months (mean: 27.7 months)	Italian	37 boys, 13 girls	Inclusion criteria: Italian as native and only language used in everyday life; an expressive vocabulary, evaluated with the Italian version of the CDI at or below 10th centile and/or absence of combinatorial language at 30 months; Absence of auditory, neurological and socio-emotional disorders, and no psychomotor delay; normal developmental history, as reported by parents; Non-verbal IQ in the normal range (mean IQ of sample=99, SD=9.4); Language assessment with at least two evaluations between 2 and 4 years

Claessen et al. (2017)	Is two too early? Assessing toddlers' phonology	Speech, Language and Hearing	Will the TPT measures of PCC, PVC, and the number of atypical phonological errors on two assessments, between four and nine months apart, be positively correlated?	Cohort with two sub groups defined by age	24	2 age groups: 6 girls & 5 boys aged 25–29 months (mean 26.6 months, SD 1.5); 5 girls, 8 boys aged 30–35 months (mean 32.6 months, SD 2.1)	English	13 boys, 11 girls	No information provided
Gilkerson et al. (2017)	Language assessment in a snap: Monitoring progress up to 36 months	Child Language Teaching & Therapy	The development and validation of the Developmental Snapshot, a 52-item parent questionnaire on child language and vocal communication development that can be administered monthly and scored automatically	Phase I: item development/ refinement and initial pilot testing of measurement tool; Phase II: analysis of tool psychometrics	Items piloted by 15 families; refinement/further development with parents of 308 typically developing children	3 - 41 months (pilot phase); 2-49 months (follow-up phase)	English	15 pilot children: no information; 308 follow-up children: 156 boys,	15 pilot parents: all monolingual English speakers; 87% White; 10 college degree; 4 associate's degree; 1 high school. 380 follow-up parents: 11% some high school; 14% high school; 28% some college; 15% college degree

								152 girls	
Gudmundsson (2015)	The Toddler Language and Motor Questionnaire: A mother-report measure of language and motor development	Research in Developmental Disabilities	This study empirically evaluates the psychometric properties of a new mother-answered developmental instrument for toddlers, the Toddler Language and Motor Questionnaire (TLMQ).	Construction of the TLMQ, item generation, item analysis of the TLMQ, content validity, concurrent validity. It also included a standardisation sample: Standardization sample.	Mothers of 1132 children.	15-38 months	Icelandic (doesn't specifically state this but sample was Icelandic mothers).	Information not provided	Marital status: Most mothers of boys (92.0%) and girls (86.8%) in the sample were cohabiting. Single mothers of boys comprised 8.0% and that of girls 13.2%. Employment status: 31% of mothers of boys had full-time jobs, 42% had part-time jobs and 27% were homemakers. 38% mothers of girls had full-time jobs 40% had part-time jobs and 22% were homemakers. 96% of the fathers of boys and 97% of girls had full-time jobs outside home. Education: About 28% of mothers in the sample had primary or lower secondary education (22% of fathers), 48% had upper secondary education (51% of fathers), and 24% had university education (27% of fathers).

Hidecker et al. (2017)	Validity of the Communication Function Classification System for use with preschool children with communication disorders	Developmental Medicine and Child Neurology	To evaluate construct and predictive validity of the Communication Function Classification System (CFCS) for use with preschool children with a range of speech and language disorders	Cohort study	77	Sample for secondary analysis (n=77). Mean = 32.48 months (2.7 years) SD = 12.26 months (1.02 years)	English (not specifically stated but study took place in Canada)	50 boys, 27 girls	Original sample (n=97): CFCS level of communicative function Level I = 7 (7%) Level II = 8 (8%) Level III = 16 (16%) Level IV = 44 (45%) Level V = 22 (23%); Medical diagnoses Global dev. delay = 28 (29%) Syndromes = 8 (8%) Hearing impairment = 8 (8%); Communication disorder Speech & language = 81 (84%) Language only = 8 (8%) Speech only = 8 (8%); Treatment goals: Expressive language (71) Receptive language (44); Articulation/phonology (39); Amount of treatment: Mean = 8.6 hours. Treatment type: Individual (50); HP/Parent consultation (29); Group (25); Parent training (10) Sample for secondary analysis (n=77) CFCS level of communicative function: Level I = 5 (6%) Level II = 6 (8%) Level III = 11 (14%) Level IV = 40 (52%) Level V = 15 (20%) Medical diagnoses: Global dev. delay = 26 (34%) Syndromes = 19 (25%) Hearing impairment = 3 (4%); Communication disorder: Speech & language = 61 (79%)

									Language only = 10 (13%) Speech only = 6 (8%); Treatment goals: Expressive language (57) Receptive language (44) Articulation/phonology (31) Amount of treatment: Mean 7.19 hours; Treatment type: Individual (33); HP/Parent consultation (15); Group (16); Parent training (7)
Johnson, & Bountziouka (2020)	Using the PARCA-R to assess children's cognitive and language development at two years of age	Infant	The main aims of the project were to: 1) use existing data to identify a large standardisation sample that was representative of the UK population 2) develop standardised scores that have a normative mean of 100 and SD of 15 3) test the validity of the new standardised scores.	Statistical analysis of large data set	6402	INFANT trial: 6,196 children born at 35-42 weeks of gestation LAMBS cohort study: 186 children born at 32-36 weeks of gestation in PANDA study: 20 children born <32 weeks' gestation	English	3,321 boys, 3,081 girls	Ethnic background n (%): White 5,009 (78.2); Non-white 508 (7.9); missing data 885 (13.8)

Marcotte et al. (2016)	Examining the classification accuracy of a vocabulary screening measure with preschool children	Assessment for Effective Intervention	Investigated the classification accuracy of the Dynamic Indicators of Vocabulary Skills (DIVS) as a preschool vocabulary screening measure	Analysis of existing dataset	240	4 years old	English	122 boys, 118 girls	Preschool in large urban neighbourhood, 23% of individuals in area living below poverty line. 69.6% of students qualified for free or reduced-price lunch. Children qualifying for special education services not included in sample
Moeller et al. (2019)	Validation of a Parent Report Tool for Monitoring Early Vocal Stages in Infants	Journal of Speech, Language & Hearing Research	To examine the validity of the clinical version of the VDLI through three research questions. 1. Does the VDLI capture the expected developmental progression from precanonical to canonical to word stages in infants and toddlers who are typically developing? 2. Are VDLI scores significantly related to a concurrent measure of early communication skills? 3. Is there internal consistency among the items in each of the VD	Cohort	160	6-21mths (mean: 13.51 months, SD: 4.6)	American English	85 boys, 75 girls	Ethnicity: the majority of parents identified as non-Hispanic/Latino (n = 155). For race, the majority of parents, including the five who identified as Hispanic/Latino for ethnicity, reported identifying as Caucasian (n = 148), with the remaining parents identifying as Black (n = 1), other (n = 2), or more than one race (n = 9). Education: Only 8% of participants had a high school education or less.

Oryadi-Zanjani (2018)	Production of Infant Scale Evaluation (PRISE) in Persian normal hearing children: A validation study	International Journal of Pediatric Otorhinolaryngology	(a) translation and back-translation of the PRISE questionnaire to develop the Persian version of the questionnaire; (b) determination of the questionnaire's reliability; (c) determination of the development of pre-first word vocalizations of Persian speaking children using PRISE-P to develop a clinical criteria; and (d) determination of the questionnaire's capability to identify vocalizations progression accompanied by increase in age to introduce it as a screening tool.	Cohort	640	3-18 months	Persian	345 boys, 295 girls	No information provided
Ryu & Sim (2019)	The validity and reliability of DDST II and bayley iii in children with language	Neurology Asia	To analyse and compare the performance of the language sectors in the Denver Developmental Screening Test II (DDST II) and the Bayley Scales of Infant Development III (Bayley III) test with that of	Cohort study	30 DLD; 5 without DLD	Mean: 25.3 months	Korean	21 boys, 9 girls	No information provided

	development delay		the Sequenced Language Scale for Infants (SELSI)						
Seager & Abbot-Smith (2017)	Can early years professionals determine which preschoolers have comprehension delays? A comparison of two screening tools	Child Language Teaching & Therapy	To see whether EYPs are able to accurately assess the receptive language of children aged 30–35 months when using their current procedures (e.g. EYFS:UCCS).	Cohort study	82 Children (12 excluded from analysis)	Mean: 33months (range 30-35 months)	Monolingual British-English speakers	34 boys; 36 girls	No information provided

English language screening tools: Table of information about screening tools

Authors & Year	Main Screener/Tool	Aspect(s) of speech/language assessed	Type of screening tool/who completed screening tool
Chilosi et al. (2019)	Italian version of the MacArthur-Bates Infant and Toddler Communication Development Inventories (PVB, Caselli & Casadio, 1995; Caselli et al., 2007). Expressive grammar was evaluated by analysing language samples collected in our laboratory during a standardized play situation involving the child and his/her parents. Syntactic comprehension was examined with an acting out test that includes 56 items of increasing complexity, consisting of verbal commands that the child is required to act out with a set of toys and familiar objects	Expressive vocabulary, expressive grammar, syntactic/verbal comprehension	Combination of indirect (parental report) and direct procedures
Claessen et al. (2017)	Toddler Phonology Test (TPT: McIntosh and Dodd, 2011)	Phonology	Direct assessment
Gilkerson et al. (2017)	MacArthur–Bates Communicative Development Inventory: Words and Gestures / Words and Sentences (MBCDI; Fenson et al., 2007), a checklist asking about words the child understands and/or says The Child Development Inventory (Ireton, 1992). Parents of children 2–6 months of age completed a phone interview with a certified speech language pathologist (SLP) using the Receptive–Expressive Emergent Language Test, 3rd edition (REEL-3; Bzoch et al., 2003). Children were evaluated by a certified SLP utilizing up to three standard observational assessments - REEL-3, the Preschool Language Scale, 4th edition (PLS-4; Zimmerman et al.,	Receptive and expressive language scores from the REEL-3 and the Child Development Inventory, as well as the verbal production (vocabulary) score from the MBCDI were used for analyses., expressive and receptive language. PLS-4 and REEL-3 receptive and expressive language scores and CAT and CLAMS cognitive and language scores were used for analyses.	Parent report/questionnaire, SLT observations

	2002), and the Cognitive Adaptive Test / Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS; Accardo and Capute, 2005).		
Gudmundsson (2015)	Toddler Language and Motor Questionnaire (TLMQ)	Language Comprehension & Language Expression	Parent report
Hidecker et al. (2017)	FOCUS & Communication Function Classification System (CFCS)	The Communication Function Classification System (CFCS) is a validated discriminative tool that allows clinicians and parents to categorize children's communication skills into five mutually exclusive levels of everyday communicative function. The CFCS, adults classify children's communication by how they communicate on a day-to-day basis. The levels vary by the familiarity of the communication partner, the child's successful sending and receiving of messages, and the pace of communicative interactions.	Clinician assessment and parent report
Johnson, & Bountziouka (2020)	Parent Report of Children's Abilities - Revised (PARCA-R).	The PARCA-R comprises scales to assess non-verbal cognition and language development.	Parent questionnaire
Marcotte et al. (2016)	Dynamic Indicators of Vocabulary Skills (DIVS). Peabody Picture Vocabulary Test–Third Edition (PPVTIII).	A vocabulary knowledge estimate distinct from other language and literacy assessments. The DIVS consists of two subtests. Picture naming fluency (PNF) is used to assess children's expressive vocabulary knowledge by measuring the rate in naming a series of pictures	Direct asessement

		representing common nouns. Reverse Definition Fluency (RDF) is used to assess children's receptive and expressive language skills.	
Moeller et al. (2019)	Vocal Development Landmarks Interview (VDLI) - 3 subscales (Precanonical, Canonical, and Word) and a total score	Precanonical, Canonical, and Word	Interview, conducted by PhD (Psychologist)
Oryadi-Zanjani (2018)	Production of Infant Scale Evaluation (PRISE)	Pre-verbal	Questionnaire administered by Speech and language therapist, either in person or face to face
Ryu & Sim (2019)	Korean versions of - Language elements in the Denver Developmental Screening Test II (DDST II). The Bayley Scales of Infant Development III (Bayley III) test. Sequenced Language Scale for Infants (SELSI)	DDST II - language development. Bayley III - receptive/expressive language. The SELSI test was designed to assess the overall language ability of infants younger than 3 years and can assess the infant's level of understanding and expression in semantic-cognitive, phonological, syntactic ability, and pragmatic competence terms.	Direct assessment by occupational or speech therapist
Seager & Abbot-Smith (2017)	Early Years Foundation Stage Unique Child Communication Sheet (EYFS:UCCS) & WellComm (Early Years)	Receptive language, Listening, attention, understanding and speaking (EYFS:UCCS), WellComm (Early Years) (receptive, expressive, play skills)	Language measure completed by Early Years Practitioners

English language screening tools: Table of study results

Authors & Year	Outcomes (statistical)	What does this tell us about screening tools?
Chilosi et al. (2019)	Predictive value of T1 measures by a discriminant function analysis (stepwise method) with Expressive Vocabulary, Verbal Comprehension z-scores and GASS level as independent variables and outcome groups (G1, G2 and G3) as grouping variables. Verbal comprehension significantly predicted G1 and G3 outcome (Wilxs lambda=0.67, X2=18.37, p < .0001) as 91.7% of children from G1 and 92.6% from G3 were correctly classified. Conversely, according to discriminant analysis, verbal comprehension did not identify children from G2, assigning four children to G1 and seven children to G3 (with zscores between -1.54 and -0.72). In order to further analyze possible predictors of G2 outcome at T1, two separate discriminant analyses, the first on G1 and G2, and the second on G2 and G3, were performed. When discriminant analysis was conducted on G1 and G2, expressive vocabulary sorted out as a significant predictor (Wilxs lambda=0.8, X2=3.9; p < .05) with 75% of correct predictions of children from G1 and 45.5% of children from G2. When discriminant analysis was conducted on G2 and G3, verbal comprehension sorted out as a significant predictor (Wilxs lambda=0.81, X2=7.4, p < .01) of G3 (92.6%), but not of G2 (27.3%). At T2, lexical quotient and GASS level sorted out as significant predictors (Wilxs lambda=0.35, X2=29.4, p < .001) of G1 (81.8%) and G3 (92%), but not of G2 (20%).Verbal comprehension no more did sort out as a discriminant variable. At T3 the GASS level correctly classified 92% of G2, and 89% of G3 (Wilxs lambda=0.62, X2=16.9, p < .001).	This study provides data based on late-talking children's early language evaluation which includes not only vocabulary measures, but also direct assessment of expressive and receptive grammar; the latter has not been investigated before with a standardized test in Italian Late Talkers.
Claessen et al. (2017)	Younger group produced fewer consonants correctly at both assessments. Younger group's PVC score improved more than did the older group's PVC score, between assessments due to a ceiling effect for the older group's PVC score.	The older group's PCC score was higher than the younger group's at both assessments. Both age groups' PCC scores were higher at second time point than at the first. Similarly, the older group had a higher PVC score than the younger group at both assessments. The

		younger group's PVC score improved more than that of the older group between assessments due to a ceiling effect for older children.
Gilkerson et al. (2017)	Criterion validity (r = .67–.97) and test–retest reliability claims of the Snapshot (r = .95). Sensitivity and specificity for language delay detection also were good at 87%.	
Gudmundsson (2015)	Positive correlation between age (in months) and developmental language and motor skills raw scores. The correlation coefficients for the five subtests ranged from .81 to .84 ($p < .001$). An overall MANOVA revealed a significant difference between age intervals on the five subtests: Pillai's trace = .872, $F(30, 5625) = 39.617$, $p < .0001$. A MANOVA with gender and age intervals as independent variables and the five subtests as dependent variables found significant main effect in three areas: gender, Pillai's trace = .151, $F(5, 1111) = 39.506$, $p < .0001$; age, Pillai's trace = .889, $F(30, 5575) = 40.193$, $p < .0001$; and gender by age interaction, Pillai's trace = .068, $F(30, 5575) = 2.578$, $p < .0001$. Reliability was estimated by using Cronbach's alpha coefficients. For both boys and girls reliability coefficients of the three composites were high and ranged from .83 to .97. Seventeen out of 21 coefficients were .90 and higher for boys. For girls 18 out of 21 coefficients were .90 and higher. The TLMQ's Language composite had clear convergent partial correlations (controlling for age) with the MSCA scales as expected, but unexpectedly had a significant partial correlation ($r_{partial} = .35$, $p < .05$) with the Perceptual–Performance scale of the MSCA. The motor composite had clear divergent partial correlations with four of the MSCA scales as expected but non-significant convergent partial correlations ($p > .05$) with the Motor and Perceptual–Performance scale of the MSCA. The mothers' inventory responses on Self Help and Language Comprehension yielded clear convergent–divergent partial correlations with the MSCA scales. Self Help correlated significantly and convergently ($r_{partial} = .32$, $p < .05$) with the MSCA Motor scale and unsignificantly ($p > .05$) and divergently with the other five scales of the MSCA (r's range: .02 to .18,$p > .05$). The Language Comprehension subtest correlated convergently and significantly with the MSCA Verbal scale ($r_{partial} = .37$, $p < .05$) and the General Cognitive Index ($r_{partial} = .36$, $p < .05$). Divergent and non-significant partial correlations emerged	Evidence of content, criterion-related and construct validity of the TLMQ indicates that it is a sound instrument and may be a useful addition to individually administered developmental tests.

	between Language Comprehension and the other four MSCA scales (r's range: .18–.27, p > .05). The Fine Motor subtest on the TLMQ yielded significant and convergent partial correlation with the MSCA Motor scale (rpartial = .34, p < .05) but not with the Perceptual–Performance scale as expected (rpartial = .13, p > .05). Clear divergent partial correlations emerged as expected between four MSCA scales and the Fine Motor subtest (r's range: .13–.20, p > .05). The Language Expression subtest had significant partial correlations with all of the MSCA scales. However, a clear convergent–divergent pattern emerged within the correlation matrix. The highest convergent partial correlations were between Language Expression and MSCA's Verbal scale (rpartial = .63, p < .01), General Cognitive Index (rpartial = .60, p < .01), Memory (rpartial = .42, p < .05) and Quantitative (rpartial = .38, p < .05). Divergent but significant partial correlations emerged between the Motor scale (rpartial = .31, p < .05) and the Perceptual–Performance scale (rpartial = .36, p < .05) on the MSCA and Language Expression. The Gross Motor subtest on the TLMQ had low and non-significant divergent partial correlations with five of the six MSCA scales (r's range: .14 to .04, p > .05). However, a convergent partial correlation between MSCA'S Motor scale and the TLMQ's Gross Motor was not found in this study (rpartial = .03, p > .05).	
Hidecker et al. (2017)	Construct validity - There were statistically significant negative correlations between SLPs' (correlation coefficient, rs[77]=0.76, p<0.001) and parents' (rs[77]=0.65, p<0.001) total FOCUS scores and CFCS classifications at time 1. Predictive validity - There were also statistically significant negative correlations between CFCS classifications at time 1 and SLPs' total FOCUS scores at time 2 (rs[77]=0.72, p<0.001), and time 3 (rs[77]=0.68, p<0.001); and between CFCS classifications at time 1 and parents' total FOCUS scores at time 2 (rs[77]=0.63, p<0.001), and time 3 (rs[77] =0.60, p<0.001). FOCUS change scores did not differ significantly by CFCS level as reported by SLPs or parents, indicating that mean FOCUS change scores were relatively similar for children at each of the CFCS levels.	Scale can be used to measure the "vocal behavior" as a developmental milestone

Johnson & ountziou ka (2020)	External validity - Using data from 709 termborn children LMS models demonstrated good external validity. Moreover, the proportions of children with developmental delay were close to the proportions expected in the general population. Clinical validity - Using the norms tables, the standard scores were applied to PARCA-R data for 692 children born very preterm (<32 weeks' gestation; PANDA Study) and 764 children with suspected or confirmed sepsis during the neonatal period (UK arm of the INIS Trial).	PARCA-R provides a cost-efficient, standardised outcome measure with robust psychometric properties that can be used to assess children's cognitive and language development and quantify delay at two years of age, for both clinical and research purposes.
Marcotte et al. (2016)	Intercorrelations among DIVS measures ranged from .77 to .87. Logistic regression found that DIVS variables were significant predictors of vocabulary outcomes, and EL status was not. Both PNF and RDF were significant predictors of spring PPVT status in the winter model, indicating that both DIVS variables accounted for unique variance in the prediction of subsequent vocabulary outcomes.	Results supported the use of the DIVS as a brief and inexpensive tool for preschool vocabulary screening.
Moeller et al. (2019)	Regression results indicate that Precanonical, Canonical, and Word subscales were significantly different, $F(2, 151) = 85.62$, $p < .001$, but the significant interaction of the VDLI subscale and age group, $F(6, 151) = 12.22$, $p < .001$, indicates that the amount of difference between them depends on age. Post hoc comparisons with Tukey–Kramer–adjusted p values were conducted to compare the VDLI subscales at each age bin. Precanonical and Canonical subscales were significantly different from each other at 6–9 months, difference = −40.95, $t(151\ df) = -10.81$, $p < .001$; 10– 13 months, difference = −25.58, $t(151\ df) = -9.51$, $p < .001$; and 14–17 months, difference = −13.32, $t(151\ df) = -4.90$, $p < .001$; but not at 18–21 months, difference = −12.24, $t(151\ df) = -3.30$, $p = .053$. Precanonical and Word subscales were significantly different from each other at 6– 9 months, difference = 61.05, $t(151\ df) = 15.97$, $p < .001$; 10–13 months, difference = 67.28, $t(151\ df) = 24.81$, $p < .001$; 14–17 months, difference = 51.87, $t(151\ df) = 18.91$, $p < .001$; and 18–21 months, difference = 34.43, $t(151\ df) = 9.21$, $p < .001$. Canonical and Word subscales were significantly different from each other at 6–9 months, difference = 20.10, t(151 df) = 4.13, p = .003; 10–13 months, difference = 41.70, t(151 df) = 12.09, p38.55, t(151 df) = 11.04, p < .001; and 18–21 months, difference = 22.19, t(151 df) = 4.67, p < .001.	The VDLI was found to be sensitive to age in a cross-sectional sample of 6- to 21-month-old infants; older children consistently demonstrated higher scores than younger children, and age was positively correlated with VDLI subscale scores. The three subscales (Precanonical, Canonical, and Word) were differentially sensitive to age ranges.
Oryadi-Zanjani (2018)	PRISE Cronbach's alpha was 0.88. PRISE corrected item-total correlations were significantly high ranging from 0.86 to 0.89. Positive correlation between the PRISE total scores and the children's age ($r = 0.791$, $n = 640$, $p = 0.000$).	Scale can be used to measure the "vocal behavior" as a developmental milestone

Ryu & Sim (2019)	Analysis against the SELSI: DDST II - Sensitivity (%) 93.3 (76.5-98.8), Specificity (%) 60 (17-92.7), Pos Pred Value (%) 93.3 (76.5-98.8), Neg Pred Value (%) 60 (17-92.7). Bayle III expressive - Sensitivity (%) 93.1 (75.8-98.8), Specificity (%) 83.3 (36.5-99.1), Pos Pred Value (%) 96.4 (79.8-99.8), Neg Pred Value (%) 71.4 (30.3-94.9). Bayle III Receptive - Sensitivity (%) 71.4 (51.1-86.1), Specificity (%) 100 (56.1-100), Pos Pred Value (%) 100 (80-100), Neg Pred Value (%) 46.6 (22.3-72.6).	DDST II had high sensitivity and was a potentially useful screening test. In contrast, the BRLS and BELS were not suitable replacements for the SELSI because of variations in their sensitivity, specificity, positive predictive value, and negative predictive value
Seager & Abbot-Smith (2017)	"Correlation between the WellComm and the PLS-4-auditory (rs(61) = .563, p < .001). Correlation between a composite of the receptive language items of the WellComm and the auditory component of the PLS-4 (rs(61) = .441, p < .001). Negative correlation between the WellComm and EYP experience (rs(49) = –.340, p = .017)."	The EYFS:UCCS was found to have poor sensitivity; using this format, EYPs failed to detect the majority of children who scored on or below the 15th percentile on the PLS-4. The WellComm showed good levels of sensitivity and borderline acceptable levels of specificity in discriminating 'typically-developing' children from those who are delayed. Children who did not score in the 'green' range on the WellComm were more likely than not to score on or below the 15th percentile on the PLS-4 (and vice versa). Although far from perfect, this result is better than that of many language screening instruments (for a review, see Law et al., 2000).

Bilingualism review - Table of demographic information in studies

Reference			Study design		Study population					
Authors & Year	**Title**	**Journal**	**Aim**	**Methodology**	**No. participants**	**Child age range**	**Home language(s)**	**Type of bilingualism**	**Child gender**	**Demographic information**
Armon-Lotem & Ohana (2017)	A CDI study of bilingual English-Hebrew children-Frequency of exposure as a major source of variation	International Journal of Bilingual Education and Bilingualism	(1) Vocabulary – Will the children show balanced bilingualism due to the relatively high prestige of the two languages? (2) Exposure variables – Which of the exposure variables best predicts variability among children? (3)	Cohort study (group comparison by age and language dominance)	40	24-45 months	Hebrew/English	Simultaneous bilinguals	23 boys, 17 girls	All mid-high SES; all but 3 fathers university educated

			Risk for SLI – Can bilingual use of the CDI contribute to identifying children at risk for SLI?							
Cattani et al. (2014)	How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests?	International Journal of Language & Communication Disorders	To investigate whether the scores of UK bilingual children on various production and comprehension measures could be predicted by the proportion of exposure to English; to identify a minimal cut-off point of English exposure	Cohort study (group comparison monolingual/bilingual)	71 (36 monolinguals & 35 bilinguals)	28-32 months (monolingual mean: 30.47, SD: 1.2; bilingual mean: 30.21; SD: 1.2)	36 English only; 35 English & additional language (Arabic, French, Punjabi, Italian, Spanish, Catalan, German, Greek, Irish Gaelic, Dutch, Finnish, Polish, Albanian,	Simultaneous bilinguals (and monolingual reference group)	33 boys, 38 girls	81% of parents degree level education, mean occupation score 3-9, mean 7.6

			above which bilingual children would perform similarly to monolingual peers and over which at-risk children could be easily identified				Czech, Kurdish, Afrikaans, Swahili or Mandarin)			
Core et al. (2013)	Total and Conceptual Vocabulary in Spanish-English Bilinguals From 22 to 30 Months: Implications for Assessment	Journal of Speech, Language & Hearing Research	(a) to compare total vocabulary and conceptual vocabulary measures in a group of bilingual children, (b) to compare measures of total vocabulary and conceptual vocabulary for young bilingual	Cohort study (group comparison monolingual/bilingual)	103 (56 monolinguals & 47 bilinguals)	22-30 months (3 time points)	56 English only; 47 English and Spanish	Simultaneous bilinguals (and monolingual reference group)	55 boys, 48 girls	Majority of parents had 4-year or advanced degree (75% of monolingual mothers/61% of fathers; 87% bilingual mothers/60% of fathers); parent ethnicity

			children with monolingual performance on a single-language measure, and (c) to determine whether total and conceptual vocabulary identify the same proportion of bilingual children performing below a cut-point for the low average range for each measure							information in article

DeAnda et al. (2016)	The Language Exposure Assessment Tool: Quantifying Language Exposure in Infants and Children	Journal of Speech, Language and Hearing Research	To propose an assessment tool - the LEAT - to measure and quantify relative language exposure that can be used across languages and settings and to assess its internal validity, criterion validity and utility.	Pilot study	98	Mean 17;14 months; SD 1.01; Range 14;22 - 19;24	English and Spanish/English and French	Simultaneous bilinguals	57 boys; 41 girls	Average maternal and paternal education: college level (maternal mean 15.20 years, SD 2.60; paternal mean 14.71 years, SD 2.75); on average lived with four family members (mean 3.84, SD 1.3, range 2-8)
De Houwer, Bornstein and Putnick (2012)	A bilingual-monolingual comparison of young children's vocabulary	Applied Psycholinguistics	Using longitudinal data from bilingual and monolingual children, to	Cohort study (group comparison monolingual/bilingual)	61 (31 bilingual and 30 monolingual)	13 - 20 months (2 time points)	30 Dutch only; 31 Dutch and French	Simultaneous bilinguals (and monolingual reference group)	Monolingual group = 53.33% female; bilingual group =	Average maternal education = monolingual mean 5.56 on 7 point scale;

	size: Evidence from comprehension and production		examine vocabulary sizes for both word comprehension and word production and to compare groups for relations between comprehension and production.						45.16% female	bilingual mean 5.45; average paternal education = monolingual mean 5.31; bilingual mean 5.16
Floccia et al. (2018)	Vocabulary of 2-year-olds learning English and an additional language: Norms and effects of linguistic distance	Monographs of the Society for Research in Child Development	1) to investigate the effect of linguistic distance on vocabulary development; 2) to build a model fo the bilingual lexicon at age 2 using a range	Pilot study - development and standardisation of a test	430	Mean 23.89 months, SD 0.39, range 23.0-25.0	English and one of: Bengali; Cantonese; Dutch; French; German; Greek; Hindi/Urdu; Italian; Mandarin; Polish;	Simultaneous - age of acquisition data not available, but all toddlers born in UK	237 boys; 193 girls	Mean income bracket 3.59, SD 0.66 on 4-point scale from 1 to 4; Mean maternal education 6.36, SD 0.94 on 7-point scale from 1 to 7; Mean

			of situational variables and linguistic distance for the target Additional Lanugages learners, resulting in production of UKBTAT tool; 3) to test UKBTAT generalisability for the non-target learners				Portuguese; Spanish; Welsh; or an additional non-target language			paternal education 5.96, SD 1.28 on 7-point scale from 1 to 7
Goh et al. (2017)	Analysis of Item-Level Bias in the Bayley-III Language Subscales: The Validity and Utility of Standardize	Journal of speech, language, and hearing research : JSLHR	To investigate bias on the BSID-III language subscales using differential item functioning to assess: 1) if	Statistical analysis of existing tool	459	24 months	English, Mandarin, Bahasa Melayu, Tamil (Singapore)	Range - some monolinguals, some balanced bilinguals, varied language experiences/exposures	239 boys, 216 girls	Mother-toddler dyads participating in 'Growing up in Singapore Towards Healthy Outcomes'

	d Language Assessment in a Multilingual Setting		items are biased towards different levels of English exposure; 2) items are biased across different levels of bilingual exposure; c) utility of subscales through correlation with behavioural emotional problems and subsequent receptive ability							study - no specific demographics reported
Hardin, Scott-Little and Mereoiu (2013)	Developing the BIO Questionnaire: A Bilingual Parent	Journal of Research in Childhood Education	To describe the process for developing the Formulario Familiar Bilingue de	Pilot study	23	Not given (only states all pre-school)	Spanish and English	Not reported	13 boys, 11 girls	66.7% high school education, 16.7% primary school, 16.7%

	Report Tool for Prekindergarten English Learners of Latino Heritage		Información Formulario y Observación/Family Bilingual Information & Observation (BIO) Questionnaire and to share the results of research conducted to refine and pilot the questionnaire.							some college; all Latin American Spanish speakers
Legacy et al. (2017)	Dog or chien? Translation equivalents in the receptive and expressive vocabularies of young	Journal of Child Language	Experiment 1: 1) to gain a better understanding of how French-English bilinguals acquire translation equivalents	Cohort study	Experiment 1: 34 Experiment 2: 22	Experiment 1: 1;3 - 2;11 (3 time points) Experiment 2: 2;0 - 2;5 (mean 2;2)	French and English	Simultaneous	Experiment 1: 19 boys, 16 girls Experiment 2: 9 boys, 11 girls	Experiment 1: 82% of mothers had university degree; Experiment 2: no information provided

	French-English bilinguals		over time, by examining changes in the proportion of TEs on the CDI 2) to determine the roles that linguistic input and vocabulary growth play in shaping TE acquisition. Experiment 2: to compare a direct, touch-screen measure of infants' TE comprehension with parent report of the same subset of words							

Mancilla-Martinez et al. (2016)	Parent Reports of Young Spanish-English Bilingual Children's Productive Vocabulary: A Development and Validation Study	Language, speech, and hearing services in schools	Study 2: To examine the validity of the parent report short forms of children's productive vocabulary for use with Spanish-speaking families from low-income homes. Data taken from Study 2 only as Study 1 involved only piloting and developing the tool	Pilot study - development and standardisation of a test	194 (Study 2)	24-48 months (M = 35.97, SD = 6.78)	English, Spanish	Range - on average children mostly heard Spanish from family but also exposed to English	98 boys, 96 girls	97% Latino origin; 80% of mothers born outside of USA; 52% of fathers and 50% of mothers less than high schol education; 69% income lower than $20,000
Marchman & Martínez-	Concurrent validity of caregiver/parent report	Journal of Speech, Language &	To evaluate whether the CDI and IDHC provide valid	Pilot study - development and standardisation of a test	26	23-34 months, M = 27.8 months	English and Spanish	Majority simultaneous	13 boys, 13 girls	Average maternal education some college

Sussman (2002)	measures of language for children who are learning both English and Spanish	Hearing Research	estimates of early lexical and grammatical skills in children with regular exposure to both English and Spanish							(Mean 14.4 years); 54% born outside of USA, primarily Mexico; 58% of parents Hispanic, 27% mixed ethnicities
Nayeb et al. (2021)	Identifying language disorder in bilingual children aged 2.5 years requires screening in both languages	Acta Paediatrica	To investigate four models of language screening for ability to detect DLD in bilingual children aged 2.5 years, and whether bilingual children with DLD could be identified using only one of their languages	Pilot study - development and standardisation of a test	111	29-33 months	Range of languages – Swedish and one of Somali, Arabic, Kurmanji, Farsi, Sorani & Turkish.	Sequential	49% boys, 51% girls	Maternal education: 25% college/university; 33% secondary school; 32% elementary school, 10% illiterate

			in the screening.							
O'Toole et al. (2016)	Parent report of early lexical production in bilingual children: a cross-linguistic CDI comparison	International Journal of Bilingual Education and Bilingualism	To examine: 1) how vocabulary scores in each language for bilinguals compare within the language pairs, and when comparing across languages using TCV? 2) the role of language exposure and other home/family variables in predicting development in the two languages 3)	Cohort study (group comparison by languages spoken)	250	24-36 months	Six research groups: Maltese & English; Irish & English; Polish & English; French & Portuguese; Turkish & German; English & Hebrew	A range depending on research group and family factors	58.4% boys, 41.6% girls	Groups included a range of SES, although parental education and occupation varied significantly across language groups

			criteria for identifying 'late talking' in bilinguals?							
Patterson, Rodríguez and Dale (2020)	Dynamic Assessment Language Tasks and the Prediction of Performance on Year-End Language Skills in Preschool Dual Language Learners	American Journal of Speech-Language Pathology	To determine if performance on Dynamic Assessment language tasks administered with graduated prompting would be related to year-end performance on language achievement testing among young dual language learners from Spanish-	Pilot study - development and standardisation of a test	20	4 years (Mean age 54 months)	Spanish and English	Mixed sequential/simultaneous (most children lived in homes where Spaish was spoken more than English)	10 boys, 10 girls	Parent education ranged from sixth grade - some college; among mothers 7 had not completed high school, 7 completed high school, 5 completed some college

			speaking homes							
Thordardottir et al. (2006)	Bilingual assessment: can overall proficiency be estimated from separate measurement of two languages?	Journal of Multilingual Communication Disorders	To investigate the effect of bilingual exposure on children's scores on measures of language development, including expressive and receptive measures of lexical and syntactic development, in a relatively homogeneous group of bilingual	Cohort study (group comparison monolingual/bilingual)	28	2;6 to 3;0	3 groups - monolingual English, monolingual French and French-English bilingual	Range - monolinguals & simultaneous bilinguals	16 boys, 12 girls	Took place in Montreal Quebec; average education of parents: monolingual English group = mothers 17.7 yrs; fathers 17.6 yrs; monolingual French= mothers 15.1 yrs, fathers 16.3 yrs; English-French mothers =

			children with a direct comparison to monolingual children of the same age and of comparable background characteristics							16.4 yrs; fathers = 17.3 yrs
Vagh, Pan and Mancilla-Martinez, (2009)	Measuring Growth in Bilingual and Monolingual Children's English Productive Vocabulary Developmen t: The Utility of Combining Parent and Teacher Report	Child Development	To investigate: 1) how growth in the developing English vocabularies of children from low-income bilingual families compares with children from low-income monolingual families across toddlerhood; 2)	Cohort study (group comparison parent/teacher report)	85	24-36 months (range of time points)	English and Spanish	Range - some monolingual English/Spanish, some bilingual	Not reported	Annual median income $10,000-$19,999; average education 13 yrs (fathers); 12 yrs (mothers); 57% of children Hispanic/Latin o; 27% Black; 5% White; 11% mixed

			whether multiple reporters provide a more comprehensive estimation of English vocabulary knowledge compared to single parent reports; 3) whether parent reports, teacher reports, and parent teacher composite reports provide valid estimations of English vocabulary knowledge when at child ages 30 and 36 months; 4) the							race; 98% of children born in USA, 50% of mothers and 55% of fathers in other countries

			utility of parent-teacher composite reports							

Bilingualism review - Table of information about screening tools

Authors & Year	Main Screener/Tool	Reference Tools used	Aspect(s) of speech/language assessed	Type of screening tool	Who completed screening tool?	Was person completing tool a native speaker?
Armon-Lotem & Ohana (2017)	MacArthur Bates CDI (English & Hebrew versions) PaBiQ-IT (Questionnaire for parents of bilingual children: infants & toddlers version)	None	Productive vocabulary Receptive vocabulary Grammar Language exposure	Parent report	Parents	Yes - parents (though some parents had less knowledge of Hebrew so only filled out the English questionnaires)
Cattani et al. (2014)	Oxford CDI, Additional language CDI where available; Parent English exposure questionnaire; BPVS III; Auditory component of PLS 4; Adapted English SETK-2	None	Productive vocabulary Receptive vocabulary Auditory skills Object-naming Language exposure	Parent report (questionnaire & CDI) Direct assessment	Parents/ Research assistants	No - monolingual British English-speaking research assistants
Core et al. (2013)	CDI Words & Sentences Task/ IDHC (Inventorio del Desarollo de Habilidades-Comunicativas: Palabras y Enunciados)	ASQ-3	Productive vocabulary	Parent report	Parents	Yes
DeAnda et al. (2016)	Language Exposure Assessment Tool	MCDI (English, Spanish and French versions) and Computerized	Language exposure	Parent report	Parent	Yes

		Comprehension Task				
De Houwer, Bornstein and Putnick (2012)	CDI in Dutch (Woorden en Gebaren: N-CDI) and French (Mots et Phrases: F-CDI)	None	Productive vocabulary Receptive vocabulary	Parent report	Parents and one additional caregiver (e.g. childminder)	Yes
Floccia et al. (2018)	UKBTAT (a combination of Plymouth Language Exposure Questionnaire; Oxford Short Form CDI (British English) and CDIs in additional languages)	None	Productive vocabulary Receptive vocabulary	Parent report	Parents	Yes (completed by parents according to language ability; if not proficient e.g. in English, an English-speaking caregiver would complete a version).
Goh et al. (2017)	BSID-III Language exposure questionnaire	CBCL, PEDS and PPVT-IV (at 48 months)	Productive vocabulary Receptive vocabulary	Direct assessment	Trained examiners who spoke 2 or more study languages	Yes
Hardin, Scott-Little and Mereoiu (2013)	BIO Questionnaire (Formulario Familiar Bilingue de Información Formulario y Observación/Family Bilingual Information & Observation)	None	Language exposure	Parent report	Parent	Yes

Legacy et al. (2017)	Experiment 1: Language Exposure Assessment Tool; CDI: Words & Gestures (American English & French Canadian adaptations); CDI: Words & Sentences (English & French Canadian adaptations) Experiment 2: Language Exposure Assessment Tool; Computerized Comprehension Task (CCT); CCT Checklist	None	Productive vocabulary Receptive vocabulary Language exposure	Experiment 1: Parent report only Experiment 2: Parent report and direct assessment	Parents/experimenter for CCT task	Yes (CDI was filled out by expert speaker who spoke that language to child regularly)
Mancilla-Martinez et al. (2016)	English CDI Short Form & Short Form Extension; Spanish CDI Short Form & Spanish Vocabulary Extension (researcher developed)	English and Spanish versions of Woodcock Johnson III Tests of Achivements & PPVT-IV	Productive vocabulary	Parent report	Parents (interviewed by research assistants)	Yes
Marchman & Martínez-Sussman (2002)	CDI in English and IDHC (Spanish version)	Language samples collected in laboratory setting	Productive vocabulary Receptive vocabulary Grammar	Parent report	Parents	Yes

Nayeb et al. (2021)	Researcher-developed screening tool (items specified in article)	Swedish version of Reynell Development Language Scales III	Productive vocabulary Receptive vocabulary	Parent report & direct assessment	Child health nurses/bilingual preschool staff/parents (for parent report section)	Yes (bilingual preschool staff were trained to perform screening in child's mother tongue)
O'Toole et al. (2016)	CDI (bilingual Maltese-English and Irish-English instruments, monolingual checklists for Polish, British English, Turkish, German, American English, Hebrew, French and Portuguese); background questionnaire PaBiQ-IT (Parents of Bilingual Children: Infants and Toddlers Version)	None	Productive vocabulary Language exposure	Parent report	Parents	A range (parents completed the questionnaires and had range of abilities in each language- some lived in bilingual communities, some were recent immigrants so spoke a minority language at home)
Patterson, Rodríguez and Dale (2020)	Range of dynamic language assessment tasks - NAL (novel adjective learning); Prediction (ability to predict results of action); SF (Similarity in Function - explaining how two similar items are alike)	Learning Accomplishment Profile (LAP-3)	Productive and receptive language (tested via word-learning ability/understanding of abstract language)	Direct assessment (including graduated prompts following incorrect responses & modelling of correct responses)	Speech-language pathology graduate student	No (student's primarylanguage was English and Spanish proficiency described as 'advanced-low' but sufficient for accurate administration)

Thordardottir et al. (2006)	MacArthur CDI & Quebec French adaptation; PPVT & Canadian French adaptation; Reynell Developmental Language Scales & translation	None	Productive vocabulary Receptive vocabulary Receptive grammar	Parent report & direct assessment	Parents/trained research assistants	Yes (examiners for each language were native speakers of English or Quebec French)
Vagh, Pan and Mancilla-Martinez, (2009)	MacArthur Bates CDI (English only)	PPVT and Picture Vocabulary subtest of WLPB-R	Productive vocabulary	Parent/teacher report	Parent (aided by bilingual research assistant) and teachers	Mixed (some parents were not native English speakers; information not provided on teachers)

Bilingualism Review - Table of study results

Authors & Year	Outcomes (information relevant for screening - sensitivity, specificity, reliability, validity, cut-off points)	Summary of language outcome results	What does this tell us about screening for bilingual children?
Armon-Lotem & Ohana (2017)	No sensitivity, specificity, validity or reliability information provided	**Production-** children had similar number of words in English and Hebrew in both age groups with no significant difference between the means. Comprehension - children comprehended significantly more words in English (M = 519.9, SD = 152.49) than in Hebrew (M =431.16, SD = 149.34), F(1,76) = 6.734, p = .011, μ2 = .081. **Correlations with individual variables -** Spearman correlations conducted - L1 vocabulary production was positively correlated with higher chronological age, later exposure to L2. A negative correlation was found with more exposure to L2 on all four Frequency of Exposure (FoE) to L2 variables. Only the FoE to L2 variables correlated negatively with L1 comprehension. Negative correlations with exposure to L2 entail positive correlation of both L1 comprehension and production with exposure to L1 as the exposure measures complement	Bilingual children's scores were variable, but differences in exposure to each language and age of onset were correlated with vocabulary. Use of CDI in each language combined with parental report of concern permits identification of children at risk of language delay. Considering the prestige of languages and extent that children are balanced bilinguals is important in predicting potential variation in bilingual language development.

		each other. For L2 production and comprehension, positive correlations were found with higher chronological age, longer exposure to L2 and more frequent exposure to L2. Yet, while total FoE to L2 correlated positively with both L2 production and comprehension, the frequency of L2 use by the mother correlated significantly only with production in L2, while the frequency of L2 directed at the child in general correlated only with comprehension, and the frequency of L2 use at home correlated with neither comprehension nor production. Identifying risk of SLI - 2 children scored more than 1 SD below the mean in both languages	

Cattani et al. (2014)	**Cut-off points**: independent-samples t-tests with unequal variance assumed between the 35 bilingual children and the 36 monolingual children on each of the following tests: BPVS III, English SETK-2, Oxford CDI Comprehension and Oxford CDI Production. The two groups of children systematically differed on these tests. Progressively removed bilingual children who had lowest amount of exposure to English from the model until there was no longer any significant difference between the bilingual and the monolingual groups.This stage was reached when bilingual children with exposure to English above 54% were compared with the monolinguals for the BPVS III measure, above 58% for the English SETK-2, above 53% for the Oxford CDI Comprehension and above 62% for the Oxford CDI Production. In other words, a bilingual child with a percentage of exposure to English at 60% (as a rounded cut-off point) or above is very likely to score similarly to a monolingual child on all four measures, whereas a child exposed to English less than 60% of the time is likely to score less than a monolingual child on the four standardized tests	**Receptive vocabulary** - Bilingual children's receptive vocabularies in English as assessed by the BPVS III were on average significantly lower than those of the monolingual children, $t(56.4) = 3.94$, $p < 0.001$; mean diff. = 0.89, Cohen's $d = 0.98$. **Expressive vocabulary** - expressive vocabulary scores as assessed by the English SETK-2 were significantly lower for the bilingual group as a whole than for the monolingual controls, $t(33.4) = 4.36$, $p < 0.001$; mean diff. = 1.0, $d = 1.12$. Both the Oxford CDI comprehension and production scores were also significantly lower for the bilingual children than for the monolingual children, $t(35.3) = 3.93$, $p < 0.001$; mean diff. = 86.82, $d = 0.95$ and $t(42.0) = 5.36$, $p < 0.001$; mean diff. = 130.1, $d = 1.30$, for word comprehension and production respectively. Finally, for the PLS 4 test, which assesses English comprehension, the bilingual children scored lower on average than the monolingual children but this difference was not significant, $t(40.9) = 1.29$, $p = 0.2$; mean diff. = 0.37).	A combination of a language exposure questionnaire and CDI provides a potential assessment method for early years workers in the UK. Study found that toddlers with minimum 60% exposure to English as a cut-off point performed like monolingual peers on CDI measure. Where possible children who are below 60% cut-off should be assessed in additional language as well, e.g. using another CDI measure (although some variability was found in both availability of translated CDIs and total word scores in different measures).

Core et al. (2013)	**Cut-off points:** When total vocabulary was the basis for percentile assignment with bilingual children, the proportion of bilingual children at or below the 25th percentile was not different from the proportion of monolingual children at or below the 25th percentile at any age: at 22 months, z = 0.87, p = .38; at 25 months, z = –0.54, p = .58; at 30 months, z = 0.95, p = .33. In contrast, using conceptual vocabulary, a greater proportion of bilingual children than monolingual children fell at or below the 25th percentile at 22 and 30 months, but not at 25 months: at 22 months, z=–2.158, p= .03; at 25 months, z = –0.17, p = .865; at 30 months, z = –4.25, p < .0002).Comparing the two vocabulary measures within bilingual children showed that at 22 and 30 months, a significantly greater proportion of bilingual children scored at or below the 25th percentile when conceptual vocabulary was used for percentile classification than when total vocabulary was used: at 22 months, p = .031; at 25 months, p = .250; and at 30 months, p = .001 (McNemar's test).	**Comparison of total & conceptual vocabulary with monolingual single-language vocabulary:** 3 (age) × 2 (language group) ANOVA used to compare bilinguals' total vocabulary scores to monolinguals' single-language scores - no difference between groups found (p = .54). To compare the bilingual children's conceptual vocabulary scores with monolingual English vocabulary sizes, a 2 (language group) × 3 (age) mixed ANOVA was used; it revealed a significant main effect of language group, F(1, 101) = 4.15, p =. 04, hp 2 = .04; a significant main effect of age, F(2, 202) = 461.79, p < .001, hp 2 = .82; and a significant Language Group × Age interaction, F(2, 202) = 4.05, p = .02, hp 2 = .04. Follow-up analysis using independent samples t tests revealed that group means for conceptual vocabulary and CDI scores did not differ significantly at 22 and 25 months (p = .44 and p = .07, respectively) but were significantly different at 30 months (p = .004). On average, group means for the monolingual children's single-language vocabularies increased more with age than did the group	Bilingual children's conceptual vocabulary scores were significantly lower than their total vocabulary scores. When total vocabulary scores were used in reference to English monolingual norms, proportion of bilingual children in at-risk range was not different to proportion of monolingual children; when using conceptual vocabulary a significantly higher proportion of bilingual children fell into at-risk range. Important to note that the study used CDI and IDHC which have some overlapping items (around 70%) but not fully overlapping - so parents can report on more words unique to a particular language. Authors suggest that measuring total vocabulary- the totality of what children know when they know a word (e.g. when they can produce it in either language) is most appropriate representation of bilingual language abilities.

		means for bilingual children's conceptual vocabularies	
DeAnda et al. (2016)	**Internal consistency** was α = .96 calculated by Cronbach's alpha; **Utility** and **criterion validity** calculated using linear regression against CCT relative vocabulary size: LEAT calculation predicted significant variance in vocabulary size, evincing a moderate effect size, F(1, 96) = 4.85, p = .03, R = .22; utility and criterion validity measured using linear regression with MCDI relative vocabulary size: model that included maternal education, age, language, the overall parent estimate, and the LEAT's calculation was significant in predicting MCDI relative vocabulary size, F(4, 78) = 2.81, p = .02, R2 = .13. Only the LEAT calculation explained significant variance	No information provided	LEAT had high internal consistency, criterion validity and additional explanatory power above simply asking parents for estimate of language exposure. In this sample exposure to two languages was relatively balanced, which was reflected in average vocabulary sizes across languages. Findings suggest the importance of consistency in the way that bilingual children's language exposure is measured, in order to inform whether to examine proficiency in one or both languages

	above the other predictor variables (R2Δ = .07, p = .008)		
De Houwer, Bornstein and Putnick (2012)	No sensitivity, specificity, validity or reliability information provided	**Comprehension - 13 months:** bilinguals understood as many Dutch words as monolinguals t (59) = 0.64, ns, d = 0.17; comparing overall word comprehension (Dutch & French), bilinguals understood significantly more words than monolinguals, t (45.91) = –3.36, p ≤ .01, d = –0.90. **20 months:** monolinguals understood similar numbers of words to bilinguals, t (52.58) = 1.50, ns, d = 0.39. **Production - 13 months:** bilingual and monolingual production did not differ, t (59) = 1.51, ns, d = 0.39; **20 months**: bilingual and monolingual toddlers produced equal numbers of Dutch words, t (58) = 1.26, ns, d = 0.32. **Total conceptual vocabulary** - 13 months: no difference between bilingual and monolingual TCVs, t (59) = 0.39, ns, d = 0.10. On average bilinguals knew 250 words,	Overall, no differences were found between bilingual and monolingual children for overall vocabulary production; when including comprehension, bilingual children at 13 months understood 60% more words - however, when taking into account translation equivalents, this advantage disappeared, suggesting children knew the same number of lexicalized meanings. There were also no differences in bilingual/monolingual production when measuring only Dutch language production. However, there was wide interindividual variation in both bilingual and monolingual children. The study suggests that if bilingual children appear slow in lexical development compared to monolingual norms, professionals should try to understand what may cause a delay rather than attributing it to bilingualism, and they should consider children's language input experiences.

		representing 144 different lexical meanings. Therefore their TCV was much lower than the total number of words in their vocabulary, $t(30) = 7.92$, $p \leq .001$, $d = 0.95$. 20 months - No differences found between the bilinguals and the monolinguals in the number of lexicalized meanings produced, $t(56) = -0.53$, ns, $d = -0.14$.	

Floccia et al. (2018)	**Cut-off points:** Distribution of observed–predicted residuals in the mixed models - SD: English comprehension 21.19, English production 22.28, Additional Language comprehension 6.04, Additional Language production 7.60. These provide a basis for converting an observed–predicted difference scored in items into a percentile, which is what UKBTAT reports for a child screened by this tool. **Testing validity** of these equations on non-target language children on English scores: For both comprehension and production, strong correlations were seen between predicted and observed scores (comprehension: r = .60; production: r = .59). No systematic underprediction or overprediction of scores in these novel data, as established by t tests of the means of observed and predicted scores (t < 1). For each child in the 58 nontarget Additional Language learners, using the standard deviations of residuals calculated above (UKBT predictive equations derived from the 372 children leaning the target Additional Languages), difference converted between each predicted score and the observed score as a percentile. Out of the 58 children, 9 were at or below the 16th percentile	**Productive & receptive vocabulary** - 24-month old monolinguals understood approximately 81.8% of words and produce 53.7% on Oxford Short Form CDI; bilinguals understood 67.9% and produced 41.2% which is significantly less than monolinguals (comprehension: t (553) =5.84, p= .0001; production: t(553) = 4.74, p = .0001). **Correlations with individual variables** - after running ANCOVAs to find significant predictors, linear mixed models were run with those predictors. For comprehension there was a significant effect of Language exposure (x2 (2) = 18.02, p < .001) and Overheard speech (x2 (2) = 18.62, p < .001) but no effect of Income (x2 (1) = 1.60, p = .21). For production there was a significant effect of LEQ (x2 (2) = 18.75, p < .001), Overheard speech (x2 (2) = 23.59,p < .001), and Gender (x2 (1) = 13.26, p < .001). Secondary variables (mode of exposure & linguistic distance) were then added to ANCOVAs containing predictors shown to be significant in Step 1 analysis (language exposure, overheard speech & gender). Only two variables achieved significance: Consistency	Normed scores on English vocabulary CDI were gathered for bilingual children speaking 13 target languages; the study found that these scores were highly predictive of the vocabulary of 58 bilingual children speaking additional (non-target) languages, suggesting feasibility for use in assessing bilingual toddlers in majority language only. Linguistic distance was found to be a predictive factor for vocabulary outcomes; for example, children's production of additional language words was improved when the language was phonologically close to English, and receptive vocabulary was greater for children who had languages with the same word order typology/morphologically close languages to British English. The most robust predictor of vocabulary was the relative amount of exposure to child-directed English vs the Additional Language, as well as proportion of English/Additional Language spoken between parents. Welsh-English toddlers did not have significantly higher vocabulary scores than other bilinguals, although they were defined as 'societal bilinguals' and have more balanced exposure. Suggests that findings apply equally to bilingual children learning language with equivalent status and learning a minority language. Norms are

	(15.5%) in comprehension and 11 in production (19.0%). Five out of the nine children with low comprehension scores had production scores at or below the 16th percentile, with the other four scoring also relatively low in production. Two children scoring very low on production had normal comprehension scores. Given the prevalence of 7–15% of experiencing delayed language acquisition (Kohnert, 2010) in monolingual children, these results suggest a very satisfactory sensitivity for the UBTAT equations	and Proportion of parental English spoken. Consistency interacted significantly with TestLang in determining production scores (F1,355 = 3.94, p = .047, n2= .01), due to English vocabulary being boosted by a decreasing consistency in parents' use of the two languages (F1,355 = 6.07, p = .014, n2= .017). Linguistic distance variables - significant effect of Phonological Overlap on Additional Language production (x2 (1) = 4.61, p = .032), a significant effect of Word Order typology on Additional Language comprehension (x2 (1) ¼ 6.02, p = .014), and a significant effect of Morphological Complexity on Additional Language comprehension (x2 (1) = 4.80, p = .028)	available for all children in British English and an additional language if available.

Goh et al. (2017)	**Validity & Reliability**: When removed of DIF Items 18, 20, and 22, the receptive scale showed excellent internal validity ($\chi2$ = 4.76, df = 6, p = .58; RMSEA = .00; CFI = 1.00; TLI = 1.00). It was reliable from z = −2.6 to −0.8 SD as only this range of receptive language proficiency reached information levels of 5 and above, equivalent to an alpha reliability of .80 (DeMars, 2010). In a similar manner, removing DIF Items 24 and 26 and fixing Item 23's factor loading to a plausible value resulted in an expressive scale with excellent internal validity ($\chi2$ = 9.22, df = 5, p = .10; RMSEA = .04; CFI = 1.00; TLI = .99). It was reliable from z = −1.5 to −0.9 SD as only this range of expressive language proficiency had information levels of 5 and above. Table 3 presents the planned correlations of BSID-III scales without DIF items with other measures of language, behavioral, and emotional problems. To enable comparison, traditional scaled scores are also presented. Patterns of association were generally similar for both DIF-free and scaled scores. Language delay (≤ −1.33SD on BSID-III scales) negatively correlated with English vocabulary at 48 months. All categorical language delay and continuous language ability	No information provided	Five of 16 tested items showed possible bias - in some cases this favoured higher English language exposure, in others bilingual children were favoured (e.g. children who were bilingual found it easier to identify 3 items of clothing, possibly because translation equivalents were easier than English words); bilingual children did worse on likelihood of using a 2-word utterance. Observed no differences in language ability between monolingual and bilingual children but did find relations between bilingualism and test items, which highlights the need to consider type of language task in bilingual language assessment

	scores were null with English, bilingual exposure, and externalizing problems at 24 months. However, scaled but not DIF-free scores, namely receptive ability, were associated with internalizing problems. Also, scaled scores were associated with PEDS scores, whereas surprisingly, DIF-free scores either did not correlate with PEDS or predicted higher PEDS receptive concern.		
Hardin, Scott-Little and Mereoiu (2013)	No sensitivity, specificity, validity or reliability information provided	Section 1 - asked about language exposure. Key findings: 79.2% reported speaking Spanish only, 4.2% English only, 12.5% both Spanish and English, and 4.2% Spanglish; English primarily used for reading books but Spanish for watching TV; Section 2 - asked about language development prior to age 4. Key findings: 60.9% said child spoke first word by 11 months; 34.8% by 18 months; 95.7%	Report on preliminary pilot results of a parent report tool on language exposure and experiences of bilingual children. Study highlighted the importance of understanding language development in both of child's languages as differences were seen in how language had developed in each.

		said first word was in Spanish. Section 3 - current language use. Key findings: 78.3% said child could use complete sentence in Spanish but only 37.5% in English	
Legacy et al. (2017)	No sensitivity, specificity, validity or reliability information provided	**Experiment 1:** mean proportion of translation equivalents (TEs) in bilinguals' expressive vocabulary at 1;4 was 48.53% (with cognates and semi-cognates removed M = 40%). Children significantly increased the proportion of TEs in their vocabularies by 12.62% between 1;4 and 2;6 ($F(2,32) = 6.91$, $p = .003$). Relative exposure and relative vocabulary size were correlated with the total proportion of translation equivalents in vocabularies at Wave 2 and 3. **Experiment 2:** Parents reported that children had significantly more TEs on the CCT checklist than children showed knowledge of on the CCT ($t(19) = 2.49$, $p = .02$) - parents reported children knew on average 16% more TEs than they showed comprehension of in direct measures.	The study shows that most bilingual children have translation equivalents in their vocabularies by age 1;4, although this varies across children. Children in this study had quite balanced rates of exposure to each language, which might lead to more TEs - in turn, children with more balanced exposure developed more balanced vocabulary sizes in each language. Parents were found in this study to over-report infants' vocabulary comprehension, when compared to a direct assessment of comprehension. The authors suggest that parents may be more likely to report on words that children have only partially mapped, since in the direct comprehension test words are taken out of context. It also may vary depending on parental definitions of 'comprehension' and cultural factors. The authors suggest it is important to use direct measures of vocabulary alongside parent report.

Mancilla-Martinez et al. (2016)	**Concurrent validity**: calculated by relating scores from parent report to direct measures. Correlation coefficients were run for scores across the four parent report forms for the full sample, which were significantly and positively correlated with children's directly assessed productive and receptive vocabulary scores in both English ($p < .001$) and Spanish ($p < .001$). **Discriminant validity:** calculated by comparing magnitude of correlation coefficients between productive and receptive vocabulary. For productive vocabulary, the significance and direction of the correlations for each age bracket were similar to those of the full sample, but a less straightforward pattern was found for receptive vocabulary. In Spanish no significant correlation was found from child ages 38 to 43 months, whereas these same associations in English were significant. In English, there was no significant correlation from child ages 44 to 48 months, whereas the Spanish scores were significantly correlated at this age bracket. **Floor & ceiling effects:** The low number of children who scored at floor and ceiling suggests that the parent report forms appear to be valid for use from age 24 through 48 months. **Developmental**	No information provided	Positive and significant associations were found between parent report measure and direct assessment of children's vocabulary, suggesting parent report has concurrent validity, in spite of varying home language environments. There were stronger associations with measurement of productive than receptive vocabulary, suggesting parents were able to discriminate words their children could produce from those they could understand. Authors recommend use of forms as a measure of change over time - continuity in assessments used is optimal when growth over time is documented.

	sensitivity: There was a positive association for the S-CDI SF and CDI Level II (Spanish: $r = .33$, $p < .001$, $N = 176$; English: $r = .24$, $p = .005$, $N = 136$) and the Spanish Vocabulary Extension and CDI Level III (Spanish: $r = .31$, $p < .001$, $N = 181$; English: $r = .33$, $p < .001$, $N = 155$), indicating that children's performance increased with age.		

Marchman & Martínez-Sussman (2002)	**Concurrent validity:** calculated using Pearson correlation coefficients for CDI/IDHC parent report and behavioural measures of vocabulary. Both English and Spanish reported vocabulary scores were significantly correlated with no. of objects named in the language sample (English CDI/English session: .72**; Spanish IDHC/Spanish session: .78**). Reported vocabulary size in English did not significantly predict number of objects named in Spanish, and reported vocabulary in Spanish was not significantly related to laboratory naming in English. Size of vocabulary in English was significantly correlated with no. of different words produced and total no. of words produced in the English but not the Spanish sessions. Conversely, reported vocabulary in Spanish was significantly correlated with Spanish, but not English, word production. Composite Vocabulary was significantly positively correlated with naming in the English sessions, with positive but weaker relationships in the Spanish sessions. More significantly, composite vocabulary was more strongly related to naming behavior in general, as indexed by the number of objects named in any language, than to either of the individual naming	The majority of the children (n = 16, 61%) had at least 50 reported words in each language, and only three younger children (23, 25, and 26 months) had fewer than 50 reported words in both English and Spanish. Composite scores indicated that only two children were reported to express fewer than 50 concepts when both languages were taken into account. Size of reported vocabularies was significantly correlated with the proportion of English ($r = .56$, $p < .003$) and Spanish ($r = .49$, $p < .05$) input.	The study assessed composite vocabulary in English and Spanish assessments, looking at no. of concepts reported in both English and Spanish. Overall, significant correlations were found between reported and observed measures of word production, suggesting that the CDI and IDHC could provide valid information about language development in bilingual toddlers. It was sometimes necessary for different reporters to fill in different sections of the form (i.e. if one parent had stronger skills in one language, or if a child was spoken to in one language more by another person); the use of multiple reporters did not affect results and may be advantageous in multicultural/multilingual family situations. Relationships between reported and observed grammatical abilities were slightly lower than word production.

	scores. For receptive grammar reporters were 72% accurate in English and 86% accurate in Spanish, with some children incorrectly reported as not yet using combinations. **Factors affecting validity**: Results indicated that correlations remained consistently strong (all $p < .05$) between reported vocabulary size and laboratory naming after the following factors were taken into account: Home Language (English: $r = .69$, Spanish $r = .77$), Proportion of English-to-Spanish input (English: $r = .69$, Spanish: $r = .73$), Mother's Years of Education (English: $r = .66$, Spanish: $r = .82$), Mother's Acculturation Level (English: $r = .53$, Spanish: $r = .76$). Correlations examined according to person completing the form showed that correlations remained strong in both the single- and multiple-reporter groups (rs = .84 and .75) for the CDI. For the IDHC, correlations were higher for the multiple- ($r = .97$) than the single-reporter group ($r = .59$); although both correlations were statistically reliable ($p < .05$). For the grammar measures, correlations were consistently stronger in the multiple-reporter (rs = .93 and .61) than the single-reporter group (rs = .39 and .50) for both the CDI and IDHC.		

Nayeb et al. (2021)	**Sensitivity and specificity of models:** Model 1 - sensitivity: 91% (75-98); specificity = 72% (61-82); PPV = 57% (48-66); LR+ = 3.3 (2.2-4.7) Model 2 - sensitivity: 97% (84-100); specificity = 34% 24-46); PPV = 37% (33-41); LR+ = 1.5 (1.2-1.8) Model 3 - sensitivity: 88% (71-96); specificity = 82% (82-90); PPV = 67% (55-77); LR+ = 4.9 (3.02 - 8.08) Model 4 - sensitivity: 30% (14-50); specificity = 99% (93-100); PPV = 89% (51-98); LR+ = 22 (2.9-167) The clinical examination identified DLD in 32 children (19 boys), corresponding to a sample prevalence of 29% (95% CI 21-38). A total of 28 children (25%) found to have DLD were suspected of having difficulties caused by developmental disabilities, for example neurodevelopmental diagnoses. Severe difficulties were confirmed in 21 children. 10% of children did not cooperate with the screening in either language. Model 2 (screening only in home language) yielded false positives resulting in low specificity; Model 4 yielded few true positives but many false negatives, with low sensitivity overall.	No information provided	Use of screening in only one of the children's languages led to many false positives; screening in both languages was the most effective method, leading to highest specificity and sensitivity. Screening in Swedish with information about language development in mother tongue from parents led to high specificity but low sensitivity, i.e. identified children as not having language impairment but failed to identify those who did. The study found that bilingual children were at high risk of DLD, making up 29% of sample (however it is necessary to query diagnosis of DLD as children are screened at just 2 years)

O'Toole et al. (2016)	**Relevant for test validity:** Different variations of CDI had different numbers of vocabulary items and there was a low but significant correlation between no. of words on checklist and no. of words used by children (rp = .26, p < .001). This relationship was not observed for number of concepts on checklist and TCV scores. **Cut-off points:** children performing in lowest 10th percentile, -1.3SD below mean. At 24-29 months this was a TCV score of 9, or 44 concepts (48 total words). At 30-36 months this was a TCV score of 15, or 96 concepts (105 total words). Overall large variability was found across groups in terms of risk and demographic factors.	**Expressive vocabulary:** children as a group were using significantly more words in their L1 than L2 (t(249) = 10.96, p < .001, d = .85; means 269 and 135, respectively) and had significantly higher TV than TCV scores (t(249) = 12.43, p < .001, d = .09; means 402 and 329, respectively). To compare across groups, raw TCV scores were converted to a percentage of the maximum TCV score possible for each language pair and then converted into z-scores. A 2 (age group) × 6 (language pair) mixed model ANOVA for the group revealed a significant main effect of language pair (F (5, 238) = 8.1, p < .001, h2 p = .15) and age group (F (1, 238) = 29.50, p < .001, h2 p = .11) on %TCV z-scores with no interaction. Post hoc comparisons with Bonferroni corrections for multiple comparisons across the groups indicated that the vocabulary sizes of the Maltese and Polish groups were significantly lower than those of the Irish and Hebrew groups. The French-Portuguese group also had significantly lower vocabulary sizes than the Irish-English group. **Risk factor regression analysis:** Model 1 (age, gender, birth order) - age only significant	Children had significantly greater total vocabularies than total conceptual vocabularies, as has been found for bilingual children in other studies. However, the numbers of words on each adaptation of the CDI were vastly different, which must be considered in any cross-linguistic comparison. Wide variation was found in vocabulary scores both within and between language pairs. Looking at children in the 10th percentile, researchers found that risk factors were similar to monolingual children - most had a vocabulary size of fewer than 48 total words at 24 months and lacked two-word combinations. At 31-36 months they said fewer than 105 words and were also not combining words. There was a high rate of parental concern in the at-risk group, showing the relevance of asking parents if they are concerned about their child's speech and language development. In contrast, effects of other risk factors like SES and family history were inconclusive when comparing children in the 10th and 90th percentile.

		predictor of vocabulary size (p<.001). Model 2 (adding child-external factors) - maternal education also had a significantly positive effect on vocabulary scores (p<.05), and parental concern had a significantly negative effect (p<.001). Model 3 (adding language exposure variables) - maternal education no longer significant (p=.076), but age and parental concern remain significant (p<.001; p<.01). Being exposed 'sometimes' or 'always' to L2 also had significant negative effect on vocabulary scores.	

Patterson, Rodríguez and Dale (2020)	**Validity:** Correlations of the year-end LAP3-L with the last two item scores for the three graduated prompting DA tasks (NAL, Prediction, and SF) were examined using one-tailed tests for significance. Findings were mixed; correlations were positive and significant for two out of the three DA tasks. Children's performance on the last two items of NAL tasks was not significantly correlated with LAP3-L year-end scores, but scores on the last two items of the Prediction and SF DA tasks were significantly and positively correlated with LAP3-L scores: for Prediction, $r = .49$, $p = .015$, and for SF, $r = .58$, $p < .005$. The relationship between children's performance on the SF task and LAP3-L scores remained significant when controlling for age and initial LAP3-L scores. This result is essentially a correlation between SF and change in LAP3-L over the year. Initial LAP3-L scores accounted for 50% (.712) of the variance in year-end LAP3-L scores. SF, the DA task that did have a significant and unique relationship with year-end LAP3-L scores, accounted for an additional 23% (.482) of variance. Combined, the initial LAP3-L score and performance (last two items) on the SF task accounted for a total of 73% of the variance	No information provided	Two of the three Dynamic Assessment tasks used in the study were related to children's subsequent performance on standardised measures; these required children to identify similarity in function between two objects and predict the result of an action. In contrast performance on the novel adjective learning task did not improve significantly across the group as a whole. Children showed within-task improvement during tasks when given prompts and feedback; perhaps contributing evidence of DA procedures tapping modifiability/learning potential. Use of bilingual scoring and development of DA tasks with both languages in mind means that children were credited for both languages and recognised that scores on language tasks are related to a wide variety of variables. Mean times for administration of the tasks were between 4 and 6 minutes, so have potential for use in screening. Future research should explore the cut-off point for item scores on DA tasks that would yield acceptable accuracy of identification of children at risk of language impairment.

	in year-end LAP3-L scores. Children's responses to the first two items of NAL, Prediction, and SF were also examined to determine if there were significant differences early in the tasks. There were significant differences ($F = 8.39$, $p < .01$) among the task scores, and post hoc contrasts indicated scores on the first two items of the SF task were significantly lower than scores on the first two items for the NAL and Prediction tasks ($p < .01$ for SF vs. NAL and for SF vs. Prediction).		

Thordardottir et al. (2006)	No sensitivity, specificity, validity or reliability information provided	**Expressive vocabulary:** Significant difference found in English vocabulary scores between the monolingual English (ME) and English-French (EF) groups (F(1,15) = 23.09, p = 0.000, Z2 ¼ 0.61), with lower scores obtained by the EF group and with no overlap between the normal ranges (+1 SD) of the group. The difference between the monolingual French (MF) and bilingual groups in French vocabulary scores did not reach significance (F(1,15) = 4.41, p = 0.051). A significant group difference was found for Total Vocabulary (TV) (F(2,25) =3.44, p = 0.048, Z2¼ 0.22) and for TCV (F(2,25) = 9.30, p = 0.001, Z2= 0.43). Fischer LSD post hoc tests revealed, for TV, that the only significant difference was between the ME and MF groups (p =0.004), with a higher score obtained by the ME group. For Total Conceptual Vocabulary, the EF group scored significantly lower than the ME group (p =0.001), but did not differ significantly from the MF group (p = 0.332). **Syntax**: The EF group differed significantly from the ME group in English sentence complexity (F(1,15) = 19.39, p = 0.001, Z2= 0.56). The EF group's English	When bilingual children's English scores were compared with monolingual English speakers, they scored significantly lower on almost all language measures, except for receptive vocabulary. As compared with monolingual French speakers, scores were not significantly lower. Bilingual childrens' total vocabulary scores exceeded monolingual children, while Total Conceptual Vocabulary results did not always reach the monolingual 'normal' range; suggesting that use of conceptual vocabulary scores in comparison to monolingual norms may need to be reconsidered. For example, compared to the English-speaking group, bilingual children's TCV was significantly lower, although there was large variability within the bilingual group with some scoring within the normal range. Measures of syntax appeared to be more similar to monolingual groups, in contrast. The authors suggest that use of TCV may be appropriate for children who have an unequal command of languages, but less appropriate for children with very equal exposure. Overall findings suggest that composite scores are 'not immune' to differences in bilingual exposure patterns.

		scores were substantially lower than those of the ME group, with the mean score falling far below the normal range of the monolingual children, and with the SD indicating that children in this group were unlikely to score within the normal range of the monolingual children. In contrast, no significant group difference was found between the EF and MF groups for French sentence complexity (p = 0.275). TSC (Total Sentence Complexity) and TCSC (Total Conceptual Sentence Complexity), ANOVA analysis detected no significant group differences. **Receptive vocabulary:** the mean score of the EF group (88.9, SD 10.5) was lower than that of the ME group (101.1, SD 14.9), but did fall within the normal monolingual range. Over half of the EF children scored within the normal monolingual range in English on this measure of receptive vocabulary. A paired-samples t-test was used to compare the English and French PPVT/EVIP scores of the children in the EF group, revealing no significant difference (p = 0.47). **Receptive syntax:** Between-groups comparison of the English scores revealed a significant difference between the ME and EF	

		groups (F(1,16) = 8.32, p = 0.011, Z2 = 0.34), with the mean score of the EF group falling below the normal range.	

Vagh, Pan and Mancilla-Martinez, (2009)	Comparison of parent and teacher CDI reports show that on average there was greater agreement between parent and teacher reports for children from monolingual families than for children from bilingual families at each time point. Also, with increasing child age and consequently increasing communicative ability, there was an increase in the % agreement between parent and teacher reports. For both groups, agreement increased from child ages 24-30 months after which agreement somewhat stabilized. Yet, for both groups of children agreement was on average less than 50%. Despite low levels of % agreement, parent and teacher reports for the full sample are significantly and positively associated at all time points except at child age 33 months. **Concurrent validity -** Correlation coefficients indicate that parent reports were significantly associated with children's productive (WLPB-R) and receptive (PPVT-III) vocabularies at concurrent time points. Parent-teacher composite reports were significantly associated with children's WLPB-R and PPVT-III scores at child age 36 months. Teacher reports were not significantly associated with either of the reference test scores. General linear mixed	Based on parent reports, children from monolingual English-speaking families had higher productive vocabularies at 24 months than bilingual families; this difference was not found according to teacher reports. Except at age 36 months, teachers reported higher productive vocabularies than parents for bilingual children who were reported to speak mostly Spanish with parents. Teachers report a greater number of unique words beyond age 27 months for children from bilingual families than for children from monolingual families; these differences are significant at child ages 30 months $f(30.3) = -2.61$, $p < .01$; 33 months, $f(20.4) = -3.47$, $p < .001$; and 36 months, $f(43.3) = -2.10$, $p < .05$. No significant difference was noted at the younger ages of 24 and 27 months, $f(22) = 0.29$, $p = .78$, and ?(23) = -0.97, $p = .34$, respectively. Of the children from bilingual families, those who were reported to speak mostly Spanish had lower scores on the WLPB-R than those who were reported to speak mostly English. This finding parallels the finding for the parent CDI reports and lends further credibility to the parent CDI reports.	Scores of children's monolingual and bilingual children's vocabulary exhibited differences according to whether assessments were completed by teachers or parents. Teacher reports did not distinguish significantly between monolingual and bilingual children, suggesting they are not reliable if reported in isolation, although teachers did report on words that parents didn't, suggesting they may contribute unique information to estimation of productive vocabularies. Overall there was insufficient evidence to back up combined teacher-parent report - parent report alone was most associated with scores on direct assessment and provided a better fit with models predicting scores than the composite reports. This indicates that parents are capable of reporting on children's productive vocabulary use at 30 and 36 months, even when they are non-native speakers of English and their children have more than one language.

	regression models showed that parent CDI reports remained significantly associated with WLPB-R scores even after controlling for child age, home language status, and child Spanish-English use, although no longer significantly associated with PPVT-III scores. Similarly the parent-teacher composite reports, like the parent reports, were significantly associated with WLPB-R scores but not with PPVT-III scores after accounting for additional variables. Predictive validity - taxonomy of fitted general linear mixed regression models show that parent reports provide a better fit in predicting WLPB-R scores in comparison to parent-teacher composite reports.		

Welsh Language review - Table of demographic information in studies

Reference			Study design		Study population					
Authors & Year	**Title**	**Journal**	**Aim**	**Methodology**	**No. participants**	**Child age range**	**Home language(s)**	**Type of bilingualism**	**Child gender**	**Demographic information**
Borsley and Jones (2001)	The development of finiteness in early Welsh	Journal of Celtic Language Learning	To examine the development of finiteness in early Welsh on the basis of a large corpus of natural speech from 7 children	Longitudinal	7	1;7-2;3	Welsh	Single home language, bilingual community	2 boys, 5 girls	4 participants from the Bangor area and 3 from the Aberystwyth area
Chondrogianni & John (2018)	Tense and plural formation in Welsh–English bilingual children with and without	International Journal of Language & Communication Disorders	1) To assess the morphosyntactic abilities of Welsh-speaking children with typical development	Cross-sectional	28	4-6 years, 49-82 months	Predominantly Welsh-speaking homes. Mean quantity of exposure: TD group - 86.9%, DLD	Welsh-English successive bilinguals (Welsh L1)	18 boys, 10 girls	Attending Welsh-medium schools located in Bangor or within a 20-mile radius (most Welsh-

	language impairment		on school entry 2) To examine the diagnostic potential of prominent error types made by children with DLD				group - 92.5%			dominant area in the UK)
Chondrogianni & Kwon (2019)	The development of English tense and agreement morphology in Welsh–English bilingual children with and without specific language impairment	Applied Psycholinguistics	Examine the production of English inflectional morphology in Welsh-English bilingual children (Welsh L1) - To explore whether the productivity problems found in the participants' L1 Welsh also occur in their L2 (English);	Cross-sectional	52	Younger groups - 4-6 years; Older group - 7-9 years	Predominantly Welsh-speaking homes. Mean quantity of exposure: Younger age-group -TD group - 86.9%, DLD group - 92.5%; Older age-group - TD group 85.3%	Welsh-English successive bilinguals (Welsh L1)	Not reported	Attending Welsh-medium schools located in Bangor or within a 20-mile radius (most Welsh-dominant area in the UK); Younger group - same participants as Chondrogianni and John (2018)

			examine whether school-age Welsh-English bilingual children with TD and with language impairment have problems with the production of third person singular -s and past tense in English							
Mayr et al. (2015)	Asymmetries in phonological development : the case of word-final cluster acquisition in Welsh–	Journal of Child Language	To provide the first systematic account of word-final cluster acquisition in Welsh-English bilingual children, who	Cross-sectional	40	2;6-5;0	Two groups: only Welsh at home (80-100% input in Welsh), only English at home (80-100% input in English)	Welsh-English bilinguals. Participants categorised according to language dominance defined in	20 boys and 20 girls	All attending Welsh-medium nurseries or primary schools. Residing in Pembrokeshire , west Wales (amongst the

	English bilingual children		differ in language dominance					terms of children's language use within their homes		areas with the highest concentration of Welsh speakers, 52% of population able to speak Welsh
Munro et al. (2005)	Phonological acquisition in Welsh-English bilingual children	Journal of multilingual communication disorders	To obtain normative trends for the development of phonology in Welsh-speaking children	Cross-sectional	83	2;6-5;0	Welsh (2;6-5;0); English (3;0-5;0)	Simultaneous and sequential	42 boys, 41 girls	Attending Welsh-medium nurseries or primary schools. Residing in Pembrokeshire, west Wales (amongst the areas with the highest concentration of Welsh speakers, 52% of population able to speak Welsh

Sharp & Gathercole (2013)	Can a novel word repetition task be a language-neutral assessment tool? Evidence from Welsh–English bilingual children	Child Language Teaching & Therapy	1) To explore the performance of typically developing Welsh-English bilingual children on a Welsh non-word repetition task. 2) To examine whether differences in exposure to Welsh affects performance on non-word repetition tasks in simulaneous and early sequential bilingual children	Cross-sectional	45	4;11-5;11	Three groups: only Welsh at home (80-100% input in Welsh), Welsh and English at home (40-60% input in each language), only English at home (80-100% input in English)	Simultaneous and sequential	24 boys and 21 girls	Attending Welsh-medium schools in Gwynedd and Anglesey.

Thomas (2007)	Natur prosesau caffael iaith gan blant: marcio cenedl enwau yn y Gymraeg (The nature of language acquisition processes in Welsh: marking noun gender in Welsh)	Gwerddon (Welsh language journal)	To examine the acquisition of the ability to mark noun gender in Welsh	Cross-sectional	45	Group 1 (n=11) 4;1-5;5; Group 2 (n=11) 5;9-6;7, Group 3 (n=12) 6;9-7;9, Group 4 (n=11) 8;3-9;7	Welsh	Single home language, bilingual community	Not reporte d	Attending schools in Anglesey or the Bangor area
Thomas et al. (2014)	Acquiring complex structures under minority language conditions: Bilingual acquisition	Bilingualism: Language and Cognition	1) to investigate plural usage in naturalistic adult-adult speech corpora; 2) to analyse children's	Cross-sectional	88	7-11 years (2 age categories : 7;4-8;11 n=40; 9;0-11;2 n=48)	3 groups - Welsh L1, Welsh-English simultaneous bilinguals, English L1	Sequential (Welsh L1, English L1); simultaneou s (2L1)	40 boys and 48 girls	Attended primary schools in Gwynedd, north Wales - long-established bilingual education

	of plural morphology in Welsh		production of plurals in elicited samples							policy and the highest proportion of Welsh speakers in Wales

Welsh Language Review - Table of study results

Authors & Year	Outcomes (information relevant for screening - sensitivity, specificity, reliability, validity, cut-off points)	Summary of language outcome results	What does this tell us about screening for bilingual children?
Borsley and Jones (2001)	Not included in relation to screening	Gradual emergence of finiteness with the earliest clausal utterances having no finite element. Then, finite clauses do appear (especially those with forms of the copula 'bod') but sentences with no finite element continue to occur. Tentative conclusions drawn regarding the development of children's grammars across this time period, the main one being persistence of sentences with no finite verb (termed by the authors as missing copula sentences) after the first appearance of finite sentences. The authors posit that changes in the grammars of the children are likely to be a result of the children's processing ability i.e. children are more likely to produce sentences with fewer words	Some examples of early productions included. Points made regarding relationship between patterns in the adult input and the structures produced by the children - missing copula sentences not uncommon in adult Welsh but there are differences between these types of sentences and those produced by children in this study
Chondrogianni & John (2018)	Diagnostic accuracy investigated by 1) sensitivity and specificity analysis - , 2) ROC curve analysis, 3) Likelihood ratio analysis. Periphrasic past tense - poor diagnostic	Paper reports on the diagnostic potential of grammatical morphology in the identification of Developmental Language Disorder, giving examples from other languages. Children with	Some information on which grammatical markers IN WELSH that differentiate between children with typical development and children with language disorder

	accuracy - high overall accuracy for both groups, overlap in performance, increasing with age. Synthetic past tense showed better differentiation between the groups - children with an accuracy score of 35% and below can be classified as having DLD. Plural task - low overall scores for both groups, some overlap; DLD group did not improve with age. Area under the curve was excellent .95 for synthetic past tense and .92 for plural formation. Production of singular nouns in the plural task had the best diagnostic value (sensitivity i.e. % accuracy DLD -100; specificty i.e. % accuracy TD - 94.4)- error types can help us distinguish between clinical and typical groups. The combination of the two tasks tapping into concatenating morphology (synthetic past tense and plural formation) increased their potential of correctly classifying children with typical language development, as demonstrated by specificity values and likelihood ratios.	DLD are tested on their production of two different types of past-tense and plurals in Welsh. PAST TENSE **-** Results show that both groups produced similar types of past tense forms, with the periphrasic form produced more often than the synthetic form of the verb in the elicitation task (as expected given that it is a more frequent form). Typically developing children at 6 years are almost at ceiling in their production of the periphrasic past verb forms, also able to produce synthetic past tense. Children with DLD exhibited problems with tense morphology at varying degrees across the two tense tasks. For the periphrasic form, there was lower accuracy and higher proportion of errors for DLD children but this form was not as problematic as reported in other languages such as English (could be linked with the properties of Welsh auxiliaries - obligatory for the formation of periphrasic tenses, past and present; remain syllabic forms even when reduced; less likely to be omitted in the context of 3rd singular forms - this contrasts with the contracted auxiliary in English). Greatest group difference shown through sentence completion task where the production of the synthetic form of the verb was targetted - for DLD group, lead-in sentence did not help (contrast	CLINICAL IMPLICATIONS **-** Nominal plural formation is late acquired in Welsh-speaking children, periphrasic use of the past tense is acquired early. For TD group, the synthetic past tense is part of their grammatical repertoire and they can produce it withe appropriately prompted; Disagnostic potential of some of the tasks and identification of error types (see outcomes cell); children with DLD are less likely to benefit from prompting when it involves a complex structure that is not yet part of their grammatical abilities.

		to TD peers). PLURALS **-** both groups still in process of acquiring Welsh plural formation patterns at 6 years. This study showed that plural formation was particularly difficult for children with DLD. Minority bilingual setting for Welsh is cited as a factor to consider (other languages with complex plural systems e.g. Iclandic, Hungarian are early acquired in TD children). Different error patterns also seen with DLD group more likely to use singular noun or an over-regularisation of the most frequent suffix.	
Chondrogianni & Kwon (2019)	Not included in relation to screening	Older children outperformed the younger children on regular past tense formation and produced more errors with regular verbs. TD groups did not differ on accuracy of irregular verbs, but did differ on error types with older children producing more overregularisations and fewer unmarked forms with irregulars. Production rates of the third person singular -s also improved with age i.e. increased exposure to English. Children with DLD performed more poorly (lower accuracy and different error types) than age matched TD peers on the two morphemes involving concatenating morphology (third person singular and regular past) when differences in vocab skills were controlled for.	Some information on which grammatical markers IN ENGLISH that differentiate between children of different ages and typical development vs children with language disorder

		Seems that the problems with children with DLD have with productive morphology in their L1 (Chondroginanni and John, 2018) transfer to their L2 - also compounded by lack of exposure to their L2. Vocabulary size is a significant child-internal predictor of children's performance on tense and agreement morphemes	
Mayr et al. (2015)	Not included in relation to screening	This paper provides percent correct performance on the word-final clusters, broken down by age and language dominance. Tables included on acquisition stages for Welsh and English depicting which clusters reached the 75% criterion by 3;0, 3;6, 4;0, 4;6 and 5;0 FINDINGS **-** English clusters produced more accurately than the Welsh ones overall (both dominance groups); main effect of cluster type in Welsh and English; main effect of dominance in Welsh dataset with Welsh-dominant outperforming English-dominant participants - no difference in English dataset showing that there was not a concomitant lag; investigation of error patterns showed that error rates decreased with age as did variability - dominance a factor with variability with increased variability seen for English-dominant participants attempting Welsh clusters; analysis of error types also included with	CLINICAL IMPLICATIONS **-** benchmark for normal acquisition, findings support the notion that monolingual norms cannot be readily be applied to bilinguals - rate and pattern of development of English word-final clusters in this study differed substantially from age-matched English monolinguals reported elsewhere; underlines the need to take into account diverse input patterns - norms needed to represent different groups of bilinguals

		reduction and substitution being the most common in both languages.	
Munro et al. (2005)	Not included in relation to screening	3-way unrelated analysis of variance undertaken to examine the effects of the three factors (age [with the 2;6 to 3;0 group excluded], sex and langauge dominance). Effect of age: improvement in production of /p/ seen in groups 4;0-4;6 and 4;6-5;0 as compared to earlier ages in Welsh word set; for English word set, improvement in production of /t/ over time - interaction with sex with females showing more steady improvement, interaction with language dominance with Welsh-dominant subjects showing an advantage; /ŋ, θ, ʒ χ, ɬ, ʧ / show age effect on both Welsh and English words. For /θ/ age interacts with dominance with Welsh dominant children showing poorer and more erratic performance on both languages, for /ʧ/ age interacts with dominance with Welsh-dominant children showing a more erratic performance for the English words; /ʤ/ improves with age only for English words with Welsh-dominant children showing a stronger effect; /r, r̥/ shows statistically significant effect of age with a strong age-related improvement on Welsh words. Effect of dominance: /t, d, g, ŋ, v, r, r̥/ show effect	Not directly related to screening but this paper does include normative data related to speech sound acquisition. Figures are included that show the stages of acquisition of the consonants according to target language and language dominance. Sounds are deemed to be acquired when at least 75% of the children in the cohort achieved at least 75% accuracy for the sound in question. Patterns of substitution are also described giving some information about what to expect within typical development for this population. Conclusions drawn about variability and the trajectory of phonological acquisition as well as specific sounds or sound categories that show greater dominance effects. As this is one of the few papers where young Welsh-English bilinguals have been studied, the points made about separation of systems and the influence of the variety of Welsh and English they are acquiring are interesting. There is also a point made about the relationship between abilities in the two languages with a positive correlation indicating that the higher % accuracy was in English, the higher it was in Welsh too - this suggests a stong effect of

		of language dominance with Welsh-dominant children showing an advantage - in the Welsh words for /d, g, ŋ, v, r, r̥/ and in the English words for /t/; /s, ʃ, ʧ, ʤ/ show language dominance effect with English-dominant children showing an advantage for English words; /l/ showed a dominance effect where the direction was modulated by sex - in females, Welsh-dominant participants had the advantage whereas in males, English-dominant participants had the advantage. Correlations: the relationship between overall performance on Welsh and English words was investigated - positive correlations were found showing that better performance on the Welsh words meant better performance on the English words, p ≤ .01 level for the following sounds: /b, k, g, θ, ʃ , ʧ, ʤ, l/, p ≤ .05 level for the following sounds: /p, ŋ, r, r̥, h/	English on all participants regardless of language background with performance in English enhancing performance in Welsh but not the other way around.
Sharp & Gathercole (2013)	Not included in relation to screening	Accuracy **-** No group differences on shared sounds. Children with English as a home language performed less well than children with Welsh as a home language on Welsh-specific sounds. Error Types - Some consonants/clusters showed homogeneous error type patterms across home language groups e.g. /ð/ -> [v][d], /ɬ/ ->	Some discussion of non-word repetition as a universal tool that can be used across different languages but the point is made that non-word repetition measures phonological processing and memory which give rise to phonological representations.These representations are language-specific and therefore children perform

		(x)[s], reduction of the clusters /tr/, /dr/, /tl/ and /gn/ to the plosive component. Other consonants and clusters showed intermediate levels of agreement in substitutions across language groups, most of these being cluster reduction or substitution of one of the elements. A group of phonemes where seen to show marked differences in error types across home language groups, these included /r/ where the only Welsh at home group and the Welsh-English at home group produced [l] or a cluster in most cases, the only English at home group produced [ɹ] in 65% of the cases, indicating transfer from English. Differences seen in children hearing Welsh from birth but with different levels of exposure (i.e. 'only Welsh and home' and 'Welsh and English at home' groups)	differently according to their experience and exposure to the language. The authors posit that designing a language-neutral non-word repetition task is virtually impossible as it is not possible to bypass assessment of the child's proficiency in the particular language.
Thomas (2007)	Not included in relation to screening	Results revealed performances approaching L1 adult norms among L1 Welsh-speaking bilinguals, but delayed progression among the other bilingual groups, although analyses of errors revealed various levels of structural knowledge. Forms involving alterations to the root were more difficult to acquire than those where a plural suffix was to be applied. Increased levels of exposure to the language led to greater levels of accuracy.	The age range of this study is outside the range that is being considered for screening but this has some relevance in that it shows how a complex structure is acquired differently by the three groups. It also shows that the plural system of Welsh is far more complex than English and therefore highlighting cross-linguistic differences in the complexity and developmental trajectory of grammatical structures

Thomas et al. (2014)	Not included in relation to screening	Focus on Group 1 for the purposes of this review. Age 4 and a half - mean score 0.796 for proportion of accurate productions of the noun. Effect of gender was found with all groups producing masculine nouns with greater accuracy. Main effect of age was found in relation to the feminine nouns only	Welsh has a complex grammatical gender system which is linked to the mutation system (3 mutation types i.e. phonetic changes that are lexically, morphologically and syntactically motivated). This system is not fully acquired by children at age 9 although there is evidence that children make a distinction between the two genders.

Annex D: Screening tools considered

Name of Tool	Recognised Abbreviation	Country of Origin	Author(s)
Ages and Stages Questionnaire 3	ASQ3	United States of America	Squires, J. and Bricker, D.
Arizona Articulation Proficiency Scale	Arizona-3	United States of America	Fudala, J.B. and Stegall, S.
Bankson-Bernthal Test of Phonology	BBTOP	United States of America	Bankson, N.W. and Berthnal J.E.
Battelle Developmental Inventory Screening Test 3	BDI-3	United States of America	Newborg, J.
Bayley Infant Neurodevelopmental Screener	BINS	United States of America	Aylward, G.P.
Bayley Scales of Infant and Toddler Development - Third Edition	Bayley III	United States of America	Bayley, N.
Bracken Basic Concepts Scale	BBC	United States of America	Bracken, B.A.
Brigance Preschool Screen	n/a	United States of America	Brigance, A.
Bristol Language development scales	BLADES	United Kingdom	Gutfreund, M., Harrison, M. and Wells, G.
Bristol Surveillance of Children's communication	BRISC	United Kingdom	North Bristol NHS Trust
British Picture Vocabulary Scale III	BPVS III	United Kingdom	Dunn, L.M., Dunn, D.M. and Styles, B.
Broad Target Recast	BTR	n/a	concept introduced by Saxton, M.

Carrow Elicited Language Inventory	n/a	United States of America	Carrow-Woolfolk, E.
CLEAR Phonological Screening Assessment	CLEAR	United Kingdom	Keeling M. and Keeling, K.
Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale	CAT/CLAMS	United States of America	Voigt, R.G., Brown, F.R. and Fraley, J.K.
Clinical Judgement	n/a	n/a	n/a
Communication and Symbolic Behavior Scales-Developmental Profile	CSBS-DP	United States of America	Wetherby, A.M.and Prizant, B.M.
Communication Function and Classification System	CFCS	United States of America	Hidecker, M.J.C., Paneth, N., Rosenbaum, P.L., Kent, R.D., Lillie, J., Eulenberg, J.B., Chester, JR, K.E.N., Johnson, B., Michalsen, L., Evatt, M. and Taylor, K.
Davis Observation Checklist for Texas	DOCT	United States of America	Texas Department of Health
Denver Articulation Screening exam	DASE	United States of America	Drumwright, A., Van Natta, P., Camp, B., Frankenburg, W., Drexler, W.
Denver Developmental Screening Test-II	DDST	United States of America	Frankenburg, W.K. and Dodds, J.B.
Derbyshire Language Scheme	DLS	United Kingdom	Knowles, W. and Masidlover, M.
Derbyshire Language Scheme Rapid Screening Test	DLS RST	United Kingdom	Knowles, W. and Masidlover, M.
Developmental Nurse Screen	n/a	Hong Kong	Stokes, S.F.

Developmental Profile-II	DP-II	United States of America	Alpern, G., Boll, T. and Shearer, M.
Diagnostic Evaluation of Articulation and Phonology	DEAP	United Kingdom	Dodd, B., Hua, Z., Crosbie, S., Holm, A. and Ozanne, A.
Dynamic Indicators of Basic Early Literacy Skills	DIBELS	United States of America	Biancarosa, G., Kennedy, P.C., Park, S. and Otterstedt, J.
Early Language Identification Measure	ELIM	United Kingdom	Law, J., Charlton, J., McKean, C., Watson, R., Roulstone, S., Holme, C., Gilroy, V., Wilson, P. and Rush, R.
Early language Milestone Scale 2	ELM - 2	United States of America	Coplan, J.
Early Literacy and Developmental growth and development indicators	EL-IGDIs	United States of America	Early Childhood Research Institute
Early Screening Profiles	ESP	United States of America	Harrison, P., Kaufman, A., Kaufman, N. Bruininks, R., Rynders, J., Ilmer, S., Sparrow, S. and Cicchetti, D.
Early Years Foundation Stage Unique Child Communication Sheet	n/a	United Kingdom	n/a
Edinburgh Articulation Test	EAT	United Kingdom	Anthony, A., Bogle, D., Ingram, T.T.S. and McIsaac, M.W.
Expressive one word picture vocabulary test IV	EOPVT	United States of America	Martin, N.A. and Brownell, R.
First Words and First Sentences Test	FWFST	United Kingdom	Gillham, B., Boyle, J. and Smith, N.
Fluharty Preschool Speech and Language screening test 2	FLUHARTY 2	United States of America	Fluharty, N.B.

Focus on the Outcomes of Communication Under Six	FOCUS	Canada	Thomas-Stonell, N.L., Oddson, B., Robertson, B. and Rosenbaum, P.L.
Foundation Phase profile	n/a	United Kingdom	Welsh Government
General Language Screen	GLS	United Kingdom	Stott, C.M., Merricks, M.J., Bolton, P.F. and Goodyer, I.M.
Goldman-Fristoe Test of Articulation III	GFTA III	United States of America	Fristoe, M. and Goldman, R.
Hackney Early Language Screening Test	HELST	United Kingdom	Law, J.
Healthy Child Wales Developmental Milestones	n/a	United Kingdom	Welsh Government
Health/Education Professional Judgement	n/a	n/a	n/a
Hodson assessment of Phonological patterns	HAPP 3	United States of America	Hodson, B.
Illinois Test of Psycholinguistic Abilities	ITPA 3	United States of America	Kirk, S.A.
Language Development Survey	LDS	United States of America	Rescorla, L.
Language Link	n/a	United Kingdom	Speechlink Multimedia Ltd
Language Use Inventory	LUI	United States of America	O'Neill, D.
LENA Developmental Snapshot	n/a	United States of America	LENA.org
Levett-Muir Language Screening test	n/a	United Kingdom	Levett, L. and Muir, J.

MacArthur-Bates Communicative Development Inventory	MacArthur Bates CDI	United States of America	San Diego State University
MacArthur-Bates Communicative Development Inventory - Short Form	n/a	United States of America	San Diego State University
Mary Sheridan's developmental milestones	n/a	United Kingdom	Sheridan, M., Sharma, A. and Cockerill, H.
Mean Length of Utterance	MLU	n/a	n/a
New Reynell Developmental Language Scales	NRDLS	United Kingdom	Edwards, S., Letts, C. and Sinka, I.
Northwestern syntax screening test	NSST	United States of America	Ratusnick, D.L., Klee, T.M and Ratusnick, C.M.
Nuffield Early Language Intervention	NELI	United Kingdom	Snowling, M. and Hulme, C.
Nuffield Early Language Intervention Language Screen	NELI language screen	United Kingdom	OxEd & Assessment
Parent evaluation of development status	PEDS	United States of America	Glascoe, F.P.
Parental Judgement	n/a	n/a	n/a
Parent Language Checklist	n/a	n/a	n/a
Parent questionnaire with comprehension item	n/a	n/a	n/a
Parent Report of Children's Abilities – Revised	PARCA-R	United Kingdom	Johnson, S., Bountziouka, V., Linsell, L., Brocklehurst, P., Marlow, N. Wolke, D. and Manktelow, B.
Peabody Picture Vocabulary Test IV	PPVT - IV	United States of America	Dunn, L. and Dunn, D.

Pediatric Language Acquisition Screening tool for early referral	PLASTER	United States of America	Sherman, T. and Shulman, B.
Percentage Consonants Correct	PCC	n/a	Concept introduced by Shriberg, L.D. & Kwiatkowski, J. (1982)
Phonological Awareness Literacy Screening	PALS	United States of America	University of Virginia
Phonological Knowledge Profile/Productive Phonological Knowledge Profile	n/a	Australia	Victoria State Government
Photo articulation test III	PAT III	United States of America	Lippke, B., Dickey, S.E., Selmar, J.W. and Soder, A.L.
Preschool and Primary inventory of phonological awareness	PIPA	UK edition (2000)	Dodd, B., Crosbie, S., McIntosh, B., Teitzel, T. and Ozanne, A.
Preschool Clinical Evaluation of Language Fundamentals - 3	CELF-pre 3	United States of America	Wiig, E.H., Secord, W. and Semel, E.
Preschool Language Scale-3	PLS -3	United States of America	Zimmerman, I.R., Steiner, V.G. and Pond, R.E.
PRoduction of Infant Scale Evaluation	PRISE	Israel	Kishon-Rabin, L., Taitelbaum-Swead, R., Ezrati-Vinacour, R., Kronnenburg, J. and Hildesheimer, M.
Renfrew Language Scales Action Picture Test	RAPT	United Kingdom	Renfrew, C.
Renfrew Word Finding Vocab Test	RWFVT	United Kingdom	Renfrew, C. and Mitchell, P.
Renfrew Bus Story Test	RBST	United Kingdom	Renfrew, C.

Rhyme identification from the Rhyme Individual Growth and Development Indicator	n/a	United States of America	Wackerle-Hollman., A.K. Schmitt., B.A., Bradfield., T.A., Rodriguez, M.C., and McConnell., S.R.
Rigby's trial speech screening test	n/a	United Kingdom	Rigby, M. and Chesham, I.
Schedule of Growing Skills	SOGS II	United Kingdom	Bellman, M. H., Lingam, S. and Aukett, A.
Screening Kit for Language Development	SKOLD	United States of America	Bliss, L.S. and Allen, D.V.
Sentence Repetition Screening test	SR	United States of America	Sturber, R.A., Kunze, L., Funk, S.G. and Green, J.A.
Sequenced Inventory of Communication Development	SICD	United States of America	Hedrick, D.L., Prather, E.M. and Tobin, A.R.
Sequenced Language Scale for Infants	SELSI	Korea	Kim, Y.T., Kim, K.H., Yoon, H.R., Kim, W.S.
Sound Blending subtest of Woodcock-Johnson III - Revised – test of Cognitive abilities	n/a	United States of America	Woodcock, R.W., McGrew, K.S. and Mather, N.
South Tyneside Assessment of Phonology II	STAP II	United Kingdom	Armstrong, S. and Anley, M.
South Tyneside Assessment of Syntactic Structures	STASS	United Kingdom	Armstrong, S. and Anley, M.
Speech and Language Screening Questionnaire	n/a	n/a	n/a
Structured Photographic Expressive Language Test – preschool	SPELT-P 2	United States of America	Dawson, J., Eyer, J. and Fonkalsrud, J.
Surestart Language Measure	SSLM	United Kingdom	Roy, P., Kersley, H. and Law, J.
Teaching Talking checklists	n/a	United Kingdom	Unknown

Test for Examining Expressive Morphology	TEEM	United States of America	Shipley, K., Stone, T. and Sue, M.
Test of abstract language comprehension 2	TALC 2	United Kingdom	McLaughlan, H. and Elka, L.
Test of Auditory Comprehension of Language 4	TACL-4	United States of America	Carrow-Woolfolk, E.
Test of Language Development (Primary) 5	TOLD-P 5	United States of America	Newcomer, P.L. and Hammill, D.D.
Test of Phonological Awareness	TOPA	United States of America	Robertson, C. and Salter, W.
Test of receptive grammar 2	TROG -2	United Kingdom	Bishop, D.
Toddler Language and Motor Questionnaire	TLMQ	Iceland	Gudmundsson, E.
Toddler Phonology Test	TPT	United Kingdom	Dodd, B. and McIntosh, B.
Uppsala Screening Test	n/a	Sweden	Westerlund, M. and Sundelin, C.
Vineland Adaptive Behavior Scale	VABS	United States of America	Sparrow, S.S., Cicchetti, D.V. and Saulnier, C.A.
Vocal Development Landmarks Interview	VDLI	United States of America	Moeller, M.P., Thomas, A.E., Oleson, J. and Ambrose, S.E.
VroegTijdige Onderkenning Ontwikkelingsstoornissen Language Screening Instrument	VTO	Netherlands	Ridder-Sluiter, J.G. and van der Lem, G.J.
Ward Infant Language screening test assessment acceleration remediation	WILSTAAR	United Kingdom	Ward, S.
WellComm (Early Years)	WellComm	United Kingdom	Sandwell Primary Care Trust

Annex E: Features of an effective screening tool

This annex relates to objective 3 of the specification.

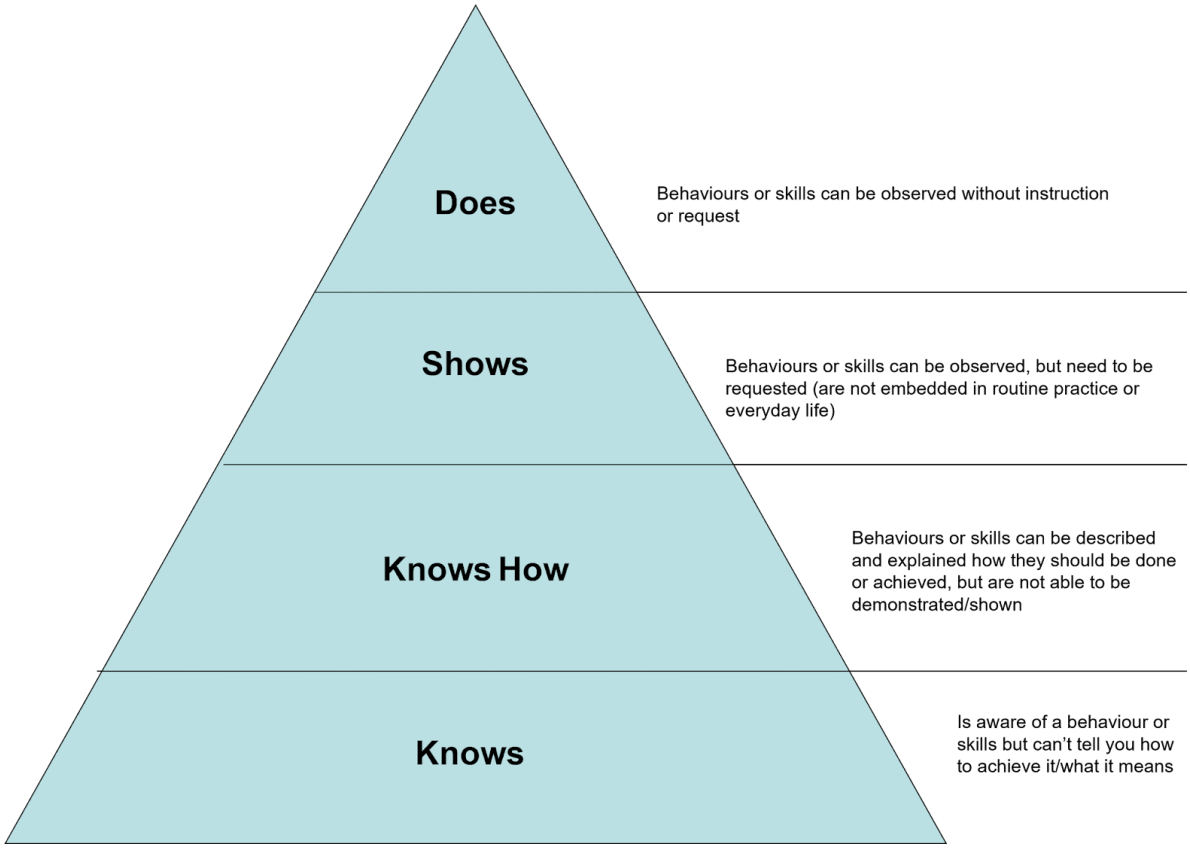
Objective 3: Assess, critique and summarise what makes an effective screening/ identification tool. Factors including validity, reliability, specificity and bi/multilingualism will need to be considered.

The purpose of a screening tool is to accurately and reliably identify those individuals who might benefit from intervention to address whatever condition is being screened for. It is not intended simply to diagnose if no treatment or support can be made available. Neither is it intended to be diagnostic in nature; that process follows for those identified from the screening process. In order to be effective as a screen however, tools which purport to carry out that function must fulfil certain key criteria. This section considers those criteria and how they relate to tools used in screening for speech, language and communication needs.

Taxonomies of assessment and screening

Screening and assessments can take many formats. As a consequence, a framework for structuring the appraisal of assessments has been proposed (Van der Vleuten 1996). The framework can be viewed as a system for describing and classifying the behaviours to be elicited by the assessment or screening task (Messick 1989). The framework most often used as a plan in medical education is Miller's Pyramid (Figure E1) as it provides a robust and useful construct to guide thinking about both service delivery and a framework within which to relate assessment type to the overall aim (Miller 1990). Miller's Pyramid has the advantage of mapping onto several of the descriptions of who competency can be demonstrated e.g. 'knows' is where a person is aware of a task, but cannot tell you how to complete it (Albanese et al. 2008; Carraccio et al. 2002; Council on Linkages between Academia and Public Health, 2008; Frank, 2005). It is also well established as providing a meaningful and successful rational system or plan to connect the content of the screening tool to the knowledge, skills and abilities judged as important (Downing & Yudkowsky, 2009).

Figure E1: Miller's Pyramid



Once the level at which the screening has been defined, e.g., is the child just meant to understand what is being asked (Knows), or to be able to demonstrate it to you when asked (Shows), the individual components should be considered in a system that satisfies the characteristics of a good screening tool (Joint Centre for Education in Medicine, 1997).

Screening Requirements

A utility equation has been suggested to structure these components: Utility = V x R x E x A x 1/C, where V is validity, R reliability, E educational impact, A acceptability and 1/C is the relationship of cost of the assessment with the other components (Van der Vleuten 1996). Although the equation provides a comprehensive guide for the factors that need consideration, the feasibility/practicality of administration is an important additional factor to be considered and is included here.

Using the utility equation, a review of screening tools was conducted. This initially focused on the validity and reliability of each tool and then its feasibility and practicability of administration and then its potential impact.

Utility

Utility or usability in respect to validity, reliability, impact, acceptability and cost, can be applied to an entire screening system or to an individual screening tool or component of the system. The concept is important in that no single element of utility should be regarded as a panacea. Instead, screening tool selection should pay attention to all the elements within the utility equation, although it is recognised that there may be a 'trade off' between the elements.

In using the utility equation, the criteria are weighted according to the importance attached to each of them by a specific user in a specific situation and this defines the utility of the method (Amin, Seng & Ang, 2006). This means that the weight of the criteria depends on how important each of the different criteria are perceived by those responsible for screening in a certain situation or screening context. An additional criterion of diagnostic power was also included where data on sensitivity, specificity and predictive values were available.

- **Validity**

Validity refers to the extent to which a measurement actually measures what it is intended to measure. Validity, unlike the classical test theory of reliability, cannot be expressed in a single coefficient; it is a conceptual term that takes several forms (Van der Vleuten 2000):

During the selection of the screening tools each was considered in turn:

- Face Validity: Acceptability of the instrument to the users (administrators) in determining its usefulness to measure what it is supposed to measure.
- Content validity: When an examination is carefully designed through good selection of the topics to be assessed it is described as having content validity.

- Construct validity: An assessment's ability to differentiate between groups with known differences in ability, such as beginners and experts in a particular area, is often called construct validity. This is because a theoretically predicted outcome of an experiment—in this case the differentiation between groups—underpins the "construct" being measured.
- Convergent, divergent, and predictive validity: Validity can also be shown by the strength with which scores for one measurement are related to other measures. When two instruments are expected to measure similar constructs, the correlation between their scores is taken as an index of convergent validity. Similarly, when the measurements should measure different aspects, the correlation is a reflection of divergent validity.
- Criterion validity: is an estimate of the extent to which a measure agrees with a gold standard (i.e., an external criterion of the phenomenon being measured).

- **Reliability**

Apart from validity, reliability is the most important quality to seek in screening tool results. Reliability refers to the consistency and precision of measurement or the reproducibility of the scores obtained with the examination. However, many forms of 'noise' can affect the measurement and therefore the reliability. Several statistical theories have been developed to estimate the reliability of constructed assessment tools.

- **Internal consistency** is a measure based on the correlations between different items on the same test (or the same subscale on a larger test). It measures whether several items that propose to measure the same general construct produce similar scores. Internal consistency is usually measured with Cronbach's alpha, a statistic calculated from the pairwise correlations between items. It is commonly accepted that an alpha of 0.6-0.7 indicates acceptable reliability, and 0.8 or higher indicates good reliability.

Note that extremely high reliabilities (0.95 or higher) are not necessarily desirable, as this indicates that the items may not be just consistent, but redundant (Cortina, 1993).

- **Interrater reliability** measures homogeneity of agreement between raters and is used when the same form is reviewed/scored by two or more raters to establish the extent of consensus on use of the instrument by those who administer it. Cohen's Kappa for interrater reliability can be used to assess interrater reliability if there are just two raters and Fleiss' Kappa can be used for multiple raters (McHugh, 2012).
- **Intraclass correlation (ICC)** is used to measure inter rater reliability for two or more raters. It may also be used to assess test-retest reliability. ICC may be conceptualised as the ratio of between-groups variance to total variance, as elaborated below (Shrout & Fleiss, 1979).

The factors of reliability and validity are elements to consider when looking at how 'good' the tools are at measuring what they claim to measure. One or more of these should be reported in publications (manuals or peer-review articles), to enable the potential users of the tools to make an informed choice.

- **Impact**

Impact focuses on a screening tool's objectives matching the objectives of the service or the service provider.

- **Acceptability**

Acceptability of the screening tool to both the family (service user) and administrator (service provider) is important. For the service user the tool must have a feeling of assessing or measuring the 'right' type of thing, as well as not being burdensome. For service providers it must not restrict their freedom as professionals. These are important factors to be considered and used in both the construction and administration of screening tools to ensure they are acceptable.

Acceptability is based on personal experience, beliefs and conceptions (Van der Vleuten 1996).

The administration of a screening tool is the most public and visible aspect of usage and can also be viewed as an element of acceptability. In selecting screening procedures and formats, practical considerations cannot be neglected. The time available for administration is almost always limited because screening is in constant competition with other important activities for time in the schedule (Linn & Miller 2005). These and other factors pertinent to the usability of assessment procedures must be taken into account when selecting screening procedures.

- **Cost**

The cost of screening tools is a variable of increasing importance: resource limitations are universal, even more so for single institutions or service providers (Van der Vleuten 1996).

- **Diagnostic power**

Sensitivity and specificity are often used to assess accuracy of scale, especially in Criterion related validity.

- Sensitivity: Instrument correctly identified the cases (true positive)
- Specificity: Instrument correctly identified the noncases (true negative)

As mentioned under criterion validity, the accuracy of a screening tool can only be directly reported in relation to how it compares to another reference test and to the expectations regarding adequate levels of sensitivity and specificity.

Selection of Assessment Formats

Overall, the selection of a screening tool or tools, requires a service to

1. Identify the area(s)/factors that they want to screen
2. Apply the Utility equations
    1. Validity
    2. Reliability

3. Impact
4. Acceptability (to staff and service users)
5. Cost
6. Diagnostic power

Annex F: Background to child language acquisition in Wales

Child language acquisition is critical to children's long-term development and wellbeing across multiple domains (Maggie et al., 2010; Marmot et al., 2010) and has a pivotal role in promoting social mobility (Irwin, Siddiqi & Hertzman; 2007; Law, Reilly & Snow, 2013; UNICEF, 2012; Wylie et al., 2014). Monitoring children's language in the early years is therefore vital to identify those who need additional support and intervention and indeed, as the Early Intervention Foundation reported, as an indicator of child wellbeing (Law, Charlton & Asmussen, 2017).

However, the development of speech and language abilities takes place over time, adding additional complexity to the process of monitoring. Nevertheless, investigations involving careful examination of children's abilities at different points along the developmental trajectory has enabled researchers and practitioners to create a set of expectations regarding what children should be able to do at different ages. This information is pivotal to our ability to monitor and screen at a universal, population and targeted level.

The purpose of screening is to identify those children whose speech and language development is below that which would be expected for their age. Their skills can be evaluated through observation and/or direct assessment by a skilled practitioner with data also being gathered from parents. The former can provide a consistent measure and, where available, comparison to population level data but require specialists trained in the use of the tool and can be problematic with young children; the latter is a practical alternative but one which relies on subjective judgements (Law et al., 2017).

A screening tool needs to be sensitive enough to identify children who are likely to have difficulties which persist and/or will impact on their progress at school, their ability to communicate, their social relationships and quality of life. It also needs to be specific enough to identify those children whose language progression is likely to lead to typical skills and avoid over-identification with a consequent drain on resources (Siu, 2015; Smith, 2018; Wilson, & Jungner, 1968). Acceptable levels of sensitivity and specificity for child language screening have been identified at 72 per cent and 88 per cent respectively (Law et al., 2000), however, relying simply on children's language output in the early years is an unreliable indicator as children's trajectories change over time (Ukoumunne et al., 2011).

Supplementing information on children's language outputs alongside other factors has been shown to increase the predictive validity of screening for speech, language

and communication needs (Law et al., 2012; Wren et al., 2016). A report for the Education Endowment Foundation suggested that a screen of children's language should include questions relating to family and parenting factors, presence of additional needs (such as speech and language difficulties) and children's social and emotional wellbeing (Law et al., 2017) while Wren et al. (2016) identified additional factors including weak sucking at age 4 weeks and a history of problems with coordination as predictive of low outcomes for speech in school-age children.

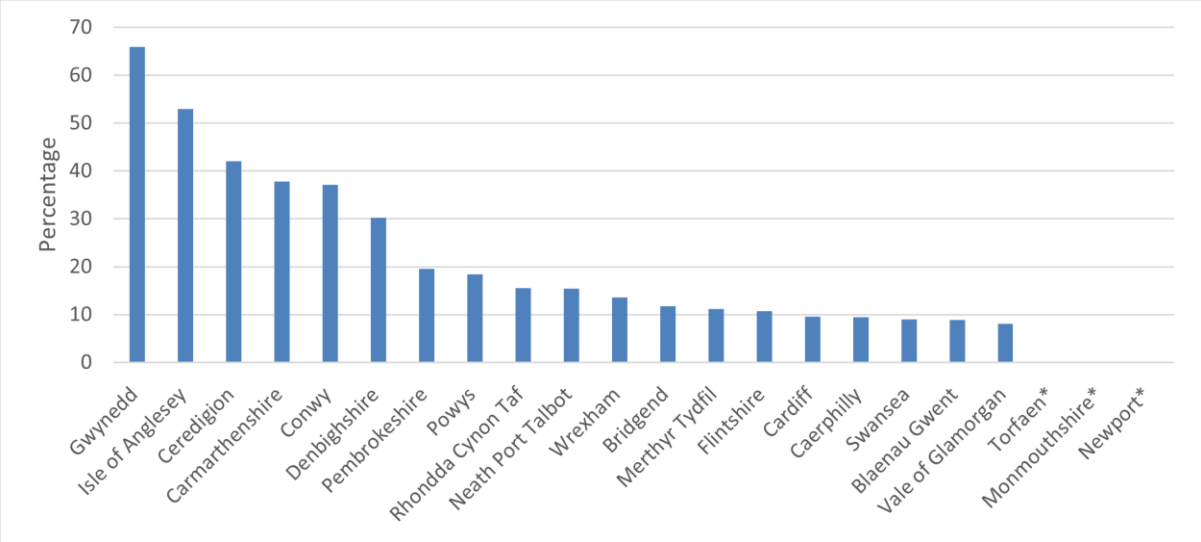
A process of universal and targeted screening of early language development has been adopted in England based on the Early Language Identification Measure (ELIM; Public Health England, 2020). This measure has been developed in accordance with current evidence which reflects the social, linguistic and cultural context in England. Some children residing in England will only be exposed to English within their homes, care and education settings as well as in the wider community. However, many children will also be exposed to other languages within their homes and communities leading to bilingual or multilingual language acquisition. In Wales, the presence of two official languages, Welsh and English, gives rise to a different linguistic landscape to England, with various patterns of language use in the home, in education and in the community. Thus it may be the case that children are exposed to Welsh or English only within the home whilst others will encounter bilingual or multilingual language exposure.

Worldwide, bilingualism is the norm rather than the exception (Kohnert and Medina, 2009), and as mobility and migration continue to grow, speech and language therapy services are increasingly called upon to provide services to bilingual and multilingual clients (Williams and McLeod, 2012). Bilingual children do not form a homogeneous group (Nayeb et al., 2015). Some children acquire two languages from birth (simultaneous bilingualism) while for others their additional language is not introduced until later (sequential bilingualism) (Pert and Bradley, 2018). Factors such as the age of onset, frequency of exposure and societal status of their languages affect the trajectory of language development for bilingual children (Floccia et al., 2018). While bilingual children experience many milestones of speech and language at around the same stage as monolingual children, there is also likely to be influence and transfer from one language to the other, meaning that stages will not always appear in a similar way to a monolingual child (Haman et al., 2017). The inclusion of a multilingual perspective within the ELIM means there is much which can be used to inform the current proposal, however it is not appropriate to simply adopt it for use in Wales given the differences in the language context and differences in how services

are delivered e.g. the age at which children receive a developmental assessment differ in Wales from England.

Welsh has been spoken in Wales throughout recorded history. According to the 2011 Census, 19% of people aged 3 and over were able to speak Welsh in Wales (Census, 2011). The number of speakers fluctuates across different parts of Wales and there is variation in the proportion of Welsh speakers in different local authorities (see Figure F1), and in individual communities. Children's levels of exposure to the Welsh language therefore varies greatly. Welsh is taught in all schools up until the end of the statutory sector (16 years). However, the nature of the linguistic provision offered between settings varies greatly. In immersion settings, pupils receive all their education through the medium of Welsh; in other settings, some pupils follow elements of the curriculum through the medium of Welsh. In English-medium schools, Welsh is taught as a subject. Early years settings also have varying degrees of Welsh-medium or bilingual provision. For children from Welsh-speaking homes, early years education provides an opportunity to build on their Welsh language skills. Children from English-speaking homes are introduced to Welsh for the first time within these nursery or school settings, leading to English-Welsh sequential bilingualism (Hickey, Lewis and Baker, 2014).

Figure F1: Percentage of adults (16+) that speak Welsh by Local Authority, National Survey For Wales (StatsWales, 2019)



*data not available

The proportion of children who present with speech and language difficulties is the same, regardless of whether they are monolingual or bilingual (Holm et al., 2013).

There are no data available regarding the number of Welsh-English bilingual children that are referred to Speech and Language Therapy services. However, there is evidence from other bilingual contexts that bilingual children are under-represented in speech and language therapy referrals before the age of five, and over-represented in referrals and special education services after five (Bedore & Peña, 2008; De Lamo White & Jin, 2011; Lazewnik et al., 2019; Salameh et al., 2002). In order to ensure efficient distribution of funding and resources, screening practices need to be appropriate for monolingual and bilingual children. Speech and language therapists have reported that the lack of appropriate tools makes assessment and screening of culturally and linguistically diverse students a challenge (Guiberson and Atkins, 2012). In addition, studies report that health and educational professionals who refer children for speech and language therapy do not always have sufficient knowledge and training about bilingual language development (De Lamo White and Jin, 2011). In one study, child health nurses reported postponing referrals for bilingual children due to a lack of confidence in their ability to interpret screening outcomes (Nayeb et al., 2015).

Screening tools in Wales need to accommodate three distinct populations:

- children who are exposed predominantly to English
- children who are exposed to both Welsh and English with a proportion of these children experiencing predominantly exposure to the Welsh language in the early years
- children who are exposed to multiple languages which include English and may include Welsh.

Within these populations, there is a wide range of levels of exposure with variations in whether children are simultaneously exposed to more than one language or whether the exposure is sequential, leading to variation in the degree of competency which might be expected in each language at varying ages. Further consideration needs to be given to those children growing up in homes which are English-speaking but where they attend Welsh-medium pre-school and childcare settings or live in communities which are largely Welsh speaking. Such ‘new speakers’ (Robert, 2009; Selleck, 2018; Williams & Cooper, 2021) have been shown to develop Welsh at different rates from those from Welsh-speaking homes. For example, Gathercole and Thomas (2009) showed that lexical and grammatical proficiency in Welsh is directly related to the input children have received in the language at home and in school, while all children, irrespective of home language, acquired English, the majority language, to high levels of proficiency. These diverse outcomes do not seem to disappear over time but can be observed over time in older children (Binks &

Thomas, 2019). If screening of children's language is not carried out in a way that is culturally and linguistically appropriate, there is the potential for long-term effects on children's outcomes and educational achievement (Wright Karem et al., 2019).

With the aim of supporting consistency in service delivery across Wales, care pathways were developed by groups of speech and language therapists with expertise in different clinical areas. These pathways were developed over many years with careful reference to the current evidence. The All Wales Speech Sound Pathway (2018) and the All Wales Early Language Pathway (2017) provide support for speech and language therapists making decisions regarding assessment and therapy for children with speech and language difficulties. There are also booklets to accompany the pathways that provide further guidance and signposting to current literature. These have been tested with the children's workforce and families and clear timelines for surveillance of development have been established into systems for public health and early education services, as detailed in the Healthy Child Wales Programme (Welsh Government, 2016). While the ELIM tool works well for England and the typical timings in most settings for monitoring of child development, these do not map onto those used in Wales.

For this reason, this work has been identified by the Welsh Government as essential to the identification of children with problems with language development which fit a pattern which suggests that they will be persistent and/or will have a long-term impact on a child's future development and life chances.