University of Groningen

Uniform screening for atypical language development in Dutch child health care

Bochane, Margot

DOI:
[10.33612/diss.171348669](https://doi.org/10.33612/diss.171348669)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Bochane, M. (2021). *Uniform screening for atypical language development in Dutch child health care*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.171348669>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Uniform screening for atypical language development in Dutch child health care

Margot Bochane

The research described in this thesis was supported by grants from the Dutch organization for Health Research and Development (ZonMw) Den Haag, Siméa Utrecht, the Hanze University of Applied Sciences Groningen, and the University of Groningen (RUG).

ISBN: 978-94-6416-494-7
Cover design: Senne Trip
Lay-out: Publiiss | www.publiiss.nl
Print: Ridderprint | www.ridderprint.nl

© Copyright 2020: Margot Bochane, Groningen, The Netherlands
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording, or otherwise, without the prior written permission of the author.

rijksuniversiteit
 groningen

Uniform screening for atypical language development in Dutch child health care

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 2 juni 2021 om 12:45 uur

door

Margot Irene Bochane

geboren op 27 juli 1973
te Leiderdorp

Promotores

Prof. dr. C.P. van der Schans

Prof dr. S.A. Reijneveld

Copromotor

Dr. M.R. Luinge

Beoordelingscommissie

Prof. dr. F.J.M. Feron

Prof. dr. P. Hendriks

Prof. dr. F.N.K. Wijnen

Paranimfen

Ellen de Wit

Kris Boyen

Contents

	Pages
Chapter 1	General introduction 9
Chapter 2	Atypical speech and language development: a consensus study on clinical signs in the Netherlands. <i>International Journal of Language & Communication Disorders</i> , 2017;52(1):10-20 25
Chapter 3	Same or different: the overlap between children with auditory processing disorders and children with other developmental disorders: a systeic review. <i>Ear and Hearing</i> , 2018;39(1):1-19 47
Chapter 4	The Dutch well-child language screening protocol at age 2 is valid to detect current and later language problems. <i>Acta Paediatrica</i> , 2021;110(2):556-562 89
Chapter 5	Identifying milestones in language development for young children ages 1-6 years. <i>Academic Pediatrics</i> , 2020;20(3):421-429 105
Chapter 6	Validation of the early language scale. <i>European Journal of Pediatrics</i> ,2021;180(1):63-71 129
Chapter 7	Towards a more family oriented and interdisciplinary screening of language development in young children. <i>Journal of Developmental and Physical Disabilities</i> , 2019;31:863-878 151
Chapter 8	General discussion 175
Appendix	Summary 195
	The Early Language Scale (ELS) 201
	Samenvatting 205
	Dankwoord 211
	About the author 217
	Research Institute SHARE: previous dissertations 219

1

General introduction

Talking in order to communicate with each other is what makes us humans unique. Our language is a basic skill that is essential for almost everything we do in life. Fortunately, everyone develops language quite automatically when growing up. Its development begins before a baby is born with listening to sounds from the womb. Following birth, children learn to understand the language of the environment and how to communicate with others. They start making sounds at approximately the age of four months, produce words around the first birthday, and commence putting two words together to generate sentences around the second birthday all in the context of communication and all seemingly effortless. However, in some children, language fails to develop which hampers their opportunities in life. The aim of this thesis was to contribute to the improvement of the identification of atypical language development in children. This chapter places the studies we performed in a broader context by providing background information on language development and screening for atypical language development in the preventive child health care setting. This chapter concludes with research questions.

Language

Speech and language are basic skills for all children and essential for participation in everyday life and, more specifically, a child's social and emotional development as well as educational success [1-5]. The terms speech and language are often used interchangeably as they are closely related, however, they have a distinct meaning [6]. Speech involves the production of speech sounds, a process that involves both motor (articulatory) and linguistic skills, whereas language refers to the comprehension and production of words, sentences, and ideas. Speech and languages are both facets of communication that include the broader context of non-verbal and verbal means of conveying information and emotions.

Children develop speech, language, and communication in interaction with other persons while meeting milestones in their development. These milestones for speech, for example, include 'babbling' and 'intelligible for the parent'. For language, they are 'says first words', 'starting to combine words', and 'begins to use pronouns'. For communication, a child 'answers simple questions'. Some of these milestones are well known, for example, combining two words at an age

of two years [7-9]. However, empirical evidence on other milestones for the age range of one to six years in which the basics for language develops is minimal and much less conclusive.

Language problems

Some children's language development does not follow typical developmental patterns. These children face problems in everyday life as language is essential for participating in our society. The so-called atypical language development can be divided into two categories: language delay and language disorder (with or without a known reason). A *language delay* is a language problem that describes children who lag behind compared to peers due to a lack of exposure to language. This is often the case for a second language or in families in which the quality of language exposure is insufficient. The prevalence of language delay is 5-19% depending on definitions used in the various studies [10].

A *language disorder* regards language development that lags behind and shows a deviant developmental pattern compared to peers. It may be associated with or result from any of the conditions hindering the development of a perceptual, motor, cognitive, or socio-emotional function. It is known that disorders such as developmental retardation, autism spectrum disorder, Down's syndrome, fragile X syndrome, traumatic brain injury, and deafness or hearing impairment increase the risk of speech and/or language disorders in children, and many children with such disorders will also have speech and language disorders [11-13]. However, in 7% of all upgrowing children, speech and language disorders occur for unknown reasons [14]. This specific type of language disorder is referred to as a *developmental language disorder (DLD)*. Children with DLD have difficulties in listening, understanding, and expressing what they want to say, and may also have difficulties with producing speech sounds [15,16]. DLD can manifest itself in one or more of the following domains: phonology (intelligibility of speech), semantics (vocabulary), syntax (grammar), pragmatics (language use), and social interaction. A DLD is a neurodevelopmental disorder that is a lifelong condition that affects how people understand and express language. It can have a significant impact on children's emotional functioning, academic success, and social relationships [1-6,17,18]. However, interventions for speech and language problems can improve a child's performance on various domains and therewith prevent or minimize the detrimental effects [19,20].

Early detection of atypical language is challenging because early signs of atypical development can manifest itself in various domains such as speech, language, and communication. In addition, these signs might refer to various disorders. Unfortunately, it is not clear which clinical signs refer to a DLD. Professionals working with children with DLD describe clinical signs such as ‘no first words’, ‘problems in turn taking’, ‘no adequate reaction to questions or commands’, and ‘ungrammatical sentences’. All clinical signs are related to the age of the child. First words emerge at approximately a child’s first birthday. If a child has no first words around that time, this can be a clinical sign. However, at the age of two years, not speaking first words is more severe and might warrant further investigation, what we refer to in this thesis as a “red flag”. Up to now, there is no consensus between professionals on which clinical signs can be considered as red flags for a language disorder at the various ages that language develops. Consensus on red flags for the various ages that language develops can contribute to early identification of language problems.

Overlap of language disorder with other neurodevelopmental disorders

Red flags in language development can indicate a developmental language disorder, however, it can also be a symptom of another neurodevelopmental disorder such as an intellectual disability, autism spectrum disorders, dyslexia, or unexplained listening difficulties [21,22]. For example, communication problems are a core diagnostic feature of autism spectrum disorders [23]; ‘no adequate reaction to questions or commands’ is present in children with unexplained listening difficulties [24,25]. Moreover, language development is considered a useful indicator of a child’s overall development and cognitive ability and is also related to school success [26-28]. The symptoms of the disorders overlap, making the boundaries between developmental language disorders and other neurodevelopmental disorders indistinguishable [29]. Insights regarding which symptoms overlap for various disorders is beneficial for understanding the nature of the disorders and in the process of early identification.

Screening in preventive child health care

Screening is defined by the World Health Organization as ‘the early identification and next treatment of unrecognized disease in an apparently healthy, asymptomatic population by means of tests, examinations or other procedures that can be applied rapidly and easily to a target population’. Early identification of developmental disorders, including developmental language disorder, is a core task of preventive child health services and is critical to the well-being of children and their families [30]. Children who are identified early can receive appropriate care, resulting in a better prognosis [19]. To ensure early identification of developmental disorders, including language disorders, preventive child health care organizations should have standardized approaches for early identification that include developmental surveillance and screening as an integral function and responsibility of all pediatric health care professionals [30-32]. In the Netherlands, 95% of the children are regularly seen in a well-child clinic. Therefore, preventive child health care provide a unique opportunity for early identification of developmental problems, including language development. However, until now, a screening instrument does not exist that can identify all children with atypical language development.

Screening principles

Screening should only be implemented if it fulfills a series of criteria, first described by Wilson & Jungner [33] and consolidated by Dobrow and colleagues [34]. In summary, the condition should be an important health problem with effective means of diagnosing it and effective methods for treating it, preventing its progression, reducing its effects or, ideally, cure it. There must also be a screening protocol that is effective, affordable, and acceptable. *Effectiveness* of a screening means that an instrument must measure what it intends to measure (reliability) and that it only detects persons in a population with the condition (validity). In the Netherlands, the quality of test instruments, including screening tests, can be assessed with the COTAN assessment system [35]. *Affordability* refers to the costs of a screening test, assessment, assessment time during well-child-visits, and for the implementation into the workflow, that includes the adaptation of software and training of professionals. Last, the screening must be *acceptable* for professionals and parents in that professionals must perform the screening, and parents need to respond or consent to the screening.

Language screening in preventive child health care

In the Netherlands, preventive child health care services provide language screening using a practice-based language screening protocol as part of their well-child care [32]. Dutch well-child care primarily focuses on prevention, more specifically the removal of causes of illness and health problems and the early detection of problems in an early and treatable stage. To do so, Dutch preventive child health care systematically follows children, vaccinates, advises, and provides extra care to vulnerable groups. This care is provided locally, in the municipality, and the neighbourhood during routine well-child visits for newborns up to the age of four years. At the age of four years, Dutch children begin primary education, therefore, preventive child health care for 4 to 18-year-old children is provided at the schools.

The Dutch language screening protocol for well-child care prescribes monitoring of language development from birth until the age of four years [32]. The protocol provides a standardized language screening at the age of two years and a reassessment at the age of two-and-a half-years [36]. The two-year language screening consists of tasks and questions to assess children's attainment of developmental language milestones. It is performed by trained well child professionals. The screening provides a distinct cut-off score with directions for referral in less than ten minutes. However, it is not known whether this protocol identifies the right children, and it is also not known to what degree the professionals adhere to the protocol in daily practice. Deviations from it might help or hinder the process of identification. Therefore, it is needed to assess the accuracy of the protocol and the effect of possible deviation in professionals' clinical decisions.

Between the ages of four to twelve, Dutch well-child care continues to monitor the development of children during two well-child visits at primary schools. At the age of five years, language development is still ongoing and must be checked, however, this is not standardized within the language screening protocol. During the primary school years, the development of children is also monitored by the primary school teachers; they expressed the need for a standardized screening instrument.

The screening for language disorder in well-child care and education could benefit from a brief, valid, and reliable instrument that covers the age range in which language develops as the current instruments are less than perfect. In addition to the Dutch protocol, the following instruments are available for language screening in the Netherlands: Speech and language norms for the primary healthcare (SNEL) [37], Lexilists [38,39], Dutch version of the MacArthur-Bates Communicative Development Inventory (N-CDI) [40], and Ages and Stages Questionnaire (ASQ) [41,42]. The CDI and ASQ were originally constructed in English and are available in several languages. Internationally, widely used instruments are the CDI [43], ASQ [42] and the Language Development Survey (LDS) [44]. The instruments differ from each other in terms of targeted construct, targeted age group, length of assessment, who performs the screening, and accuracy of the prediction (Table 1).

The ASQ is a parental questionnaire for identifying children with suspected developmental delays in communication, gross motor, fine motor, problem solving, and the personal-social domain. However, instruments that target only language development outperform the ASQ communication domain in identifying atypical language development [46, 51]. These instruments for language screening are the CDI [43] and LDS [44]. However, both instruments are too lengthy for routine use in well-child care as they require parents to check at least 100 and 310 items, respectively. Recently, CDI short forms were introduced comprising only 25 items, confirming the need for fewer comprehensive instruments that focus on language development. However, this CDI short form is only suitable for children up to 30 months of age [52] whereas the moment of identification of DLD typically exceeds this age [30,53-56]. Currently, there is no valid instrument available that covers the full early developmental period in which language is developed, i.e. up to six years, and that is also short to administer.

Table 1. Features of the available language screening instruments

Instrument	Language	Construct	Child age (years; months)	Length (minutes)	Assessor	Reliability/ validity
2-y language screening	Dutch	Language	2;0	?	Trained CHC professional	?
ASQ	Dutch / English*	communication, gross motor, fine motor, problem solving, personal-social	0;01 – 5;06	10-15	Parent	+ / + (a)
CDI	Dutch / English*	Language	0;08 - 2;06	10	Parent	+ / ± (b)
CDI short form	Dutch / English*	Language: vocabulary	1;04 – 2;06	2-3	Parent	+ / ± (c)
LDS	English*	Language: vocabulary and word combinations	1;06 – 2;11	10	Parent	+ / + (d)
Lexilist-NL	Dutch	Language (production)	1;03 – 2;03	2-3	Parent	+ / ± (e)
Lexilist-comprehension	Dutch	Language (comprehension)	1;03 – 2;01	2-3	Parent	+ / ± (f)
SNEL	Dutch	Speech, language	1;0 – 5;11	2-3	Parent	+ / + (g)
Van Wiechen	Dutch	Language	0 – 4;06	3	Trained CHC professional	+ / + (h)

* Also available in other languages

Accuracy of the prediction expressed in reliability/validity based on a) Kerstjens et al, 2009 [41]; b) Feldman et al, 2000 [45]; c) Kim et al, 2016 [46]; d) Rescorla & Alley, 2001 [47]; e) Egberink et al, 2007 [48]; f) Egberink et al, 2007 [49]; g) Luinge et al, 2007 [37]; h) Boere-Boonekamp et al, 2009 [50].

Parents and language screening in preventive child health care

Last, the effective screening process must be acceptable for the patients, in our case, the parents of children with a possible developmental language disorder. Having a child with developmental needs is an enormous challenge for family caregivers beginning with the identification of the condition. In many cases, parents were the first to realize that there was something wrong with their child's speech and language development with approximately a fifth of them expressing concerns regarding their child's language development during well-

child visits [57]. Additionally, it is recommended that preventive child health care elicit parents' concerns [30]. Parents consider themselves as experts on their child and want to be part of the decision making process regarding their child's development [58]. However, parents reported that well-child professionals tended to underestimate speech and language problems and did not adequately address parental concerns. Additional insight into the perceptions and desires of parents with regard to language screening in preventive child health care can be beneficial for identifying target areas for improving the alignment between the parents and professionals and thereby the routine language screening in preventive child health care.

Content of the thesis and research questions

The aim of the thesis was to contribute to the improvement of the identification of atypical language development in children. This has been translated into the following research questions:

1. What is the consensus of speech language professionals on red flags for developmental language disorder?
2. Which characteristics of auditory processing disorder overlap with characteristics of language disorder?
3. What is the validity of the current preventive child health care language screening protocol at age two years, and does the clinical decision of the professional affect its validity?
4. What set of clear and distinctive milestones empirically reflects language development in children from one to six years of age?
5. What is the validity of the newly developed Early Language Scale?
6. What are the perceptions, opinions, and desires of caregivers of children with and without atypical language development regarding language screening at the preventive child health care?

Figure 1 provides an overview of the research questions within the concepts in this thesis.

Figure 1. Overview of the research questions within this thesis.

References

1. Boyle CA, Decoufle P, Yeargin-Allsopp M. Prevalence and health impact of developmental disabilities in US children. *Pediatrics*. 1994;93(3):399-403.
2. Conti-Ramsden G, Mok PL, Pickles A, Durkin K. Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Res Dev Disabil*. 2013;34(11):4161-4169.
3. Snowling MJ, Duff FJ, Nash HM, Hulme C. Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *J Child Psychol Psychiatry*. 2016;57(12):1360-1369.
4. Whitehouse AJ, Lin, EA, Watt HJ, Bishop DVM. Qualitative aspects of developmental language impairment relate to language and literacy outcome in adulthood. *Int J Lang Commun Disord*. 2009;44(4):489-510.
5. Conti-Ramsden G, Durkin K, Mok PL, Toseeb U, Botting N. Health, employment and relationships: Correlates of personal wellbeing in young adults with and without a history of childhood language impairment. *Soc Sci Med*. 2016;160:20-28.
6. Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T, Catalise Consortium. CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one*. 2016;11(7), e0158753.
7. Diepeveen FB, Dusseldorp E, Bol GW, Oudesluys-Murphy AM, Verkerk PH. Failure to meet language milestones at two years of age is predictive of specific language impairment. *Acta Paediatr*. 2016;105(3):304-310.
8. Rudolph JM, Leonard LB. Early language milestones and specific language impairment. *J Early Intervention*. 2016;38(1), 41-58.
9. Sheldrick RC, Perrin EC. Evidence-based milestones for surveillance of cognitive, language, and motor development. *Acad pediatr*. 2013;13(6):577-586.
10. Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord*. 2000 Apr-Jun;35(2):165-188.
11. Bishop DV. Overlaps between autism and language impairment: phenomimicry or shared etiology? *Behav Genet*. 2010;40(5):618-629.
12. Pennington BF, Bishop DV. Relations among speech, language, and reading disorders. *Annu Rev Psychol*. 2009;60:283-306.
13. Reilly S, McKean C, Morgan A, Wake M. Identifying and managing common childhood language and speech impairments. *BMJ*. 2015 May 14;350:h2318.
14. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res*. 1997;40(6):1245-1260.
15. Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T, Catalise-2 Consortium. Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *J Child Psychol Psc*. 2017;58(10), 1068-1080.
16. World Health Organization. *International classification of diseases for mortality and morbidity statistics* (11th Revision). 2018. Retrieved from <https://icd.who.int/browse11/l-m/en>.

17. Clegg J, Law J, Rush R, Peters TJ, Roulstone S. The contribution of early language development to children's emotional and behavioural functioning at 6 years: an analysis of data from the Children in Focus sample from the ALSPAC birth cohort. *J Child Psychol Psc.* 2015;56(1):67-75.
18. McIntyre LL, Pelham WE, Kim MH, Dishion TJ, Shaw DS, Wilson MN. A Brief Measure of Language Skills at 3 Years of Age and Special Education Use in Middle Childhood. *J Pediatr.* 2017;181:189-194.
19. Broomfield J, Dodd B. Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial. *Int J Lang Comm Dis.* 2011;46(6):628-640.
20. Siu AL. Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force recommendation statement. *Pediatrics.* 2015;136(2), e474-e481.
21. Conti-Ramsden G, Durkin K. Language development and assessment in the preschool period. *Neuropsychol Rev.* 2012;22(4):384-401.
22. Miniscalco C, Nygren G, Hagberg B, Kadesjö B, Gillberg C. Neuropsychiatric and neurodevelopmental outcome of children at age 6 and 7 years who screened positive for language problems at 30 months. *Dev Med Child Neurol.* 2006;48(5):361-366.
23. Williams D, Botting N, Boucher J. Language in autism and specific language impairment: Where are the links? *Psychol Bull.* 2008;134(6):944.
24. Dawes P, Bishop D. Auditory processing disorder in relation to developmental disorders of language, communication and attention: a review and critique. *Int J Lang Comm Dis.* 2009;44(4):440-465.
25. Sharma M, Purdy S C, Kelly AS. Comorbidity of auditory processing, language, and reading disorders. *J Speech Lang Hear Res.* 2009; 52(3):706-722.
26. Beitchman JH, Wilson B, Brownlie EB, Walters H, Lancee W. Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *J Am Acad Child Psy.* 1996;35(6):804-814.
27. Beitchman JH, Wilson B, Brownlie E, Walters H, Inglis A, Lancee W. Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *J Am Acad Child Psy.* 1996;35(6):815-825.
28. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics.* 2006;117(2):e298-e319.
29. Bishop DVM, Rutter M. Neurodevelopmental disorders: conceptual issues. In M. 591 Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), 592 Rutter's Child and Adolescent Psychiatry. 2008(pp. 32-41). Oxford: Blackwell.
30. American Academy of Pediatrics, Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, & Medical Home Initiatives for Children With Special Needs Project Advisory Committee. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, 118(1), 405-420.
31. American Academy of Pediatrics. (2014). AAP publications reaffirmed or retired. *Pediatrics*, 134(5), e1520-e1520.

32. Lanting CI, de Wolff MS, Wiefferink K, Uilenburg N. JGZ-richtlijn taalontwikkeling [well-child care Guideline Language development] 2019.
33. Wilson JMG, Jungner G, World Health Organization. Principles and practice of screening for disease. 1968.
34. Dobrow MJ, Hagens V, Chafe R, Sullivan T, Rabeneck L. Consolidated principles for screening based on a systematic review and consensus process. *CMAJ* 2018 Apr 9;190(14):E422-E429.
35. Evers AVAM, Lucassen W, Meijer RR, Sijsma K. COTAN beoordelingssysteem voor de kwaliteit van tests [COTAN assessment system for the quality of tests]. *Amsterdam, Netherlands: Nederlands Instituut van Psychologen*. 2009.
36. Carmiggelt E, Uilenburg N, Romeijn J, Stam-van den Doel H, Pijpers F. Handreiking 'Uniforme signalering van taalachterstanden bij jonge kinderen'. *Utrecht: Nederlands Centrum Jeugdgezondheid*. 2013.
37. Luinge MR, Post W, Goorhuis-Brouwer SM. The language screening instrument SNEL. *Educ Child Psychol*. 2007;24:20–30.
38. Schlichting JEPT, Lutje Spelberg HC. Lexilijst Begrip [Lexilist Comprehension]. *Amsterdam: Pearson*. 2007.
39. Schlichting JEPT, Lutje Spelberg HC. Lexilijst Nederlands [Lexilist Netherlands Production]. *Amsterdam: Pearson*. 2002.
40. Zink I, Lejaegere M. N-CDI 3. Aanpassing en hernormering van de MacArthur CDI: Level III van Dale et al. *Leuven/Leusden: Acco*. 2007.
41. Kerstjens JM, Bos AF, ten Vergert EM, de Meer G, Butcher PR, Reijneveld SA. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. *Early Hum Dev*. 2009 Jul;85(7):443-447.
42. Squires J, Twombly E, Bricker D, Potter L. ASQ-3: Users Guide. *Baltimore: Paul H. Brookes Publishing Co*. 2009.
43. Fenson, L. (2002). MacArthur Communicative Development Inventories: User's guide and technical manual.
44. Rescorla L. The Language Development Survey: A screening tool for delayed language in toddlers. *J Speech Hear Disord*. 1989;54(4):587-599.
45. Feldman HM, Dollaghan CA, Campbell TF, Kurs-Lasky M, Janosky JE, Paradise JL. Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Dev*. 2000;71(2):310-322.
46. Kim SW, Kim JY, Lee SY, Jeon HR. The Comparison of M-B CDI-K Short Form and K-ASQ as Screening Test for Language Development. *Ann Rehabil Med*. 2016 Dec;40(6):1108-1113.
47. Rescorla L, Alley A. Validation of the language development survey (LDS). *J Speech Lang Hear Res*. 2001.
48. Egberink IJL, Leng WE de, Vermeulen CSM. (21-9-2020). COTAN beoordeling 2007, Lexilijst Nederlands [COTAN review 2007, Lexilist Netherlands]. Retrieved from www.cotandocumentatie.nl.
49. Egberink IJL, Leng WE de, Vermeulen CSM. (21-9-2020). COTAN beoordeling 2007, Lexilijst begrip [COTAN review 2007, Lexilist comprehension]. Retrieved from www.cotandocumentatie.nl.

50. Boere-Boonekamp MM, Dusseldorp E, Verkerk PH. Onderbouwing van de validiteit van het ontwikkelingsonderzoek bij kinderen van 0 tot en met 4 jaar: het van Wiechenonderzoek. [Validity of the Van Wiechen developmental assessment for children aged 0 to 4 years] Leiden: TNO. 2009.
51. Berkman ND, Wallace I, Watson L, Coyne-Beasley T, Cullen K, Wood C, Lohr KN. Screening for speech and language delays and disorders in children age 5 years or younger. A Systematic Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 120. AHRQ Publication No. 13-05197-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
52. Mayor J, Mani N. A short version of the MacArthur–Bates Communicative Development Inventories with high validity. *Behav res methods*. 2019;51(5):2248-2255.
53. Armstrong R, Scott JG, Whitehouse AJ, Copland DA, McMahon KL, Arnott W. Late talkers and later language outcomes: Predicting the different language trajectories. *Int J Speech-Lang Pa*. 2017;19(3):237-250.
54. Law J, Rush R, Anandan C, Cox M, Wood R. Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*. 2012 Jul;130(1):e132-7.
55. McKean C, Wraith D, Eadie P, Cook F, Mensah F, Reilly S. Subgroups in language trajectories from 4 to 11 years: the nature and predictors of stable, improving and decreasing language trajectory groups. *J Child Psychol Psyc*. 2017;58(10):1081-1091.
56. Zambrana IM, Pons F, Eadie P, Ystrom E. Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *Int J Lang Comm Dis*. 2014;49(3):304-316.
57. Rannard A, Lyons C, Glenn S. Parent concerns and professional responses: the case of specific language impairment. *Br J Gen Pract*. 2005 Sep;55(518):710-714.
58. Marshall J, Goldbart J, Phillips J. Parents' and speech and language therapists' explanatory models of language development, language delay and intervention. *Int J Lang Comm Dis*. 2007;42(5):533-555.

2

Atypical speech and language development: a consensus study on clinical signs in the Netherlands

Published as: Visser-Bochane MI, Gerrits E, van der Schans CP, Krijnen WP, Reijneveld SA, Luinge MR. Atypical speech and language development: a consensus study on clinical signs in the Netherlands. *International Journal of Language & Communication Disorders*, 2017;52(1):10-20

Abstract

Background: Atypical speech and language development is one of the most common developmental difficulties in young children. However, which clinical signs characterize atypical speech language development at what age is not clear.

Aim: To achieve a national and valid consensus on clinical signs and red flags (i.e. most urgent clinical signs) for atypical speech language development in children from one to six years of age.

Methods & Procedures: We conducted a two-round Delphi study in the Netherlands with a national expert panel (n=24) of speech and language therapists. The panel members responded to web based questionnaires addressing clinical signs. Consensus was defined as $\geq 70\%$ of the experts agreeing on an issue.

Outcomes & Results: The first round resulted in a list of 161 characteristics of atypical speech and language development. The second round led to agreement on 124 clinical signs and 34 red flags.

Conclusions & Implications: Dutch national consensus concerns 17 to 23 clinical signs per age year for the description of an atypical speech language development in young children and three to 10 characteristics per age year being red flags for atypical speech language development. This consensus contributes to early identification and diagnosis of children with atypical speech language development, awareness, and research.

Introduction

Atypical speech and language development is one of the most common developmental problems in early childhood. We use the term speech-language to encompass problems in any component of language, i.e. phonology, semantics, morphology, syntax or pragmatics. The children we address in our study can have speech-sound disorders, isolated phonological problems and also broader language problems. The estimated prevalence of children with atypical speech-language development varies from 2% to 19% [1,2]. Speech-language problems have a significant impact on children's emotional functioning, academic success and quality of life [3]. However, which clinical signs characterize atypical speech-language development at what age is not clear [4]. This may result in under-detection and late referral of children at risk for speech-language disorder.

Identification of children at risk is one of the primary tasks of Preventive Child Healthcare (PCH), which offers routine healthcare services to the population as a whole. About 90-100% of all children in The Netherlands visit a preventive child health centre [5]. Dutch preventive child health professionals systematically monitor the development, including speech-language development, of children aged 0-4 years using the 'Van Wiechenonderzoek' [6]. This is a checklist with milestones that is filled out by the PCH professional on the basis of observation or parental report. Public health aims at early identification of all developmental problems, but unfortunately, early identification misses some children at risk, both in The Netherlands [7] and internationally [8].

Early identification is challenging in a complex domain as language acquisition, since children vary quite a bit in the age at which they reach certain language milestones [4,9]. In addition, speech and language are two entities with different aetiologies that can occur together, but they can also occur alone [10]. Language profiles of children with an atypical language development change over the years, and as a consequence clinical signs of atypical development may vary by age group and also according to underlying linguistic deficits. Moreover, children can show spontaneous recovery [1].

There is an urgent need for diagnostic criteria for children with speech-language problems [11]. A set of typical clinical signs of early speech-language problems might be useful in the identification of these children. The most urgent clinical signs, which we will call 'red flags', might be useful for clinical practice as

well as for awareness of speech-language disorders by professionals working in healthcare and education, and by parents. Red flags are symptoms that require immediate referral for further evaluation. In this study consensus on clinical signs and red flags of an atypical speech-language development will be described per age year.

Consensus on clinical signs and red flags contributes to awareness, and early identification and referral of children with early speech-language problems, regardless of the cause of the atypical development. Our aim, therefore, is to achieve a national, and valid consensus on clinical signs for atypical speech-language development in children from 1 to 6 years of age.

Methods

A two-round Delphi study was utilized to achieve consensus on clinical signs and red flags for atypical speech-language development. A Delphi survey is a group facilitation technique, which is an iterative multistage process, designed to transform individual opinions into group consensus. Consensus via a Delphi study design ensures that the contribution of each participant is equally recognized, regardless of background or years of experience. The facilitator was the first author (MVB) in cooperation with all other authors. A flow diagram of the Delphi process of this study is outlined in Figure 1 and described below.

Figure 1. Flow diagram of the consecutive steps in the Delphi study

Participants and recruitment

The Delphi panel consisted of Dutch professionals comprising primarily speech and language therapists (SLTs) and clinical linguists in the field of typical and atypical child language development. The panel members have various work backgrounds (audiology centres, preventive care, private practices, schools for special education, and research groups with a focus on child language) and are located throughout The Netherlands (Table 1). The panel members have all

assisted children with speech-language problems or have conducted research on typical or atypical language development. We used purposive snowball sampling beginning with the participants from the ‘Dutch Developmental Language Disorder Network’ initiated by Fisher, Wijnen & Fikkert in 2012. To ensure a broad knowledge base in the panel, we aimed at a panel that included professionals from all types of major work settings and thereby represented all service delivery models. Another criterion for panel composition was geographic representation across The Netherlands.

Panel members consented to participate and to anonymous publication of the results. A total of 42 experts were invited of whom 24 (57%) joined the panel. Of the other 18 invitees, 11 did not respond, and seven experts stated that, due to time constraints, they were unable to participate.

Table 1. Work setting and Location of the participating SLT/Linguists

Delphi Panel	24 (100%)
<i>Work setting*</i>	
audiological centres	3 (13%)
preventive sector	5 (21%)
private practices	7 (29%)
schools for special education	6 (25%)
research groups with child language as interest	7 (29%)
universities	2 (8%)
<i>Region of the Netherlands</i>	
North	9 (38%)
Middle	11 (46%)
South	4 (17%)

* Some panel members work at various settings, therefore, the total percentages of work setting is not 100%.

Delphi procedure and measures

We obtained data in two Delphi rounds via web-based questionnaires that were performed by SurveyMonkey. Participants received a personal link to the questionnaire via e-mail. After 1 week we sent a reminder to all participants who had not yet finished the questionnaire. Answers were analysed after

anonymization. Consensus was considered to be reached when at least 70% of the panel members agreed with a statement.

First round of Delphi procedure

The first round contained open-ended questions to explore opinions on characteristics of atypical speech-language development in children from 1 to 6 years of age. In the introduction of the questionnaire we stated that all questions were related to children with speech-language problems regardless of the cause of the atypical development. The term ‘speech-language problems’ in The Netherlands is known to encompass problems in any component of language, i.e. phonology, semantics, morphology, syntax or pragmatics.

The questions included: (1) Which characteristics describe children with difficulties in speech-language development? Describe characteristics for children of one year. (The same question was posed for children aged 2, 3, 4, 5, and 6 years) Clinical signs and red flags were first explored by inquiring about the characteristics of atypical development. In the following round, the characteristics were rated and labelled as no clinical sign, clinical sign or red flag.

Responses from the first round were collated and summarized. All exact duplicates for the response of the characteristics questions were omitted, and those closely corresponding were maintained. For example, ‘does not react’ and ‘does not react to sounds’ were construed as near duplicates, therefore they were maintained. Summarization was performed per child age group with exact duplicates in different age groups, if applicable. The list of characteristics obtained with this method was supplemented with 12 items from Dutch screening instruments: SNEL [12] and Van Wiechenonderzoek [6]. We added these items since they are utilized in daily practice to screen for speech-language problems and were not in the exact formulation evidenced in the response of our panel. This procedure resulted in a set of 173 characteristics divided over six age groups (1-7-year-olds). In addition, two independent researchers categorized the characteristics into the domains of speech-language development (phonology, semantics, morphology, syntax, pragmatics and interaction). When the two researchers had no consensus, a third researcher was consulted.

Second round of Delphi procedure

The second round aimed at consensus on the characteristics of atypical development that resulted from the first round. Participants were asked to categorize the items on a three-point scale with the options ‘no clinical sign’, ‘clinical sign’, or ‘red flag’ for atypical development. No clinical sign was defined as ‘this is not distinctive for problems in the speech-language development’; a clinical sign as ‘potential problem in speech-language development’; and a red flag as ‘definitely a problem in speech-language development’. Consensus for classification of clinical signs was defined as when at least 70% of the panel members selected a characteristic as a clinical sign or as a red flag. Moreover, when 70% of the panel members indicated a characteristic as a red flag, and thereby a definite problem in speech-language development, the characteristic was indicated as a red flag.

Results

First Delphi round

Panel members mentioned a number of characteristics of children with difficulties in speech-language development that ranged from 20 to 32 characteristics per age group. For example, for the 1-year-olds, ‘no babbling’ and ‘no joint attention’ were acknowledged. This procedure resulted in a total number of 161 characteristics. The characteristics encompassed the linguistic domains of phonology, semantics, morphology, syntax, and pragmatics and additionally included items regarding interaction. The 161 characteristics supplemented with 12 items from currently used screening instruments as described in our Method section, resulted in a total list of 173 items.

Second Delphi round: from characteristics to clinical signs

In this second round, we presented the 173 characteristics of atypical language development from the first round to the panel and asked them to categorize these characteristics as ‘no clinical sign’, ‘clinical sign’, or ‘red flag’ for atypical speech-language development (table 2). A total number of 158 characteristics were considered to be a clinical sign or a red flag. The remaining 15 characteristics were considered ‘no sign’ (see figures A1-A6 in appendix A).

Table 2. Number of characteristics rated by the panel as ‘no sign’, ‘clinical sign’ or ‘red flag’ for atypical speech-language development, per age group.

Years of age	Total number of characteristics	No sign	Clinical Sign	Red flag
1-2	27	3	21	3
2-3	36	11	21	4
3-4	28	1	22	5
4-5	29	0	23	6
5-6	26	0	20	6
6-7	27	0	17	10
Total	173	15	124	34

Table 3 presents all the characteristics that panel members consented on (>70%) as being a red flag for atypical development, which indicates a definite problem in speech and language development. The red flags for the first age group were ‘no babbling’, ‘no reaction’ and ‘no reaction to sounds’. These red flags are characteristics within the domains of phonology and interaction. In the following age group (2-3 years of age), a red flag emerged in the domain of semantics. In the later age groups, there were also red flags in the domains syntax and pragmatics.

Table 3. Characteristics chosen by more than 70% of the panel members as being a red flag.

Red Flags	Domain	Percentage of the panel
<i>1-2 years of age</i>		
- No babbling	Phonology	95%
- No reaction	Interaction	86%
- No reaction to sounds	Interaction	71%
<i>2-3 years of age</i>		
- No first words	Semantics	86%
- No interaction	Interaction	86%
- No intention to communicate	Interaction	76%
- Little or no reaction on spoken language	Pragmatics/Interaction	71%
<i>3-4 years of age</i>		
- No speech	Pragmatics	100%
- Does not put two words together	Syntax	85%
- Not intelligible for close relatives	Phonology	86%

Table 3. (Continued)

Red Flags	Domain	Percentage of the panel
- Does not understand simple commands	Semantics/Syntax	81%
- At most, two word sentences	Syntax	76%
<i>4-5 years of age</i>		
- At most, three word sentences	Syntax	91%
- No simple sentences	Syntax	91%
- Not intelligible for strangers	Phonology	86%
- Not 50-75% intelligible for parents	Phonology	86%
- Disturbed interaction	Interaction	76%
- Does not understand spoken language / cannot process spoken language quickly enough	Semantics/Syntax	71%
<i>5-6 years of age</i>		
- Poorly intelligible for teacher / grandparents	Phonology	86%
- No adequate reaction to questions or commands	Pragmatics/Interaction	86%
- Does not understand composite commands	Semantics/Syntax	81%
- At most, simple sentences	Syntax	81%
- Poor intelligibility	Phonology	71%
- Does not ask questions	Pragmatics	71%
<i>6-7 years of age</i>		
- Does not understand composite commands	Semantics/Syntax	86%
- Disturbed communication	Pragmatics / Interaction	86%
- Incomplete sentences	Syntax	81%
- No adequate reaction to questions or commands	Pragmatics/Interaction	81%
- Poor understanding of language	Pragmatics	76%
- Poor intelligibility	Phonology	76%
- Poor vocabulary	Semantics	71%
- Does not produce compound sentences	Syntax	71%
- Poor story telling	Pragmatics	71%
- Cluster reduction	Phonology	71%

Discussion and conclusions

Consensus on clinical signs was reached on 17-23 characteristics of atypical speech-language per age year for the description of an atypical speech-language development in young children. In addition, consensus was reached on 3-10 characteristics per age year as being a definite red flag for atypical speech-language development.

The clinical signs and red flags resulting from our study describe speech and language development within the linguistic domains of phonology, semantics, morphology, syntax, and pragmatics and the non-linguistic domain of interaction. Interaction is new in this context since speech and language development is often described in only the linguistic domains, for example, the early language milestones [13]. For a number of characteristics, it was debatable if a characteristic belonged to pragmatics or to the broader domain of interaction, e.g., ‘no intention to communicate’ or ‘difficulty with interaction’. However, in most cases, characteristics were definitely suggesting interaction, which can be described as an element of pragmatics or as a necessary environment for language learning and as a condition to develop language. We determined that characteristics within the domain of interaction were primarily mentioned in the first year. Interaction begins to develop during the first month of life [14] and can thereby provide the first signs of atypical development. We propose interaction to be a domain belonging with the linguistic domains. As interaction is required for language to develop, it can be a beneficial predictor for language development [15].

The term ‘speech-language problems’ adopted in this paper might be a bit ambiguous. We specifically chose not to use an existing label in our questionnaire because we did not want to direct the panel members. The consequence of using the term ‘speech-language problems’ is that the clinical signs and red flags resulting from our study refer to a heterogeneous group of children with a speech-language problem. This group includes speech-sound disorders, isolated phonological disorders, and language disorders of any known or unknown origin. All children with these disorders should, however, be identified early by primary healthcare or education.

The clinical signs and red flags, as presented in appendix A, are relatively non-specific. For instance, children showing ‘no interaction’ at age 2, could have a hearing loss, autism-spectrum disorder, intellectual disability or language problem. Therefore, children who meet one or more of these criteria need further evaluation to establish if they are in need of service and if there is a known aetiology. Furthermore, the psychometric properties of the clinical signs and red flags are not yet tested. This requires additional study.

All characteristics described in our study were labelled: ‘no clinical sign’, ‘clinical sign’ or ‘red flag’, according to the response in the second Delphi round. These labels describe a continuum from characteristics that are not specific for problems in speech language development (no clinical sign), via possible problems (clinical signs) to definite problems (red flags) in speech language development. The cut-off score of 70% of the panel members agreeing on an item was chosen in advance. We applied the cut-off on the continuum of characteristics as described in the Methods section resulting in consensus on ‘clinical signs’ and on ‘red flags’, but not on all ‘no clinical signs’. There were 15 characteristics labelled as ‘no clinical sign’, with only one characteristic classified by more than 70% of the panel as being no clinical sign. The other 14 characteristics did not achieve agreement for a label ‘clinical sign’ or ‘red flag’.

Although 12 items were taken from existing screening instruments, three were not rated as being clinically significant. This does not automatically imply that these do not have added value in the original screening instruments. In those instruments, they were phrased as abilities, e.g., for 1-year-olds: ‘says “mama” or “papa”’. In the current study, we rephrased them into signs, i.e., ‘does not say “mama” or “papa”’. An item that was not rated as being clinically significant is not automatically also ‘not an ability’ if phrased in reverse.

The strengths of this study include the range of panel members who represented all major work settings, from community and various clinical services to research settings, and regions of the Netherlands. The panel had experience based on the population they see in their clinical practice, and on research. The clinical population seen by our individual panel members may not reflect all children with speech-language disorders in need of services, because their intake is influenced by the policy of their service delivery model. By including all

major work settings, our outcomes describe a general consensus that covers the policies of various service delivery models and research. Therefore, our findings can be considered to be representative of the Dutch setting.

Relative weaknesses include that the study was limited to only SLTs and linguists. We decided to initiate our study with only SLTs and linguists because these are most involved in interventions regarding children with impaired speech-language development. We do recognize that assessment, special education and therapy are multidisciplinary in nature. Another possible limitation of the study is that the moderator of the Delphi process may not have been neutral to the outcomes. Last, we performed no literature review on the possible clinical signs and red flags. In the first Delphi round we explored the experts' opinion on clinical signs, which we supplemented with items from currently used screening instruments in the second round.

Future empirical research can focus on which clinical signs can discriminate speech-language disorders in children of different age groups, and which clinical signs are associated with spontaneous recovery. Also, the clinical signs and red flags can be used for future development of a language screening tool. Moreover, research should define inclusionary criteria for speech-language disorder that should be useful for application in daily practice and in research. Last, we hope that researchers in other countries who are performing similar studies can benefit from our study.

The clinical signs and red flags in our study were described for six age groups and were derived from consensus. Additional research is needed on the use and added value of these clinical signs and red flags in routine practice in various countries and settings.

Implications

We described a consensus on 124 clinical signs and 34 red flags of atypical speech-language development for children from 1-6 years. Consensus regarding clinical signs and red flags is an important step in developing diagnostic criteria for children with speech-language problems. It will also contribute to awareness not only of health care professionals but also of teachers and parents. In addition,

knowledge of the red flag list may contribute to the identification of children with an atypical speech-language development. This may lead to major advantages for children with developmental speech-language disorder.

Acknowledgements

This research was funded by ZonMw (Dossier Number 200330002). The authors thank their panel members for their response. Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. All authors read, edited, and contributed to the manuscript.

References

1. Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord*. 2000 Apr-Jun;35(2):165-188.
2. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res*. 1997;40(6):1245-1260.
3. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics*. 2006;117(2):e298-e319.
4. Conti-Ramsden G, Durkin K. Language development and assessment in the preschool period. *Neuropsychol Rev*. 2012;22(4):384-401.
5. Dam P. Kosteneffectiviteit van de jeugdgezondheidszorg [Cost effectiveness of youth health care]. 2012.
6. Laurent de Angulo M, Brouwers-de Jong E, Bijlsma-Schlosser J, Bulk-Bunschoten A, Pauwels J, Steinbruch-Linstra I. Ontwikkelingsonderzoek in de jeugdgezondheidszorg Het Van Wiechenonderzoek-De Baecke-Fassaert Motoriektest. Assen: Van Gorcum 2005.
7. Stephan M, Diender M, Uilenburg N, Wiefferink C. Verwijzing van kinderen met een taalachterstand naar een audiologisch centrum. *JGZ Tijdschrift voor jeugdgezondheidszorg* 2015;47(5):96-100.
8. Bishop DV, McDonald D. Identifying language impairment in children: combining language test scores with parental report. *Int J Lang Com Dis*. 2009;44(5):600-615.
9. Luinge MR, Post WJ, Wit HP, Goorhuis-Brouwer SM. The ordering of milestones in language development for children from 1 to 6 years of age. *J Speech Lang Hear Res*. 2006;49(5):923-940.
10. Shriberg LD, Tomblin JB, McSweeney JL. Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *J Speech Lang Hear Res*. 1999;42(6):1461-1481.
11. Reilly S, Bishop DV, Tomblin B. Terminological debate over language impairment in children: forward movement and sticking points. *Int J Lang Commun Disord*. 2014;49(4):452-462.
12. Luinge MR, Post WJ, Goorhuis-Brouwer, SM. The language screening instrument SNEL. *Educ Child Psychol*. 2007;24, 20-30.
13. Sheldrick RC, Perrin EC. Evidence-based milestones for surveillance of cognitive, language, and motor development. *Acad pediatr*. 2013;13(6):577-586.
14. De Schuymer L, De Groote I, Striano T, Stahl D, Roeyers H. Dyadic and triadic skills in preterm and full term infants: A longitudinal study in the first year. *Infant Behav Dev*. 2011;34(1):179-188.
15. Reilly S, Wake M, Bavin EL, Prior M, Williams J, Bretherton L, et al. Predicting language at 2 years of age: a prospective community study. *Pediatrics*. 2007;120(6):e1441-e1449.

Appendix A: Characteristics of atypical language development per age year

Figure A1. Characteristics of atypical speech-language development for children aged 1.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

*Characteristics included from screening instruments. All other characteristics were achieved from the panel members in round 1

Figure A2. Characteristics of atypical speech-language development for children aged 2.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

*Characteristics included from screening instruments. All other characteristics were achieved from the panel members in round 1

Figure A3. Characteristics of atypical speech-language development for children aged 3.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

*Characteristics included from screening instruments. All other characteristics were achieved from the panel members in round 1

Figure A4. Characteristics of atypical speech-language development for children aged 4.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

Figure A5. Characteristics of atypical speech-language development for children aged 5.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

Figure A6. Characteristics of atypical speech-language development for children aged 6.

Notes: Characteristics are plotted on the y-axis, and the percentage of respondents who indicated a certain characteristic as a red flag (dark grey), clinical sign (light grey), and no sign (grey) is plotted on the x-axis. The cut-off of 70% refers to a clinical sign, i.e. when 70% of the respondents indicated a characteristic as a red flag or clinical sign. Moreover, it refers to a red flag and thereby a definite problem in speech and language development, when 70% of the respondents indicated this characteristic to be a red flag.

3

Same or different:
the overlap between children with audi-
tory processing disorders and children
with other developmental disorders:
a systematic review

Published as: de Wit E, van Dijk P,
Hanekamp S, Visser-Bochane MI,
Steenbergen B, van der Schans CP,
Luinge MR. Same or different: the over-
lap between children with auditory pro-
cessing disorders and children with other
developmental disorders: a systematic
review. *Ear and Hearing*,
2018;39(1):1-19

Abstract

Objectives: Children diagnosed with auditory processing disorders (APD) experience difficulties in auditory functioning and with memory, attention, language, and reading tasks. However, it is not clear whether the behavioral characteristics of these children are distinctive from the behavioral characteristics of children diagnosed with a different developmental disorder, such as specific language impairment (SLI), dyslexia, attention-deficit hyperactivity disorder (ADHD), learning disorder (LD) or autism spectrum disorder. This study describes the performance of children diagnosed with APD, SLI, dyslexia, ADHD, and LD to different outcome measurements. The aim of this study was to determine (1) which characteristics of APD overlap with the characteristics of children with SLI, dyslexia, ADHD, LD, or autism spectrum disorder; and (2) if there are characteristics that distinguish children diagnosed with APD from children diagnosed with other developmental disorders.

Design: A systematic review. Six electronic databases (Pubmed, CINAHL, Eric, PsychINFO, Communication & Mass Media Complete, and EMBASE) were searched to find peer-reviewed studies from 1954 up to May 2015. The authors included studies reporting behaviors and/or performance of children with (suspected) APD and children diagnosed with a different developmental disorder (SLI, Dyslexia, ADHD, and LD). Two researchers identified and screened the studies independently. Methodological quality of the included studies was assessed with the American Speech-Language-Hearing Association's levels-of-evidence scheme.

Results: In total, 13 studies of which the methodological quality was moderate were included in this systematic review. In five studies, the performance of children diagnosed with APD was compared with the performance of children diagnosed with SLI; in two with children diagnosed with dyslexia, one with children diagnosed with ADHD, and in another one with children diagnosed with LD. Ten of the studies included children who met the criteria for more than one diagnosis. In four studies, there was a comparison made between the performances of children with comorbid disorders. There were no studies found in which the performance of children diagnosed with APD was compared with the performance of children diagnosed with autism spectrum disorder. Children

diagnosed with APD broadly share the same characteristics as children diagnosed with other developmental disorders with only minor differences between them. Differences were determined with the auditory and visual Duration Pattern Test, the Children's Auditory Processing Performance Scale questionnaire, and the subtests of the Listening in Spatialized Noise-Sentences test, in which noise is spatially separated from target sentences. However, these differences are not consistent between studies and are not found in comparison to all groups of children with other developmental disorders.

Conclusions: Children diagnosed with APD perform equally to children diagnosed with SLI, dyslexia, ADHD, and LD on tests of intelligence, memory or attention, and language tests. Only small differences between groups were found for sensory and perceptual functioning tasks (auditory and visual). In addition, children diagnosed with dyslexia performed poorer in reading tasks compared with children diagnosed with APD. The result is possibly confounded by poor quality of the research studies and the low quality of the used outcome measures. More research with higher scientific rigor is required to better understand the differences and similarities in children with various neurodevelopmental disorders.

Introduction

Children diagnosed with an auditory processing disorder (APD) have difficulty with listening. This is especially prominent in an unfavorable listening environment despite well-functioning peripheral hearing [e.g. 1,2]. Notwithstanding the attempts of special working groups to obtain clarification about the construct of APD [e.g., 1,3, 4], discussion continues among professionals about the diagnostic criteria for APD, the overlap of APD with other developmental disorders, and whether APD exists as a unique diagnostic entity [5-7].

Because of the lack of a clear definition and the use of multiple diagnostic criteria, different professionals approach children with listening complaints from different perspectives [8]. Different diagnostic criteria for APD are proposed in various position statements and by several researchers [1, 3,4,9-12]. The different sets of diagnostic criteria have in common that children with listening difficulties are classified as having APD based on their performance on one or more behavioral central auditory tests or checklists or questionnaires. They differ, however, in the types of tests on which they must demonstrate inadequate performance and on how abnormal the performance is actually considered to be (e.g., < 2 SD or < 3 SD below the mean). The lack of a clear definition of APD together with the variation in diagnostic criteria for APD results in a range of approximate prevalence rates from 0.5 to 1.0% to 7% of the population [13-15]. For instance, depending on which diagnostic criteria were used, Wilson and Arnott [12] identified 7.3% (diagnostic criteria by Bellis [9]) to 96% (diagnostic criteria by ASHA [5]) of the children in their study group with APD.

Children with difficulties in the processing and understanding of auditory stimuli and with normal pure-tone thresholds have been recognized from the mid-20th century. These difficulties are “characterized by poor perception of both speech and non-speech” [4, p.3]. Frequently reported symptoms are difficulty understanding speech in noisy environments; problems in locating the source of a signal; fail to respond correctly to verbal information; frequently asking for repetition of information; reduced attention to auditory information and easily distracted [1]. Since the 1970s, these difficulties are more commonly known in the field of speech-language pathologists and audiologists as APD [3,16-18]. During recent years, this group of children is also described as children with suspected APD (susAPD) or children with listening difficulties.

One major issue that has dominated the field for many years concerns the distinction between APD and other developmental disorders, such as a specific language impairment (SLI), dyslexia, learning disorder (LD), attention-deficit hyperactivity disorder (ADHD), and autism spectrum disorders (ASD). The various developmental disorders have in common that they are characterized by developmental delays which can cause impairment in personal, social, academic or occupational functioning [19]. Not all of the developmental disorders are well validated, and at the phenotypic level, the various developmental disorders overlap considerably with each other. Furthermore, research has shown that some disorders have in part the same genetic origin [20]. It is not clear whether there are different underlying mechanisms for the various disorders. Therefore, separation of the various developmental disorders from each other can be difficult for clinicians, but also for scientific research [20]. This is also the case for APD. Several researchers have reported that the characteristics of children diagnosed with APD seem to correspond to the behaviors and symptoms of children diagnosed with other developmental disorders [1,4,6,10,21-24]. For example, difficulties in comprehending and complying to verbal information are also commonly observed in children diagnosed with SLI. The concentration and attention complaints reported in children diagnosed with APD also correspond to the difficulties of children diagnosed with ADHD [10,22], and atypical processing of auditory information (e.g., difficulties listening in noise, hyperacusis, hypersensitivity to pitch), something that can be also difficult for children diagnosed with APD, is also an inherent component of ASD [10,25]. Likewise, it is contended that the diagnosis of a child may depend more on the referral route than on the symptoms of a child [24].

The overlapping symptoms of children diagnosed with APD and children with other disorders also contribute to the discussion among clinicians and scientists as to whether or not APD can be considered a distinct clinical disorder [5,21,22,24,26]. One study indicated that 94% percent of the children identified with APD also had a comorbid language impairment or reading impairment [27]. In another study, 30% of the children with susAPD also had problems with reading and writing, 90% had additional speech-language problems, and 10% had ADHD. Sixty percent of the children had two or more accompanying problems (measured with a teacher-based questionnaire [28]). Three possible explanations for the co-occurrence between APD and other developmental disorders are provided by Moore et al. [7]: (1) not all different disorders can be distinguished

from each other; the diagnosis of a child depends more on the reference route than by the symptoms; (2) SLI and dyslexia are caused by an APD; (3) a more general neurodevelopmental deficit is the cause of the various disorders.

Our recent systematic review [29] was aimed at establishing the characteristics of children with APD. The results showed that children with susAPD performed significantly poorer on tests of auditory processing when compared with typically developing (TD) children. However, significant differences between children with susAPD and TD children were not only found in the auditory domain. Children with susAPD also performed significantly poorer on tests of intelligence, memory, attention, and visual, language, and reading tests, as well as auditory brain measures. This again suggests that the characteristics of children diagnosed with APD overlap with the characteristics of children diagnosed with other developmental disorders. It is unclear whether a certain ensemble of symptoms exists that is solely attributable to difficulties with auditory processing. Such a distinctive symptom or group of symptoms could assist audiologists and speech-language pathologists in differentiating APD from other developmental disorders.

This present systematic review intends to describe the overlap between the characteristics of APD, SLI, dyslexia, ADHD, LD, and ASD. Our goal is to contribute to the discussion on whether APD is a distinct disorder that is separate from the other conditions that are mentioned. The central questions of this systematic review are the following: (1) Which characteristics of APD overlap with characteristics of other developmental disorders? and (2) Are there characteristics that distinguish children diagnosed with APD from children diagnosed with other developmental disorders?

Methods

Studies published in peer-reviewed journals from 1954 up to May 2015 were considered for inclusion in this systematic review. The four stages of the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) 2009 flow diagram [30;prisma-statement.org] of (1) identification, (2) screening, (3) eligibility, and (4) inclusion, were used for systematically going through the different phases of a systematic review. Stages 1 and 2, in which author 1, 4 and 7

of the current review were involved, are equal to those stages as described in the study of de Wit et al. [29].

After each stage, there was a consensus meeting held with the reviewers (author 1, 3, 4, and 7) who were involved in the process of assessing the studies that were found. A study was included or excluded in the subsequent stage when complete consensus was reached between reviewers.

Stage 1: Identification

To identify appropriate studies for this systematic review, the following databases were searched: Pubmed, CINAHL, Eric, PsychINFO, Communication & Mass Media Complete, and EMBASE. As described in de Wit et al. [29], the original search consisted of two separate search queries, which have been combined in this review to one search. Search strings used in the different databases are depicted in Table 1. The used search strategy can be found in Supplement 1, which demonstrates the review protocol of this systematic review. RefWorks was used to manage, store, and share the found studies and for removing duplicates.

Stage 2: Screening

The screening stage consisted of two steps, namely, (1) screening of titles and (2) screening of abstracts. For both steps in this stage, studies were screened against the inclusion and exclusion criteria by two researchers (author 1 and 4 or 1 and 7).

To be included in the review, the title and abstract of the study must meet the following criteria: (1) published in English and in a peer-reviewed journal; (2) addressed factors in title about auditory processing in combination with deficit(s), impairment(s), problem(s), difficulties, or disorder(s); (3) addressed elements in the abstract regarding the characteristics of susAPD or children at risk for APD in the presence of normal hearing. The terms/synonyms for APD that were also considered for inclusion can be found in Supplement 1. In addition to these three inclusion criteria, studies must contain data regarding participants under the age of 18 years. The age limit of 18 years was chosen, in order to search somewhat broader than the age at which the final step in structural maturation of the auditory cortex normally occurs (between the age of 6 and 12 years; [31]). Studies that included neonates or participants with neuropathy, cochlear implants, Down syndrome or another syndrome, peripheral hearing loss, chronic otitis media, or brain damage were excluded from this review.

Table 1. Search strings and databases used.

Databases	Search string
Pubmed	(“Auditory Diseases, Central”[Mesh] OR auditory process- ing[tiab] OR auditory perceptual[tiab]) AND (child[tiab] OR adolescent*[tiab])
PsycInfo, Eric, CINAHL, Communication & Mass media complete EMBASE (until March 15, 2012)	(TI “auditory processing” OR TI “auditory perception” OR TI “auditory perceptual”) OR AB “auditory perception” OR AB “auditory perceptual”) AND (AB child OR AB adolescent) “auditory processing”, “auditory perception”, “auditory perceptual” child:ab OR children:ab OR adolescent:ab OR adolescents:ab.

Stage 3: Eligibility

At the eligibility stage, it was decided which of the selected studies specifically matched the research question of the present study. Therefore, it was necessary that the focus of the study be on the comparison of the behaviors or performance of children with susAPD with children diagnosed with a different developmental disorder such as SLI, Dyslexia, ADHD, ASD or LD. This stage consisted of two sequential steps: (1) full text articles assessed for eligibility and (2) full text articles assessed for methodological quality.

The full text of each identified study was individually evaluated by one of the three reviewers (author 1, 4, and 7). The reviewer assessed whether the study was appropriate for the topic of the current systematic review.

Subsequently, the methodological quality of each included study was independently reviewed by two reviewers (author 1 and 3 or 1 and 7) with the American Speech-Language-Hearing Association’s levels-of-evidence scheme [32]. A description of this quality assessment tool can be found in Table 2. The quality indicator ‘intention-to-treat’ was removed from the scheme. Each quality indicator that complied with the highest quality level was assigned one point (see Table 2). The maximum achievable quality score for a study was seven points.

To classify the evaluated studies, the final quality score of the individual studies was used (adapted to the quality assessment tool developed by Gyorkos et al. [33]. Based on their final quality score, studies were classified as strong (5-7 points), moderate (2-4 points), or weak (0-1 points). For inclusion in this

systematic review, a study must have been appraised as a moderate or strong study. Studies rated as weak were excluded.

Table 2. Quality Indicators in the ASHA Levels-of-Evidence Scheme [32].

Indicator	Quality Marker
Study design	Controlled trial Cohort study Single-subject design or case control study Cross-sectional study or Case series Case study
Blinding	Yes = Assessors blinded No = Assessors not blinded or not stated
Sampling	Yes = Random sample adequately described No = Random sample inadequately described No = Convenience sample adequately described No = Convenience sample inadequately described or hand-picked sample or not stated
Group / participant comparability	Yes = Groups comparable at baseline on important factors (between-subject design) or subject(s) adequately described (within subject design) No = Groups/subjects not comparable at baseline or comparability not reported or subject(s) not adequately described
Outcomes	Yes = At least one primary outcome measure is valid and reliable Reasonable = Validity is unknown but appears reasonable; measure is reliable No = Invalid and/or unreliable
Significance	Yes = P value reported or calculable No = P value neither reported nor calculable
Precision	Yes = Effect size and confidence interval reported or calculable No = Effect size or confidence interval, but not both, reported or calculable No = Neither effect size or confidence interval reported or calculable

Note. Boldface indicates highest level of quality marker.

Stage 4: Inclusion

In stage 4, relevant data from the included studies with moderate and strong qualities were extracted and analyzed. The following data were extracted from the studies: (1) study characteristics; (2) subject characteristics; (3) measurement instruments; and (4) study results. Pooling the results of all of the studies was

not possible because of the substantial variation in outcome measures and the inconsistent manner of presenting the results between studies. Therefore, the results of the individual studies were summarized in a table (Table 5 - 8). The overlap and differences between groups on the outcome of various measurements used in the included studies are summarized in a nonsymmetric Venn diagram (Fig. 2).

Results

The result of stages 1 to 4 of the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2009 flow diagram [30] is illustrated in Figure 1. Database searching yielded 3317 unique studies (stage 1) of which the titles were screened independently by two reviewers. This yielded 548 studies of which the abstract was assessed (stage 2). Subsequently, three reviewers screened the eligibility of 194 full-text studies (stage 3). One hundred and eighty studies were excluded in this stage. Of these 180 studies, 53 studies matched the research question of the previously published systematic review, and 24 studies did not satisfy the research question of both reviews. In 70 studies, the auditory processing skills of children diagnosed with disorders other than APD were investigated. However, these studies were excluded because these studies are beyond the scope of our review and did not fit the research question and inclusion criteria of this review. Finally, 14 of the full-text studies were rated as appropriate for inclusion in the current systematic review (stage 4).

Methodological quality of the studies

Based on the total quality score, one study [34] was classified as a study with weak quality (see Table 3 for details). Table 4 shows the methodological quality assessment of the included studies. As indicated in the data in Table 4, none of the studies were classified as being methodologically strong (≥ 5 points). In all of the included studies, significance and precision was either reported or could be calculated from the data; however, in none of the studies were the subjects randomly enrolled or subjects or assessors blinded.

Figure 1. Process for identification of included studies. APD, auditory processing disorder. From PLoS Medicine. 2009;6:e1000097.

Table 3. Methodological quality of the excluded studies.

Study	Study Design	Assessor Blinded	Random Sample	Groups / Participants Comparable	Valid Primary Outcome Measure (s)*	Significance Reported or Calculable	Precision Reported or Calculable	Total Quality Score
Simoes & Schochat, 2010	Cross-sectional study	No	No	No	Reasonable	Yes	No	1/7

Critical appraisal ratings of excluded studies evaluated with the ASHA’s levels-of-evidence (ASHA’s LOE) scheme [32]. Based on the quality score studies awarded with one or no points were classified as weak, and were excluded from the review.

*At the criterion “Valid primary outcome measures” three answer options were possible, namely: Yes, Reasonable, and No. All other criterion had two possible outcomes, Yes or No. Boldface indicates highest level of quality in each category.

Table 4. Methodological quality of the included studies.

Study	Study Design	Assessor Blinded	Random Sample	Groups / Participants Comparable	Valid Primary Outcome Measure(s) *	Significance Reported or Calculable	Precision or Reported Calculable	Total Quality Score
Bellis et al. 2011	Cross-sectional study	No	No	No	Reasonable	Yes	Yes	2/7
Dawes et al. 2009	Cross-sectional study	No	No	No	Reasonable	Yes	Yes	2/7
Iliadou et al. 2009	Cross-sectional study	No	No	No	Reasonable	Yes	Yes	2/7
Ro-cha-Muniz et al. 2014	Cross-sectional study	No	No	No	Reasonable	Yes	Yes	2/7
Cameron & Dillon 2008	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Dawes & Bishop 2010	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Ferguson et al. 2011	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Ferguson & Moore 2014	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Riccio et al. 1994	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Riccio et al. 1996	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Ro-cha-Muniz et al. 2012	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Walker et al. 2011	Cross-sectional study	No	No	No	Yes	Yes	Yes	3/7
Miller & Wagstaff 2011	Cross-sectional study	No	No	Yes	Yes	Yes	Yes	4/7

Critical appraisal ratings of included studies evaluated with the ASHA's levels-of-evidence (ASHA's LOE) scheme [32]. Based on the quality score studies awarded with two to four points were classified as moderate, and studies awarded with five to seven points were classified as strong. Studies are arranged from low to high quality score.

*At the criterion "Valid primary outcome measures" three answer options were possible, namely: Yes, Reasonable, and No. All other criterion had two possible outcomes, Yes or No. Boldface indicates highest level of quality in each category.

General study characteristics

Thirteen studies are included in this systematic review. In one of the 13 studies, a comparison was made between children diagnosed with LD and children with APD + LD [35] and, hence, not children diagnosed with only APD. In the other 12 studies, the performance of children diagnosed with APD was compared with the performance of children diagnosed with a different developmental disorder. In five studies, a comparison was made with children diagnosed with SLI [36-40], two with children diagnosed with dyslexia [41,42], one with children diagnosed with ADHD [43], and another with children diagnosed with learning or attention disorders (LD group; [44]). In three studies, a comparison was made between children diagnosed with APD and a group of children with comorbid disorders: two studies compared children diagnosed with APD with a group of children diagnosed with APD + ADHD [45,46], and one compared children diagnosed with APD with a group of children diagnosed with APD + dyslexia [47]. No studies were found in which a comparison was made between the performance of children diagnosed with APD and children diagnosed with a disorder in the autistic spectrum.

In this systematic review, the diagnosis which primarily has been studied and was the focus of the original study was used in the comparison of the groups. It has been found that it is not always possible to fit children with developmental issues into stiff diagnosis-categories. Ten of the studies included children who met the criteria for more than one diagnosis. In the studies of Ferguson et al. [36] and Ferguson and Moore [37], more than 25% of the included children had an additional diagnosis of ADHD, ASD, or dyslexia. In the studies of Dawes et al. [42] and Dawes and Bishop [41], these percentages were even higher: 52% of the children with a diagnosis of APD also met the diagnostic criteria of SLI, dyslexia, or both, and about 20% of the children in the dyslexia group also fits the diagnosis of SLI. In addition, there was also a high percentage of abnormal cases of hyperactivity and inattention in both groups. Riccio et al. [45,46] disclosed information about comorbid psychopathology (e.g., ADHD and conduct disorder). The authors also suspect coexisting language disorders because of low scores of the subjects on language tests and measures of cognitive ability. Miller and Wagstaff [38] reported that the parents of 16 children indicated that there was also a diagnosis of attention deficit disorders (ADD) or ADHD. There

was no co-morbidity with ADHD or dysorthography in the subjects included in the study of Iliadou et al. [47]; however, three children with low IQ and three children with borderline IQ were included. Cameron and Dillon [44] and Walker et al. [35] included children with LDs. These studies, included children with different kinds of problems, like ADD, ADHD, SLI, dyslexia, and working memory deficits. In three studies, no information about additional diagnoses was given [39,40,43].

Some participants took part in more than one study. In the study of Ferguson and Moore [37] and Dawes and Bishop [41], participants were recruited from, respectively, the study of Ferguson et al. [36] and Dawes et al. [42]. For the study of Rocha-Muniz et al. [40] and Riccio et al. [45], it is not clear whether the authors used partially the same participants as in the preceding studies [39,46]. The number of subjects ranged from seven to 41 in the APD group (mean = 20.8 subjects) and from ten to 29 in the comparison group (the group of children included with a different developmental disorder: mean SLI = 23.8 subjects, mean dyslexia = 19 subjects, ADHD = 10 subjects, LD = 11 subjects, mean comorbid disorder = 13.5 subjects). The ages of the subjects in the APD group ranged from 6 years to 15 years and 11 months; in the SLI and dyslexia groups, from 6 years to 13 years; in the LD group, from 7 years and 2 months to 11 years and 8 months; in the group with comorbid disorders, from 8 years to 15 years and 11 months; and in the ADHD group, the average age was 13 years and 1 month.

Diagnostic criteria for inclusion in the APD group varied among studies. Seven studies used the diagnostic criteria of a below normal performance on at least two behavioral diagnostic tests of auditory processing [35,38,39,40,43,46,47], and three studies used a below normal performance on at least one behavioral diagnostic test of auditory processing [41,42,45]. The remaining three studies used typically APD symptoms reported by parents or a referral for an auditory processing assessment for inclusion in the APD group [36,37,44].

Five studies, which included an SLI group, used the diagnostic criteria of Leonard [48], referring to children with significant speech or language difficulties that could not be explained by factors such as hearing loss, autism, learning or physical disability, or bilingualism.

Participants in the studies that included children with ADHD met the diagnostic criteria from the Diagnostic and Statistical Manual of Mental Disorders [49,50]. In the two studies by Dawes et al. [42] and Dawes and Bishop [41], the diagnosis of dyslexia was established by an educational psychologist in the event of a below normal performance (standard score <85) on a reading or spelling tests in the presence of an average intelligence (nonverbal IQ ≥ 80). Iliadou et al. [47] used the Diagnostic and Statistical Manual of Mental Disorders-IV criteria for the diagnosis of dyslexia.

Cameron and Dillon [44] and Walker et al. [35] used a very diverse patient population in their comparison group. The LD group in the study of Cameron and Dillon [44] comprised children with various types of difficulties, such as working memory deficit, ADHD, and dyslexia. Also, the children in the LD group in the study of Walker et al. [35] had various types of learning difficulties, including ADD, language and reading impairment, and more general learning impairments. All of the children in both studies had an overall intellectual performance within normal limits (Cameron & Dillon: a scaled score of eight or above, a standard score of 90 or above, or the equivalent percentile rank of 25 or above; Walker et al.: full-scale IQ scores of 85 or greater).

Children diagnosed with APD versus children diagnosed with other developmental disorders

Table 5 to 8 show, per included study, a summary of the participant characteristics, measurement instruments used, and differences between children diagnosed with APD and children diagnosed with a different developmental disorder. It is apparent from this table that there were only marginal differences between these groups of children. Children diagnosed with SLI can be separated from children diagnosed with APD on the speech measures (weaker performance on the Children's Communication Checklist – second edition [CCC-2]; subscale speech; speech-evoked auditory brain stem response; speech-in-noise test with monosyllabic words; and the Dichotic Digit test), and children diagnosed with dyslexia can be separated from children diagnosed with APD on the reading measures. On the contrary, children diagnosed with APD can be separated from children diagnosed with SLI and dyslexia on the listening questionnaire (Children's Auditory Processing Performance Scale [CHAPPS] total score and subscales

noise, multiple inputs, and attention). Opposed to children diagnosed with ADHD, children diagnosed with APD perform poorer on auditory and visual temporal ordering measures. Finally, children diagnosed with APD may be separated from children diagnosed with a variety of LDs by their problems with auditory stream segregation.

Five studies included different types of assessment of IQ, attention, and memory abilities in their study [36,38,41,45,46]. No significant differences were found between clinical groups on one of these tests. Five studies examined the language and reading skills of children diagnosed with APD compared with children diagnosed with SLI, dyslexia, or ADHD [35,36,38,41,46]. Furthermore, all 13 studies assessed the auditory processing capabilities of the included children with auditory behavioral tests, speech-evoked auditory brain stem response, or parental/ teacher questionnaires.

Children diagnosed with APD versus children diagnosed with SLI

Table 5 shows the similarities and differences on various measurement instruments between children diagnosed with APD and children diagnosed with SLI. No differences were found on IQ, working memory, language, and reading tests between children diagnosed with APD and children diagnosed with SLI. Most studies also found no difference between these children on behavioral auditory processing tests [36-38]. Both groups experienced difficulties with speech perception. Two studies found comparably poorer performance for children with SLI in the perception of speech [39,40]. The only study included in this systematic review that used electrophysiological measurements found that the children with SLI had a more impaired brainstem encoding for speech signals than the children with APD had [39]. Also, significance variances were found with the CHAPPS [51]. Parents of children diagnosed with APD rated their children as having poorer listening skills in noisy places, poorer listening skills in multiple input, and poorer attention skills [36].

Table 5. (Continued)

Study	Participants/Groups		Measurement Instrument	Significance Difference Between Clinical Groups	Difference With TD Group	Summary of Significant Differences APD vs. SLI
	APD, n; Gender; Age Range; Mean Age (M)	SLI, n; Gender; Age Range; Mean Age (M)				
Rocha-Muniz et al. 2012	n = 18; 14 M, 4 F; 6–12; M = 9.17	n = 21; 16 M, 5 F; 6–12; M = 8.0	Speech-evoked ABR (speech syllable /da/) Timing measures: Defined as the latencies of the 7 prominent response peaks (V, A, C, D, E, F, and O)			
			V	*	O	
			A	*	U	
			C	*	O	
			D	*	E	
			E	†	O	Increase latency for the SLI group.
			F	†	O	Increase latency for the SLI group.
			O	*	O	
	No information available about additional diagnoses.		Spectral encoding measures			
			Pitch			
			F ₀	*	E	
			F ₁	*	E	
			Harmonics			
			HF	†	O	Smaller amplitude for the SLI group in encoding the higher harmonic HF.
Miller & Wagstaff 2011	n = 35; 25 M, 10 F; 8.5–12.7; M = 10.3	n = 29; 15 M, 14 F; 8.5–12.7; M = 10.0	Language tests			
			Formulating Sentences subtest of the CELF-4	*	E	
			Concepts and Following Directions subtest of the CELF-4	*	E	
			PPVT-3	*	E	
			Expressive Vocabulary Test or the Picture Vocabulary subtest of the W-J III	*	E	
			Auditory tests			
			FPT	*	NR	
			DPT	*	NR	
			DDT right	*	NR	
			DDT left	*	NR	
			SSW left competing	*	NR	
			SSW total errors	*	NR	
	ADHD was not an exclusionary criterion. Parents of 16 participants reported that their child had been diagnosed with ADD or ADHD.		Nonword Repetition test	*	NR	
			NV-IQ			
			Symbolic Memory subtest of the UNIT	*	E	
			Cube Design: subtest of the UNIT	*	E	
			Reading fluency			
			GORT-4	*	U	
			Motor speed	*	E	
			Verbal working memory			
			CLPT	*	NR	
			Visual-spatial working memory			
			SWMT	*	NR	
			Attention			
			CADS-P	*	E	

Studies are arranged in order of the total quality score.

*No significant difference between clinical groups.

†Significant difference between clinical groups.

ADHD, attention deficit hyperactivity disorder; ADD, attention deficit disorder; APD, auditory processing disorder; ASD, Autism Spectrum Disorder; F, female; SLI, specific language impairment; TD, typically developing children; M, male; NR, not reported; A, APD group underperformed significantly compared with TD children; E, both groups displayed equal performance in comparison with TD children or norm values; O, other clinical group (SLI, dyslexia, ADHD, or LD) underperformed significantly compared with TD children; U, both groups underperformed significantly compared with typically developing (TD) children or norm values; ABR, auditory brainstem response; ASS, Adaptive Sentence List (Blackwood & Summerfield 1992); BKB, Bamford-Kowal-Bench sentences (Bench et al. 1978); BM, backward masking; CADS-P, The Conners' ADHD/DSM-IV Parent Version (Conners 1999); CCC-2, Children's Communication Checklist, 2nd edition (Bishop 2003); CELF-4, Clinical Evaluation of Language Fundamentals, 4th ed. (Semel et al. 2003); CHAPPS, Children's Auditory Processing Performance Scale (Snooski et al. 1998); CLPT, The Compelling Language Processing Test (Gaulin & Campbell 1994); CPRS-RS, Conners' Parent Rating Scale, revised: Short Form (Conners 1998); DDT, Dichotic digits test (Tzavaras et al. 1991); Musiek 1983; Musiek et al. 1991); DPT, Duration Patterns test (Pinheiro & Musiek 1985; Musiek 1994); FD, frequency discrimination; FPT, frequency patterns test (Pinheiro & Ptasnik 1971; Musiek 1994); GORT-4, The Gray Oral Reading Tests—4th edition (Wiederholt & Bryant 2001); HF, high frequency; IHR-STAR software, Institute of Hearing Research STAR software (Garry et al. 2010); NEPSY, Neuropsychological Test Battery (Korkman et al. 1997, 1998); The Nonword Repetition Test (Dollaghan & Campbell 1998); NV-IQ, nonverbal intelligence quotient; FFT, pattern of frequency test; PPVT-3, Peabody Picture Vocabulary Test—3rd edition (Dunn & Dunn 1997); SM, simultaneous masking; SMO, simultaneous masking noise task-band-pass noise; SMN, simultaneous masking noise task—spectrally notched noise; SSW, Staggered Spondaic Word Test (Katz 1962, 2001); SWMT, Spatial Working Memory Test (Ellis Weismer 2008); The Spoonerisms subtest (Walton & Brooks 1995) of the Phonological Assessment Battery (Frederickson et al. 1997); TOWRE, Test of Word Reading Efficiency (Torgesen et al. 1999); TROG-E, Test for Reception of Grammar—Electronic, version 2 (Bishop 2005); UNIT, Universal nonverbal intelligence test (Bracken & McCallum 1998); VCV, vowel-consonant-vowel; WASI, Wechsler Abbreviated Scale of Intelligence (Wechsler 1997); WISC-III, Wechsler Intelligence Scale for Children, third edition (Wechsler 1991); W-J III, Woodcock-Johnson Tests of Achievement—3rd edition (Woodcock et al. 2001).

Children diagnosed with APD versus children diagnosed with dyslexia

Table 6 shows the similarities and differences on various measurement instruments between children diagnosed with APD and children diagnosed with dyslexia. No differences were found on IQ, memory and language tests between children diagnosed with APD and children diagnosed with dyslexia. Children in the dyslexia group performed poorer with reading abilities. Parents of children diagnosed with APD rated significantly more often that their child met characteristics for autism. Furthermore, the CHAPPS total score is worse in children diagnosed with APD [41]. Children with comorbid APD and dyslexia performed significantly poorer on the duration pattern sequence test [47].

Table 6. Summary of the participant characteristics, reported measurements, and differences between children diagnosed with APD and children diagnosed with dyslexia.

Study	Participants/Groups		Measurement Instrument	Significance Difference Between Clinical Groups	Difference With TD Group	Summary of Significant Differences APD vs. dyslexia
	Diagnosis Reported; n; Gender; Age Range; Mean Age (<i>M</i>)	Diagnosis Reported; n; Gender; Age Range; Mean Age (<i>M</i>)				
Dawes et al. 2009	APD; n = 22; 14 M, 8 F; 6–13; <i>M</i> = 10.1 Twelve of the 22 children (54%) in the APD group also met criteria for dyslexia.	Dyslexia; n = 19; 17 M, 2 F; 6–13; <i>M</i> = 9.8	SCAN-C or SCAN-A	NR	NR	27% of the dyslexia group and 31% of the APD group scored below –1 SD.
			Experimental Tests of auditory processing			
			Auditory: 2-Hz FM	*	E	
			Auditory: 40-Hz FM	*	U	
			Auditory: 240-Hz FM	*	A	
			Auditory: IRN detection	*	A	
			Visual: Coherent form detection	*	E	
			Visual: Coherent motion detection	*	E	
Iliadou et al. 2009	APD; n = 41; NR; 8–15.11; NR No comorbidity with ADHD or dysorthography. Six of the children in the whole sample (N=55) had low or borderline IQ.	APD + Dyslexia; n = 14; NR; 8–15.11; NR	Auditory tests:			Poorer performance for APD group (without dyslexia). Poorer performance for APD + dyslexia group. Poorer performance for APD + dyslexia group.
			Speech in babble, right ear	*	NR	
			Speech in babble, left ear	†	NR	
			DDT, right ear	*	NR	
			DDT, left ear	*	NR	
			FPT, right ear	*	NR	
			FPT, left ear	*	NR	
			DPT, right ear	†	NR	
			DPT, left ear	†	NR	
			RGDT	*	NR	
			MLD	*	NR	

Table 6. (Continued)

Study	Participants/Groups		Measurement Instrument	Significance Difference Between Clinical Groups	Difference With TD Group	Summary of Significant Differences APD vs. dyslexia
	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)				
Dawes & Bishop 2010	APD; n = 25; 15 M, 10 F; NR; M = 10.4 Thirteen of 25 (52%) children in the APD group would also fit a diagnosis of either SLI, dyslexia, or both. Eleven of 19 (58%) children in the dyslexia group would also fit a diagnosis of SLI (score < -1 SD on two or more out of 6 language tests). Hyperactivity/inattention: the proportion of abnormal cases was 37% and 46% for the dyslexia and APD groups.	Dyslexia; n = 19; 17 M, 2 F; NR; M = 10.1 Thirteen of 25 (52%) children in the APD group would also fit a diagnosis of either SLI, dyslexia, or both. Eleven of 19 (58%) children in the dyslexia group would also fit a diagnosis of SLI (score < -1 SD on two or more out of 6 language tests). Hyperactivity/inattention: the proportion of abnormal cases was 37% and 46% for the dyslexia and APD groups.	CAST	*	A	Six children with an APD diagnosis fell within the clinical range for Asperger syndrome.
			CCC-2	*	U	
			General Communication Composite	*	U	Poorer performance scores for APD group.
			Subscales	*	U	
			Speech	*	U	
			Syntax	*	U	
			Semantic	*	U	
			Coherence	*	U	
			Inappropriate initiation	*	U	
			Stereotype language	*	U	
			Context	*	U	
			Nonverbal communication	*	U	
			Social relations	*	U	
			Interest	*	U	
			CHAPPS total score	†	U	
			Language composite	*	E	
			Average of the standard scores of six language tests	*	E	
			TROG-E	*	E	
			Sentence Repetition of the NEPSY	*	E	
			Repetition of Nonsense Words of the NEPSY	*	E	
			ERRNI story telling	*	E	
			ERRNI Mean length of utterance (MLU)	*	E	
			ERRNI story Comprehension	*	E	
			Literacy composite	*	E	
			Average of the standard scores of three literacy tests:	†	U	Worse performance for dyslexia group.
			OSCC spelling test	*	E	
			Word Reading of the TOWRE	*	E	10 of 25 of the APD group and 4 of 18 of the dyslexia group scored in the clinical range (< -1 SD).
			Nonword Reading of the TOWRE	*	E	
			Matrix Reasoning and block design subtest of the WASI	*	E	
			SCAN-C or SCAN-A	*	E	

Studies are arranged in order of the total quality score.

*No significant difference between clinical groups.

†Significant difference between clinical groups.

ADHD, attention-deficit hyperactivity disorder; APD, auditory processing disorder; SLI, specific language impairment; TD, typically developing children; F, female; M, male; NR, not reported; A, APD group underperformed significantly compared with TD children; E, both groups displayed equal performance in comparison with TD children or norm values; U, other clinical group (SLI, dyslexia, ADHD, or LD) underperformed significantly compared with TD children; U, both groups underperformed significantly compared with typically developing (TD) children or norm values; CAST, Childhood Asperger Syndrome Test (Williams et al. 2009); CCC-2, Children's Communication Checklist, 2nd edition (Bishop 2003); CHAPPS, Children's Auditory Processing Performance Scale (Smolki et al. 1998); DDT, Dichotic digits test (Tzavaras et al. 1981; Musiek 1983; Musiek et al. 1991); DPT, Duration Patterns test (Pinheiro & Musiek 1985; Musiek 1994); ERRNI, Expression, Reception and Recall of Narrative Instrument (Bishop 2004); FPT, Frequency Patterns test (Pinheiro & Ptacek 1978; Musiek 1994); FM, frequency modulation; IRN, iterated rippled noise; MLD, Masking Level Difference (Auditec, St Louis); NEPSY, Neuropsychological Test Battery (Korkman et al. 1997, 1998); OSCC, Oxford Study of Children's Communication Impairments (self-developed); RGDT, Random Gap Detection Test (Keith 2000a); SCAN-A, test for auditory processing disorders in adolescence and adults (Keith 1994b); SCAN-C, test for auditory processing disorders in children—revised (Keith 2000b); Speech in babble test (Iliadou et al. 2006); TOWRE, Test of Word Reading Efficiency (Torgesen et al. 1999); TROG-E, Test for Reception of Grammar—Electronic, version 2 (Bishop 2005); WASI, Wechsler Abbreviated Scale of Intelligence (Wechsler 1998).

Children diagnosed with APD versus children diagnosed with ADHD

Table 7 shows the similarities and differences on various measurement instruments between children diagnosed with APD and children diagnosed with ADHD. No differences were found in performance on intelligence, attention, and language tests between children diagnosed with APD and children diagnosed with ADHD.

In two of the three included studies, there was, besides a group of children diagnosed with APD, a group of children with comorbid ADHD and APD included. Only one study included, besides a group of children diagnosed with APD, a group of children with a single diagnosis of ADHD. This one study found that the performance on the auditory and visual duration pattern test (DPT) can possibly distinguish children diagnosed with APD from children diagnosed with ADHD. Children with APD had significantly lower and divergent scores on the auditory and visual DPT [43].

Children diagnosed with APD versus children diagnosed with learning disorders

Two studies compared the performance of children diagnosed with APD to the performance of children diagnosed with LD. Table 8 shows the similarities and differences on various measurement instruments between the children in these two groups.

The study of Walker et al. [35] enrolled no group with children diagnosed with only APD in their study. No differences were found between these two groups on the auditory and visual temporal processing tasks and language and reading tests used in this study. However, according to the authors, it is likely that the included group is very heterogeneous and contains children with ADD, language and reading problems, and more general developmental learning impairments [35]. Detailed information regarding the problems of these children is missing.

In the study of Cameron and Dillon [44] the performance on the Listening in Spatialized Noise-Sentences test (LiSN-S; [52]) of nine children with suspected APD is compared to the performance of 11 children with various learning disabilities, such as ADHD, dyslexia, memory deficits and visual processing

problems. With the LiSN-S the auditory stream segregation skills of a child - one specific part of listening - can be examined [53]. The test assesses of a child is capable to understand speech when noise is coming from different directions. Cameron and Dillon [44] found that children with suspected APD performed poorer in test conditions where the masker and target were spatially separated.

Table 7. Summary of the participant characteristics, reported measurements, and differences between children diagnosed with APD and children diagnosed with ADHD.

Study	Participants/Groups		Measurement Instrument	Significance Difference Between Clinical Groups	Difference With TD Group	Summary of Significant Differences APD vs. ADHD
	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)				
Bellis et al. 2011	APD; n = 7; NR; NR; M = 10.9	ADHD; n = 10; NR; NR; M = 13.1	Auditory tests DDT FPT DPT	* * †	U U U	Poorer performance for APD group. APD group exhibited a significantly larger HLD.
	No cognitive, sensory, developmental or related disorder unrelated to the primary diagnosis.		Visual analogs (self-developed) Dichoptic Digits Visual High-Low Visual Duration Patterns	* * †	U U U	Poorer performance for APD group.
Riccio et al. 1994	APD; n = 15; NR; 9.0-13.3; M = NR	APD + ADHD; n = 15; NR; 9.0-13.3; M = NR	Auditory measures SSW, right competing SSW, left competing LPFS, right LPFS, left FPPS, right FPPS, left SRT WISC-R/WISC-III CELF-R PPVT-R	* * * * * * * * * *	NR NR NR NR NR NR E E E	Of the 30 subjects, 17 met criteria for one or more diagnoses. One child in the APD group had an additional diagnosis of overanxious disorder.
	Whole sample: N = 30 (25 M, 5 F); M = 10.10					
Riccio et al. 1996	APD; n = 15; NR; 9.0-12.11; M = 11.20	APD + ADHD; n = 15; NR; 9.0-12.11; M = 10.55	ACPT	*	U	Whole sample: N = 30 (24 M, 6 F); M = 10.87; No information available about additional diagnoses.
	Whole sample: N = 30 (24 M, 6 F); M = 10.87; No information available about additional diagnoses.		WISC-III	*	E	

Studies are arranged in order of the total quality score.

*No significant difference between clinical groups.

†Significant difference between clinical groups.

ADHD, attention deficit hyperactivity disorder; APD, auditory processing disorder; TD, typically developing children; F, female; M, male; NR, not reported; A, APD group underperformed significantly compared with TD children; E, both groups displayed equal performance in comparison with TD children or norm values; O, other clinical group (SLI, dyslexia, ADHD, or LD) underperformed significantly compared with TD children; U, both groups underperformed significantly compared with typically developing (TD) children or norm values; ACPT, Auditory Continuous Performance Test (Keith 1994a); CELF-R, Clinical Evaluation of Language Fundamentals-Revised (Semel et al. 1980); DDT, Dichotic digits test (Zavanas et al. 1991; Musiek 1993; Musiek et al. 1991); DPT, Duration Patterns test (Pinheiro & Musiek 1995; Musiek 1994); FPPS, Frequency (Pitch) Pattern Sequence Test (Pinheiro 1977); FPT, Frequency Patterns test (Pinheiro & Placek 1971; Musiek 1994); HLD, Humming-Labeling Differential; LPFS, Low-Pass Filtered Speech Test (Willeford 1977); PPVT-R, Peabody Picture Vocabulary Test-Revised (Dunn & Dunn 1981); SRT, Seashore Rhythm Test (Seashore et al., 1960); SSW, Staggered Spondaic Word Test (Katz 1962, 2001); WISC-R, Wechsler Intelligence Scale for Children-Revised (Wechsler 1974); WISC-III, Wechsler Intelligence Scale for Children, third edition (Wechsler 1991).

Table 8. Summary of the participant characteristics, reported measurements, and differences between children diagnosed with APD and children diagnosed with LD.

Study	Participants/Groups		Measurement Instrument	Significance Difference		Summary of Significant Differences APD vs. LD
	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)	Diagnosis Reported; n; Gender; Age Range; Mean Age (M)		Between Clinical Groups	Difference With TD Group	
Cameron & Dillon 2008	APD; n = 9; 8 M, 1 F; 6.6–11.2; M = 9.1	LD; n = 11; 7 M, 4 F; 7.2–11.8; M = 9.4	LISN-S			
			Low-Cue speech reception thresholds	*	E	
			High-Cue speech reception thresholds	†	A	Poorer performance APD group.
			Talker Advantage	*	E	
			Spatial Advantage	†	A	Worse performance for APD group.
			Total Advantage	†	A	Worse performance for APD group.
	Children in the APD group had no ADHD and IQ scores >90.					
Walker et al. 2011	LD + APD; n = 10; NR; NR	LD; n = 28; NR; NR	Temporal processing tasks			
			Auditory			
			Within-channel gap detection			
			3 ms	*	E	
			8 ms	*	E	
			24 ms	*	E	
			Between-channel gap detection			
			30 ms	*	O	
			80 ms	*	O	
			200 ms	*	E	
			Sequential TOJ			
			0 ms	*	A	
			84 ms	*	A	
			400 ms	*	A	
			Overlapping auditory TOJ			
			50 ms	*	E	
			200 ms	*	U	
			614 ms	*	U	
			Visual			
			Sequential visual TOJ			
			0 ms	*	E	
			5 ms	*	E	
			24 ms	*	E	
			Overlapping visual TOJ			
			3 ms	*	E	
			12 ms	*	E	
			38 ms	*	E	
			Random dot kinematograms—			
			Coherent motion task			
			15 ms	*	E	
			25 ms	*	E	
			35 ms	*	E	
			Random dot kinematograms—			
			Transparent motion task			
			10 ms	*	E	
			20 ms	*	E	
			40 ms	*	E	
			Language/Reading tasks			
			Phonological Awareness Quotient subtest of the CTOPP	*	U	The mean score for LD + APD was consistently lower than the LD only subjects for the Olson PHONO, Olson ORTHO, CTOPP and WRAT-3.
			Reading subtest of the WRAT-3	*	U	
			Short version of the Olson PHONO and Olson ORTHO subtest	*	U	
			The Token Test	*	E	

Studies are arranged in order of the total quality score.

*No significant difference between clinical groups.

†Significant difference between clinical groups.

ADHD, attention deficit hyperactivity disorder; ADD, attention deficit disorder; APD, auditory processing disorder; LD, learning disorder; SLI, specific language impairment; TD, typically developing children; F, female; M, male; NR, not reported; A, APD group underperformed significantly compared with TD children; E, both groups displayed equal performance in comparison with TD children or norm values; O, other clinical group (SLI, dyslexia, ADHD, or LD) underperformed significantly compared with TD children; U, both groups underperformed significantly compared with typically developing (TD) children or norm values; CTOPP, Comprehensive Test of Phonological Processing (Wagner et al. 1999); LISN-S, Listening in Spatialized Noise test-sentences (Cameron & Dillon 2007, 2009); Olson PHONO and ORTHO, short versions of the Olson Phonological and Olson Orthographic subtests (Olson 1985); The Token Test (Bolter & Vignolo 1996; Orgass & Poeck 1986); TOJ, temporal order judgment; WRAT-3, Wide Range Achievement Test 3 (Wilkinson 1993).

Overlap between groups on the outcome of various tests

To be able to make a comparison between the performance of children diagnosed with APD and children diagnosed with other developmental disorders, the results of the isolated comparisons were analyzed and combined. Because comorbid groups can demonstrate additive effects of both disorders, only the studies in which children with APD were compared with children with a different developmental disorder were used for the overall comparison. Therefore, the four studies in which comorbid groups were enrolled were excluded for this analysis [35,44,45,47]. Figure 2 illustrates the pair-wise comparisons made in this systematic review and the overlap and differences on the outcomes of various tests used in the included studies for children diagnosed with APD and children diagnosed with SLI, dyslexia, ADHD, and LD. As mentioned earlier, also some children in the studies in which the focus was on the comparison of two different diagnostic groups have multiple diagnoses. This is not taken into account in this equation. The area of the circles is equal to the total number of subjects included in the studies in this analysis. The quantity N identifies the number of studies included (APD vs. SLI, $N=5$; APD vs. dyslexia, $N=2$; APD vs. ADHD, $N=1$; APD vs. LD, $N=1$). The amount n identifies the number of children included in the studies. All included children from the various studies are added together. This means that some children are counted twice, because in a number of studies the same participants took part in more than one study (e.g., [37,41]). No differences (bold and outlined numbers in Figure 2) were found in the performance of children with APD compared with children with a different developmental disorder on the outcome of 85 of the 102 (sub)tests used in the included studies. Significantly poorer performance of children with APD (encircled italic numbers in Figure 2) was found in the case of ten (sub)tests (LiSN-S: High-Cue speech reception threshold, Spatial advantage and Total advantage, [44]; CHAPPS total score, [41]; auditory and visual DPT, [43]; CHAPPS total score and noise, multiple input and attention subscales, [36]). For seven (sub)tests, performances in children with SLI or dyslexia were poorer than in children with APD (literacy composite, [41]; CCC-2 subscale Speech, [36]; Speech-evoked auditory brain response: timing-measure E and F peak and spectral encoding measure higher harmonics, [39]; Speech-in-noise test and Dichotic Digit test, [40]).

Figure 2. Venn diagram of the comparisons made in this systematic review and the overlap and differences between groups on the outcome of various tests used in the included studies. The performance of children diagnosed with APD on various tests is compared with the performance of children diagnosed with SLI, dyslexia, ADHD, and LD. *N* identifies the number of studies included. The area of the circles is equal to the number of subjects included in the studies (*n*). The bold and the outlined numbers show the overlap (no significant differences) on the outcome of the used measurement instruments. Italic numbers illustrate the significant differences between children diagnosed with APD and children diagnosed with other disorders. The encircled italic numbers show the number of measurement instruments (test and subtest level) whereupon children diagnosed with APD perform significantly poorer in comparison with children diagnosed with other developmental disorders. The studies in which children with comorbid disorders (*N*=4) were enrolled are not captured in this figure. ADHD, attention-deficit hyperactivity disorder; APD, auditory processing disorder; CHAPPS, Children's Auditory Processing Performance Scale [51]; DPT, Duration Pattern Test; LD, learning disorder; LiSN-S, Listening in Spatialized Noise test-Sentence [44,52,53]; SLI, specific language impairment, SRT, speech reception threshold.

Discussion

The present study was designed (1) to determine which characteristics of APD overlap with the characteristics of SLI, dyslexia, ADHD, LD, and ASD and (2) to examine whether there are characteristics that distinguish children diagnosed with APD from children diagnosed with other developmental disorders. With respect to the first research question, we found that children diagnosed with APD and children diagnosed with SLI, dyslexia, ADHD, and LD have overlapping characteristics of intelligence, memory, attention and language and that the reading skills of children diagnosed with APD overlap with the reading skills of children diagnosed with SLI. No studies were found in which the performance of children diagnosed with APD were compared with children diagnosed with ASD. Compared with TD children, children diagnosed with APD and children diagnosed with other disorders show broadly similar results. Unfortunately, not all studies reported how the performance on several tests of children in the research groups was compared to TD children (for details, see Table 5 – 8). In the case of 57 (sub) tests (45,6%), both groups underperformed significantly compared with TD children or norm values, and in the case of 46 (sub) tests (36,8%), both groups displayed equal performance in comparison with TD children or norm values. With regard to the second research question, possible distinguishing characteristics for children diagnosed with APD were found in only four studies [36,41,43,44]. Ferguson et al. [36] compared the performance of 25 children diagnosed with APD to the performance of 22 children diagnosed with SLI. Dawes and Bishop [41] also included 25 children diagnosed with APD and compared their performance to that of 19 children diagnosed with dyslexia. Bellis et al. [43] included 7 children diagnosed with APD and 10 children with an ADHD diagnosis, and Cameron and Dillon [44] compared the performance of 9 children diagnosed with APD to the performance of 11 children diagnosed with LD. These studies found that children with APD had: (1) auditory and visual temporal ordering difficulties opposed to children diagnosed with ADHD, (2) poorer outcome on subtests of the LiSN-S compared to children diagnosed with LD, and (3) weaker listening skills opposed to children diagnosed with SLI and children diagnosed with dyslexia according to the parental evaluation. These characteristics may suggest a distinction between children diagnosed with APD and children diagnosed with SLI, dyslexia, ADHD, and LD. However, it is important to acknowledge that generalization to a broader

population is limited because of the fact that these characteristics were found in single studies with small sample sizes, heterogeneous groups with comorbidity, and moderate methodological quality.

Overlap between APD and SLI, dyslexia, and ADHD

In this review, we found that children diagnosed with APD share overlapping intelligence, attention, memory and language characteristics with children diagnosed with other developmental disorders. Executive functions could be a potentially useful construct for understanding the overlapping symptoms observed in different disorders, such as APD, ADHD, SLI, and dyslexia [54,55]. Executive functions appear to be related to the processing of language. Children diagnosed with SLI have demonstrated deficits in cognitive control related to deficits in language processing [56]. Such deficits are frequently incorporated into the general category of executive functions [56]. The overlapping symptoms of children diagnosed with APD, SLI, ADHD, and dyslexia fits the idea that it is difficult to say that the various psychiatric categories, as stated in the Diagnostic and Statistical Manual of Mental Disorders-V, are separate diagnosis with clear boundaries and a clear underlying cause. Our results support the idea that many of the individual categorical diagnoses are associated with each other in clusters, groups, or on a spectrum rather than that they are truly distinct disorders [7,57-62]. A model that fits this idea is the multiple deficit model proposed by Pennington [60]. The multiple deficit model assumes that a developmental disorder is caused by a combination of underlying specific and shared components [63]. Pennington proposes with the model that “the etiology of complex behavioral disorders is multifactorial and involves the interaction of multiple risk factors and protective factors, which can be either genetic or environmental” [60, p. 404]. As a result of the shared etiologic, comorbidity among neurodevelopmental disorders can be expected. Specific components ensure that the various developmental disorders differ slightly from each other on the surface. According to Moore and Hunter [7], the various psychiatric categories can be conceptualized as a more general neurodevelopmental syndrome, wherein the behavioral difficulties (e.g., auditory, language, and attentional) of children serve as markers that can be expressed along a continuum of severity [7,24,64].

Differences between APD and SLI, dyslexia, and ADHD

The APD-group demonstrated significantly lower performance on the CHAPPS questionnaire [36,41] and the subtests of the LiSN-S in which the noise was spatially separated from the target speech [44]. This may indicate that difficulties with the ability to listen in noisy conditions in challenging environments is a distinctive characteristic for children with suspected APD. Difficulties in noisy environments are one of the most mentioned symptoms of children with susAPD [3,65]. However, there are a number of concerns regarding the interpretation of the experimental evidence that children diagnosed with APD experience “difficulties with listening in noise”.

Information regarding the validity and reliability of the CHAPPS questionnaire [51] is lacking, whereby it is not certain what is actually measured [1,24,66]. A number of studies have examined the clinical utility of the CHAPPS [67-70]. Hereby, varying results are reported. Lam and Sanchez [69] and Wilson et al. [70] found that the CHAPPS has no significant predictive capability for APD, while others [67,68] reported that a number of subsections (Listening condition: ideal, auditory memory/sequencing, and auditory attention span) of the CHAPPS may be clinically useful for identifying listening difficulties in children; however, further research is needed. The CHAPPS is originally designed to be completed by the teacher of the child [51]. The teacher answers the questions by comparing the listening skills of the individual child with the listening skills of other children of the same age group and background [51]. However, in both studies included in this systematic review, the CHAPPS was completed by parents instead of the teacher. This may plausibly cause a referral bias because parents of children with a diagnosis of APD could complete the questionnaire from a different perspective than parents of children with a diagnosis of SLI or dyslexia [36]. No clear relationship between CHAPPS scores and the diagnosis APD were found in studies in which the CHAPPS questionnaire was completed by the child’s teacher [69,70].

The results of this systematic review suggest that the LiSN-S [52] could perhaps be a valuable instrument to differentiate between children diagnosed with APD and children diagnosed with other developmental disorders. The LiSN-S was designed to assess auditory stream segregation skills in children with susAPD

[53], and can be used to examine spatial processing difficulties [71]. With this test, one specific part of the whole range of listening abilities children need to have to be able to pick up information from what they hear, can be examined. The LiSN-S has a high test-retest reliability and normative data are provided for people aged 6 to 60 years [53,72,73]. However, the LiSN-S is only used in one of the 14 included studies. In addition, the study group of this one study was very small (9 children diagnosed with APD versus 11 children diagnosed with learning or attention disorders). Studies that previously have used the LiSN-S to examine the differences between children with susAPD and TD children found varying results. Cameron and Dillon [44,74] found that the inability to use spatial information may be one important cause of the listening difficulties in children. Others found no differences in spatial processing between children with susAPD and TD children on the LiSN-S Test [66,75]. As also explained by the authors themselves, it would be beneficial to conduct studies that are more large-scale in the future with the LiSN-S to investigate whether the test can be used to distinguish between children with and without spatial processing difficulties.

The significant underperformance of children diagnosed with APD on the auditory and visual DPT in the study of Bellis et al. [43] indicates that a nonmodal-specific temporal ordering deficit may also be a possible characteristic that distinguish children diagnosed with APD from children diagnosed with other developmental disorders. Bellis et al. [43] used an auditory DPT [76,77] in their study with triads of 1000 Hz tone bursts differing in short (250 msec) and long (500 msec) duration. For the auditory DPT [78], the levels of diagnostic accuracy (sensitivity, 0.86; specificity, 0.92) are known [62,79]. The visual DPT (self-developed) consisted of a black rectangle presented on a white screen in short (250 msec) and long (800 msec) triads. Bellis et al. [43] reported robust effect sizes despite the small number of participants. The differences between children diagnosed with APD and children diagnosed with ADHD in auditory and visual DPT suggest that the deficits of children with listening difficulties are not specific to the auditory modality because children diagnosed with APD underperformed on both the visual and auditory version. Some authors proposed that problems in auditory processing are modality-specific if the problems are only situated in the auditory modality [80,81]. When there are also problems in other modalities, such as problems with visual tasks, they would rather not speak of APD. Furthermore,

there is no evidence that the DPT also differentiates between children diagnosed with APD and children diagnosed with SLI or dyslexia. In the study of Iliadou et al. [47], children with a diagnosis of APD and dyslexia scored poorer on the DPT compared with children with a diagnosis of only APD. Also, no differences between children diagnosed with APD and children diagnosed with SLI in performance on the DPT were found by Miller and Wagstaff [38]. Studies in which auditory pattern recognition is examined found similar results, that is, children diagnosed with dyslexia or SLI perform poorer than TD children in the detection of duration patterns [82-85]. Therefore, it cannot be argued that a temporal ordering deficit is a distinguishing characteristic for APD.

In general, there is substantial overlap between children with various developmental disorders. The performance of children diagnosed with APD on three outcome measurements (CHAPPS, LiSN-S, and DPT) suggested that there are differences between children diagnosed with APD and children diagnosed with other developmental disorders. However, the results are not consistent and the findings seem insufficiently researched; hence, it remains ambiguous whether these results are replicable. As previously mentioned by Moore and Hunter [7], the various developmental disorders could be conceptualized as a general neurodevelopmental syndrome wherein “auditory, speech, language, attention, memory and behavioral difficulties (markers) in children are expressed along a continuum of severity” [7, pp. 165]. Perhaps it is possible to use the three outcome measurements found in this review to indicate where children are located on this continuum.

Methodological quality

All included studies in this systematic review had moderate methodological quality because none of the studies used a random sample, because none of the assessors were blinded, and because it was not clear if the participants in the different groups were comparable or not at baseline. In most studies, it was also not evident whether the children in the various groups (APD, SLI, ADHD, and dyslexia) had additional deficits besides the primary diagnosis they had received. The studies did not describe a broad assessment at the beginning in order to: (1) check the specified diagnosis and (2) investigate all important skills covering all developmental areas. It could be that the number of pure cases of children with APD, SLI, ADHD, or dyslexia is negligible if a detailed assessment battery

was used at the beginning of the studies [59]. The absence of valid auditory processing measurement instruments is another factor that is of influence on the methodological quality of the studies. There are no valid auditory processing tests available because there is no reference standard for the assessment of APD [62].

Clinical implications and future research

Children with various developmental disorders perform similar on the outcome of 85 of the 102 (sub)tests (see Tables 5 – 8 and Fig. 2 for details). No significant differences were found between groups on most auditory tests, and on questionnaires and tests of intelligence, attention, memory, language, and reading. With the current measurement instruments used in clinical practice, it is difficult to distinguish the various disorders from each other. From the results of our previous systematic review [29], we know that children with susAPD perform significantly lower, compared to TD children, on auditory behavioral tests. Clinicians should take into account the fact that, if children have abnormal performance on a test for auditory processing, this does not automatically signify that the child has APD. Otherwise stated, the tests may distinguish abnormal from normal performance, however, they are not able to distinguish the various conditions from each other. It is also known that failure on a specific test, not necessarily mean that the child has a problem [86]. Also, TD children show considerable variability on at least some of these measures. From this and our previous review [29], it is evident that listening difficulties experienced by children are multifaceted and that there is substantial overlap between various developmental disorders. As Moore [87] stated, “Hearing necessarily involves the ear, the central auditory nervous system, and other brain systems, including attention, memory and vision.” Therefore, it is crucial that various professions work together and use a multidisciplinary approach not only in the assessment of children with listening complaints but also in the event of children who satisfy the diagnostic standards of SLI, ADHD, and dyslexia. This accords with the latest suggestion of Chermak et al. [71] who stated that an audiologist, speech-language pathologist, and psychologist must collectively decide whether the listening difficulties arise from an auditory issue, problems with processing language, cognitive deficiencies such as attention, executive function, or working memory, or an aggregation of them. It is important to note that the similarities found in this systematic review between children with various disorders are valid

on a group level and does not reveal individual differences. Audiologists, speech-language pathologists, and psychologists should focus on the performance of the individual child. To apply appropriate treatment, the symptoms and complaints in everyday life of the individual child must be focused on and not necessarily the diagnostic label.

The findings across the 13 studies should be interpreted according to study quality and design characteristics. Heterogeneity of the participants, inadequate reporting of the profile of participants, large variability in the measures and procedures used across studies, and a minimal number of studies with sometimes a small sample size should be taken into account when making comparisons between studies. There is a clear need for higher quality, well-designed studies in which a comparison is made between different groups of children with various disorders. We recommend that these studies include multiple groups of children with different diagnoses and the use of a detailed assessment battery incorporating the entire range of symptoms of neurodevelopmental disorders so that not only the performance of children diagnosed with APD is compared to the performance of children diagnosed with some other disorder but that the performance of children with all various developmental disorders can be compared with each other and that the entire continuum of developmental disorders can be identified.

Limitations of the current review

The limitations of this systematic review are the same as the restrictions listed in our previous study (de Wit et al. 2016), specifically (1) only articles written in English were included; (2) case studies were excluded; (3) studies in which auditory processing or one of the used synonyms was not mentioned in the title were excluded; and (4) another way of classifying the methodological quality of the studies could lead to other studies included or excluded in this review. Unfortunately, conducting a meta-analysis was not possible because of the diversity among outcome measures and the wide variation in the descriptive of participants.

Conclusion

Only marginal differences were found between children diagnosed with APD and children diagnosed with a different developmental disorder. Children diagnosed with APD and children diagnosed with SLI, dyslexia, ADHD, and LD have overlapping characteristics in terms of intelligence, memory, attention, and language. The reading skills of children diagnosed with APD overlap with the reading skills of children diagnosed with SLI. The results suggest that the CHAPPS questionnaire and the subtests of the LiSN-S in which noise is spatially separated from target sentences could possibly differentiate between children with difficulties in auditory functioning and children with language, reading, and attention disorders. However, this result is possibly confounded by the generally poor quality of the research studies included in this review and the quality of the used outcome measures. Additional research is required to better understand the different profiles of children with various complaints or disorders.

Acknowledgements

This research was funded by the Netherlands Organization for Scientific Research (NWO). Preliminary results of this study were presented by poster presentation on the British Society of Audiology, Annual Conference 2016 and by oral presentation on the 30th World Congress of the International Association of Logopedics and Phoniatrics (I.A.L.P.) 2016. The authors declare no conflict of interest.

References

References marked with an asterisk (*) indicate studies included in the systematic review.

1. American Academy of Audiology. *Clinical practice guidelines. Diagnosis, treatment and management of children and adults with central auditory processing disorders*. Retrieved from http://audiology-web.s3.amazonaws.com/migrated/CAPD%20Guidelines%208-2010.pdf_539952af956c79.73897613.pdf. 2010
2. Geffner D, Ross-Swain D. In Geffner D., & Ross-Swain D. (Eds.), *Auditory Processing Disorders: Assessment, Management, and Treatment* (2nd ed.). San Diego, CA, US: Plural Publishing. 2013
3. American Speech-Language-Hearing Association. (Central) auditory processing disorders [Technical report]. doi:10.1044/policy.PS2005-00114. Retrieved from <http://www.asha.org/policy/TR2005-00043/>. 2005.
4. British Society of Audiology. Position statement on auditory processing disorder. Auditory Processing Disorder (APD) Steering Committee. Retrieved from ftp://ftp.phon.ucl.ac.uk/pub/andyf/BSA_APD_Position_Consultation.pdf. 2011.
5. Cacace A, McFarland DJ. *Current Controversies in Central Auditory Processing Disorder (CAPD)*. San Diego, CA, US: Plural Publishing. 2009
6. DeBonis DA. It is time to rethink central auditory processing disorder protocols for school-aged children. *Am J Audiol*. 2015;24, 124-136.
7. Moore DR, Hunter LL. Auditory processing disorder (APD) in children: a marker of neurodevelopmental syndrome. *Hearing, Balance and Communication*. 2013;11, 160-167.
8. McFarland DJ, Cacace AT. (2006). Current Controversies in CAPD: From Procrustes' Bed to Pandora's Box. In T. K. Parthasarathy (Ed.), *An Introduction to Auditory Processing Disorders in Children* (pp. 247-263). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.
9. Bellis TJ. *Assessment & Management of Central Auditory Processing Disorders in the Educational Setting: From Science to Practice* (2nd ed.). San Diego, CA: Plural Publishing. 2003.
10. Dawes P, Bishop DVM. Auditory processing disorder in relation to developmental disorders of language, communication and attention: a review and critique. *Int J Lang Comm Dis*. 2009;44(4), 440-465.
11. McArthur GM. Auditory processing disorders: can they be treated? *Curr Opin Neurol*. 2009;22(2), 137-143. doi:10.1097/WCO.0b013e328326f6b1.
12. Wilson WJ, Arnott W. Using different criteria to diagnose (central) auditory processing disorder: how big a difference does it make? *J Speech Lang Hear Res*. 2013;56(1), 63-70.
13. Bamiau DE, Musiek FE, Luxon LM. Aetiology and clinical presentations of auditory processing disorders - a review. *Arch Dis Child*. 2001;85(5), 361-365.
14. Chermak GD, Musiek F. *Central Auditory Processing Disorders: New Perspectives*. San Diego, CA, US: Singular Publishing Group. 1997.

15. Hind SE, Haines-Bazrafshan R, Benton CL. et al. Prevalence of clinical referrals having hearing thresholds within normal limits. *Int J Audiol.* 2011;50(10), 708-716.
16. Bellis TJ. Historical foundations and the nature of (central) auditory processing disorder. In F. E. Musiek, & G. D. Chermak (Eds.), *Handbook of (Central) Auditory Processing Disorder. Volume 1: Auditory Neuroscience and Diagnosis.* (1st ed., pp. 119-136). San Diego: Plural Publishing. 2007.
17. Jerger JF. The concept of auditory processing disorder. A Brief History. In A. T. Cacace DJ, McFarland (Eds.), *Controversies in Central Auditory Processing Disorder (CAPD)* (pp. 1-14). San Diego, CA, US: Plural Publishing. 2009.
18. Lucker JR. The history of auditory processing disorders in children. In D. Geffner, D. Ross-Swain (Eds.), *Auditory Processing Disorders: Assessment, Management, and Treatment* (2nd ed., pp. 33-58). San Diego, CA, US: Plural Publishing. 2013
19. American Psychiatric Association. *Handboek voor de classificatie van psychischestoornissen (DSM-5)* [Dutch version of the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition]. Amsterdam: Boom. 2014
20. Rutter M, Pine DS. Diagnosis, diagnostic formulations, and classification. In A. Thapar, D. S. Pine, J. F. Leckman et al. (Eds.), *Rutter's Child and Adolescent Psychiatry* (6th ed., pp. 17-30). Oxford, UK: Wiley Blackwell. 2015.
21. Kamhi AG. What speech-language pathologists need to know about auditory processing disorder. *Lang Speech Hear Ser.* 2011;42(3), 265-272.
22. Levy S, Parkin CM. Are We Yet Able to Hear the Signal Through the Noise? A Comprehensive Review of Central Auditory Processing Disorders: Issues of Research and Practice. *Can J Sch Psychol.* 2003;18(1-2), 153-182. doi:10.1177/082957350301800108.
23. Miller CA. Auditory processing theories of language disorders: Past, present, and future. *Lang Speech Hear Ser.* 2011;42(3), 309-319.
24. Moore DR, Rosen S, Bamiou D. et al. Evolving concepts of developmental auditory processing disorder (APD): A British Society of Audiology APD Special Interest Group 'white paper'. *Int J Audiol.* 2013;52(1), 3-13.
25. O'Connor K. Auditory processing in autism spectrum disorder: a review. *Neurosci Biobehav R.* 2012;36, 836-854. doi:10.1016/j.neubiorev.2011.11.008.
26. Medwetsky L. Spoken Language Processing Model: Bridging Auditory and Language Processing to Guide Assessment and Intervention. *Lang Speech Hear Ser.* 2011;42, 286-296. doi:10.1044/0161-1461.
27. Sharma M, Purdy SC, Kelly AS. Comorbidity of auditory processing, language, and reading disorders. *J Speech Lang Hear Res.* 2009;52(3), 706-722. doi:10.1044/1092-4388(2008/07-0226).
28. Neijenhuis K, Snik A, van den Broek P. Auditory processing disorders in adults and children: evaluation of a test battery. *Int J Audiol.* 2003;42(7), 391-400.
29. Wit de E, Visser-Bochane MI, Steenbergen B, et al. Characteristics of Auditory Processing Disorders: A Systematic Review. *J Speech Lang Hear Res.* 2016;59, 384-413.

30. Moher D, Liberati A, Tetzlaff J, et al. PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*. 2009;6(7), e1000097. doi:10.1371/journal.pmed1000097.
31. Moore JK, Linthicum FH. The human auditory system: a timeline of development. *Int J Audiol*. 2007;46, 460-478.
32. Mullen R. The state of the evidence: ASHA develops levels of evidence for communication sciences and disorders. *ASHA Lead*, 2007;12(3), 8-9, 24,25.
33. Gyorkos TW, Tannenbaum TN, Abrahamowicz M, et al. An approach to the development of practice guidelines for community health interventions. *C J Public Health*. 1994;85(suppl 1), S8-S13.
34. Simões, M. B., & Schochat, E. (Central) auditory processing disorders in individuals with and without dyslexia. *Pro Fono*. 2010;22(4), 521-524.
35. *Walker KMM, Brown DK, Scarff C, et al. Temporal processing performance, reading performance, and auditory processing disorder in learning-impaired children and controls. *Can J Speech-Lang Pathology and Audiology*. 2011;35(1), 6-17.
36. *Ferguson MA, Hall RL, Riley A, et al. Communication, listening, cognitive and speech perception skills in children with auditory processing disorder (APD) or Specific Language Impairment (SLI). *J Speech Lang Hear Res*. 2011;54(1), 211-227.
37. *Ferguson MA, Moore DR. Auditory Processing Performance and Nonsensory Factors in Children with Specific Language Impairment or Auditory Processing Disorder. *Seminars in Hearing*. 2014;35(1), 1-14.
38. *Miller CA, Wagstaff DA. Behavioral profiles associated with auditory processing disorder and specific language impairment. *J Comm Dis*. 2011;44(6), 745-763. doi:10.1016/j.jcomdis.2011.04.001.
39. *Rocha-Muniz CN, Befi-Lopes DM, Schochat E. Investigation of auditory processing disorder and language impairment using the speech-evoked auditory brainstem response. *Hearing Res*. 2012;294(1-2), 143-152.
40. *Rocha-Muniz CN, Zach EC, Teixeira RA, et al. Association between language development and auditory processing disorders. *Braz J Otorhinolaryngology*. 2014;80(3), 231-236.
41. *Dawes P, Bishop DV. Psychometric profile of children with auditory processing disorder and children with dyslexia. *Arch Dis Child*. 2010;95(6), 432-436.
42. *Dawes P, Sirimanna T, Burton M. et al. Temporal auditory and visual motion processing of children diagnosed with auditory processing disorder and dyslexia. *Ear Hearing*. 2009;30(6), 675-686. doi:10.1097/AUD.0b013e3181b34cc5.
43. *Bellis TJ, Billiet C, Ross J. The utility of visual analogs of central auditory tests in the differential diagnosis of (central) auditory processing disorder and attention deficit hyperactivity disorder. *J Am Acad Audiol*. 2011;22(8), 501-514. doi:10.3766/jaaa.22.8.3.
44. *Cameron S, Dillon H. The Listening in Spatialized Noise-Sentences Test (LiSN-S): comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *J Am Acad Audiol*. 2008;19(5), 377-391.
45. *Riccio CA, Cohen MJ, Hynd GW, et al. Validity of the Auditory Continuous Performance Test in differentiating central processing auditory disorders with and without ADHD. *J Learn Disabil*. 1996;29(5), 561-566.

46. *Riccio CA, Hynd GW, Cohen MJ, et al. Comorbidity of central auditory processing disorder and attention-deficit hyperactivity disorder. *J Am Acad Child Psy.* 1994;33(6), 849-857. doi:10.1097/00004583-199407000-00011
47. *Iliadou V, Bamiou DE, Kaprinis S, et al. Auditory Processing Disorders in children suspected of Learning Disabilities--a need for screening? *Int J Pediatr Otorhi.* 2009;73(7), 1029-1034. doi:10.1016/j.ijporl.2009.04.004.
48. Leonard LB. *Children with Specific Language Impairment*. Cambridge, MA: MIT Press. 1998.
49. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)* (3rd ed.). Washington, DC: American Psychiatric Association. 1980.
50. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (4th ed.). Washington, DC: American Psychiatric Association. 1999.
51. Smoski WJ, Brunt MA, Tannahill JC. *Children's Auditory Processing Performance Scale (CHAPPS)*. Tampa Florida: The Educational Audiology Association. 1998.
52. Cameron S, Dillon H. Development of the Listening in Spatialized Noise Sentences Test (LiSN-S). *Ear Hearing.* 2007;28(2), 196-211. doi:10.1097/AUD.0b013e318031267f.
53. Cameron S, Dillon H. The listening in spatialized noise-sentences test (LiSN-S): test-retest reliability study. *Int J Adiol.* 2007;46, 145-153. doi:10.1080/14992020601164170.
54. Chermak GD, Hall JW, Musiek FE. Differential diagnosis and management of central auditory processing disorder and attention deficit hyperactivity disorder. *J Am Acad Audiol.* 1999;10, 289-303.
55. Thapar A, Pine DS, Leckman JF, et al. (Eds.). *Rutter's Child and Adolescent Psychiatry* (6th ed.). Oxford, UK: Wiley Blackwell. 2015.
56. Victorino KR, Schwartz RG. Control of auditory attention in children with specific language impairment. *J Speech Lang Hear Res.* 2015;58, 1245-1257. doi:10.1044/2015_JSLHR-L-14-0181.
57. Bishop DVM, Snowling MJ. Developmental dyslexia and specific language impairment: same or different? *Psychol Bull.* 2004;130(6), 858-886. doi:10.1037/0033-2909.130.6.858.
58. Bishop DVM. The interface between genetics and psychology: lessons from developmental dyslexia. *Proc R Soc: B.* 2015;282, 20143139. doi:10.1098/rspb.2014.3139.
59. Bishop DVM, Rutter M. Neurodevelopmental disorders: conceptual issues. In M. Rutter, D. V. M. Bishop, D. S. Pine et al. (Eds.), *Rutter's Child and Adolescent Psychiatry* (5th ed., pp. 32-41). Oxford, UK: Blackwell Publishing. 2008.
60. Pennington BF. From single to multiple deficit models of developmental disorders. *Cognition.* 2006;101, 385-413. doi:10.1016/j.cognition.2006.04.008.
61. Pennington BF, Bishop DVM. Relations among speech, language, and reading disorders. *Annu Rev Psychol.* 2009;60(1), 283-306. doi:10.1146/annurev.psych.60.110707.163548.
62. Vermiglio AJ. On diagnostic accuracy in audiology: central site of lesion and central auditory processing disorder studies. *J Am Acad Audiol.* 2016;27, 141-156. doi:10.3766/jaaa.15079.
63. McGrath LM, Pennington BF, Shanahan MA. et al. A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: searching for shared cognitive deficits. *J Child Psychol Psc.* 2011;52(5), 547-557.

64. Wallach GP. Peeling the onion of auditory processing disorder: a language/curricular-based perspective. *Lang Speech Hear Ser.* 2011;42, 273-285. doi:10.1044/0161-1461(2010/10-0008).
65. Jerger J, Musiek F. Report of the Consensus Conference on the Diagnosis of Auditory Processing. *J Am Acad Audiol.* 2000;11(9), 467-474.
66. Barry JG, Tomlin D, Moore DR. et al. Use of questionnaire-based measures in the assessment of listening difficulties in school-aged children. *Ear Hearing.* 2015;36(6), e300-e313. doi:10.1097/AUD.0000000000000180.
67. Ahmmed AU, Ahmmed AA. Setting appropriate pass or fail cut-off criteria for tests to reflect real life listening difficulties in children with suspected auditory processing disorder. *Int J Pediatr Otorhi.* 2016;84, 166-173. doi:10.1016/j.ijporl.2016.03.014.
68. Iliadou V, Bamiou DE. Psychometric evaluation of children with auditory processing disorder (APD): comparison to a normal and a clinical non APD groups. *J Speech Lang Hear Res.* 2012;55, 791-799. doi:10.1044/1092-4388(2011/11-0035).
69. Lam E, Sanchez L. Evaluation of screening instruments for auditory processing disorder (APD) in a sample of referred children. *Aust NZ J Audiol.* 2007;29(1), 26-39.
70. Wilson WJ, Jackson A, Pender A, et al. The CHAPS, SIFTER, and TAPS-R as predictors of (C)AP skills and (C)APD. *J Speech Lang Hear Res.* 2011;54, 278-291.
71. Chermak GD, Bamiou DE, Iliadou V, et al. Practical guidelines to minimize language and cognitive confounds in the diagnosis of CAPD: a brief tutorial. *Int J Audiol.* 2017;1-8, published online: <http://dx.doi.org/10.1080/14992027.2017.1284351>.
72. Cameron S, Brown D, Keith R, et al. Development of the Nort American Listening in Spatialized Noise-Sentences Test (NA LiSN-S): Sentence Equivalence, Normative Data, and Test-Retest reliability studies. *J Am Acad Audiol.* 2009;20, 128-146. doi:10.3766/jaaa.20.2.6.
73. Cameron S, Glyde H, Dillon H. Listening in Spatialized Noise – Sentences Test (LiSN-S): normative and retest reliability data for adolescents and adults up to 60 years of age. *J Am Acad Audiol.* 2011;22, 697-709. doi:10.3766/jaaa.22.10.7.
74. Cameron S, Dillon H, Newall P. The Listening in Spatialized Noise Test: An auditory processing disorder study. *J Am Acad Audiol.* 2006;17, 306-320.
75. Sharma M, Dhamani I, Leung J, et al. Attention, memory, and auditory processing in 10- to 15- year-old children with listening difficulties. *J Speech Lang Hear Res.* 2014;57, 2308-2321. doi:10.1044/2014_JSLHR-H-13-0226.
76. Pinheiro ML, Musiek FE. Sequencing and temporal ordering in the auditory system. In M. L. Pinheiro, & F. E. Musiek (Eds.), *Assessment of Central Auditory Dysfunction: Foundations and Clinical Correlates* (pp. 219-238). Baltimore: Williams and Wilkins. 1985.
77. Musiek FE, Baran JA, Pinheiro, M. L. Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology.* 1990;29, 304-313.
78. Musiek FE. Frequency (pitch) and duration pattern tests. *J Am Acad Audiol.* 1994;5, 265-268.
79. Friberg JC, McNamara TL. Evaluating the reliability and validity of (central) auditory processing tests: a preliminary investigation. *J of Educational Audiology.* 2010;16, 4-17.

80. Cacace AT, McFarland DJ. The importance of modality specificity in diagnosing central auditory processing disorder. *Am J Audiol*. 2005;14(2), 112-123.
81. Cacace AT, McFarland DJ. Factors influencing tests of auditory processing: A perspective on current issues and relevant concerns. *J Am Acad Audiol*. 2013;24(7), 572-589.
82. King WM, Lombardino LJ, Crandell CC, et al. Comorbid auditory processing disorder in developmental dyslexia. *Ear Hearing*. 2003;24(5), 448-456.
83. Stollman MH, van Velzen EC, Simkens HM, et al. Assessment of auditory processing in 6-year-old language-impaired children. *Int J Audiol*. 2003;42(6), 303-311.
84. Walker MM, Givens GD, Cranford JL, et al. Auditory pattern recognition and brief tone discrimination of children with reading disorders. *J Comm Dis*. 2006;39(6), 442-455.
85. Walker MM, Shinn JB, Cranford JL, et al. Auditory temporal processing performance of young adults with reading disorders. *J Speech Lang Hear Res*. 2002;45(3), 598-605.
86. Dillon H, Cameron S, Glyde H, Wilson W, et al. An opinion on the assessment of people who may have an auditory processing disorder. *J Am Acad Audiol*. 2012;23(2), 97-105. doi: 10.3766/jaaa.23.2.4.
87. Moore DR. What's new in auditory processing? *ENT & Audiology News*. 2016;25(2), 2-3.

References belonging to Table 5, Table 6, Table 7, Table 8.

- Barry JG, Ferguson MA, Moore DR. Making sense of listening: the IMAP test battery. *J Vis Exp (JoVE)*. 2010;(44), e2139-e2139. doi:10.3791/2139.
- Bench J, Kowal Å, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Brit J Audiol*. 1979;13(3), 108-112.
- Bishop DVM. *The Children's Communication Checklist: CCC-2* (2nd ed.). London, UK: Pearson Assessment. 2003.
- Bishop DVM. *Expression, Reception and Recall of Narrative Instrument (ERRNI)*. Oxford, UK: Pearson assessment. 2004.
- Bishop DVM. *Test for Reception of Grammar, Version 2—Electronic (TROG-E)*. London, UK: Psychological Corporation. 2005.
- Boller F, Vignolo LA. Latent sensory aphasia in hemisphere-damaged patients: an experimental study with the Token Test. *Brain*. 1966;89(4), 815-830.
- Bracken BA, McCallum RS. *Universal Nonverbal Intelligence Test*. Itasca, IL: Riverside. 1998.
- Conners CK. *Conners' Rating Scales-Revised (CRS-R)*. London, UK: Psychological Assessment Resources. 1996.
- Conners CK. *Conners' Attention Deficit Scale - Parent Version*. North Tonawanda, NY: Multi-Health Systems. 1999.
- Dollaghan C, Campbell TF. Nonword repetition and child language impairment. *J Speech Lang Hear Res*. 1998;41(5), 1136-1146.
- Dunn LM, Dunn LM. *Manual for the Peabody Picture Vocabulary Test Revised*. Circle Pines, MN: American Guidance Service. 1981.

- Dunn LM, Dunn LM. *Peabody Picture Vocabulary Test (PPVT-III)* (3rded.). Circle Pines, MN: American Guidance Service. 1997.
- Ellis Weismer S. Living on the edge of communicative competence: Processing capacity limitations. *Ann Meeting Am Speech-Lang-Hear Association*, Chicago, IL. 2008.
- Frederickson N, Frith U, Reason R. *Phonological Assessment Battery (Manual and Test Materials)*. Windsor, UK: NFER-Nelson. 1997.
- Gaulin CA, Campbell TF. Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Percept Motor Skill*. 1994;79(1 Pt 1), 55-64.
- Iliadou V, Fourakis M, Vakalos A, et al. Bi-syllabic, Modern Greek word lists for use in word recognition tests: Listas de palabras bisilábicas del Griego moderno para uso en pruebas de discriminación. *Int J Audiol*. 2006;45(2), 74-82.
- Katz J. The Use of Staggered Spondaic Words for Assessing the Integrity of the Central Auditory Nervous-System. *J Aud Res*. 1962;2, 327-337.
- Katz J. *Central Test Battery*. Vancouver, WA: Precision Acoustics. 2001.
- Keith RW. *SCAN-A: A Test for Auditory Processing Disorders in Adolescents and Adults*. San Antonio, TX: Psychological Corporation. 1994a.
- Keith RW. *The Auditory Continuous Performance Test*. San Antonio, TX: Psychological Corporation. 1994b.
- Keith RW. *SCAN-C: Test for Auditory Processing Disorders in Children Revised*. San Antonio, TX: Psychological Corporation. 2000a.
- Keith RW. *Random Gap Detection Test*, AUDITEC, St. Louis, MO. 2000b.
- Korkman M, Kirk U, Kemp S. *NEPSY. Neuropsychological Assessment of Children*. Oxford, UK: Pearson Assessment. 1997.
- Korkman M, Kirk U, Kemp, S. *NEPSY: A Developmental Neuropsychological Assessment*. San Antonio, TX: Psychological Corporation. 1998.
- MacLeod A, Summerfield Q. A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *Brit J Audiol*. 1990;24(1), 29-43.
- Musiek FE. Assessment of central auditory dysfunction: the dichotic digit test revisited. *Ear Hearing*. 1983;4(2), 79-83.
- Musiek FE, Collegly KM, Kibbe KS. et al. Proposed screening test for central auditory disorders: Follow-up on the Dichotic Digits Test. *Oto Neurotol*. 1991;12(2), 109-113.
- Olson RK. (1985). Disabled reading processes and cognitive profiles. In U. B. Gray, & J. Kavangh (Eds.), *Behavioural Measures of Dyslexia* (pp. 215-244). Parkton, MD: York Press.
- Orgass B, Poeck K. Ein neuer Aphasie-Test zur Diagnose von Sprachverständnisstörungen [A new aphasia-test for the diagnosis of speech comprehension disorders]. *Nervenarzt* 1966;37(3), 124-126.
- Pinheiro M. Tests of central auditory function in children with learning disabilities. In R. W. Keith (Ed.), *Central Auditory Dysfunction* (pp. 223-256). New York: Grune & Stratton. 1977.
- Pinheiro ML, Ptacek PH. Reversals in the perception of noise and tone patterns. *The J Acoust Soc Am*. 1971;49(6B), 1778-1782.

- Seashore CE, Lewis D, Saetveit JG. *Seashore Measures of Musical Talents*. New York: Psychological Corporation. 1960.
- Semel EM, Wiig EH, Secord W. *Clinical Evaluation of Language Fundamentals—Revised (CELF-R)*. New York: Psychological Corporation. 1980.
- Semel EM, Wiig EH, Secord W. *Clinical Evaluation of Language Fundamentals (CELF-4) Observational Rating Scale* (4th ed.). San Antonio, TX: Psychological Corporation, Harcourt Brace. 2003.
- Smoski WJ, Brunt MA, Tannahill JC. Listening characteristics of children with central auditory processing disorders. *Lang Speech Hear Ser.* 1992;23(2), 145-152.
- Torgesen JK, Wagner RK, Rashotte CA. *Test of Word Reading Efficiency*. London, UK: Psychological Corporation. 1999.
- Tzavaras A, Kaprinis G, Gatzoyas A. Literacy and hemispheric specialization for language: Digit dichotic listening in illiterates. *Neuropsychologia.* 1981;19(4), 565-570.
- Wagner R, Torgesen J, Rashotte, C. *Comprehensive Test of Phonological Processes (CTOPP)*. Austin, TX: Pro-Ed. 1999.
- Walton D, Brooks P. The Spoonerism Test. *Educational and Child Psychology.* 1995;12, 50-52.
- Wechsler D. *Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation. 1974
- Wechsler D. *Wechsler Intelligence Scale for Children (WISC-III)* (3rd ed.). San Antonio, TX: Psychological Corporation. 1991.
- Wechsler D. *Manual for the Wechsler Abbreviated Intelligence Scale (WASI)*. San Antonio, TX: Psychological Corporation. 1999.
- Wiederholt J, Bryant B. *Gray Oral Reading Test (GORT-4)* (4th ed.). Austin, TX: Pro-Ed. 2001.
- Wilkinson GS. *WRAT-3 Wide Range Achievement Test, administration manual*. Wilmington, DE: Wide Range. 1993.
- Willeford J. Assessing central auditory behavior in children: A test battery approach. In R. Keith (Ed.), *Central Auditory Dysfunction* (pp. 43-72). New York: Grune and Stratton. 1977.
- Williams J, Scott F, Stott C. et al. The CAST (Childhood Asperger Syndrome Test): test accuracy. *Autism.* 2005;9(1), 45-68.
- Woodcock RW, McGrew KS, Mather N. et al. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing. 2001.

4

The Dutch well-child language screening protocol at age 2 is valid to detect current and later language problems

Published as: Visser-Bochane MI, Luinge MR, Dieleman LE, van der Schans CP, Reijneveld SA. The Dutch well-child language screening protocol at age 2 is valid to detect current and later language problems. *Acta Paediatrica*, 2021;110(2):556-562

Abstract

Aim: Little is known about predictive validity of and professionals' adherence to language screening protocols. This study assessed the concurrent and predictive validity of the Dutch well child language screening protocol for two-year-old children and the effects of protocol deviations by professionals.

Methods: A prospective cohort study of 124 children recruited and tested between October 2013 and December 2015. Children were recruited from four well child clinics in urban and rural areas. To validate the screening, we assessed children's language ability with standardized language tests following the two-year screening and one year later. We assessed the concurrent and predictive validity of the screening and of protocol deviations.

Results: At two years, the sensitivity and specificity of the language-screening were 0.79 and 0.86, and at three years 0.82 and 0.74, respectively. Protocol deviations by professionals were rare (7%) and did not significantly affect the validity of the screening.

Conclusion: The language-screening protocol was valid for detecting current and later language problems. Deviations from the protocol by professionals were rare and did not affect the concurrent nor predictive validity of the protocol. The two-year language screening supports professionals working in preventive child health care and deserves wider implementation in well child care.

Introduction

Problems in speech and language are one of the most reported developmental problems in children with an estimated prevalence of 7% [1]. These problems may impact children's emotional functioning, academic success and social relationships [2-3] and early detection and subsequent treatment may significantly reduce their impact [4]. Therefore, the American Academy of Pediatrics has recommended developmental surveillance and regular screening of children in order to identify developmental problems early [5].

According to the American Academy of Pediatrics' developmental screening recommendations, well child care professionals, such as nurses and pediatricians, require valid language screening protocols. Some of these are specific for language, such as the MacArthur-Bates Communicative Development Inventory [6] and Language Development Survey [7]. There are also screening tools for the whole development of young children, such as the Ages and Stages Questionnaire [8], but, none of them include guidelines for the referral of children who failed the screening. Additionally, follow-up was not evaluated for any of the screening protocols. Evidence on the concurrent validity of such screening protocols is frequently reported and expressed in sensitivity and specificity in a research setting [9,10]. Sensitivity is the percentage of individuals with a health condition, in this case children, who experience speech and language problems and are correctly identified as having the condition [11]. Specificity is the percentage of healthy people who are correctly identified as not having the condition [11]. Concurrent validity regards the degree to which the results of the screening test are parallel with a reference standard that is administered at the same time. However, evidence is scarce for the performance of a screening instrument regarding health outcomes in the long term, the so called predictive validity (10). This lack of evidence implies that the efficacy of protocols to monitor the development of a child over time is not known. Moreover, there is insufficient evidence on the effect of protocol deviations of professionals in daily practice, a setting in which the sensitivity and specificity of the screening may differ from the research settings [12]. For example, a professional refers a child with a negative screening outcome because the professional has raised concerns regarding the language development of the child or is requested to do so by the parents. A clinical decision by the professional that deviates from the screening protocol

might improve or worsen the validity of the protocol. Evidence regarding the concurrent and predictive validity of a language screening protocol in routine practice is thus highly needed.

In the Netherlands, well child care professionals monitor the development of children using the Van Wiechenschema [13]. It is a brief schedule that aids a professional in monitoring a child's development regarding communication, fine motor skills, gross motor skills, adaptive behavior, social behavior, and personality from birth to the age of 54 months. The skills that the child is expected to have mastered at specific ages are elicited in 2.5- 3.0 minutes by using a small amount of materials. However, guidelines for referral are only available for children from birth to 12 months. For children aged 12 months and over, a referral depends on the interpretation of the assessor [14]. To assist assessors in the identification and referral of children with possible language problems, a practice-based language screening protocol was implemented in well child care by the Dutch preventive child healthcare services [15]. This protocol consists of tasks and questions to assess children's attainment of developmental language milestones and is performed by trained well child paediatricians and nurses. It provides a clear cut-off score with directions for referral in less than 10 minutes. However, the concurrent and predictive validity of this language screening protocol as well as the frequency and effect of deviating professionals' clinical decisions are not known. Evaluation of this protocol and the effect of protocol deviations provide insights that can contribute to the optimization of language screening in well child care in other countries. Therefore, the aim of this study was to assess the concurrent and predictive validity of the Dutch well child language screening protocol for two-year-old children and the effects of protocol deviations by professionals.

Methods

Sample and procedure

This was a prospective cohort study. We obtained the sample from four well child clinics that had implemented the screening protocol in their routine procedures. The clinics were selected from urban and rural areas in order to obtain a representative sample for the country as a whole. Well child paediatricians and nurses of these clinics asked all parents of children failing the screening to

participate in the study. For each child failing the screening that was recruited, a gender-matched child that passed the screening was recruited as well. All of the participating children had Dutch as their first language. The study was conducted between October 2013 and December 2015.

Protocol language screening at 2.0 years of age

The purpose of the two-year screening is to detect language delays in children by assessing the children's ability to use two-word sentences and comprehend basic, common nouns. It is administered by trained well child nurses and paediatricians and concerns the observation of two milestones, specifically, production of a two-word sentence and pointing out five body parts on a doll. If the professional cannot observe the desired behaviour during the well child visit, the professional asks the parent about this in a standardised manner. Both items are scored with a maximum of two points for a child's correct response [15]. There will be no referral if a child achieves four points. If a child scores less than four points, an additional question is asked with regard to the child's playing behavior. If the child plays together with the parent and can also play alone, one additional point is added. Any child with fewer than four points is referred. A total score of 0-1 points results in a referral to an audiological center for diagnostic assessment. A total score of 2-3 points results in a referral for guidance by a preventive speech language pathologist or well child nurse and a follow-up consultation is offered when the child is 2.5 years.

Procedure and measures

For the validation of the two-year language screening protocol, we assessed the language development of the child at age two years (concurrent validity) and at age three years (predictive validity). We used age appropriate standardized tests on language comprehension, word production, and sentence production. We defined atypical language as a deviant score on two or three tests or a moderate to severe deviation on one test. A reference standard was operationalised as follows: two or more test scores below minus one standard deviation of the norm score or one test score below minus one and a half standard deviation of the norm score resulted in atypical language. We used the Schlichting tests for Language Comprehension (SLC) [16], Word Production (SWP) [16] and Sentence Production (SSP) [17] as reference tests. These are language tests for

children from approximately two to seven years of age. The SLC is an 85-item test assessing comprehension of grammatical constructions using toys, pictures, and tokens. The SWP is a 70-item test to evaluate expressive vocabulary using a stimulus booklet with pictures. The SSP is a 40-item test to determine expressive grammatical constructions by using imitation of expressions visualized in a stimulus booklet with, in some cases, associated toys. Age-standardized scores for each test (mean = 100; standard deviation = 15) can be calculated according to the manuals in which entry levels per age and cut-off rules are also described [16,17]. The SLC, SWP and SSP have excellent internal consistency ($\lambda = 0.93, 0.89$ and 0.90 , respectively) and demonstrate significant association with subtests of the Dutch version of the Clinical Evaluation of Language Fundamentals ($0.63, 0.47$ and 0.59 respectively) [16,17].

A speech language pathologist (LD), who was blinded for the screening outcome, tested all of the children within four months, 90% within two months, of the two-year screening and one year later. Testing occurred during a home visit of, on average, two hours. Parents provided background information regarding birthweight and pregnancy duration.

The well child professionals reported the screening outcome and their clinical decision to the first author (MVB). We defined protocol deviations by professionals as the degree of adherence of the professional to the protocol and influence of deviations from the protocol on concurrent and predictive validity.

Analyses

First, we described the background characteristics of the sample. Next, we assessed the concurrent and predictive validity of the two-year language screening protocol and the effect of protocol deviations. We did so by calculating the sensitivity and specificity of the screening against the reference standard at the age of two years (concurrent validity) and at the age of three years (predictive validity). Next, we calculated the sensitivity and specificity of the clinical decision of the professional against the reference standard at the age of two years and at the age of three years. In addition, logistic regression analyses were performed with the reference standard at age two and at age three as the dependent variables and the screening result as the independent variable. These analyses were repeated using the clinical decision by the professional as the independent variable.

Ethics

The Medical Research Ethics Committee of Groningen (METc2013/103) approved the study and written, informed consent was obtained from parents or guardians of all children participating in this study.

Results

Sample

In total, 124 children participated in the study at the age of two years. This regarded 61 children that were identified by the language screening protocol and 63 children that were not identified by the screening protocol. In advance, 132 parents agreed to participate in the study. However, eight parents withdrew from the study before the first test moment. This regarded five withdrawals from parents with children that were identified by the screening and three withdrawals from parents with children that were not identified by the screening. One child from the group that was identified by the protocol was absent in the follow up, resulting in a sample of 123 children at the age of three years of which 78% were males. The children identified by the protocol did not significantly differ from the children that were not identified by it on age at the first test moment ($p = 0.85$) and birthweight ($p = 0.081$) (Table 1). Groups differed for length of pregnancy ($p = 0.021$) as all of the three children born before 36 weeks were in the group with identified problems.

Table 1. Characteristics of the sample at baseline for children identified, and not identified by the Dutch two-year language screening protocol, and total sample.

	Identified by screening protocol <i>n</i> = 61	Not identified by screening protocol <i>n</i> = 63	Total <i>n</i> = 124
Boys/girls (%)	50/11 (82/18)	47/16 (75/25)	97/27 (78/22)
Age at first test moment, months, mean (sd)	26 (1)	26 (1)	26 (1)
Birthweight, grams, mean (sd)	3300 (573)	3440 (514)	3370 (546)
Length pregnancy, weeks, mean (sd)	39 (2)	40 (2)	40 (2)

Concurrent and predictive validity of the two-year language screening protocol

A total number of 67 children (54% of the total sample) had atypical language at the age of two years according to the reference standard; 53 of these children (43% of the total sample) were identified by the screening protocol, and 14 children (11% of the total sample) were missed by the screening protocol (Table 2). One year after the screening, 50 children (41% of the total sample) had atypical language of which 41 children (33% of the total sample) were identified by the two-year screening protocol, and nine children (7% of the total sample) were missed (Table 2). The two-year screening compared to the reference standard at two years had a sensitivity of 0.79 and a specificity of 0.86 (concurrent validity). Regarding predictive validity, with reference testing one year later at age three, these values were 0.82 and 0.74, respectively (Table 3).

Table 2. Results of the two-year screening protocol and the speech language therapists’ (SLT) reference standards at age two years and at age three years. Percentages refer to total percentage of total sample of n = 124 at age two years and n = 123 at age three years.

		Identified by screening protocol n = 61	Not identified by screening protocol n = 63
Results of SLT reference standard at age two	Atypical language	53 (43%)	14 (11%)
	Typical language	8 (6%)	49 (40%)
Results of SLT reference standard at age three	Atypical language	41 (33%)	9 (7%)
	Typical language	19 (15%)	54 (44%)

Effect of protocol deviations

The professionals demonstrated strong adherence to the protocol. Their clinical decision deviated from it in only nine cases (7%). They did not refer seven children that were identified by the screening protocol and referred two children who were not identified by it (Figure 1). Deviations from the protocol had no added value regarding concurrent validity. We found minimal difference in specificity and sensitivity for the screening protocol compared to the clinical decision of the professional (Table 3). In accordance with findings for sensitivity and specificity, the logistic regression analyses also yielded quite similar findings for detection by

the screening and detection by the clinical decision of the professional, slightly in favour of the screening (Table 4)

Regarding predictive validity, one year after screening, again, we found hardly any difference in specificity and sensitivity for the screening protocol compared to the (deviations in) clinical decisions of the professional (Table 3). The odds ratios for detection by the screening protocol decreased from 23 to 13 and remained 15 for the clinical decision of the professional. Although all odds ratios differed significantly from 1, the 95% confidence intervals almost completely overlapped, indicating that the screening protocol and the professional's clinical decision had comparable predictive validity (Table 4). At an individual child level, the deviation of the professional was confirmed by the reference test after one year for six children and was not confirmed for three children.

Figure 1. Screening, referral, and outcomes on the reference standards at age two years and age three years of all children (n = 124). Dark grey refers to atypical language established by the reference tests; white refers to typical language development according to these tests.

*one child missing on follow up.

Table 3. Sensitivity, specificity for screening protocol with and without professionals' clinical decision and the reference standard at age two years (concurrent validity) and at age three years (predictive validity).

	two-year standard: protocol	two-year standard: protocol + professional	three-year standard: protocol	three-year standard: protocol + professional
sensitivity	0.79	0.72	0.82	0.80
specificity	0.86	0.86	0.74	0.79

Table 4. Results from univariate logistic regression analysis on atypical language development at age two years and at age three years.

Dependent: Language	Language problem at age two years Univariate (n=124)		Language problem at age three years Univariate (n=123)	
Covariables	OR (95%CI)	p-value	OR (95%CI)	p-value
2-y protocol	23.19 (8.95-60.05)	<0.01	12.95 (5.31-31.56)	<0.01
2-y protocol + professional	15.47 (6.19-38.71)	<0.01	15.47 (6.32-37.89)	<0.01

Discussion

The aim of this study was to assess the concurrent and predictive validity of the Dutch well child language screening protocol for two-year-old children and the effects of protocol deviations by professionals. We found that the two-year language screening protocol had good concurrent and predictive validity. The professionals showed strong adherence to the protocol, and any deviations from it that were made by the professionals did not improve its sensitivity and specificity.

We found that the two-year language screening protocol had good concurrent validity, a sensitivity of 0.79; and a specificity of 0.86, confirming previous studies concluding that atypical language can be identified by using a combination of milestones at the age of two [18-19]. The two language milestones and one milestone regarding social interaction in play in the protocol thus seem to be adequate for a screening in well child care. The first language milestone in the protocol, says two-word sentences, aligns with the international evidence that this milestone is strongly related to language problems [18-20].

The second language milestone in the protocol, pointing out five body parts on a doll, accords with the red flag for immediate referral for evaluation at age two years, does not point to pictures or body parts when named [21]. The finding on the protocol including a third milestone regarding the child's playing behaviour, concerning social interaction, confirms that interaction is part of language development [22]. These three milestones are thus pivotal for language screening at age two years.

We further found a high predictive validity of the screening, in particular a sensitivity of 0.82 and a specificity of 0.74 after one year, confirming the few studies that reported that results of language screening at age two years indeed predicts later language status [10]. However, in this meta-analysis of Sim et al, the strongest overall predictive validity was reported for parent-report measures compared to direct-child-assessment, whereas we found high predictive validity for a screening based on direct-child-assessment in combination with parent report [10]. Reported predictive sensitivity and specificity for the MacArthur Bates Communicative Development Inventories one year after the screening was 0.61 and 0.94 respectively [23]. For the Language Development Survey, the predictive sensitivity and specificity was 0.67 and 0.96 respectively, at an average of 23 days after the screening [20]. Both parent-report instruments showed high specificity with a moderate sensitivity. Our protocol had poorer specificity than the CDI and LDS but had better sensitivity one year after screening, indicating that our protocol showed a higher proportion of correct detections and had more over-referrals than did the LDS and CDI. With our study we provide additional evidence for good predictive validity of language screening at two years based on a screening protocol performed by professionals. However, higher predictive specificity is desirable to minimise over-referrals.

We found that the professionals showed strong adherence to the protocol and that deviations from it could not improve the sensitivity and specificity at ages two and three; but, a higher predictive specificity is desirable for community-based screening. Such an improvement might be realized in two ways. A first option would be to add an extra step to the screening protocol that might improve its performance. Such an additional step could be to include the assessment of risk indicators, such as parental concerns or a family history of language problems [12,24] or to include a more extensive language screening

with high specificity for flagged children. In the Netherlands, that could be a first-stage diagnostic instrument, such as the Language Standard [25]. Specificity should exceed 0.90 in order to make the protocol suitable for population based screening as this minimises over-referrals with its negative effects, e.g. discomfort for parents and children as well as costs [26]. A second option to improve specificity could be to extensively train professionals and increase the amount of time that they have for screening. Costs for training as well as time and implementation of a protocol in the workflow are barriers for successful screening that were noted in earlier studies [27,28]. The professionals participating in our study were already trained, and the protocol was implemented in the workflow which contributed to the success of our two-year screening protocol. Further research will be required to determine whether an extra step could improve sensitivity and specificity of the language screening protocol.

Our study had a number of major strengths, in particular that the actual referral by the professional could be assessed which provided insights into the validity of the protocol in routine practice. Also, the same extensive language tests were used in a longitudinal design, providing a sound reference standard for validation. Last, we achieved almost 100% retention, highly restricting potential bias in our study. A limitation of our study was that the sample was too small to differentiate the referred group into referral for further assessment or referral to guidance groups. This limited the potential to determine sensitivity and specificity for these graded referral rules, therefore, overall results should be interpreted with caution. Moreover, sensitivity and specificity were calculated within a predetermined sample and not one that was population based. Therefore, the positive and negative predictive values could not be estimated. Another limitation was that the well child professionals were not asked what, according to them, could improve sensitivity and specificity of this screening. Last, the protocol included a parent report in the event that professionals could not observe the child's ability regarding two-word sentences and pointing out body parts. The reliability of the answers of the parents was not investigated, but, parent report is generally accepted as reliable for assessing language abilities in young children [20]. Therefore, a screening protocol based on professional observation supplemented with parent report, if needed, appears to provide a valid overview of a young child's language abilities.

Multiple studies have reported the positive effect of interventions to support the communication skills of children [6,29]. One way to promote language development at an early age could be to deliver the Hanen Program in which certified speech therapists assist parents in stimulating the speech and language development of their child [30]. The potential benefits of adding this intervention to the protocol should be investigated. Moreover, several studies show relationships between the development of children on domains, such as language, feeding, psychosocial and motor development. Future research could relate the outcomes of children on the language protocol to outcomes of assessments on these other developmental domains.

Conclusion

Our findings on concurrent and predictive sensitivity and specificity indicate that the two-year screening protocol was valid for the detection of language problems in children. However, the specificity should be further improved for population based screening. Professionals demonstrated strong adherence to the protocol and any deviations from it made by the professionals did not improve its sensitivity and specificity. The protocol can provide support to professionals working in preventive child health care and thus deserves wider implementation in well child care.

Acknowledgements

The authors thank participating well child professionals, parents, and children.

Funding:

This work was supported by the Netherlands Organization for Health Research and Development (ZonMw; grant number 200330002).

References

1. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear R.* 1997;40(6):1245-1260.
2. Conti-Ramsden G, Mok PL, Pickles A, Durkin K. Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Res Dev Disabil.* 2013;34(11):4161-4169.
3. Snowling MJ, Duff FJ, Nash HM, Hulme C. Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *J Child Psychol Psychiatry.* 2016;57(12):1360-1369.
4. Broomfield J, Dodd B. 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.* 2011;46(6):628-640.
5. Bright Futures Steering Committee, & Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics.* 2006;118(1):405-420.
6. Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS. *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual.* PB Brookes 2014.
7. Rescorla L. The Language Development Survey: A screening tool for delayed language in toddlers. *J Speech Hear Disord.* 1989;54:587-599.
8. Bricker D, Squires J, Mounts L, Potter L, Nickel R, Twombly E, Farrell, J. *Ages and stages questionnaire.* Baltimore: Paul H. Brookes, 1999.
9. Wallace IF, Berkman ND, Watson LR, et al. Screening for speech and language delay in children 5 years old and younger: a systematic review. *Pediatrics.* 2015;136(2):e448-e462.
10. Sim F, Thompson L, Marryat L, Ramparsad N, Wilson P. Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISMA systematic review. *PLoS one.* 2019;14(2):e0211409.
11. Rothman KJ, Greenland S, Lash, T. L. (Eds.). *Modern epidemiology.* Lippincott Williams & Wilkins, 2008
12. Sheldrick RC, Frenette E, Vera JD, Mackie TI, Martinez-Pedraza F, Hoch N, Eisenhower A, Fettig A, Carter AS. What Drives Detection and Diagnosis of Autism Spectrum Disorder? Looking Under the Hood of a Multi-stage Screening Process in Early Intervention. *J Autism Dev Disord.* 2019;49:2304-2319 doi.org/10.1007/s10803-019-03913-5
13. Laurent de Angulo MS, Brouwers-de Jong EA., Bijlsma-Schlösser JFM, Bulk-Bunschoten AMW, Pauwels JH, Steinbuch-Linstra I. Ontwikkelingsonderzoek in de jeugdgezondheidszorg; het Van Wiechenonderzoek—De Baecke-Fassaert Motoriektest. Assen: *Van Gorcum*, 2005.
14. Boere-Boonekamp MM, Dusseldorp E, Verkerk PH. Onderbouwing van de validiteit van het ontwikkelingsonderzoek bij kinderen van 0 tot en met 4 jaar: het van Wiechenonderzoek. [Validity of the Van Wiechen developmental assessment for children aged 0 to 4 years] 2009.

15. Carmiggelt EC, Uilenburg NN, Romeijn JE, Stam-van den Doel HH, Pijpers FIM. Handreiking uniforme signalering van taalachterstanden bij jonge kinderen. Nederlands Centrum Jeugdgezondheid. [Guideline Uniform Screening for language delays in young children. Dutch Preventive Child Health Services Centre] Utrecht, 2013.
16. Schlichting JEPT, Spelberg HC. Schlichting test voor taalbegrip [Schlichting test for language comprehension]. Houten: Bohn Stafleu van Loghum 2010.
17. Schlichting JEPT, Spelberg HC. Schlichting Test voor Taalproductie-II: voor Nederland en Vlaanderen [Schlichting test for Language Production]. Houten: Bohn Stafleu van Loghum, 2012.
18. Diepeveen FB, Dusseldorp E, Bol GW, Oudesluys-Murphy AM, Verkerk PH. Failure to meet language milestones at two years of age is predictive of specific language impairment. *Acta Paediatr.* 2015;105(3), 304-310.
19. Sheldrick RC, Perrin EC. Evidence-based milestones for surveillance of cognitive, language, and motor development. *Acad Pediatr.* 2013;13(6):577-586.
20. Rescorla L, Alley A. Validation of the Language Development Survey (LDS): A Parent Report Tool for Identifying Language Delay in Toddlers. *J Speech Lang Hear Res.* 2001;44:434-445.
21. McLaughlin MR. Speech and language delay in children. *Am Fam Physician.* 2011;83:1183-8.
22. Visser M, Bochane MI, Reijneveld SA, Krijnen, WP, van der Schans CP, Luinge MR. Identifying Milestones in Language Development for Young Children Ages 1-6 Years. *Acad. Pediatr.* 2020;20(3):421-429.
23. Sachse S, Von Suchodoletz W. Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr.* 2008;29(1):34-41.
24. Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T. CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one.* 2016;11(7), e0158753.
25. Slofstra-Bremer CF, Meulen S van der, Lutje Spelberg HC. De Taalstandaard [The Language standard]. Amsterdam: Pearson, 2006.
26. Wolff MS de, Theunissen MH, Vogels AG, Reijneveld SA. Three questionnaires to detect psychosocial problems in toddlers: A comparison of the BITSEA, ASQ: SE, and KIPPI. *Acad Pediatr.* 2013;13(6):587-592.
27. Gellasch P. Developmental screening in the primary care setting: A qualitative integrative review for nurses. *J Pediatr nurs.* 2015;31:159-171.
28. Morelli DL, Pati S, Butler A, Blum NJ, Gerdes M, Pinto-Martin J, Guevara JP. Challenges to implementation of developmental screening in urban primary care: a mixed methods study. *BMC Pediatr.* 2014;14:16.
29. Girolametto L, Pearce P, Weitzman E. The effects of focused stimulation for promoting vocabulary in children with delays: A pilot study. *J of Child Commun Dev.* 1996;17:39-49.
30. Pepper J, Weitzman E. *It Takes Two to Talk®: A practical guide for parents of children with language delays* (2nd ed.). Toronto: The Hanen Centre, 2004.

The background is a light gray abstract design featuring various letters in different fonts (serif, sans-serif, script) and shapes (circles, wavy lines, organic forms) scattered across the page. The letters include W, p, C, K, N, B, G, t, J, h, F, B, I, O, y, U, z, s, Z, a, m, P, w, e, A, o, S, and D.

5

Identifying milestones in language
development for young children ages
1-6 years

Published as: Visser-Bochane MI, Reijneveld
SA, Krijnen, WP, van der Schans CP, Luinge
MR. Identifying milestones in language
development for young children ages 1-6
years. *Academic Pediatrics*,
2020;20(3):421-429

Abstract

Objective: Language development is important for children's success in life. Therefore, language is monitored by child health care professionals and parents, but a uniform set of milestones in language development is lacking. Our aim was to identify a set of clear and distinctive milestones that empirically reflect language development in children aged one to six years of age.

Methods: We obtained a community-based sample of 1,381 parents reporting on milestones derived from clinical signs and currently used language screening instruments. We used nonparametric Item Response Theory analysis to identify milestones belonging to one unidimensional scale.

Results: Twenty-six milestones were excellently scalable (item H coefficients 0.62-0.90) and formed a strong scale (total H coefficient 0.83). The final set of identified milestones covered vocabulary, grammar, and communication, with an item ordering that holds for all children.

Conclusion: This unidimensional set of 26 clear and distinctive milestones reflects language development in young children and can be used as instrument to monitor language development.

Introduction

Language development is a basic skill for all children, and essential for participation in everyday life, more specifically: a child's social and emotional development, and educational success [1-3]. Atypical language development has a major impact on a child and its environment, and needs to be identified as early as possible [4]. Therefore, language development of children is monitored with various instruments by professionals working in child health care and professionals working in educational settings, and by parents [4-7]. In order to adequately identify atypical language development, clear milestones of typical language development are required. Various milestones, e.g. the well-known 'puts two words together at age two', are used for early assessment. However, empirical evidence lacks on the best milestones for identification of language development from age 1 to 6 years.

Milestones in language development should represent the underlying construct of language development, including its intertwined aspects such as speech, language, and communication. Speech involves the production of speech sounds. Children start babbling from the age of 6 months, next speech sounds develop until the age of 7 years [8]. Language involves the comprehension and production of words, sentences, and ideas. The first words emerge when a child is about 12 months, up from that age vocabulary increases [9]. Next, grammar development starts when children put two words together, typically occurring around a child's second birthday [10,11]. Thereafter, various elements of grammar, such as syntax and morphology develop [12]. Communication involves nonverbal and verbal use of language in interaction with other persons, and is seen as a necessary condition for language development [13]. Up till now, it is unclear whether milestones in speech, language, and communication reflect the same underlying construct, that is language development.

Children develop language with great variability in the onset time and rate of vocabulary, and grammar, e.g. mean length of utterance, proportion of verbs in total vocabulary, and onset of use of irregular nouns and verbs [12,14,15]. Because of the variation in onset time and rate, it is difficult to describe average language development that holds for all children [14,15]. However, it may be possible to identify more general milestones in speech, language, and

communication that have the same order for all children. To our knowledge, two previous studies identified a general set of ordered language milestones [11,16]. These milestones covered language in different ways, namely, in the study of Luinge et al [16] as vocabulary, grammar, and speech, but not communication, and in the study of Sheldrick et al [11] as a broader construct than language, namely vocabulary, grammar, and communication but not speech, and motor and cognitive development in addition. Also, milestones in language development are used in developmental screening questionnaires, such as the widely implemented Ages and Stages Questionnaire (ASQ) [17]. In contrast with Luinge and Sheldrick, constructors of the ASQ did not identify ordered milestones for the whole continuum of language development. Instead, the ASQ includes six milestones on language development that partially vary for 21 specific age windows (ages 1 to 5.5 years). These milestones regard comprehension and production of words and sentences, also grammar, but not speech and communication (use of language in interaction). Further research is necessary to find out what general milestones reflect the construct of language development.

Furthermore, such general milestones should be feasible for parents and professionals in well-child care and education in order to monitor language development. Therefore, we need to identify clear milestones that can be observed by parents and professionals. A set of clear distinctive milestones that have the same order for all children can contribute to the empirical description of typical language development, and to the monitoring of language development by parents and by professionals working in community-based settings. Therefore, the aim of our study was to identify a set of clear distinctive milestones that empirically reflect language development in children from one to six years of age.

Methods

Study Design

We conducted a cross-sectional study to identify a set of milestones (items) that reflect language development in children from 1 to 6 years of age, and to assess the 1) feasibility, concerning use of the items for a certain age range; 2) the comprehensibility of the items for parents; and 3) the construct validity of the resulting set. The construct validity of the set was evaluated by interpreting the results of Mokken scale analysis that builds upon nonparametric Item Response

Theory. The consecutive steps for the identification of the set of milestones that reflect language development are shown in Figure 1.

The study was approved by the Medical Ethical Committee of the University Medical Center of Groningen, Netherlands (number NL45253.042.13), and registered at trialregister.nl (number 5746).

Figure 1. Consecutive steps in the process of identification of milestones.

Statistical approach – Mokken scaling Procedure

A way to identify a set of clear distinctive items that reflect an underlying ability such as language development is Mokken scaling [18], which belongs to nonparametric Item Response Theory (IRT). The identified items as identified in a Mokken scaling procedure belong to one underlying ability and are individually

distinctive for a specific range of the ability (Figure 2). The advantage of Mokken analysis is that items can be identified that contribute to the rank ordering of persons on the same underlying ability [19], where the identified items reflect the development of language. Mokken analysis is not the only IRT model. Other frequently used models are parametric IRT-models, such as the Rasch model. Rasch analysis provides logistic item characteristic information, highly similar to Mokken. Moreover, the basic IRT assumptions are similar for Mokken and Rasch. However, the Rasch framework does not include an automated item selection procedure or a general procedure to check invariant item ordering. It consists of an iterative estimation procedure to fit the logistic item characteristic curves. This works only fine for a selected set of items sharing the same measurement properties, making the Rasch model unsuited for the selection of items.

Figure 2. Mokken scale model for language development. An ordering of distinctive items (milestones) that hold for all children reflecting the same underlying language development.

Sample

We used two-step sampling to recruit parents/caregivers of children from one to six years of age. First, we took a random selection of municipalities. Second, from these, we randomly selected daycare centers (to include the parents of 1-2-3 year old children) and kindergartens (to invite parents of 4-5 year old children). We also recruited well-child clinics throughout the Netherlands (1-2-3 year olds). By recruiting children via daycare centers and well-child clinics for the 1-2-3 year olds, and kindergartens for the 4-5 year olds, we obtained a community-based sample representative for the general population in the Netherlands. There were no exclusion criteria. Parents provided written informed consent.

Measures and Procedure

We constructed parent questionnaires for the ages 1-2, >2-3, >3-4, >4-5, >5-6 years to describe language development based on items derived from various sources. Candidate items were derived from clinical signs for atypical language development [20], and from all currently used Dutch screening instruments: Van Wiechen [21]; Speech and Language Norms for Primary Healthcare (SNEL) [16]; the questionnaire on speech and language problems from the Dutch national association for the education of children with speech and language problems retrieved from <http://www.simea.nl>; and from a Dutch well-child clinic retrieved from <https://www.ggddrenthe.nl/scholen/logopedie>. During the data collection, all items were phrased in Dutch as yes/no questions on language ability and, if applicable, supplemented with an example. Face validity (item is specific, measurable, relevant, clear for parents) of the items was discussed by eight professionals in child language, resulting in an initial set of 75 items in total. The items were administered in Dutch. We used forward-only translation of the original items to English for international publication. Development of the instrument was finalized by dividing the items over five partially overlapping questionnaires, each of which pertained to one age group (for example: 'Does your child ask questions with 'why'?' was asked in the age groups 3, 4, and 5-year olds but not in the age groups 1, and 2 year olds).

We measured feasibility as whether the items were appropriate for a certain age range according to the response of parents on the items and whether all of the items were appropriate for use in daily practice. Comprehensibility was measured by transcribing parents' reactions to the items other than a 'yes' or 'no' answer, and asking them whether they actually understood each of them.

We further measured background characteristics of the child and family. These concerned age, gender, birthweight, length of pregnancy, language situation at home, and highest level of education achieved by the mother. Frequently reported risk factors for language disorders are gender, birthweight, length of pregnancy, educational level of the mother [5]. Therefore, we have included these factors as background variables. Education level was classified into three categories: low (primary school or less, and pre-vocational education), middle (secondary education), and high (higher vocational education and university).

We collected data on language and background characteristics from the parents using a telephone interview of, on average, 20 minutes that was conducted by speech and language pathology students who were specifically selected and trained for this study.

Analyses

First, we described the background characteristics of the sample. Second, we assessed feasibility by calculating the proportion of 'yes' and 'no' answers of parents on the items for the entire group. Items that answered 100% 'yes' were removed. Third, we assessed the comprehensibility of the items. Parental responses on the items, other than 'yes' or 'no', were transcribed and coded followed by group discussions per item by three examiners regarding specificity and clarity for parents. Items that appeared multi-interpretable were then labeled as not meeting the criteria (i.e. not specific or not clear for parents) and were therefore removed from the item set.

Fourth, we identified items and investigated the construct validity of the identified items by comparing the results of Mokken scale analysis with the assumptions of the Mokken model, viz scalability, monotonicity and item ordering. For the missing observations, we applied a substitution by the item

mean rounded to the nearest integer. Non-discriminating items with proportions equal to one were omitted from further analysis. We then applied Mokken scale analysis to identify items and evaluate the (1) scalability; (2) monotonicity; and (3) item ordering based on the responses on the items by the parents. We assessed scalability by evaluating unidimensionality and homogeneity. Unidimensionality was investigated using an automated item selection procedure (AISP), taking a lower bound of 0.3 for homogeneity and significance level α 0.05 [22,23]. Homogeneity was quantified with Loevinger's coefficient for the total scale (H) and each item (H_i). The scale coefficient H is a weighted mean of the quality of the items (H_i s). We adopted the generally accepted interpretation on scalability as: $0.3 \leq H < 0.4$ weak, $0.4 \leq H < 0.5$ moderate, and $0.5 \leq H < 1.0$ strong [22]. (2) We assessed monotonicity by evaluating the item response function (IRF). A monotonically non-decreasing IRF indicated a positive relation of an item to the latent trait of language development. Items violating monotonicity were removed. We evaluated (3) item ordering for all pairs of items with a backward selection procedure [22]. From the item pair with the largest number of ordering violations, we removed the item with the lowest H coefficient. The procedure identifies a set of items without such violations. All items were plotted (proportions of children with a 'yes'-score for each item against total scores on the final set of milestones), to visualize the relationship between the indicator as a function of language development. The steepness of the plotted curves indicates the distinctiveness of an item, with steeper slopes indicating better discrimination. Moreover, the set of items should cover the developmental period. Last, we determined the age of acquisition for each item based on fitted curves, for percentages of the children passing of 15%, 50% and 85%. Data were analyzed using R version 3.3.1 using the Mokken-package version 2.7.7 [22].

Results

Sample

We collected data on 1,381 children attending 146 daycare centers, 128 kindergartens, and 10 well-child clinics. The first 150 parents of children from one to six years of age (73 boys, 77 girls) were assigned to the sample for feasibility analysis. Next, a sample consisting of 1,231 parents of children ranging from one to six years of age was used in the analysis of the comprehensibility, and construct validity. Characteristics of our sample are presented in Table 1. The sample was representative for the Dutch population with respect to birthweight, pregnancy duration, and educational level of the mother, three factors closely related to the outcomes. For comparison: 6.2% of the Dutch children is born with a birthweight below 2500 grams, 7.2% of the children is born preterm, and 56% of the females (25-35 years old) has a higher educational level, vs 4.4% low birthweight and 5.9 % preterm (Volksgezondheidszorg.info, 2019), and 60% high educational level (Statistics Netherlands, 2019) in our sample. Also, age groups did not differ on birthweight and pregnancy duration. However, one way analysis of variance showed that age groups slightly differed regarding educational level [$F(4,1223) = 2.70, p = 0.029$]. Specifically, two year olds had the highest percentage of higher educated mothers (64.2%), and three year olds had the lowest percentage of higher educated mothers (53.3%).

Feasibility and comprehensibility

We removed 22 out of 75 items for the following reasons: 100% yes-responses (100% no-responses did not occur), i.e. the item measures an ability that all children already have acquired (eight items), the item has no learning curve and is thereby not specific for an age group (eight items), the items have duplicate meaning (four items), and the item did not fit the construct (two items). The remaining items were distributed over five overlapping questionnaires corresponding to the five age groups. Because some items were assigned to other age groups, we began a new sample for analysis of comprehensibility and psychometric properties. The analysis of the comprehensibility resulted in the exclusion of five items (Supplemental Table A) because these items were not clear to parents.

Table 1. Characteristics of the sample used for the identification of the milestones.

	12 - 23 months (n=215)	24 - 35 months (n=303)	36 - 47 months (n=263)	48 - 59 months (n=239)	60 - 72 months (n=211)	Total (n=1,231)
Characteristics	n (%)	n (%)	n (%)	n (%)	n (%)	
Gender						
Male	110 (51.2)	160 (52.8)	143 (54.4)	115 (48.1)	107 (50.7)	635 (51.6)
Female	105 (48.8)	143 (47.2)	120 (45.6)	124 (51.9)	104 (49.3)	596 (48.4)
Birthweight (grams)*						
High (>5000)	0 (0.0)	1 (0.3)	1 (0.4)	2 (0.8)	1 (0.5)	5 (0.4)
Average (2500-5000)	206 (98.1)	276 (93.6)	242 (94.2)	225 (95.4)	196 (95.6)	1145 (95.2)
Low (<2500)	4 (1.9)	18 (6.1)	14 (5.4)	9 (3.8)	8 (3.9)	53 (4.4)
Pregnancy duration						
Serotine (>42 weeks)	11 (5.1)	17 (5.6)	13 (4.9)	17 (7.1)	12 (5.7)	70 (5.7)
Term (37-42 weeks)	198 (92.1)	260 (85.8)	229 (87.1)	208 (87.0)	193 (91.5)	1088 (88.4)
Preterm (<37 weeks)	6 (2.8)	26 (8.6)	21 (8.0)	14 (5.9)	6 (2.8)	73 (5.9)
Language						
Monolingual	183 (85.1)	249 (82.2)	224 (85.2)	210 (87.9)	180 (85.3)	1046 (85.0)
Bi-, Multilingual	32 (14.9)	54 (17.8)	39 (14.8)	29 (12.1)	31 (14.7)	185 (15.0)
Education of mother*						
High	135 (62.8)	194 (64.2)	139 (53.3)	145 (60.6)	129 (61.1)	1228 (60.4)
Middle	73 (34.0)	92 (30.5)	99 (37.9)	79 (33.1)	73 (34.6)	416 (33.9)
Low	7 (3.2)	16 (5.3)	23 (8.8)	15 (6.3)	9 (4.3)	70 (5.7)

* Numbers do not always add up to N = 1,231 due to missing data; Birthweight (n = 28), Education of mother (n = 3).

Construct validity

The responses on the remaining 48 items were generally of good quality. On the 98 (0.17%) missing observations on items over all of the respondents, we applied a substitution to the nearest integer of the column mean. Five items were excluded because they appeared to be too easy, having an item proportion of endorsement of .995 or higher (Supplemental Table A).

Testing scalability (unidimensionality) on the remaining 43 items with automated item selection procedure (AISP) resulted in the selection of 40 items. Three items were unscalable, meaning that they did not fit the dimension of the selected 40 items, and had to be excluded (Supplemental Table 1). The set of 40 items had a total H coefficient of 0.77. The items showed strong scalability with item coefficients H_i ranging from 0.51 to 0.91. The lowest item H coefficient (0.51) regarded the item *'Is the pronunciation of your child (with exception of the letter r) good?'*, and the highest H coefficient (0.91) regarded the item *'Does your child understand tasks consisting of two words? For example: 'coat on' or 'look there.'*

The assessment of monotonicity revealed no violation by any of the items, therefore, no items were excluded. That implied that the probability of a positive response for each item increased with a more extensive language ability of the child. Investigating invariant item ordering by backward selection resulted in 15 steps and thereby the exclusion of 14 items (Supplemental Table A). The frequencies of violations are reduced by subsequent removal of the worst item until no violations on item ordering were found. This procedure resulted in a set of items where all of the children have the same ordering of items.

Final set of items

The final set of items consisted of 26 items regarding milestones in vocabulary, grammar, and communication, reflecting language development covering the age range from one to six. The items were ordered covering language development (expressed in the total score on the final set of items), with item curves that increased (ascending slopes) at increasing levels of language development (Figure 3). The item H coefficients were between 0.62 and 0.90 which indicated excellent scalability (Supplemental Table B). The scale H coefficient was 0.83 (SE=0.01) which indicated a strong scale. Items proportion of endorsement

ranged from 0.36 to 0.99, i.e., 36 to 99% of the full sample of children had acquired the language ability that was involved. A percentage of 99% indicates that the corresponding item does not discriminate well over all age groups, but rather only in the younger age groups. The age of acquisition for each item based on fitted curves showed increasing ages of acquisition for the various items. Relatively many items were easy, i.e. feasible for younger children, and fewer were difficult, i.e. only suitable for older children. Only four items (item 23-26) covered the full language development of children aged 40 to 71 months (Figure 4).

Figure 3. The fitted proportion of positive responses of the parents on each item (y-axis) by the total score on the final set of 26 items, i.e. increasing language development (x-axis). Each curve represents a single item.

Figure 4. Ages at which each of the 26 milestones is achieved for 15% (light grey), 50% (grey), and 85% (dark grey) of the children.

Discussion

In this study, we identified 26 milestones in one dimension that reflect typical language development in children from one to six years of age. All milestones belonged to one, unidimensional, scale that covers vocabulary, grammar, and communication, with an item ordering that holds for all children. All identified milestones were feasible for the age range (1-6) and were comprehensible for parents.

The identified milestones reflect language development in children from one to six years of age, belonging to one scale, covering vocabulary, grammar,

and communication, and that does not include the production of speech sounds. To our knowledge, only two previous studies also identified a set of milestones in language development on a unidimensional scale [11,16], but none of these covered vocabulary, grammar, and communication in a concise way. The removal of all milestones on the production of speech sounds is consistent with the idea that speech and language should be considered to be different dimensions [24]. Our study empirically demonstrates that milestones in vocabulary, grammar, and communication form a unidimensional scale that reflects language development in children aged one to six, and that production of speech sounds is not part of that unidimensional scale for children aged one to six.

We found an ordering of milestones that holds for all included children, with sentence comprehension (item 3,4,5) preceding sentence production (e.g. item 11,15). This ordering aligns with the current knowledge on language acquisition that comprehension precedes production [9]. It has also been suggested that the inverse pattern could occur in some children or for detailed linguistic forms, e.g. object pronouns, subject-verb-object word order [25]. Our findings suggest that these deviations at least do not occur if assessing a general level of comprehension or production using parental report. Only such more general milestones are more useful for community-based use as detailed linguistic forms require professional knowledge about linguistics and experimental assessment strategies. Our study thus shows that on a more general level that parents perceive that language comprehension indeed precedes language production.

All milestones in the final set were comprehensible for parents supporting the previous conclusion that language development can be monitored by parents when using easily observable items [4,11,16,26,27]. The easiness of observation can be attributed to our process of item formulation that resulted in comprehensible, and positively formulated milestones in language development. We used no negatively formulated milestones, which may have minimized response bias as reversing some milestones has been shown to result in errors due to inattention or confusion by the respondent [28]. Moreover, milestones did not focus on deficits which could raise unnecessary concerns. It is thus feasible to measure language development in children from one to six years of age derived from parental observation.

In contrast with other studies [29,30], we found that items on comprehension (item 3,4,5) could be quantified based on parent report. In our study, the included comprehension items consist of short tasks such as ‘Does your child understand tasks consisting of two words? For example: ‘coat on’ or ‘look there’’. Parents can judge whether their child understands this, because they can observe the current response behavior of the child [31]. Other studies assess the comprehension based on a checklist with words checked by the parents, e.g. the MacArthur Bates Communicative Development Inventory Words and Gestures (CDI-WG) [32] and the Language Development Survey (LDS) [33]. The validity of such checklists with words is questioned, because it is difficult for parents to judge whether a child really understands a word. For example, the words in the checklists may be used at home, but that does not necessarily mean the child comprehends these words [30]. Yet, assessing comprehension is important as deficits in language comprehension are a clear indicator for persisting language disorders [9,34]. Our study shows that language comprehension can be measured based on parent-report regarding the performance of short tasks by their child that show that the child understands the task.

The scale has been developed in Dutch, but is likely to be of use in other Germanic languages, such as English and German, as the items represent milestones in language that are very common in related languages and are not restricted to the Dutch language, e.g. ‘says two-word sentences’ (item 11). Use of the scale in other languages requires a language specific verification of the presented norms.

A major strength of our study is our use of a large, community-based sample. In addition, we had a very small number of missing values, and prevented information bias by using trained interviewers. Our study also had some limitations, the most important one being that high educated mothers were slightly overrepresented in our sample for two-year old children, i.e., 64% compared to the national average being approximately 56% (Statistics Netherlands, 2019). This may have slightly affected the estimated ages of acquisition as children of these parents have been shown to have slightly faster language development [35].

Conclusions

In conclusion, our findings show that language development in children aged one to six can be measured with one scale. Our final set of empirical milestones encompasses milestones in vocabulary, grammar, and communication and reflects increases in one underlying construct of language development in children aged one to six. The properties of this scale, i.e., consisting of parent-reported, clear and distinctive milestones, make the set feasible as well as promising for monitoring language development in community-based settings. The scale was constructed for quick community based developmental screening. The ages of attainment can be used as indicative norms of typical language development during well child visits. However, cut-off norms for screening purposes can best be determined against a gold standard. Our study has several implications for further research. First, the final scale should be validated in order to establish its test characteristics in terms of sensitivity and specificity, and predictive values in routine care. Second, the use of the scale in community settings deserves further study regarding feasibility and added value in routine care. Last, further research is needed to determine whether this unidimensional scale can also be used in other languages.

Acknowledgements

The authors thank the participating teachers, health nurses, speech language therapy students, and parents.

Financial Disclosure: This work was supported by the Netherlands Organization for Health Research and Development (ZonMw; grant number 200330002).

References

1. Boyle CA, Decoufle P, Yeargin-Allsopp M. Prevalence and health impact of developmental disabilities in US children. *Pediatrics*. 1994;93(3):399-403.
2. Conti-Ramsden G, Durkin K, Mok PL, Toseeb U, Botting N. Health, employment and relationships: Correlates of personal wellbeing in young adults with and without a history of childhood language impairment. *Soc Sci Med*. 2016;160:20-28.
3. Roulstone S, Law J, Rush R, Clegg J, Peters T. Investigating the role of language in children's early educational outcomes. Research Report on Department for Education, UK; 2011.
4. Bright Futures Steering Committee, & Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*. 2006;118(1):405-420.
5. Berkman ND, Wallace I, Watson L, Coyne-Beasley T, Cullen K, Wood C, Lohr KN. Screening for speech and language delays and disorders in children age 5 years or younger. A Systematic Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 120. AHRQ Publication No. 13-05197-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
6. Williams C. Teacher judgements of the language skills of children in the early years of schooling. *Child Lang Teach Ther*. 2006;22(2):135-154.
7. Glascoe FP, Dworkin PH. The role of parents in the detection of developmental and behavioral problems. *Pediatrics*. 1995;95(6):829-836.
8. Dodd B. *Differential diagnosis and treatment of children with speech disorder*. John Wiley & Sons; 2013.
9. Simonsen HG, Kristoffersen KE, Bleses D, Wehberg S, Jørgensen RN. The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*. 2014;34(1):3-23.
10. Conti-Ramsden G, Durkin K. Language development and assessment in the preschool period. *Neuropsychol Rev*. 2012;22(4):384-401.
11. Sheldrick RC, Perrin EC. Evidence-based milestones for surveillance of cognitive, language, and motor development. *Acad Pediatr*. 2013;13(6):577-586.
12. Jiang H, Logan JA, Jia R. Modeling the Nature of Grammar and Vocabulary Trajectories From Prekindergarten to Third Grade. *J Speech Lang Hear Res*. 2018;61(4):910-923.
13. Kuhl PK. Early language acquisition: cracking the speech code. *Nat Rev Neurosci*. 2004;5(11):831.
14. Bates E, Dale PS, Thal D. Individual differences and their implications for theories of language development. *The handbook of child language*. 1995;96-151.
15. Fenson L, Dale PS, Reznick JS, et al. Variability in early communicative development. *Monogr Soc Res Child Dev*. 1994;i-185.
16. Luinge MR, Post W, Goorhuis-Brouwer SM. The language screening instrument SNEL. *Educational and Child Psychology*. 2007;24:20-30.
17. Squires J, Twombly E, Bricker D, Potter L. ASQ-3: User's guide. Baltimore: Paul H. Brookes Publishing Co.; 2009.

18. Sijtsma K, Molenaar IW. *Introduction to nonparametric item response theory* (Vol. 5). Sage, 2002.
19. Meijer RR, Baneke JJ. Analyzing psychopathology items: a case for nonparametric item response theory modeling. *Psychol Methods*. 2004;9(3):354.
20. Visser-Bochane MI, Gerrits E, van der Schans CP, Reijneveld SA, Luinge MR. Atypical speech and language development: a consensus study on clinical signs in the Netherlands. *Int J Lang Commun Disord*. 2017;52(1):10-20.
21. Laurent de Angulo MS, Brouwers-de Jong EA., Bijlsma-Schlösser JFM,
22. Bulk-Bunschoten AMW, Pauwels JH, Steinbuch-Linstra I. *Ontwikkelingsonderzoek in de jeugdgezondheidszorg; het Van Wiechenonderzoek—De Baecke-Fassaert Motoriektest*. Assen: Van Gorcum; 2005.
23. Van der Ark LA. Mokken scale analysis in R. *J Stat Softw*. 2007;20(11):1-19. URL <http://www.jstatsoft.org/v20/i11/>.
24. Van der Ark LA. New developments in Mokken scale analysis in R. *J Stat Softw*. 2012;48(5):1-27.
25. Shriberg LD, Tomblin JB, McSweeney JL. Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *J Speech Lang Hear Res*. 1999;42(6):1461-1481.
26. Hendriks P, Koster C. Production/comprehension asymmetries in language acquisition. *Lingua*. 2010;120(8):1887-1897.
27. Miller LE, Perkins KA, Dai YG, Fein DA. (2017). Comparison of parent report and direct assessment of child skills in toddlers. *Res Autism Spectr Disord*. 2017;41: 57-65.
28. Sachse S, Von Suchodoletz W. Early identification of language delay by direct language assessment or parent report? *J Dev Behav Pediatr*. 2008;29(1):34-41.
29. Van Sonderen E, Sanderman R, Coyne JC. Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS one*. 2013;8(7):e68967.
30. Feldman HM, Dollaghan CA, Campbell TF, Kurs-Lasky M, Janosky JE, Paradise JL. Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Dev*. 2000;71(2):310-322.
31. Tomasello M, Mervis CB. The Instrument is great, but measuring comprehension is still a problem. *Monogr Soc Res Child Dev*. 1994;59(5):174-179.
32. Dale PS, Bates E, Reznick JS, Morisset C. The validity of a parent report instrument of child language at twenty months. *J Child Lang*. 1989;16(2):239-249.
33. Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS. *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*. PB Brookes; 2014
34. Rescorla L. The Language Development Survey: A screening tool for delayed language in toddlers. *J Speech Hear Disord*. 1989;54:587-599.
35. Vehkavuori SM, Stolt S. Screening language skills at 2;0. *Infant Behav Dev*. 2018;50:174-179
36. Rice ML, Hoffman L. Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2; 6 to 21 years of age. *J Speech Lang Hear Res*. 2015;58(2):345-359.

Supplemental Table A. Reasons for exclusion of items and excluded items per reason: poor comprehensibility for parents, the item being too easy for all children, the item not fitting unidimensional scale, or the item violating the item ordering (with numbers of violations between brackets).

Reason removed	Item removed
Poor comprehensibility	<p>Does your child say more than 30 words?</p> <p>Does your child ask questions about ‘who’, ‘what’, ‘where’, ‘how’, for example, ‘Where’s my teddy bear?’</p> <p>Can other people understand your child well? Like teacher, grandfather or aunt?</p> <p>Does your child talk like an adult in terms of language?</p> <p>Is your child able to talk with you and others in turns?</p>
Too easy	<p>Does your child understand you when you say ‘no’?</p> <p>When you call your child’s name, does he/she react to it? For example, by looking or by saying something?</p> <p>Does your child wave ‘bye bye’ when someone leaves?</p> <p>Does your child make contact by themselves? For example, by holding something up for you or by making a sound to draw your attention?</p> <p>Does your child imitate you sometimes? For example, by making noises or by waving?</p>
Not scalable	<p>Does your child babble with different sounds? For example, ‘dada’, ‘baba’ or ‘gaga’?</p> <p>Does your child look at your face when you are talking?</p> <p>Can your child alternate making sounds with you?</p>
Violations on item ordering	<p>Is the pronunciation of your child (with exception of the letter /r/) good? (18 violations)</p> <p>Does your child use three word sentences? For example: ‘Also have candy’, or ‘Ben not hurt’ (13)</p> <p>Does your child use four word sentences? For example: ‘I put in bowl’ or ‘can I build blocks?’ (12)</p> <p>Is your child able to ask a question? For example: ‘What’s that?’ or ‘who’s that?’ (9)</p> <p>Do you understand almost everything your child says? (6)</p> <p>Is your child able to clarify what he/she wants to say or ask? (5)</p> <p>Can your child easily remember new words? (4)</p>

Supplemental Table A. (Continued)

Reason removed	Item removed
	Can your child say words with a consonant cluster in it, like /st/ in stick? (3)
	Does your child understand it when you give him/her two tasks in one sentence? For example: 'put on your shoes and then pick up your drawing' (3)
	Is your child able to express his/her thoughts verbally? (2)
	Do you understand about 75% of what your child is telling you? (2)
	Do you understand about half of the things your child says? (2)
	Does your child make very long sentences? For example: 'When I grow up, I want to be a cook'. (2)
	Does your child call themselves 'me' or 'I'? (1)

Supplemental Table B. Final scaled items that reflect language development for children from 1 to 6 years of age.

	Items Final Scale	Domain	Proportion endorsement	H _i (SE)
1	Does your child say any 'words'? For example: 'mama', 'papa', cookie'? It doesn't have to be pronounced correctly.	V	0.99	0.81 (0.05)
2	When you play with your child, with a ball for example, does your child have attention for you and the ball?	C	0.99	0.62 (0.08)
3	Does your child understand tasks consisting of two words? For example: 'coat on' or 'look there'	V	0.99	0.90 (0.02)
4	Does your child understand when you ask him/her something? For example: 'Shall we read a book together?'	V/G	0.98	0.89 (0.02)
5	Does your child understand 3 word sentences? For example: 'on the chair' or 'to the hallway?'	V	0.98	0.87 (0.03)
6	Can your child point out something you ask them to? For example: 'where is your nose?' or 'where is the ball?'	V	0.98	0.88 (0.03)
7	Can your child say about ten words in total?	V/G	0.94	0.88 (0.02)
8	Is your child able to point 5 pictures in a book when they are verbally offered to him/her?	V	0.93	0.89 (0.01)

Supplemental Table B. (Continued)

	Items Final Scale	Domain	Proportion endorsement	H_i (SE)
9	Is your child able to point out 6 body parts on a doll/themselves? Where are the eyes, mouth, tummy, foot, hair, hand?	V	0.92	0.86 (0.02)
10	Does your child ask you when he/she wants to have some food, or wants to play with toys?	G/C	0.89	0.82 (0.02)
11	Is your child able to combine two words? For example: 'daddy ball' or 'look cat' ?	G	0.89	0.88 (0.01)
12	Can your child name four or more pictures of animals? For example : 'dog', 'cat', 'horse', 'cow'	V	0.89	0.88 (0.01)
13	Is your child able to talk with you in turns?	C	0.86	0.71 (0.03)
14	Is your child able to initiate a conversation?	C	0.85	0.77 (0.02)
15	Are the words in the sentences of your child mostly in the right place?	G	0.82	0.80 (0.02)
16	Does your child ever spontaneously tell you a story? For example, about what they did that day.	C	0.80	0.81 (0.02)
17	Does your child use words that say something about other words (adjectives)? For example, 'big' in 'a big house'	G	0.78	0.85 (0.01)
18	Is your child able to name a couple of colours correctly?	V	0.76	0.85 (0.01)
19	Can your child retell a story with the help of pictures? For example when reading a book together.	C	0.75	0.83 (0.01)
20	Does your child use words like 'we', 'he', and 'she' in a sentence? For example: 'He is really happy'	G	0.69	0.87 (0.01)
21	Does your child make sentences with words like 'when' or 'and'? For example : 'when we have finished dinner, we are going to play with clay' or 'they put their coat on and they put their shoes on'	G	0.65	0.86 (0.01)
22	Does your child ask questions beginning with 'why'?	G	0.65	0.85 (0.01)
23	Does your child use the proper plural form? For example feet instead of foots	G	0.57	0.82 (0.01)

Supplemental Table B. (Continued)

	Items Final Scale	Domain	Proportion endorsement	H_i (SE)
24	Can your child complete the following sentences: Not black but.... Not high but....	V	0.54	0.82 (0.01)
25	Does your child make sentences with 'because'?	G	0.48	0.84 (0.01)
26	Does your child use the correct form of the past? For example : 'went' instead of 'goes' or 'had' instead of 'haved'	G	0.36	0.83 (0.02)

Note. Domain: V = Vocabulary, G = Grammar, C = Communication; Proportion of endorsement = proportion of our sample agreeing on the item; H_i = scalability coefficient of the item. The original items in Dutch are available on request. The authors hold the intellectual property right and they welcome any further use that is aligned with them.

The background of the page is a light gray with a pattern of darker gray, wavy, organic shapes. Scattered throughout these shapes are various cursive letters in different sizes and orientations, including 'W', 'p', 'C', 'K', 'S', 'L', 'B', 'G', 't', 'J', 'h', 'F', 'B', 'A', 'O', 'S', 'z', 'y', 'U', 'e', 'm', 'a', 'D', 'P', 'w', 'I', 'f', 'g', 'l', 'N', 'E', and 'J'.

6

Validation of the early language scale

Published as: Visser-Bochane MI, van der Schans CP, Krijnen WP, Reijneveld SA, Luinge MR. Validation of the early language scale. *European Journal of Pediatrics*, 2021;180(1):63-71

Abstract

Aim: The aim of this study was to assess the criterion-validity of a new screening instrument, the Early Language Scale (ELS) for the identification of young children at risk for developmental language disorder (DLD), and to determine optimal age-adjusted cut-off scores.

Methods: We recruited a community-based sample of 265 children aged 1 to 6 years of age. Parents of these children responded on the ELS, a 26-item “yes-no” questionnaire. The children were assessed with extended language tests (language comprehension, word production, sentence production, communication). A composite score out of these tests (two tests below -1SD or one below -1.5SD) was used as reference standard. We assessed the validity of the ELS, measured by sensitivity, specificity, predictive values, and AUC.

Results: The optimal sensitivity/specificity age dependent cut-off ELS score was at 15th percentile. Sensitivity and specificity were 0.62 and 0.93, respectively. Positive predictive value was moderate (0.53), negative predictive value was high (0.95), the positive likelihood ratio was 9.16, and negative likelihood ratio 0.41. The area under the ROC curve was 0.88. The items covered the increasing language development for the ages from one to six.

Conclusion: The ELS is a valid instrument to identify children with DLD covering an age range of one to six years in community-based settings.

Introduction

Developmental language disorder (DLD) is one of the most common developmental problems in children and has a negative effect on children's emotional functioning, academic success and social relationships [1-5]. Its estimated prevalence is 7% [6]. Early identification and treatment of DLD can prevent or reduce its detrimental effects [7], and preventive child health services may offer an excellent setting for early identification.

Early identification of DLD in well-child care requires short, valid and reliable instruments, that preferably cover the age-range in which language develops. A widely used short instrument for delineated age-categories in a large age-range is the Ages and Stages Questionnaire (ASQ) [8]. The ASQ is a simple parental questionnaire to identify children with suspected developmental delays in communication, gross motor, fine motor, problem solving, and the personal-social domain. However, instruments that target only language development outperform the ASQ communication domain in identifying atypical language development [9,10]. These instruments for language screening, also listed in the US Preventive Task Force (USPSTF) update [9,11,12], are the MacArthur-Bates Communicative Development Inventory (CDI) [13], and Language Development Survey (LDS) [14]. Both instruments are too lengthy for routine use in well-child care as they require parents to check at least 100 and 310 items, respectively. Recently, CDI short forms were introduced, comprising only 25 items, confirming the need for shorter instruments that focus on language development. However, this CDI short form is only suitable for children up to 30 months of age [15], whereas the moment of identification of DLD typically exceeds this age [16-20]. Currently, there is no valid instrument available that covers the full early developmental period in which language is developed, i.e. up to 6 years, and that is also short to administer.

We developed an Early Language Scale (ELS) - that measures language development in 26 easy observable items by parents, covering the full range of early language development (1-6 years). However, evidence on its validity and scoring are lacking. The aim of this study was therefore to assess the criterion-validity of the ELS, and to determine optimal age-adjusted cut-off scores.

Methods

Study design

This study was the validation part of a cross-sectional study on the development of an Early Language Screening instrument (ELS) for the measurement of language development in children (registered at [trialregister.nl](https://www.trialregister.nl): 5746). In the current study, we compare the results of the ELS with a reference standard in a community-based cohort of children aged 1 to 6 years. This study was approved by the Medical Ethical Committee of the University Medical Center of Groningen (M13.134252 / NL45253.042.13).

Sampling and Participants

We recruited parents and their children from 1 to 6 years of age in a two-step procedure. In the first step, we recruited well-child clinics, kindergartens, and schools. Second, these institutions recruited parents by distributing our folder, and collecting and returning reply forms. We asked them to do so according to predefined guidelines on the sample required (e.g. parents of the five youngest three-year old boys), in order to create a well-balanced sample regarding age and sex distribution. Exclusion criteria were significant sensory impairments (e.g. blindness, deafness), or mental disorders (e.g. mental retardation). In total, 1,231 parents of children from 1 to 6 were recruited via 162 institutions and provided normative data. In this study, we validated the Early Language Scale (ELS) in a subsample of the normative sample. For this subsample, all participating parents were invited to let their child participate in the validation part of the study (Figure 1). In total, 919 (74%) parents agreed. Out of these, a random sample for validation regarding children with Dutch as their first language was invited for testing, stratified by gender and age-year of the child. All parents provided written, informed consent. Characteristics of the sample are presented in Table 1. The sample was representative for the Dutch population with respect to birth weight, and pregnancy duration. Participating children were more likely to have mothers who were more highly educated. For comparison, 56% of the females (25-35 years old) has a higher education level, vs 68% in our sample (Statistics Netherlands, 2019).

Figure 1. Flowchart of participant recruitment

Procedure

Data for validation were collected by speech language pathologists from March 2015 to July 2016 during a home visit of, on average, two hours. During the home visit, we first obtained informed consent. Thereafter, we collected from the parent: the filled-in parent forms belonging to the reference language tests, background characteristics of the child and the family, and responses to the ELS questionnaire. During the collection of these data, the children could get used to the speech language pathologist. Last, we administered reference tests to the children, in order to provide a reference standard for the validation of the ELS. Children of 2 years of age and older were assessed with specific language tests on language comprehension, followed by language production. Children younger than two years of age are too young for these tests. Therefore, we assessed their language comprehension and language production with parental questionnaires. Finally, we assessed communication with a standardized observational instrument for children up to four years of age. Communication of children aged four and five was assessed with a standardized parental questionnaire. The tests used are described below.'

Table 1. Characteristics of the sample by age category

	15 - 23	24 - 35	36 - 47	48 - 59	60 - 72	Total
	months	months	months	months	months	
	(n = 35)	(n = 58)	(n = 59)	(n = 58)	(n = 55)	(n = 265)
Characteristics	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Gender						
Male	16 (46%)	26 (45%)	35 (59%)	29 (50%)	25 (45%)	131 (49%)
Female	19 (54%)	32 (55%)	24 (41%)	29 (50%)	30 (55%)	134 (51%)
Birth weight (grams)*						
High (>5000)	-	-	-	-	-	-
Average (2500-5000)	35 (100%)	55 (95%)	54 (92%)	51 (88%)	52 (95%)	247 (93%)
Low (<2500)	-	3 (5%)	2 (3%)	6 (10%)	2 (4%)	13 (5%)
Pregnancy duration						
Serotine (>= 42 weeks)	5 (14%)	3 (5%)	3 (5%)	1 (2%)	9 (16%)	21 (8%)
Term (37-42 weeks)	30 (86%)	52 (90%)	53 (90%)	48 (83%)	44 (80%)	227 (86%)
Preterm (<37 weeks)	-	3 (5%)	3 (5%)	9 (16%)	2 (4%)	17 (6%)
Language						
Dutch only	31 (89%)	52 (90%)	52 (90%)	50 (86%)	49 (89%)	235 (89%)
Dutch and other language(s)	4 (11%)	6 (10%)	6 (10%)	8 (14%)	6 (11%)	30 (11%)
Maternal education*						
High	28 (80%)	41 (71%)	41 (70%)	36 (62%)	35 (64%)	181 (68%)
Middle	7 (20%)	14 (24%)	16 (27%)	20 (35%)	18 (33%)	75 (28%)
Low	-	3 (5%)	2 (3%)	1 (2%)	2 (4%)	8 (3%)
Maternal age*						
21-30	11 (31%)	20 (35%)	23 (39%)	27 (47%)	26 (47%)	107 (40%)
31-35	16 (46%)	27 (47%)	28 (48%)	20 (35%)	20 (36%)	111 (42%)
>36	7 (20%)	7 (12%)	8 (14%)	11 (19%)	6 (11%)	39 (15%)

* Numbers do not always add up to n = 265 due to missing data, Birthweight: 5 missing, Maternal Education: 1 missing, Maternal Age: 8 missing

Measurements

Early Language Scale (ELS)

The ELS is a parental questionnaire consisting of 26 yes/no questions on the language development of children in the age of 1 to 6 years old (Appendix 1). The items were administered in Dutch. We used forward-only translation of the original items to English for international publication. A yes-score results in one point, the ELS total score represents the sum of all items a child has acquired, this score varies from 0 to 26. The ELS consists of items on language development, within the domains of vocabulary (semantics), sentences/grammar (syntax and morphology), and communication (language use). In a prior study, using the data from a community-based sample ($n = 1,231$), items were selected from an item bank of 75 items based on the criteria of Mokken scaling, which is based on the nonparametric Item Response Theory [21]. Items were identified via automated item selection procedure [22], evaluated on monotonicity and item ordering, resulting in a final ELS scale consisting of 26 items. The constructed ELS was a strong scale (scalability coefficient $H=0.83$) with Item H coefficients between 0.62 and 0.90 [23]. This shows that the scale consists of the most distinctive items that describe the development of language in children aged one to six years.

Reference tests

For the validation of the ELS, we assessed the language development of the child with age appropriate reference tests of language development, i.e. the core concept to be assessed by the ELS. The ELS encompasses items regarding language comprehension and production (of vocabulary and grammar), and communication. The reference tests measured the same aspects of language development, therefore the reference standard is a combination of language tests that assess the aspects language comprehension and production (of words and sentences), and communication. Table 2 gives an overview of tests used per age-category. We constructed a reference standard as the composite score out of these tests, calculated as follows: two or more test scores below -1 SD of the norm score or 1 test score (on language comprehension or production) below -1.5 SD of the norm score resulted in a deviant language score. We denoted children with a deviant language score as children with atypical language development; this includes DLD and language delays. The tests used are described below.

The *LexiList Comprehension (LLC)* [24] and *LexiList Production (LLP)* [25] to measure language comprehension and language production in children younger than two years of age, comprise lists of words and phrases (190 words and 25 phrases for comprehension, and 263 words and 11 phrases for production) for parents to tick the words and phrases of which they think their child understands, and uses. Results vary from 0 to 225 (LLC), and 0 to 274 (LLP), and an age-standardized score (mean = 100; SD = 15) were calculated according to the manuals [24,25]. Internal consistency of the LLC is good (Cronbach's alpha of .98). The LLP shows good reliability and sufficient validity [26].

Table 2. Tests used per language domain and age year

Age (months)	Language Comprehension	Language Production		Language Use / Communication
12 - 23	LexiList Comprehension (LLC)	LexiList Production (LLP)		Language Standard – Communication
24 - 35	Schlichting test for Language Comprehension (SLC)	Schlichting test for Word Production (SWP)	Schlichting test for Sentence Production (SSP)	Composite Score (LS-CCS)
36 - 47				CCC-2-NL- Pragmatic Composite Score (CCC-PCS)
48 - 59				
60 - 72				

The *Schlichting tests for Language Comprehension (SLC)* [27], *Word Production (SWP)* [28], and *Sentence Production (SSP)* [28] to measure language comprehension, and language production in children from 2 to 7 years of age. The SLC is a 85-item test assessing comprehension of grammatical constructions using toys, pictures and tokens. The SWP is a 70-item test assessing expressive vocabulary using a stimulus booklet with pictures. The SSP is a 40-item test assessing expressive grammatical constructions, using imitation of expressions, visualized in a stimulus booklet with in some cases associated toys. Age-standardized scores for each test (mean = 100; SD = 15) were calculated according to the manuals, in which also entry levels per age, and cut-off rules are described [27,28]. The SLC, SWP, and SSP have excellent internal consistency (lambda-2 = .93, .93, and .90 respectively) [27,28].

The *Language Standard (LS)* [29] is a 20-item observational instrument, providing information on general language ability. The trained professionals observed the child while playing with the parent and with the professional on standardized observation items, and scored findings on a 5-point scale according to the manual ‘5-points: clear evidence for normal; 4-points: between 5 and 3; 3-points: possible evidence for problem; 2-points: between 3 and 1; 1-point: clear evidence for problem’. [29]. We extracted a composite score on communication (LS-CCS), based on consensus of six professionals, including the authors MVB and ML. Consensus on items that regard communication was reached in two consensus rounds prior to testing. In the first round all professionals indicated which items from the Language Standard regarded communication. In the second round, professionals received their own scores and the scores of all other professionals. They were then asked to indicate which items regarded communication again. This resulted in consensus on 8 items, regarding the observation of the child’s communication skills while playing with the parent. Observed items were for example: initiation of contact, responds in communication, attention for language, nonverbal communication. Counting the scores on these items (5, 6, 7, 8, 12, 17, 18, and 20) results in a maximum score of 40, with a cut-off at 32 or lower indicating a problem on communication.

The *Children’s Communication Checklist – 2 – NL (CCC)* [30] is a 70-item parental checklist that assesses children’s communication behaviors over 10 subscales. The sum of the scales E-H (E. inappropriate initiation, F. stereotyped language, G. use of context, H. nonverbal communication) provides a composite score on Pragmatics (CCC-PCS). This frequently reported score identifies children likely to have significant pragmatic language problems. Psychometric properties for the Dutch adaptation of the CCC-2 are satisfactory; with internal consistency ranging from .53 to .75 [30]. The subscale pragmatics shows sufficient reliability, but insufficient criterion validity [31].

Background characteristics

We obtained data by parent report on the following background characteristics: age of the child, gender, birthweight, length of pregnancy, language situation at home, highest level of education achieved by the mother, maternal age at birth of the child, and family history (sibling, parent, grandparent) of language delay. Education level was classified into three categories: low (primary school or less,

and pre-vocational education), middle (secondary education), and high (higher vocational education and university).

Analyses

First, we selected children with two or more reference test scores -1 SD or more below the mean or one reference test score (on language comprehension or production) -1.5 SD or more below the mean in order to identify these children as children with atypical language development. Second, we assessed the criterion validity of the ELS. To visualize our data, we plotted the ELS Total score against age, and added smoothed splines, representing group means; one for the children with typical language development and one for children with atypical language development, based on the combined reference standard. Next, we calculated sensitivity and specificity indices, positive predictive value and negative predictive value, positive likelihood ratio and negative likelihood ratio, for various age-adjusted cut-offs of the ELS against the reference standard. The age-adjusted cut-off values of the ELS were calculated for 3-months age groups regarding 5th, 10th, 12th, 15th, 25th, 50th, 75th, and 100th percentiles within our norm sample ($n=1,231$). We chose 3-months age groups to be able to differentiate within an age-year, as children achieve several language milestones per age year. We also performed ROC analyses determining the area under the curve (AUC), by calculating sensitivity and specificity for the various cut-off values. Finally, we assessed whether the performance of the ELS differed by background characteristics.

Results

Sample

All children spoke Dutch as first language. Atypical language development, based on the combined reference standard was found in 29/265 (11%) of the children. These 29 children were from all ages (five 1-year olds; nine 2-year olds; eight 3-year olds; four 4-year olds; three 5-year olds). One child spoke Dutch as first language and another language at home, one child had a birthweight below 2500 grams, five children were born at a pregnancy duration of less than 37 weeks, and two children had mothers with a low education level.

Outcomes on the Early Language Scale and the reference standard

ELS Total scores for children with a typical language development, based on the reference standard were higher than for children with an atypical language development. ELS Total scores for both groups increased across ages, with a ceiling effect at the maximum score of 26 on the ELS (Figure 2).

Figure 2. Early Language Scale - Total scores by age in months, for children with typical language development in blue ($n = 236$) and children with atypical language development in red ($n = 29$). Lines refer to group means.

The median scores on the ELS-test were 11 for one year old, 20 for two year old, 24 for three year old, 25 for four year old, and 26 for five year old.'

Validity

The receiver-operating-curve (ROC) was built using various age-adjusted cut-offs of the ELS. The area under the receiver-operating-curve (AUC) in Figure 3 was 0.88, indicating that the ELS adequately differentiated children with atypical from those with typical language development. Taking a cut-off at the 15th percentile, to provide an optimal balance between sensitivity and specificity while minimizing over-referral rates, resulted in a sensitivity of .62, a specificity of .93, a positive predictive value of .53, a negative predictive value of .95, a positive likelihood ratio of 9.16, and negative likelihood ratio of .41 (Table 3). This regarded the following cut-off values per age: 12-14 months: 2; 15-17 months: 6; 18-20 months: 7; 21-23 months: 9; 24-26 months: 14; 27-29 months: 15; 30-32 months: 16; 33-35 months: 18; 36-38 months: 20; 39-41 months: 21; 42-44 months: 22; 45-53 months: 23; 54-71 months: 24. We applied the cut-off values as follows: for a child of 23 months old, the cut-off value of 9 means: endorsement of a maximum of 9 items results in a deviant score, endorsement of 10 or more items results in a normal score.

Figure 3. Receiver Operating Curve regarding the ability of the Early Language Scale (ELS) to distinguish children with atypical language development and children with typical language development.

Table 3. Results of the screening instrument (ELS) and the reference standard.

	Identified by screening n = 34	Not-identified by screening n = 231	Total n = 265
Result reference standard			
Atypical language	18	11	29
Typical language	16	220	236

Differences in the performance of the ELS by background characteristics

Finally, we assessed whether the performance of the ELS differed by background characteristics, i.e. for the four subgroups: true positives, false positives, false negatives and true negatives of the ELS. We found no significant effect of birthweight [$F(3, 256) = 1.19, p = 0.315$], Length pregnancy [$F(3, 261) = 2.47, p = 0.06$], maternal age [$F(3, 253) = 1.32, p = 0.27$] assessed using one-way ANOVA. There was also no significant effect of gender [$\chi^2(3, N = 265) = 0.69, p = 0.88$]. Maternal education and Language could not be tested as these had too little observations in one or more subgroups. Maternal education low is more prevalent in the true positive group, confirming that children from low educated mothers have slower language development. Regarding language situation at home, most of the children with Dutch and another language spoken at home had typical language development and were not identified by the screening, indicating that the language situation did not affect the performance of the ELS. Finally, results were similar if the 17 included preterm children were excluded ($p = 0.248$, Fisher's exact test). In general, results confirmed expectations.

Discussion

This study examined the validity of the Early Language Scale (ELS) in children aged one to six using a composite score of language tests as reference standard. We found that the ELS has a good validity making use of only a limited set of items feasible for an age range of one to six years.

We found that the ELS has a good validity, i.e. an AUC of .88, comparable to the CDI and LDS with reported AUC's of .86 and .85 respectively [32,33]. However, in contrast to these instruments, the ELS covers a large age range (one to six); the AUC holds for the entire age range in which the basics for language develops, meaning that the ELS measures this developmental ability. Moreover, the ELS is valid instrument to measure language over the developmental period of one to six years of age.

We found a sensitivity of .62 and a specificity of .93, which is suitable for community-based screening. A high specificity, of at least .90, minimizes over-referrals with its negative effects, e.g. discomfort for parents and children, and costs [34]. On the other hand, the sensitivity of the ELS is moderate, but increasing sensitivity by larger cut-off values would decrease specificity. A higher sensitivity is desirable, however not at the expense of lower specificity. This contrasts with screening of high-risk populations in which the balance between sensitivity and specificity should shift towards a higher sensitivity, to minimize under referrals with its negative effects, e.g. progressive language problems and associated long-term effects. The ELS is thus suitable for community-based screening of children in a rather large age-range.

The properties of the ELS, e.g. a limited set of 26 items reflecting a broad definition of language development that is also feasible for a large age range, make the ELS an addition to other valid language screening instruments. The ELS exceeds the age range of the MacArthur-Bates Communicative Development Inventory (CDI) short forms [15] and the Language Development Survey (LDS) [14]. Both CDI and LDS regard vocabulary checklists filled by parents. However, LDS assesses only expressive language and not receptive language. The ELS included items on expressive and receptive language. Atypical language development in children aged one to six are easy to detect with only a small number of items, however some children will still be missed.

The scale has been developed in Dutch, but consists of items that are very common in related languages, such as English and German. For instance, an item like ‘says two-word sentences’ (item 11) is rather similar in those languages. Use of the scale in other languages requires additional research to obtain language specific norms and a validation in that language.

A major strength of our study is that we used a large, community-based sample, equally distributed over gender, and ages groups of 1 to 6 years, with state of the art reference tests. We tested the power of our sample based on the assumption of a prevalence of 11% atypical language development and 89% typical language development (according to the reference standard) in our sample; i.e. the prevalence under the independence. Both resulting p-values were $< .001$, indicating that the results of the ELS are not based on coincidence, and therefore the sample of $n=265$ was sufficient to assess validity. However, a limitation of this study is that despite its size, the number of cases was still relatively small. This limited the potential to analyze outcomes per age-year. Another limitation of our study is that the sample had some overrepresentation of mothers with a high education level. This resulted in cut-off norms that might be slightly low for children from disadvantaged families, as evidence suggests that children of highly educated mothers (and fathers) have better language [35]. However, the norms are feasible for our sample, including the full range of educational levels. Moreover, a better language development is associated with successful developmental and educational outcomes. Therefore all children with atypical development are entitled to identification of this crucial deficit. Last, children develop language with individual variability and speed. Therefore, it is recommended to monitor language development over time. The ELS regards ordered language items, and can be used repeatedly.

The results of this study show that the ELS is a valid instrument for the detection of children at risk for developmental language disorder in a population-based sample. The ELS can provide support to professionals working in well-child care and early educational settings to quickly identify children (1-6) with an atypical language development. This identification of children supports timely referral to proper diagnostics and intervention. This may improve these children’s emotional functioning, academic success and social relationships.

The validity of the 26-item ELS is satisfactory for a broad age range. Further research is recommended to investigate its age-specific validity and long term predictive value, as well as the applicability of its set-up to other languages. The use of the ELS in routine community settings deserves further study regarding feasibility and effects in a clinical setting.

Conclusion

The ELS is a valid instrument to identify children at risk for DLD covering an age range of one to six years in community-based settings. The ELS is promising for identifying children with atypical language development in a public health setting as it does not yield a high proportion of false-positives.

Acknowledgements

The authors thank participating speech-language therapists, parents and children.

Funding:

This work was supported by the Netherlands Organization for Health Research and Development (ZonMw; grant number 200330002).

References

1. Boyle CA, Decoufle P, Yeargin-Allsopp M. Prevalence and health impact of developmental disabilities in US children. *Pediatrics*. 1994;93(3):399-403.
2. Conti-Ramsden G, Mok PL, Pickles A, Durkin K. Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Res Dev Disabil*. 2013;34(11):4161-4169.
3. Snowling MJ, Duff FJ, Nash HM, Hulme C. Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *J Child Psychol Psychiatry*. 2016;57(12):1360-1369.
4. Whitehouse AJ, Lin, EA, Watt HJ, Bishop DVM. Qualitative aspects of developmental language impairment relate to language and literacy outcome in adulthood. *Int J Lang Commun Disord*. 2009;44(4):489-510.
5. Conti-Ramsden G, Durkin K, Mok PL, Toseeb U, Botting N. Health, employment and relationships: Correlates of personal wellbeing in young adults with and without a history of childhood language impairment. *Soc Sci Med*. 2016;160:20-28.
6. Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res*. 1997;40(6):1245-1260.
7. Broomfield J, & 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*. 2011;46(6):628-640.
8. Squires J, Potter L, Bricker D. *The ASQ User's Guide*. 2nd ed. Baltimore, MD: Paul H. Brookes Publishing Co;1999
9. Berkman ND, Wallace IF, Watson L, et al.. *Screening for speech and language delays and disorders in children age 5 years or younger: A Systematic Review for the U.S. Preventive Service Task Force*. Evidence synthesis No. 120. Rockville, MD: Agency for Healthcare Research and Quality; 2015. AHRQ Publication 13-05197-EF-1
10. Kim SW, Kim JY, Lee SY, Jeon, HR The Comparison of M-B CDI-K Short Form and K-ASQ as Screening Test for Language Development. *Ann Rehabil Med* 2016;40(6):1108-1113
11. Siu AL, US Preventive Services Task Force. Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force recommendation statement. *Pediatrics*. 2015;136(2):e474-e481
12. Wallace IF, Berkman ND, Watson LR, et al. Screening for speech and language delay in children 5 years old and younger: a systematic review. *Pediatrics*. 2015;136(2):e448-e462.
13. Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS. *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*. PB Brookes; 2014
14. Rescorla L. The Language Development Survey: A screening tool for delayed language in toddlers. *J Speech Hear Disord*. 1989;54:587-599.
15. Mayor J, Mani N. A short version of the MacArthur-Bates Communicative Development Inventories with high validity. *Behav Res Methods*. 2018;1-8.

16. Armstrong R, Scott JG, Whitehouse AJ, Copland DA, McMahon KL, Arnott W. Late talkers and later language outcomes: Predicting the different language trajectories. *Int J Speech Lang Pathol*. 2017;19(3):237-250.
17. Bright Futures Steering Committee, & Medical Home Initiatives for Children With Special Needs Project Advisory Committee. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*. 2006;118(1):405-420.
18. Law J, Rush R, Anandan C, Cox M, Wood R. Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*. 2012;130(1):e132-e137
19. McKean C, Reilly S, Bavin EL, et al. Language outcomes at 7 years: Early predictors and co-occurring difficulties. *Pediatrics*. 2017;139(3) e20161684.
20. Zambrana IM, Pons F, Eadie P, Ystrom E. Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *Int J Lang Commun Disord*. 2014;49(3):304-316.
21. Sijtsma K, Molenaar IW. *Introduction to nonparametric item response theory*. Vol 5. Sage, 2002.
22. Straat JH, Van der Ark LA, Sijtsma K. Comparing optimization algorithms for item selection in Mokken scale analysis. *J Classif*. 2013;30:72-99.
23. Visser-Bochane MI, Reijneveld SA, Krijnen WP, van der Schans CP, Luinge MR. Identifying Milestones in Language Development for Young Children Ages 1-6 Years. *Acad Pediatr*. 2019 [Prepub, doi: 10.1016/j.acap.2019.07.003]
24. Schlichting & Lutje Spelberg, Lexilijst Begrip [Lexilist Comprehension]. Amsterdam: Harcourt Test Publishers; 2007.
25. Schlichting JEPT, Lutje Spelberg HC. Lexilijst Nederlands [Lexilist Netherlands Production]. Amsterdam: Pearson; 2002.
26. Egberink IJL, Leng WE de, Vermeulen CSM. COTAN beoordeling 2009 Lexilijst Nederlands [COTAN review 2009, Lexilist Netherlands Production] Retrieved from www.cotandocumentatie.nl; 2019.
27. Schlichting JEPT, Spelberg HC. Schlichting test voor taalbegrip [Schlichting test for language comprehension]. Houten: Bohn Stafleu van Loghum; 2010.
28. Schlichting JEPT, Spelberg HC. Schlichting Test voor Taalproductie-II: voor Nederland en Vlaanderen [Schlichting test for Language Production]. Houten: Bohn Stafleu van Loghum; 2012.
29. Slofstra-Bremer CF, Meulen S van der, Lutje Spelberg HC. De Taalstandaard [The Language standard]. Amsterdam: Pearson; 2006.
30. Geurts HM. *Handleiding CCC-2-NL*. [Manual CCC-2-NL]. Amsterdam: Harcourt Test Publishers; 2004.
31. Egberink IJL, Leng WE de, Vermeulen CSM. COTAN beoordeling 2007 Nederlandse bewerking van de Children's Communication Checklist-2 [COTAN review 2007, CCC-2-NL] Retrieved from www.cotandocumentatie.nl; 2019.
32. Duyme M, Capron C. (2010). L'Inventaire du Développement de l'Enfant (IDE). Normes et validation françaises du Child Development Inventory (CDI). *Devenir*. 2010;22(1):13-26.
33. Beaulieu-Poulin C, Simard MN, Babakissa H, Lefebvre, Luu, TM. Validity of the language development survey in infants born preterm. *Early Hum Dev*. 2016; 98:11-16.

34. Wolff MS de, Theunissen MH, Vogels AG, Reijneveld SA. Three questionnaires to detect psychosocial problems in toddlers: A comparison of the BITSEA, ASQ: SE, and KIPPI. *Acad Pediatr*. 2013;13(6):587-592.
35. Rice ML, Hoffman L. Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2; 6 to 21 years of age. *J Speech Lang Hear Res*. 2015;58(2):345-359.

Appendix 1. The Early Language Scale (ELS)

Name child:		Age child (in months):		
ELS item		yes	no	
1	Does your child say any 'words'? For example: 'mama', 'papa', 'cookie'? It doesn't have to be pronounced correctly.			
2	When you play with your child, with a ball for example, does your child have attention for you and the ball?			
3	Does your child understand tasks consisting of two words? For example: 'coat on' or 'look there'.			
4	Does your child understand when you ask him/her something? For example: 'Shall we read a book together?'			
5	Does your child understand 3 word sentences? For example: 'on the chair' or 'to the hallway'.			
6	Can your child point out something you ask them to? For example: 'where is your nose?' or 'where is the ball?'			
7	Can your child say about ten words in total?			
8	Is your child able to point 5 pictures in a book when they are verbally offered to him/her?			
9	Is your child able to point out 6 body parts on a doll/themselves? <i>Where are the eyes, mouth, tummy, foot, hair, hand?</i>			
10	Does your child ask you when he/she wants to have some food, or wants to play with toys?			
11	Is your child able to combine two words? For example: 'daddy ball' or 'look cat'.			
12	Can your child name four or more pictures of animals? For example: 'dog', 'cat', 'horse', 'cow'.			
13	Is your child able to talk with you in turns?			
14	Is your child able to initiate a conversation?			
15	Are the words in the sentences of your child mostly in the right place?			
16	Does your child ever spontaneously tell you a story? For example, about what they did that day.			
17	Does your child use words that say something about other words (adjectives)? For example, 'big' in 'a big house'			
18	Is your child able to name a couple of colours correctly?			
19	Can your child retell a story with the help of pictures? For example when reading a book together.			
20	Does your child use words like 'we', 'he', and 'she' in a sentence? For example: 'He is really happy'.			

Appendix 1. (Continued)

Name child:		Age child (in months):			
ELS item				yes	no
21	Does your child make sentences with words like ‘when’ or ‘and’? For example: <i>‘when we have finished dinner, we are going to play with clay’</i> or <i>‘they put their coat on and they put their shoes on’</i> . Does your child ask questions beginning with ‘why’? Does your child use the proper plural form? For example ‘feet’ instead of ‘foots’. Can your child complete the following sentences: <i>Not black but.... Not high but....</i> Does your child make sentences with ‘because’? Does your child use the correct form of the past? For example: ‘went’ instead of ‘goes’ or ‘had’ instead of ‘haved’.				
22					
23					
24					
25					
26					
	Total ‘yes’:				

Note. The original items in Dutch are available on request. The authors hold the intellectual property right and they welcome any further use if aligned with them.

The background of the page is a light gray with a complex, organic pattern of darker gray shapes and lines. Scattered throughout this pattern are various letters in different fonts, including serif, sans-serif, and script styles. Some letters are large and prominent, while others are small and subtle. The overall effect is a textured, artistic backdrop for the text.

7

Towards a more family oriented and interdisciplinary screening of language development in young children: perceptions, opinions and desires of caregivers of children with and without atypical language development

Published as: Luinge MR, Visser-Bochane MI, Dieleman LEP, Coster FW. Towards a more family oriented and interdisciplinary screening of language development in young children: perceptions, opinions and desires of caregivers of children with and without atypical language development. *Journal of Developmental and Physical Disabilities*, 2019;31:863–878

Abstract

Aim: Only a small number of children with atypical language development who can benefit from early intervention are actually identified before reaching school age. Our aim was to explore perceptions of caregivers of monolingual and multilingual children with and without atypical language development in order to provide relevant themes for improving language screening and referral in Well Child Care (WCC).

Methods: We interviewed 38 Dutch caregivers of monolingual and multilingual children with and without atypical language development. They were recruited via kindergarten, nurseries, and speech and language pathologists. First, we analyzed the interviews of caregivers of monolingual children with atypical language development and caregivers of multilingual children. Second, we adopted data source triangulation with the interviews of caregivers of monolingual children without atypical language development for further exploration of the results of the first round. Third, we categorized emerging codes across identified subthemes and themes.

Results: We identified six subthemes: (1) Perception WCC; (2) Perception consult at the WCC; (3) Opinion of the consult at the WCC; (4) Desires regarding WCC; (5) Parent factors, and (6) Child factors (first and second round).

Conclusion: Target themes, relating to the themes and subthemes, that can be used to enhance structural language screening in WCC are: a more family oriented consult, better interprofessional collaboration, and more specific knowledge about language development and different referral routes for children with atypical language development.

Introduction

The prevalence estimates of children with atypical language development (ALD) vary between and within specific countries with published estimates ranging from 2 to 19% [1]. To date, only a small number of children with ALD who can benefit from early intervention services are actually identified before reaching school age [2]. However, children with ALD need to be identified for referral as early as possible because it can have a significant impact on every day social interaction, wellbeing, and school success [3]. These effects can only be prevented or minimized with intervention services and family assistance [4,5]. For this reason, it is of major importance to determine which practices may improve the quality of the early identification and referral to intervention services and family assistance.

The goal of structural language screening in a well child clinic (WCC) is to detect ALD and to direct children with this to appropriate follow-up at a young age. Youth doctors and nurses, systematically screen the (language) development of young children. However, the efficacy of language screening in WCC's is hampered by unsuccessful identification and referral of ALD in a sufficiently early phase [6]. The early identification of ALD in monolingual children is difficult and is even more difficult in multilingual children because the language acquisition of multilingual children differs from monolingual children depending on the languages they speak, the amount and level of exposure to the languages, and the opportunities to use their languages.

To our knowledge, caregivers' perceptions about facilitators and barriers as well as desires for better speech and language screening in Dutch WCCs have not yet been evaluated. How the speech and language screening at a WCC is perceived depends on whether or not ALD is identified. Having a child with a developmental problem is an enormous challenge for family caregivers [7]. Caregivers worry that having a label might adversely affect their children, although such a label might also confirm caregivers' concerns [8]. Additional insight into the perceptions and desires with regard to language screening at a WCC of caregivers of children with and without ALD can be beneficial for identifying target areas for improving the alignment between the caregivers and child healthcare professionals and, therefore, the routine screening of ALD in WCCs. Therefore, the purpose of this study is to explore the perceptions, opinions, and

desires of caregivers of monolingual and multilingual children with and without ALD towards WCCs in order to provide relevant themes for improving structural language screening in them.

Methods

To explore the perception with regard to the language screening at WCCs, we interviewed caregivers of monolingual children with diagnosed ALD and caregivers of multilingual children without ALD (and one multilingual child with ALD). Through inductive content analysis, we identified caregivers' experiences, opinions, and desires until saturation was reached. First, we analyzed the interviews of caregivers of mono-lingual children with ALD and caregivers of multilingual children. In the second round, we adopted data source triangulation with the interviews of caregivers of monolingual children without ALD for further exploration of the results of the first round. We assumed that the inclusion of different perspectives of caregivers – monolingual and multilingual children with and without ALD - would result in data that is more comprehensive and, therefore, would provide more insight into speech language screening in WCC. In the third round, we categorized the emerging codes across identified subthemes and themes with regard to the perception, opinions, and desires of the interviewed caregivers. The medical ethics review committee of University Medical Center Groningen (METc2013/103) approved the study for interviewing caregivers included in this study.

Sample

We selected caregivers with children aged 1;0 to 4;9 years by convenience sampling. They were recruited via kindergarten, nurseries, and SLPs. Participants were included if the last consult at a WCC was no more than a year ago and if they understood the Dutch language. The participants were living in villages ($n = 24$) and midsized towns ($n = 14$) in the south and southwest of the Netherlands. As determined by the Dutch educational level standards according to the Central Statistical Office, they had low ($n = 7$), middle (24), and high ($n = 7$) socioeconomic backgrounds.

The total sample ($n = 38$) included ten caregivers of monolingual children with diagnosed ALD determined by a SLP, seven caregivers of multilingual children of which one child had an ALD (first round), and 21 caregivers of monolingual children without ALD (second round). Multilingual was defined as a child who speaks two or more languages in daily life. The multilingual children spoke, besides the Dutch language, also Turkish, Romanian, Hungarian, German, Czech, English, Papiamentu, or Spanish.

Data Collection

The design of the questions in the interview guide reflected the concepts that are embedded in the research question, literature on language screening (see interview guide) and the ecological framework of Bronfenbrenner [9]. Bronfenbrenner assumed that a child's development is affected by their environment. The ecological paradigm can offer new insights into the child, the family, and the system in which the child lives. In the interviews, we asked family caregivers about the child (the individual), about the perspective of the family and about the health services (microsystem), and about the local policies (exosystem). Based on the research question and this theoretical framework, the first three authors prepared a semi-structured interview, also applying the clinical experiences in the WCC of the third author. The interview guide was semi-structured with open-ended questions and probes in variable sequences about the topics of perceptions, opinions, and desires with regard to language screening in WCC. The interviews were conducted in the Dutch language.

Data collection was ended when no new information related to the interview questions was emerging. The third author as well as students of speech language therapy of the Fontys University of Applied Sciences Eindhoven performed the interviews. The interviews were conducted by telephone or face-to-face at the participants' home, public locations, and at the Fontys University. The time and place for the interviews were selected by the participants. Interviews were recorded with a digital voice recorder and transcribed verbatim.

Interview guide (main questions and probes (indicated by *))

1. When did you last go to the WCC? (microsystem)
*How do you experience the WCC? (microsystem, research question).
2. How does the doctor / nurse pay attention to your child's learning to talk??
(microsystem, research question)
*How?
*A clear indication of the language development of your child? [10]
3. How would you like speech language problems to be detected? (exosystem, research question)
*Who, how, how long, what outcome? (research question)
4. How did you deal with the outcome of the language screening? [8]
5. What are your experiences with the speech language screening? (microsystem, research question)
*What are your experiences with the detection of a language problem?
(individual).
*How did you notice that your child had a language problem? (individual, [11])
6. How did you experience the approach of the WCC-nurse or pediatrician?
(microsystem, [12], research question)
*How were you involved? ([13], clinical experience third author)
7. What do you enjoy most in your child learning to talk? (individual, clinical experience third author)

Data Analyses

The first author analyzed the interviews using the principles of grounded theory, as described by Strauss et al. [14]. The codebook was checked and applied independently to transcripts from the first round by the third author. Disagreements were resolved through consensus discussions between the first and third author. Subsequently, these codes were merged into six subthemes that were evaluated and adjusted by the second author. In the second round, the first author organized all text fragments of the 21 transcripts of caregivers of children without ALD to all of the codes and subthemes of the first round. In the third round, the first and the last author, independently grouped all inductive codes into a matrix of the six inductive subthemes and three identified themes. Both agreed on the themes and mostly about the categorization of the codes in the matrix. Disagreements about the codes were solved by further defining the three themes.

Results

We identified six subthemes relating to the caregivers' perceptions concerning language screening in WCC's: (1) Perception WCC; (2) Perception consult at the WCC; (3) Opinion of the consult at the WCC; (4) Desires regarding WCC; (5) Parent factors, and

(6) Child factors. These subthemes were based on the codes in round one and were confirmed and enriched with extra codes from round two. The codes and subthemes are presented in Figure 1.

Figure 1. Emerging codes categorized into six inductive subthemes based on the interviews of caregivers of monolingual and multilingual children with - and without - atypical language development

Perception WCC

The subtheme *Perception WCC* contains information about the WCC in general and consist of the codes: general view, mean scores, the basics (in vivo), specialist, referral, alternation of the professionals, and other countries. The participants indicated that the WCCs-professionals have a general view on child development and use average scores to determine if a child might have

a developmental problem. Most participants indicated that the WCC-consult is primarily focused on physical aspects such as the vaccinations as well as measuring growth and weight. They refer to these tasks as ‘the basics’. “Yes, so with the basic things really you can fall back on them. What they [the children] need to do on average. If you continue with a really specific problem, I think you should not go right there.” Mother of a child with ALD.

Some participants had doubts about the screening by a WCC-professional because she or he has no overview on the development of young children like those that a kindergarten-professional has. “I also think that the WCC is not the right authority to pick up any signals immediately. I think that the playgroups and nurseries can be much stronger in picking up these things. I think they should advise a lot more. Yes, those who know about the ‘manual’ of the child.” Mother of a child without ALD. Especially caregivers of children with ALD and multilingual children, mentioned the difficulty of the alternation of healthcare professionals at the WCCs. Due to this alternation, the healthcare professionals do not have a good overview of the child’s development. “It is then quite difficult if you have three times another doctor. Especially in the beginning when you come more often. Every time the doctor and then the nurse again.” Mother of a multilingual child without ALD. Additionally, caregivers also discussed the involvement of a pediatrician in the developmental screening in Germany, the Czech Republic, and Austria and about having group consults with changing healthcare professionals in England.

Perception Consult at the WCC

The subtheme *Perception consult at the WCC* is about the actions that are performed during a WCC consult. This subtheme contains the codes: action at the WCC, mapping language development, perception language screening, and offered help. Participants indicated that the WCC-professional mapped the language development of their children by asking the caregiver a few questions. Some participants stated that no attention was paid to the language development at all. Many caregivers from the second round stated that most children were very quiet during the consultation at the WCC, making it difficult to map language, and making WCC professionals dependent on the parents’ report. “With X it was easy, because nothing was wrong, but whether they can

really discover if a child has a speech or language disorder... I dare not say whether they really can do that in such a short time. Because our X is pretty quiet when he is with a stranger, so he was pretty quiet there [at the consult], he doesn't talk that much." Mother of a child without ALD.

Sometimes, the language disorder was discovered in the playgroup or sometimes the participants themselves, family, friends, or neighbors indicated that it lagged behind. Many caregivers stated that they found the questions of the language screening difficult to answer. *"For example, you are asked; does your child babble, but what is babbling? And there comes sound from your child, so you have something like; that is babbling. And then I was at the audiological center in the waiting room. Then, for the first time, I heard babble language. And then I thought; oh, this is babbling. My children have never done that."* Mother of a child with ALD.

Some caregivers were happy with the offered help and indicated that they could always ask the WCC professionals questions. However, especially caregivers of children with ALD felt that WCC professionals did not always listen carefully to their request for help. *"And then she actually said quite shortly and without further explanation, I would think to call a speech therapist. Here is the phone number and see what you do with it. Then I was really a bit,. and now, is this it? Or why?"* Mother of a child with ALD.

Furthermore, several caregivers of multilingual children mentioned that they received too little information on how to deal with multilingualism.

Opinion about the WCC

The subtheme *Opinion about the WCC* describes the experiences and opinions of caregivers regarding the WCC in general (code: opinion of the WCC), the approach by the WCC professional during the consult (code: approach WCC professional), the language screening during the consult (codes: expectation language screening, opinion language screening, snap shot, coping with the outcome emotionally, taking action).

Generally, the parents/caregivers were quite positive about the WCC. Most participants indicated that they were pleased with the support for the basic

things. Some participants were less satisfied, because of the quick judgement about their child or because they became insecure about the development of their child after a WCC consult. *“Occasionally, I go home with a bad feeling. Am I doing well? Yes, you usually do well, but still...”* Mother of a child with ALD.

Most participants liked the approach of the pediatrician and nurse. *“Actually good [the approach]. Yes, very directly, but clearly.”* Mother of a child with ALD. Some participants indicated that some WCC pediatricians were somewhat curtly. They also indicated that they were not given the opportunity to react on the given advice by the WCC-professional. *“And then, in that sense, I do not really feel the space to say I do not agree or uh”* Mother of a child without ALD.

Almost no caregiver believed that the WCC professional obtained a clear picture of the language development of their child during the language screening at a consult. They added that children act differently at home compared to a WCC-setting which makes it difficult to acquire a comprehensive overview of language development during a consult. Many participants missed the inclusion of the environment of the child during the language screening. *“Yes, I always think that, at the clinic, when they [the child] have to show something then you will just see that it does not happen. And they may not be able to make a home visit, but then they would see how things are going in the home situation.”* Mother of a child with ALD. Caregivers also used the word ‘snapshot’; professionals see the child for a short moment. However, some participants indicated that ALD was well detected despite the short screening time. Participants of the second round added that parental report could compensate for the measurement at only one point in time because they can report on the entire development of their child. Furthermore, some participants mentioned that the questions of the language screening posed to the child did not fit the child’s experience world. *“And the first words that they apparently have to know first is ‘ball’, but our X has nothing with a ball so he cannot say that word either. Our X world is made up of tractors and cows because that is also daddy’s world.”* Mother of a child with ALD.

Participants of children with ALD also shared their feelings about being confronted with the outcome. Some participants needed time for accepting that their child had an ALD. *“With her, yes with her it is very clear that she has a developmental delay of one year and five months. That is not nothing. In the*

beginning, it is that confrontation, but then it comes the point that you have to start accepting. That costs a lot of energy.” Mother of a child with an ALD. Other participants had a more ‘wait-and-see’ approach or took immediate action by contacting a speech language therapists for example. Participants of the second round added that they would like to make an informed decision themselves based on good advice. *“Yes, if that judgment is well-founded, I would be able to go along well with that. But then I also think about it myself first. As a parent, you have the best overview of your child and about the development of your child. I would not see it as binding.”* Mother of a child without ALD.

Desires Regarding the WCC

This subtheme describes the desires with respect to the WCC. This subtheme consists of the following codes: procedure WCC, screening professional, desired way of screening, practical tips, outcome and duration language screening. Almost all of the participants said that the nurse should screen together with the caregivers and the playgroup leader, i.e., a professional who frequently interacts with the child. Another desire was setting up a group consultation with changing healthcare professionals. Participants of the second round added that it would be beneficial to have a consultation hour with professionals who can give advice with ‘not so urgent issues’. Participants also suggested making a special language appointment for consultation at a critical age in language development. *“When X was still a baby, you could come to a consultation hour every week and if you had any questions you could ask if you were uncertain. Unfortunately, this is abolished because of the retrenchment.”* Mother of a child without ALD.

Furthermore, caregivers emphasized the importance of including the parental perspective in language screening. Only with other professionals and caregivers, a comprehensive overview of the development of a child can be obtained. *“I think, for example, if they see a problem. Yes then I would always inform the parents first. I would always be for cooperation. That a WCC, for example, also has contact with a teacher, a pediatrician, or the day nursery. That would actually be necessary now, as this society is now.”* Mother of a child without ALD.

Nearly all of the participants desired the inclusion of the home environment of the child in the language screening, and advised that a WCC nurse should make a home visit. Furthermore, they suggested that WCC professionals should have additional specialized knowledge on (language) development and about approaching young children and on playful and child-oriented screening.

Caregivers gave also many practical tips regarding the language screening: using a book with pictures in order to provoke language from the child, sending a questionnaire about language development to the caregivers prior to the health care visit and then focus on the questions that arose, or to use modern media such as recording the language of the child and thereby assessing the language not just in one moment but over time. *“I would think join the modern media. Go look for more; just make some recordings and take these with you. It remains a snapshot. You enter that room, and you have to have a conversation with someone. That I would say, you have to record some things of your child and bring that with you.”* Mother of a child without ALD.

Most participants prefer to that the outcomes of the screening to be expressed in growth curves instead of absolute scores. One of the reasons they provided is that it visualizes the development of their child. *“Well, I think those curves are fine. Just like a growth curve. There is no right or wrong. Where is the average? Is he above or under the line? Good or bad is so definite.”* Mother of a child with ALD.

Finally, participants desired information regarding language development and about what they can do best if their child has ALD or is multilingual. They indicated that a WCC-professional must listen to them carefully and that a caregiver has to decide what is best for their child. They emphasized that they think there are different good choices for further referral and treatments. *“And then, they advise you to go there or there. But, as a parent, you will always remain the final responsible person. What you will do with it. You decide how to deal with that. But they certainly may have an advisory role in that.”* Mother of a child without ALD.

Parent Factors

Within this subtheme, caregivers discussed their own background and way of thinking about child development. The subtheme ‘parent factors’ includes: the parent is the expert, go their own way, character of the parent, fun to learn to talk. The participants of children with ALD spoke about the importance of language development already from birth and about the importance of early detection of ALD. Participants of children with ALD or participants of multilingual children found it difficult to detect ALD while participants of children without ALD indicated that they were the expert and that they know the best about their child’s development. *“And then the talking, yes. I always had to think. Because some children talk so easily. Because then I think, yes, do they talk like that? Do they put two words in a row?... Yes, he says bye bye. But is that a two- word sentence?”* Mother of a child with ALD.

Many participants discussed the importance of going their own way. Some caregivers had to do so because of a lack of information they were given at the WCC. They also reflected on their own personality and the influence of their personality on their coping strategy. A number of participants indicated that they were down-to-earth and did not let themselves become overwhelmed. Others stated that they were insecure about raising their child while others were flexible or very performance-orientated.

Most participants indicated that they really liked having their children learn to talk. Participants of children with ALD indicated that they liked having their children less frustrated as they learned to speak better and that they, as well as others, now understand their children better. However, one participant said that she no longer liked the language development of her child because she had to deal with it so much because of the ALD of her child.

Child Factors

This subtheme contains the codes language behavior and characteristics of the child. Participants of children with ALD told about how they noticed their child had an ALD. Mostly, their children said nothing by the age of three, i.e., only a few words like ‘daddy’ and ‘mommy’, or could not express short sentences at the age of two and a half. *“Because she actually said nothing about six months ago.*

Yes, only mom and dad and nothing else, and she makes everything clear to me by taking my hand and taking me with her and shows it to me." Mother of a child with ALD. Caregivers also stated their child was unintelligible for themselves and/or for others. In contrast to participants of children with ALD, caregivers of children without ALD told about the speed with which children acquired the language automatically. *"Yes, it's just amazing how he develops everything in such a short time. He only says a few words and words are added from one day to the next."* Mother of a child without ALD.

In addition, caregivers of children with ALD told about possible child factors that are related to the ALD, such as born preterm or having ear tubes. Caregivers of children with ALD also often stressed what, besides the ALD, went well in the development of their child. The walking and crawling went well. Only the speech development didn't go well. Mother of a child with ALD. Furthermore, participants talked about the consequences of having an ALD. *"Especially the frustrations because he can't express himself. In turn, we also became frustrated, because we did not understand him. I have always said; if he can talk then, it will be much easier."* Mother of a child with ALD.

Codes, Subthemes and Themes

Further analyses revealed three themes related to the emerging codes and subthemes: (1) A family centered approach which is about the involvement of families, listening to family caregivers, and including the perspective of caregivers; (2) Interprofessional collaboration which is about the role of the professional and the collaboration between professionals in different settings, and (3) Knowledge that focusses on language development, multilingual development, and children with ALD, about possible referral routes, and about ways to involve young children in the screening, see Table 1.

Table 1. Three themes related to the emerging codes and subthemes based on the interviews of parents of monolingual and multilingual children with- and without atypical language development. The numbers behind the codes refer to the codes that emerged in analysis of the first round (1); the interviews with caregivers of children with ALD, and the second round (2); the interviews with caregivers of children without ALD

Themes / Subthemes	A family centered approach	Interprofessional collaboration	Knowledge
Perception WCC	Other countries (1)	Alternation of the WCC professional (1)	General view (1) Referral (1) Specialist (1) Mean scores (1, in vivo) The basics (1, in vivo)
Perception consult at the WCC	Action at the WCC (1) Offered help (1) Mapping language development (1)	Perception language screening (1)	
Opinions about the WCC	Taking action (1) Opinion about the WCC (1) Approach WCC professional (1) Coping with outcome emotionally (1)		Expectation language screening (1) Opinion language screening (1) Snap shot (1, in vivo)
Desires regarding WCC	Duration WCC (1)	Screening professional (1)	Outcome screening (1) Desired way of language screening (1) Procedure WCC (2) Practical tips (2)
Parent factors	The parent is the expert (1) Fun to learn to talk (1) Go their own way (1, in vivo) Character of the parent (1)		
Child factors	Characteristics of the child (1) Language behavior (2)		

Discussion

The target themes that emerged from the interviews with caregivers of monolingual and multilingual children with and without ALD that can be used to enhance language screening in WCC are: a more family oriented consult, better interprofessional collaboration, and more specific knowledge about (multilingual) language development and different referral routes for children with ALD.

A Family Centered Approach

A target issue towards a more family centered approach is listening to caregivers. Understanding of caregivers' beliefs and experiences are important to effectively promote caregivers' health [13]. A family centered approach in which the perspective of caregivers is taken into account in a WCC is widely recognized but is sometimes overshadowed by the direct services for the child [15]. The focus on the direct services also emerged in our study; caregivers referred to this as the 'basics' or the 'medical' perspective. Listening to caregivers is important because caregivers cope differently. Caregivers require emotional and informational support from healthcare professionals in order to be encouraged to ask questions, to express their feelings, and to empower them in guiding their young children [16,17].

To discuss the concerns of caregivers more thoroughly, caregivers of our study gave the suggestion to offer group consultations. A group consultation accords with the idea of centering parenting. Centering parenting mostly consists of six to seven caregivers and infants as well as WCC professionals during the first year of life addressing different health issues [18]. Centering parenting might be an opportunity to encourage caregivers to engage in their care, to raise issues of importance to them, to discuss their concerns with other caregivers and WCC professionals, and to have a more family centered approach in a WCC.

Another important issue that emerged from our study was the use of parental report in language screening. Caregivers play an important role in the process of early detection of ALD [3,19]. The use of the parental perspective in language screening can provide valuable information [20] because caregivers are the experts regarding their child's functioning in different settings. Parental

concerns about child development have been found to be accurate indicators of delays in development [21]. Also, the US Preventive Services Task Force (USPSTF) found that parent-administered screening tools performed better than other tools such as trained-examiner instruments [19]. However, in our study, most of the caregivers of children with ALD or multilingual children do not feel confident about their view on the development of their child. This is in accordance with the study of Peng Kiat Pua et al. [22] who ascertained that caregivers of bilingual children were much less consistent in their judgment about their child's first language development compared to kindergarten teachers.

Interprofessional Collaboration

Interprofessional collaboration is another important theme that emerged from the interviews with caregivers. A collaboration between WCC-professionals who has a general view on development and kindergarten professionals who frequently see the child facilitates the possibility to obtain a more comprehensive overview of the development of a child. It is recognized that interprofessional collaboration between health services, children's services, and schools have become increasingly important in a well-functioning healthcare system. No single agency can deliver all of the help that is necessary for children with language and communication needs [5].

Insufficient collaboration between different agencies can lead to misunderstandings and can be bothersome to caregivers. The study of Ødegård and Strype [23] showed that key elements of effective interprofessional collaboration as perceived by the professionals working with children with mental health problems were: motivation, group leadership, social support, and organizational culture. In order to enhance interprofessional collaboration between WCCs, kindergartens, and school, it is important to motivate the different professionals for it by providing evidence-based reasons why it is better to do so. Furthermore, it has to be clear who is the case manager who has an overview of which care pathways a child has taken and what the outcomes were.

Finally, it is important to facilitate interprofessional collaboration by creating opportunities for this in the collaborating organizations as well. Characteristics of an excellent screening management are family centered, use interdisciplinary collaboration, and are responsive to the diverse needs and important issues of both children and families [24].

Knowledge

A last important theme that emerged from the interviews was the lack of knowledge about (multilingual) language development and referral routes for children with ALD. In accordance with our findings, the BCRP in the UK revealed that caregivers and children desire that professionals, family, and friends acquire more knowledge and awareness about language disorders [5,25]. Caregivers need clear information that makes it easier for them to take responsibility for the upbringing of their child themselves [17]. Reaching consensus regarding clinical signs was an important contribution to the awareness of ALD among healthcare professionals, teachers, and caregivers [3,11]. However, clinical signs for multilingual children are challenging because norms for multilingual children are mostly not available and because the features in language development that might refer to an ALD are very similar to the features in the language development of children who acquire a second language [22].

A recognized way to support caregivers using a more family centered approach and providing knowledge and support is Shared Decision Making (SDM) [26]. Using SDM, WCC-professionals and caregivers are involved and come together to provide the best guidance for the child. The strength of SDM is that caregivers are invited to think along (which gives them empowerment), that the options for different care pathways are discussed, and that caregivers can make a decision based on complete information. This implies that the WCC professionals must listen carefully to the needs of caregivers and have sufficient knowledge about ALD and care pathways in order to support caregivers making their own choices. Adequate information and support from the WCC professionals can help caregivers make their own choices instead of giving top down advice. In this way, a caregiver is more invited to actually take action because it fits with their own perspective and thoughts about raising a child [27]. A good relationship between caregivers and WCC professionals facilitates the alignment with the desires of caregivers and to build trust with the family [28].

Limitations of the Study

A limitation of this study was that the caregivers of multilingual children were mostly multilingual themselves; however, the interviews were conducted in the Dutch language. Consequently, the multilingual caregivers may have had

difficulties expressing their experiences with a WCC. The transcripts of the interviews with multilingual caregivers showed that these caregivers used shorter sentences and that they had more word-finding problems.

Although we have collected interesting data in the individual system (perception, opinions and desires of caregivers) and in the micro system (health services) of Bronfenbrenner [9], it would have been of added value to also interview WCC professionals and professionals from other children's services (e.g. kindergarten) and schools. In this way it would have been possible to check whether these professionals recognize the themes mentioned by the care providers. Moreover, the micro system and the exosystem of Bronfenbrenner [9] could have been investigated more thoroughly in this way. Additionally, more insight into the perception, opinions and wishes of professionals in other children's services and schools can contribute to better interdisciplinary cooperation. And this is important because we assume that one agency cannot provide all the help needed for children with language and communication needs. In addition, it would be of added value to interview healthcare professionals in other countries to gain more insight into the pros and cons of other healthcare systems. The obtained advantages and disadvantages of the different forms of screening procedures in the WCC could be submitted to the local policy to jointly decide which changes need to be made to improve the structural language screening in WCC.

Conclusions

The target themes that emerged from the interviews about the perception, opinions, and desires of caregivers and that can be used to enhance language screening in WCC are: a more family oriented consult, better interprofessional collaboration, and more specific knowledge about (multilingual) language development and different referral routes for children with ALD. A more family centered approach in which WCC professionals caregivers is very important as not all caregivers feel confident about their own view on their child's development. WCC professionals should use information provided by caregivers when screening language development and involve caregivers when providing advice. In addition, caregivers suggest that making home visits in which the WCC professional can playfully observe the child and talk with the family can help to obtain a better estimation of the language functioning of the child. Furthermore, interdisciplinary

collaboration is recommended in the structural WCC screening for ALD as no single discipline can have an overview of language development and can deliver all of the help that is necessary for children with ALD. Finally, additional adequate knowledge about language development, ALD, and referral routes can help caregivers to make their own choices is needed.

Acknowledgements

This work was supported by the Netherlands Organization of Health Research and Development (ZonMw) [grant number: 200330002, 2017/10900/ZONMW]. We are also very grateful to Anke Bankers and Matthijs van Pelt for their support in the data collection.

References

1. Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord*. 2000 Apr-Jun;35(2):165-188.
2. Boh A, Johnson L. Universal screening to promote early identification of developmental delays: Exploring childcare providers' beliefs and practices. *Early Child Development and Care*. 2017;16, 1-15.
3. Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T, Catalise Consortium. CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one*. 2016;11(7), e0158753.
4. Broomfield J, Dodd B. Is speech and language therapy effective for children with primary speech and language impairment? Report of a randomized control trial. *Int J Lang Comm Dis*. 2011;46(6):628-640.
5. Bercow J. The Bercow Report: A review of services for children and young people (0-19) with speech, language and communication needs. 2008.
6. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics*. 2006;117(2):e298-e319.
7. Oberklaid F, Efron D. Developmental delay-identification and management. *Aust Fam Physician*. 2005;34, 739-742.
8. Huneke A, Lascelles L. Terminology mayhem: Why it matters-the ramifications for family caregivers and families. *Int J Lang Comm Dis*. 2014;49(1), 410- 412.
9. Bronfenbrenner U. The ecology of human development: Experiments by nature and design. Cambridge: Harvard University Press. ISBN 0-674-22457-4. 1979
10. Conti-Ramsden G, Durkin K. Language development and assessment in the preschool period. *Neuropsychol Rev*. 2012;22(4):384-401.
11. Visser-Bochane MI, Gerrits E, van der Schans CP, Krijnen WP, Reijneveld SA, Luinge MR. Atypical speech and language development: a consensus study on clinical signs in the Netherlands. *Int J Lang Comm Dis*. 2017;52(1), 10-20.
12. Swift JK, Callahan JL. The impact of client treatment preferences on outcome: a meta-analysis. *J Clin Psychol*. 2009;65, 368-381.
13. Erlingsson CL, Magnusson L, Hanson E. Family caregivers' health in connection with providing care. *Qual Health Res*. 2011;22, 640-655.
14. Strauss C, Quinn N, Corbin J. Basics of qualitative research: Techniques and procedures for developing grounded theory. SAGE Publications. 1998.
15. King G, Williams L, Goldberg H. Family-oriented services in pediatric rehabilitation: A scoping review and framework to promote parent and family wellness. *Child Care Hlth Dev*. 2017;43, 334- 347.
16. Roche B, Cowley S, Salt N, Scammell A, Malone M, Savil P, et al. Reassurance or judgement? Family caregivers' views on the delivery of child health surveillance programmes. *Fam Pract*. 2005;22, 507-512.
17. Fägerskiöld AM, EK AC. Expectations of the child health nurse in Sweden: Two perspectives. *Int Nurs Rev*. 2003;50, 119-128.

18. Bloomfield J, Rising SS. Centering parenting: An innovative dyad model for group mother- infant care. *J Midwifery Wom Heal.* 2013;58, 683–689.
19. Siu AL. Screening for speech and language delay and disorders in children aged 5 years or younger: US preventive services task force recommendation statement. *Pediatrics.* 2015;136(2), e474–e481.
20. Miller LE, Perkins KA, Dai YG, Fein DA. Comparison of parent report and direct assessment of child skills in toddlers. *Res Autism Spec Dis.* 2017;41, 57–65.
21. Glascoe FP. Evidence-based approach to developmental and behavioural surveillance using family caregivers' concerns. *Child Care Hlth Dev.* 2000;26, 137–149.
22. Peng Kiat Pua E, Lay Choo Lee M, Rickard Liowa SJ. Screening bilingual preschoolers for language difficulties: Utility of teacher and parent reports. *J Speech Lang Hear Res.* 2017;60, 950–968.
23. Ødegård A, Strype J. Perceptions of interprofessional collaboration within child mental health care in Norway. *J Interprof Care.* 2009;23, 286–296.
24. Kahn L, Hurth J, Kasprzak CM, Diefendorf MJ, Goode SE, Ringwalt, SS. The national early childhood technical assistance center model for long-term systems change. *Top early child spec.* 2008;29(1), 24–39.
25. Dockrell J, Lindsay G, Roulstone S, Law J. Supporting children with speech, language and communication needs: An overview of the results of the better communication research Programm. *Int J Lang Commun Disord.* 2014;49, 543–557.
26. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Loyd A, Kinnersley P, et al. (2012). Shared decision making: A model for clinical practice. *J Gen Intern Med.* 2012;27, 1361–1367.
27. Eronen R, Pincombe J, Calabretto H. The role of child health nurses in supporting family caregivers of young infants. *Collegian.* 2010;17, 131–141.
28. De La Cuesta. Relationships in health visiting: Enabling and mediating. *Int J Nurs Stud.* 1994;31, 451–459.

The background of the page is a light gray with a complex, organic pattern of wavy, flowing lines in various shades of gray. Scattered throughout this pattern are numerous calligraphic letters and symbols in a variety of styles, including serif, script, and sans-serif. Some letters are large and prominent, while others are small and subtle. The overall effect is a rich, textured backdrop that complements the central text.

8

General discussion

The aim of the present thesis was to contribute to the improvement of the identification of atypical language development in children. We have gained additional insight into typical and atypical language development (Chapters 2, 3, and 5), developed a new language screening instrument (Chapters 5 and 6), and obtained perspective into and future directions for language screening in well-child care (Chapters 4 and 7).

In this general discussion, a brief overview of the main findings as presented in the previous chapters will be provided. Next, these findings will be addressed in a broader perspective to explicate how they contribute to the concept of atypical language development and the improvement of language screening in public health. Additionally, the methodological strengths and limitations will be discussed. The chapter concludes with implications for research and practice and an overall conclusion.

Main findings

We aimed at contributing to improving the identification of children with atypical language by answering six research questions, as summarized below.

The first research question (Chapter 2) was: What is the consensus of speech language professionals on red flags for developmental language disorder? In a national sample of speech language professionals, we ascertained that consensus regarding atypical language development concerned a list of 124 clinical signs of atypical speech and language development in children from one to six years of age. The 34 most urgent clinical signs, which we referred to as ‘red flags’, are symptoms of severe difficulties that may indicate a range of underlying concerns including a developmental language disorder, dyslexia, attention-deficit hyperactivity disorder, or autism spectrum disorder.

The second research question (Chapter 3) was: Which characteristics of auditory processing disorder overlap with characteristics of language disorder? In a systematic review, 13 studies with moderate methodological quality met the inclusion criteria. These studies demonstrated that children diagnosed with various neurodevelopmental disorders (developmental language disorder, auditory processing disorder, dyslexia, attention-deficit hyperactivity disorder, and learning disorder) perform equally on tests of intelligence, memory or

attention, and language. There were only minimal differences in performance for sensory and perceptual functioning tasks between groups of children with these various neurodevelopmental disorders. The following characteristics overlap for groups of children with an auditory processing disorder and children with a language disorder: IQ, working memory, language, and reading tests. Hence, it is important to have an interprofessional focus on an individual child's performance to determine appropriate treatment.

The third research question (Chapter 4) was: What is the validity of the current preventive child health care language screening protocol at age two years, and does the clinical decision of the professional affect the validity? In a matched sample of children that failed and children that passed the 2-year language screening, we found that the current language screening protocol for two-year old children was able to detect current and later language problems with rather few protocol deviations by the professionals. We assessed the children's language development at the age of two years and one year later at the age of three years to determine the concurrent and predictive validity of the screening and of protocol deviations. The concurrent and predictive validity of the screening was sufficient for population-based screening. The professionals showed high adherence to the protocol, therefore, the rather few deviations did not affect the validity of the screening.

The fourth research question (Chapter 5) was: What set of clear and distinctive milestones empirically reflects language development in children from one to six years of age? A community-based sample of 1,381 parents reported on milestones in their child's language development. We used nonparametric item response theory analysis to identify the most distinctive milestones that could describe language development in children of one to six years. This analysis resulted in a set of 26 milestones in language development that we titled the Early Language Scale (ELS).

The fifth research question (Chapter 6) was: What is the validity of the newly developed Early Language Scale? We assessed a community-based sample of 265 children aged one to six years with extended language tests to validate the Early Language Scale (ELS). We ascertained a sensitivity of .62 and a specificity of .93, which is suitable for community-based screening. The ELS was valid for detecting atypical language in children aged one to six in community-based settings.

The sixth research question (Chapter 7) was: What are the perceptions, opinions, and desires of caregivers of children with and without atypical language development regarding language screening at the preventive child health care? We interviewed a total sample of 38 caregivers of monolingual and multilingual children with and without atypical language development. We aimed to explore their perceptions, opinions, and desires regarding language screening in order to provide relevant themes for improving structural language screening. The themes that emerged from the interviews to enhance structural screening were: implement a more family centered approach, implement interprofessional collaboration in the structural screening, and provide a greater amount of specific knowledge to parents about language development and possible referral routes for children with atypical language development.

Below, we will address these results in more detail and will explicate how these findings contribute to more insight into symptoms and causes of atypical language development in young children and to the improvement of language screening in public health. Three overarching themes emerged from the findings of the thesis, that is, knowledge regarding typical and atypical language development, the new screening instrument and its value, and language screening in well child care.

Reflection on main findings

Language development:

Milestones in language development, such as producing two-word sentences, are indicators of active change in language development. Although the language a child develops depends on the context in which the child grows up, each child normally begins with babbling followed by a first word and, thereafter, two-word phrases, that is, the language milestones (Chapter 4). The fact that every child, regardless of the language it is learning, shows the same sequence in the development of language milestones could indicate an innate mechanism that makes this possible. However, the age at which children achieve the next milestone may differ depending on a child's level of development and on environmental stimuli in the form of linguistic variables (e.g. words) as well as gains from communication with others [1]. In summary, language development

is achieved through an interplay between innate mechanisms, the child's level of development, and the environment in which a child grows up.

The milestones in language development were identified as representing a one dimensional scale (Chapter 5). Thus, for any cause of an atypical language development, the expression in the effects on passing of milestones is rather similar. All children achieve the language milestones in this same order, indicating a typical language development. However, absence or late achievement of a language milestone signaling an atypical language development may not be specific for a developmental language disorder but could, for example, also be indicative of other underlying cognitive impairments. Moreover, problems of children with one impairment are often multifactorial, and characteristics of various neurodevelopmental disorders show substantial overlap (Chapter 3) [2]. An explanation could be that the various disorders have a partially overlapping etiology, i.e., disorders that appear similar at the behavioral level may have different causal origins. The causal model, proposed by Bishop and Snowling [3], illustrates four distinct levels: the level of etiology, neurobiology, cognition, and behavior that are interrelated (Figure 1). The etiological level includes both genetic and environmental factors that influence neurobiology. The (neuro) biological level regards abnormalities of the brain's structure and function. The level of cognition regards cognitive processes such as phonological processing or grammatical morphology. The level of behaviour regards all observed behavior such as speaking or understanding. Two children with the same observed red flag may present with different profiles of cognitive disorders. For example, 'no first words' could reflect a lack of communication intent due to an autism spectrum disorder or could reflect problems with speech perception or production [4-6]. Therefore, the one dimensional scale of language milestones detects atypical language development but indicates nothing about the underlying cause of the delay.

Figure 1 Levels of Causation for developmental disorders [3]

Many red flags resulting from the consensus study by the professionals (Chapter 2) correspond with the language milestones included in the ELS (Chapter 5), supporting its validity. However, compared with the empirical evidence on language milestones, the professionals are reluctant to mark a milestone as a red flag. Milestones at a certain age, according to the ELS, become a red flag at an older age based on an expert opinion. For example, milestone 11 ‘Is your child able to combine two words?’ was achieved by 85% of the children at an age of 24 months. Next, the professionals consented on ‘not saying two word sentences’ being a red flag at an age of three years (36-47 months old). The expert opinion regarding red flags has a strong evidence base that could empirically underpin and refine the ages of children at which a milestone becomes a red flag and strongly confirms the validity of the ELS. Next, the ELS will generally detect an atypical development earlier than detection that is based on the professionals’ consensus.

The Early Language Scale as new screening instrument

We constructed an instrument to screen for atypical language development that we titled the Early Language Scale (ELS). The instrument performed well on several psychometric properties and can serve as a basis for language screening in child health care (Chapters 5 and 6). However, an assessment of the quality of the ELS using COTAN criteria [7] revealed two points that could be improved. First, the ELS lacks a formal manual per age category. We provided a brief instruction for assessment and a score form including instructions for scoring and interpretation of the result (Appendix) but no manual that thoroughly describes the test instruction. More specifically, the ELS aims to identify atypical language development in children aged one to six years. However, not all 26 items of the ELS are applicable for all ages of children. The items of the ELS describe language development, therefore, the first items are applicable for young children, and the last items are only relevant for older children. We could adjust the test instruction for specific ages by using the ages in months at which 15 – 50 – 85 percentage of the children in our sample pass a specific milestone. We provided a cut-off score for passing or failing the screening, but we did not yet standardize rules for start or end points for the various ages. Second, the specificity of the ELS was high, but the sensitivity was moderate. For population-based screening and for repeated measurement with the ELS, we believe the effectiveness is sufficient, however, attempts to improve the sensitivity while maintaining the high specificity are recommended. The ELS thus performs sufficiently on psychometric properties and can also meet all COTAN criteria with relatively less additional effort.

Language screening in well child care

The currently used language screening in Dutch well-child care can easily be supplemented with the ELS to improve the identification of children with atypical language development. The current language screening protocol in Dutch well-child care prescribes monitoring the language development of children from birth and to perform a standardized language screening at age two years [8]. The protocol also includes a standardized reassessment at age two-and-a-half years. Recently, Diepeveen [9] concluded that ‘a combined approach using their concise tool of language milestones at age visit 24, 35 and 45 months of age in combination with developmental surveillance (the currently used van Wiechen)

is the best approach to identify children as early as possible'. The concise tool refers to five language milestones that are currently assessed at well-child visits. It expands the currently used protocol with a referral rule for the 24 months visit and for the 45 months visit. What could the ELS contribute to the current language screening?

We found that excellence of a screening management for language delay requires that it is family centered, based on interdisciplinary collaboration, and provides a greater amount of specific knowledge about language development and different referral routes for children with atypical language development (Chapter 7). Next, the contribution of the ELS within these three themes will be discussed. *Family centered care* provides that the provision of health care occurs in a partnership in which families and professionals work together for the child [10,11]. The ELS is a parent-report instrument that has several positive aspects within the paradigm of family centered care. First, by filling in the ELS, parents are involved in the monitoring of their child's development. Repeated assessment using the ELS will engage parents in their child's language development which contributes to a healthy language development and to early identification of any problems in this development. Second, parents are prepared with knowledge regarding language development for the well-child visits, this will enhance the process of shared decision-making [12]. Last, the currently used language screening protocol is valid for the detection of language problems at the well-child visit for two-year old children. However, the ELS is valid for detection in the age range of one- to six- year-old children. Thereby, the ELS can be used in child health care organizations with various or flexible well-child contact moments. The ELS may be a tool to facilitate the partnership of parents and professionals for identifying children with atypical language development.

Regarding *interdisciplinary collaboration*, the ELS could contribute to the current language screening protocol in Dutch well-child care by offering a uniform tool that can be applied in both daycare centers and early educational settings because of the wide age range of the ELS. Use of the same instrument in well-child care, preschool care, and early educational settings facilitates interdisciplinary collaboration because the same instrument may allow for smoother communication between disciplines. It is recognized by parents/ caregivers that interprofessional collaboration between health services, children's

services, and schools has become increasingly important in a well-functioning healthcare system. No single agency can deliver all of the help that is necessary for children with language and communication needs [13]. A child's language is uniformly monitored by using the ELS in various settings that collaborate. This might contribute to early identification of atypical language development.

In a family centered approach, parents and professionals work together for the child, and it is important that the *parents are well-informed* in order to be able to make decisions. Therefore, the professional needs to provide information to them regarding language development and various possible referral routes. Monitoring of language development and decision making can be supported with a visualization of the development. For example, for the development of height and weight, the absolute scores of a child are routinely visualized in growth curves that consist of a series of percentile curves that illustrate the distribution of height or weight in children. In these curves, an individual child can be plotted. Individual child curves that follow the typical developmental path show that a child develops in a typical manner. Atypical development becomes visual when a child's curve does not follow the typical curves, for example, a curve that flattens or shows dips. Such a visualization is helpful for parents and professionals for obtaining insights into the development of a child. Parents are now expressing a need for a growth curve of language development (Chapter 7). The milestones of the ELS can provide input for this development and inform parents about language development.

Methodological considerations

We used both qualitative and quantitative research methods to contribute to the improvement of language screening. Below, we will discuss the quality of the sample, the quality of information, and the causal inferences that can be made.

Quality of the sample

In total, we used four different samples in our studies. All of the samples were community-based, fitting the purpose of the thesis. The two samples for our qualitative studies (Chapters 2 and 7) were both representative for the national population. The sample for our Delphi study in Chapter 2 included language

professionals working with children with atypical language development from all over the country and from all types of settings where children with atypical language development are seen for identification, diagnostics, or guidance. The second qualitative study included parents (Chapter 7). By enriching the results from parents with typically developing children (monolingual and bilingual) with results from parents with atypically developing children, we provide the broad range of experiences of our national community.

An evident strength of our quantitative studies is the inclusion of parents and children from the community who were recruited via an extensive procedure in community services. The cohort sample for the evaluation of the current language screening was recruited via professionals working in four well-child clinics selected from urban and rural areas. An evident strength of the cohort study was the high adherence of the parents in this sample. We only had one withdrawal at the reassessment after one year. The sample for the identification of language milestones was recruited by professionals working in randomly selected daycare centers as well as well-child clinics and schools, resulting in a large sample ($N=1,231$).

The samples of our studies also had a number of limitations. The sample used for the validation of the early language screening instrument was a weighted subsample of the sample we used for the identification of language milestones. In both the sample and the subsample, some response bias occurred as there was a slight overrepresentation of mothers with a high education level compared to national averages. We could not compare our sample with the non-response as our sampling procedure proved unsuitable for obtaining information on non-response and response rate. The screening instrument resulting from our research is therefore based on a sample with some form of response bias. When implementing the instrument, it is advisable to check validity in the target population. Another limitation of our community-based sampling is that we could not analyze outcomes for specific subgroups of the heterogeneous population including, for example, monolingual and multilingual children as well as children from parents with low and high educational backgrounds because these subgroups were too small. We increased the generalizability of our studies by recruiting participants via different institutions and from all parts of the country. However, variations in subgroups might be present within the population and cannot be fully charted given this design.

Quality of information

Regarding the quality of the obtained qualitative data, we were able to profoundly study the experience of professionals by performing a consensus study and of parents by conducting interviews. A strength regarding the quality of the findings was our use of parents from various backgrounds (Chapter 7) for which the results from a first sample were enriched with the results of parents with other backgrounds until saturation was achieved. In our other qualitative sample (Chapter 2), the professionals were all speech-language professionals. Other professionals that also work with children with atypical language development might have enriched the results regarding red flags in language development. For example, professionals in (early) educational settings such as teachers or kindergarten teachers observe social communicational behavior of a child interacting in the educational setting. They might have inserted red flags regarding social communicational behavior that are applicable in their setting.

Regarding the quality of our quantitative data, first we reflect on our systematic review (Chapter 3) and then on the data for the construction of the ELS (Chapters 5 and 6). The information from the literature in our systematic review stemmed from the perspective of an auditory processing disorder. The research question in this study was to determine which characteristics of an auditory processing disorder overlap with characteristics of children with developmental language disorder and not the other way around. A study that began with characteristics of children with a developmental language disorder compared to children with other (neuro)developmental disorders would enrich the results. However, the findings on most auditory tests and questionnaires as well as tests of intelligence, attention, memory, language, and reading of the various groups of children did not significantly differ thus might be central in a new systematic review.

Last, a limitation may be that we determined cutoffs based on parents' reports for use in a clinical setting from data obtained in a research setting. However, a research setting differs from the clinical health care setting in which the results of the children are discussed by professionals and are not anonymous as they are in the research setting. Cut-offs based on an anonymous research setting may be too high for use in the individual health care setting [14]. Hoofs

and colleagues found lower mean scores on the Strength and Difficulties Questionnaire (SDQ) indicating fewer problems in the individual health care setting than in the research setting. If we begin using the ELS in the individual health care setting, parents might also be overly positive in scoring it. Therefore, children that fail the ELS in the clinical setting are even more likely to have language problems and to need a referral.

Causality and confounding

Our studies had several strengths with regard to causality. The first strength is that the construct of the ELS was based on a process of evidence-based deduction. Instruments designed for language screening need to sensitively identify children with atypical language. Thus, the items of the screening instrument need to reflect the underlying construct of language development. The construct was predefined as including speech, language, and communication (Chapter 2). However, during our study on the identification of language milestones using a non-parametric item response theory (IRT) analysis, the construct was narrowed to language and communication (Chapter 4). The IRT study thereby confirmed that speech and language are separate constructs and that speech items would not contribute to the identification of children with atypical language development [9].

Another strength is that the analyses we used for the item selection of the ELS resulted in specific information on items that were selected and items that were rejected. This IRT analysis is helpful for selecting a short subset of items with adequate precision [15]. Additionally, an IRT analysis provides information on the validity and difficulty of all of the individual items [16,17]. In contrast, a classical test theory provides reliability and validity information of a scale (i.e. a set of items). Using IRT analysis, we identified items that contributed to the rank ordering of children on the construct 'language development'. We used the Mokken-model and not the more widely known Rasch model because the later model is more stringent. In more stringent models, fewer items will fit. The current ELS includes valid items that are suitable for further development, such as adaptive testing.

A limitation of our study regarding the validity of the currently used language screening protocol and the adherence of the professional was that participating institutions had recently implemented the protocol in their workflow. At the time of our research, half of the clinics in the country had implemented the protocol. The professionals participating in our study had recently had a protocol training prior to the implementation of the screening protocol. It is known that adherence rates improve after implementation and particularly if training is part of the implementation strategy [18]. As for adherence, the high adherence we found may have been flattered by the moment of our study. In order to maintain this high adherence, the protocol must be properly implemented into the workflow.

Implications

Implications for practice and policy

Our findings have several implications for practice and policies on preventive child health care and early educational settings. First, the easily observable characteristics of typical and atypical language development in children aged one to six years should become common knowledge for all professionals working with children in the aforementioned age range. To support the health professionals in well-child care, the evidence for language screening should be included in undergraduate and continuing medical and nursing education. The evidence regarding typical and atypical language development and the overlap of developmental disorders provided in Chapters 1, 2, and 4 is essential basic knowledge for professionals working in preventive child health care but also for professionals working in educational settings with children from one to six years of age. Additionally, public awareness and education campaigns can target a wider group of stakeholders including the general public, parents, and decision-makers [19].

Second, the currently used language screening is adequate and can be enriched with the ELS. This new instrument, already mentioned in the national protocol, should be implemented in the workflow of all well-child clinics and also in early educational settings. Parents should complete the ELS before a well-child visit so that results can be discussed. This should be mandatory for every child

at the age of one to six years and could be listed as such in future professional guidelines for well-child care.

Third, the insight in the overlap of red flags for various (neurodevelopmental) disorders (Chapter 3) confirmed that red flags in language development might be useful in the early identification of developmental problems such as an autism spectrum disorder or intellectual disability [20,21]. The identified red flags for atypical language development were found to be valid for monitoring the language development of children aged one to six years old (Chapters 5 and 6). Therefore, it is wise to monitor the language development of children and refer children with an atypical development as soon as possible.

Finally, parents indicated that they require knowledge regarding language development. They express the need more often if they experience concerns regarding the development of their child. Other cases are also known; parents do not worry when they should be worried. Therefore, common knowledge regarding milestones and red flags might help in gaining awareness on typical and atypical development. Parents also use growth curves for height and weight which is a visualization of two other developmental paths. Language should be the third curve!

Implications for future research

Our findings have several implications for future research. To begin with, the knowledge from this thesis regarding typical and atypical language needs to be available in courses for all professionals (educated for) working with children and for the parents of young children. Further studies are needed to evaluate the effectiveness of these training courses for the various target groups.

Next, our newly developed instrument also requires further research regarding some of its aspects. First, the ELS is developed and applied in research settings in which the assessors were trained. They developed a short assessment instruction. A next step would be to develop a manual for the use in daily practice and develop a computer aided version of the ELS. Second, the cut-off values are based on data from a research setting; these should also be validated in the clinical setting. Third, the sensitivity of the ELS could be improved while maintaining its high specificity. This might be established by longitudinal use of

the instrument. Another option would be to include risk factors such as parental or professionals' concerns [4], family history of language problems [23,24], and protective factors such as no parental or professional concerns. Fourth, the items were assessed on clarity in use, however, their inter-rater reliability has not yet been assessed. This could be done by, e.g., asking both parents to fill in the various items. Fifth, the ELS was developed for the Dutch language. Further studies are needed to assess it in other languages such as English, German, or French.

To conclude, the monitoring and screening of language in Dutch well-child care needs continuous attention. The overarching goal of the screening is the improvement of language outcomes of children. We must continuously evaluate and improve the performance of our language screening related to the language outcomes of children in the longer term. Further research is required to investigate the possible added value of the ELS to the existing protocol and to determine the best ages of a child at which a parent needs to fill in the ELS.

Conclusion

We conclude that screening for delays in language development in the Dutch setting already performs well but might be improved further by including the ELS. This may also hold for other countries and settings. The utilization of the results of this thesis in routine practice may increase the number of capable and confident children from early school age onwards and offer a positive start by largely reducing language disorders and accelerating their early development.

References

1. Hohenberger A, Peltzer-Karpf A. Language learning from the perspective of nonlinear dynamic systems. *Linguistics*. 2009;47(2):481-511.
2. Pennington BF. From single to multiple deficit models of developmental disorders. *Cognition*. 2006;101(2):385-413.
3. Bishop DV, Snowling MJ. Developmental dyslexia and specific language impairment: Same or different? *Psychol Bull*. 2004;130(6):858.
4. Bishop DVM, Snowling MJ, Thompson PA, Greenhalgh T, Catalise Consortium. CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLOS one*. 2016;11(7), e0158753.
5. Dockrell J, Ricketts J, Lindsay G. Understanding speech, language and communication needs: Profiles of need and provision. London: Department for Education. 2012.
6. Eigsti I, Bennetto L, Dadlani MB. Beyond pragmatics: Morphosyntactic development in autism. *J Autism Dev Disord*. 2007;37(6):1007-1023.
7. Evers A, Lucassen W, Meijer R, Sijsma K. COTAN beoordelingssysteem voor de kwaliteit van tests [COTAN assessment system for the quality of tests]. Amsterdam, Netherlands: Nederlands Instituut van Psychologen. 2009.
8. Lanting CI, de Wolff MS, Wiefferink K, Uilenburg N. JGZ-richtlijn taalontwikkeling [well-child care Guideline Language development] 2019.
9. Diepeveen FB, van Dommelen P, Oudesluys-Murphy AM, Verkerk PH. Concise tool based on language milestones identifies children with specific language impairment at 24–45 months of age. *Acta Paediatr*. 2018;107(12):2125-2130.
10. Arango P. Family-centered care. *Acad pediatr*. 2011;11(2):97.
11. Kuo DZ, Houtrow AJ, Arango P, Kuhlthau KA, Simmons JM, Neff JM. Family-centered care: current applications and future directions in pediatric health care. *Matern Child Health J*. 2012;16(2):297-305.
12. Molenaar J, Korstjens I, Hendrix M, de Vries R, Nieuwenhuijze M. Needs of parents and professionals to improve shared decision-making in interprofessional maternity care practice: a qualitative study. *Birth*. 2018;45(3):245-254.
13. Bercow J. The Bercow Report: A review of services for children and young people (0-19) with speech, language and communication needs. 2008.
14. Hoofs H, Jansen N, Mohren D, Jansen M, Kant I. The context dependency of the self-report version of the Strength and Difficulties Questionnaire (SDQ): A cross-sectional study between two administration settings. *PLoS One*. 2015;10(4):e0120930.
15. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual life res*. 2007;16(1):5.
16. Reise SP, Henson JM. A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *J Pers Assess*. 2003;81(2):93-103.
17. Bortolotti SLV, Tezza R, de Andrade DF, Bornia AC, de Sousa Júnior, Afonso Farias. Relevance and advantages of using the item response theory. *Qual Quan*. 2013;1-20.
18. Ista E, van Dijk M, van Achterberg T. Do implementation strategies increase adherence to pain assessment in hospitals? A systematic review. *Int J Nurs Stud*. 2013;50(4):552-568.

19. Sagan A, McDaid D, Rajan S, Farrington J, McKee M. Screening: When is it appropriate and how can we get it right?[Internet]. 2020.
20. Horwitz SM, Irwin JR, Briggs-Gowan MJ, Heenan JMB, Mendoza J, Carter AS. Language delay in a community cohort of young children. *J Am Acad Child Psy.* 2003;42(8):932-940.
21. Reilly S, McKean C, Morgan A, Wake M. Identifying and managing common childhood language and speech impairments. *BMJ.* 2015 May 14;350:h2318.
22. Snowling MJ, Duff FJ, Nash HM, Hulme C. Language profiles and literacy outcomes of children with resolving, emerging, or persisting language impairments. *J Child Psychol Psc.* 2016;57(12):1360-1369.
23. Zambrana IM, Pons F, Eadie P, Ystrom E. Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *Int J Lang Comm Dis.* 2014;49(3):304-316.

The background of the page is a light gray with a complex, organic pattern of wavy lines and irregular shapes. Scattered throughout this pattern are various letters in a serif font, including 'y', 'W', 'A', 'p', 'C', 'X', 'R', 'f', 'g', 'N', 'E', 'L', 'B', 'S', 'G', 't', 'J', 'h', 'F', 'B', 'I', 'O', 'y', 'U', 'z', 'A', 'O', 'S', 'p', 'C', 'K', 'Q', 'S', 'Z', 'a', 'm', 'p', 'w', 'e', 'y', 'U', 'z', 'A', 'O', 'S'. The letters are in different sizes and orientations, creating a textured, artistic effect.

APPENDICES

Summary

The Early Language Scale (ELS)

Samenvatting

Dankwoord

About the author

Research Institute SHARE: previous dissertations

Summary

Language is a basic skill that is essential for almost everything we do in life. Unfortunately, speech and language problems are one of the most reported developmental problems in early childhood. These problems can have major impact on a child, its family, and the society as a whole. Early identification of these problems and referral to adequate treatment may significantly reduce the detrimental effects. Therefore, all children with language problems need to be detected as early as possible. Early identification of developmental disorders, including language disorders is a core task of preventive child health services. To ensure early identification, preventive child health care organizations recommend standardized approaches for developmental surveillance and screening as an integral function and responsibility of all pediatric health care professionals. However, up till now, not all children who could benefit from early intervention services are identified. Therefore, the general aim of this thesis was to contribute to the improvement of the identification of atypical language development in children. The aim has been translated into six research questions. In this summary the rationale, main methods and results of these studies will be shortly reviewed.

In chapter 2 we describe our study regarding the first research question: *What is the consensus of speech language professionals on red flags for developmental language disorder?* We conducted a Delphi study to achieve a national and valid consensus on clinical signs and red flags (i.e. most urgent clinical signs) for atypical speech language development in children from one to six years of age. In two Delphi rounds, speech language professionals consented on 124 clinical signs of atypical speech and language development in children from one to six years of age. The 34 most urgent clinical signs, which we called ‘red flags’ are symptoms of severe difficulties that may indicate a range of underlying concerns, including developmental language disorder, dyslexia, attention-deficit hyperactivity disorder, or autism spectrum disorder. These findings contribute to the identification of atypical language development, as they make clear which language characteristics were seen as red flags requiring immediate referral for further evaluation. The red flags were identified for children from one to six years, and can thereby be used by professionals in care and education.

In chapter 3 we discuss our systematic review of the literature regarding research question two: *which characteristics of auditory processing disorder overlap with characteristics of language disorder?* It is known that various

neurodevelopmental disorders have overlapping characteristics, making it difficult for clinicians and researchers to separate the disorders. Insights in overlapping and in distinguishing characteristics would be helpful for clinicians from various professions that need to refer, diagnose, or treat children with developmental disorders. We found 13 studies that met the inclusion criteria, all with moderate methodological quality. These studies demonstrated that children diagnosed with various neurodevelopmental disorders (developmental language disorder, auditory processing disorder, dyslexia, attention-deficit hyperactivity disorder, and learning disorder) perform equally on tests of intelligence, memory or attention, and language. There were only minimal differences in performance for sensory and perceptual functioning tasks between groups of children with these various neurodevelopmental disorders. There were no significant differences found regarding IQ, working memory, language, and reading tests between children with language disorders and children with auditory processing disorders. This study showed that there is substantial overlap between children with various developmental disorders. Therefore it is important to have an interprofessional focus on a child's performance to determine which treatment is appropriate. These findings were based on a small number of studies with moderate methodological quality. More high-quality research is needed to better understand the differences and similarities of children with various neurodevelopmental disorders.

In chapter 4 we describe our cohort study that aimed to answer the following research question: *What is the validity of the current preventive child health care language screening protocol at age two years, and does the clinical decision of the professional affect its validity?* We recruited a matched sample of children that failed and children that passed the two-year language screening, resulting in a sample of 124 children. We assessed the children's language development at the age of two years and one year later at the age of three years. The language scores resulting from the language assessment and from the screening at the age of two years (with and without deviations from the protocol by the professionals) were used to assess the concurrent and predictive validity of the screening. We found that the current language screening protocol was able to detect current and later language problems, with rather few protocol deviations made by the professionals. The outcomes of the two-year screening compared

to those on the reference standard at two years (concurrent validity) had a sensitivity of 0.79, and specificity of 0.86. Compared to three-years reference testing (predictive validity), these values were 0.82 and 0.74, respectively. Clinical decisions by the professional that deviated from the protocol were rare (7%) and did not significantly affect the validity of the screening. This study showed that the current language screening at the age of two years performs sufficiently. However, for population-based screening a higher specificity is desirable to minimize over referrals. Moreover, this screening protocol regards two-year old children. Problems in language development might occur after the age of two years, as language develops up to six years. These children will not be identified with the current language screening.

In the following two chapters (5 and 6), we describe the construction and validation of a new language screening instrument for children in the age of one to six years. In the first study, we constructed the instrument using the following research question: *What set of clear and distinctive milestones empirically reflects language development in children from 1 to 6 years of age?* A community-based sample of 1,381 parents reported on milestones in their child's language development. The milestones were derived from clinical signs for atypical language development we found in the Delphi study described in chapter 2, and from currently used screening instruments. All milestones were phrased as yes/no questions on language ability and their comprehensibility for parents was checked. We used nonparametric item response theory analysis to identify the most distinctive milestones that described language development in children of one to six years. The analysis resulted in a set of 26 milestones in language development that we titled the Early Language Scale (ELS).

The second study regarded the validity of the ELS. The research question was *What is the validity of the newly developed Early Language Scale?* To validate the ELS, we assessed 265 children aged one to six years with extended language tests. This sample was a sub sample of the community-based sample in the previous study. We found a sensitivity of 0.62 and a specificity of 0.93. This is suitable for community-based screening, as the high specificity leads to a relatively lower number of children that are referred. This study showed that the ELS was valid for detecting atypical language in children aged one to six in community-based settings. The ELS might provide support to professionals

working in well-child care and early educational settings to quickly identify children (aged one to six years) with an atypical language development. A limitation of the sample in both studies was that there was some overrepresentation of mothers with a high education level. This resulted in norms that would identify higher percentages of children from mothers with a low educational level than from children from mothers with a high educational level, as evidence suggests that children of highly educated mothers (and fathers) have better language. However, the norms are applicable for our sample, including the full range of educational levels. Moreover, better language development is associated with higher school success and better quality of life. Therefore all children with atypical development deserve identification of this crucial deficit.

In chapter 7 we describe a study that explored themes for improvement of structural language screening from the parents' point of view. The research question was: *What are the Perceptions, Opinions and Desires of Caregivers of Children with and without Atypical Language Development regarding language screening at the preventive child health care?* We interviewed a total sample of 38 parents of monolingual and multilingual children with and without atypical language development. The themes that emerged from the interviews to enhance structural screening were: implement a more family centered approach, implement interprofessional collaboration in the structural screening, and provide a greater amount of specific knowledge to parents about language development and possible referral routes for children with atypical language development.

Conclusions

We conclude that screening for delays in language development in the Dutch setting already performs well but might be improved further by including the ELS. This may also hold for other countries and settings. The utilization of the results of this thesis in routine practice may increase the number of capable and confident children from early school age onwards and offer a positive start by largely reducing language disorders and accelerating their early development.

The Early Language Scale (ELS)

Signalering van problemen in de taalontwikkeling
voor kinderen van 1 tot 6 jaar

Naam kind
Datum afname
Leeftijd kind (jaar en maanden)

Bij 'soms' doorvragen: vaker wel dan niet = 1, vaker niet dan wel = 0.
Instap-, afkapwaarden nog niet gestandaardiseerd.

	Vragen over het leren praten van kinderen van 1 tot 6 jaar. (ja=1, nee=0)	1 / 0
1	Zegt uw kind al 'woordjes', Bijvoorbeeld: 'mama', 'papa', 'koek'? Het hoeft nog niet helemaal goed uitgesproken te zijn.	
2	Als u met uw kind speelt met bijvoorbeeld een bal, heeft uw kind dan ook aandacht voor u en de bal?	
3	Begrijpt uw kind opdrachtjes van twee woorden? Bv: 'jas aan' of 'kijk daar'.	
4	Begrijpt uw kind het als u hem/haar iets vraagt? Bv: 'Zullen we een boekje lezen?'	
5	Begrijpt uw kind zinnetjes van 3 woorden? Bv: 'op de stoel' of 'naar de gang'.	
6	Kan uw kind iets aanwijzen dat u noemt? Bv: 'Waar zit je neus?' of 'Waar is de bal?'	
7	Zegt uw kind in totaal ongeveer 10 woordjes?	
8	Kan uw kind 5 benoemde plaatjes in een boek aanwijzen?	
9	Kan uw kind 6 lichaamsdelen aanwijzen bij een pop/zichzelf? Waar zijn de oogjes, mond, buik, voet, haar, hand?	
10	Vraagt uw kind het aan u als hij/zij iets te eten wil, of met speelgoed wil spelen?	
11	Kan uw kind twee woordjes combineren? Bijvoorbeeld: 'papa bal' of 'kijk poes'.	
12	Kan uw kind vier of meer plaatjes van dieren benoemen? Bijvoorbeeld: 'hond', 'poes', 'paard', 'koe'.	
13	Kan uw kind met u om de beurt praten?	
14	Kan uw kind uit zichzelf een gesprekje beginnen?	
15	Staan de woordjes in de zinnen van uw kind meestal op de goede plek?	
16	Vertelt uw kind weleens spontaan een verhaaltje? Bijvoorbeeld over wat uw kind die dag heeft gedaan.	
17	Gebruikt uw kind woorden die iets zeggen over andere woorden? Bijvoorbeeld: 'groot' in 'een groot huis'.	
18	Kan uw kind al een paar kleuren correct benoemen?	
19	Kan uw kind een verhaaltje navertellen bij een aantal plaatjes? Bijvoorbeeld als u een boekje leest met uw kind.	
20	Gebruikt uw kind ook woorden als 'wij', 'hij' en 'zij' in een zin? Bijvoorbeeld: 'Wij hebben een torentje gemaakt.' of 'Hij is helemaal kapot'.	

21	Maakt uw kind zinnen met 'als' of 'en'? Bijvoorbeeld: <i>'Als wij het eten op hebben, gaan we kleien.'</i> of <i>'Die ging een jas aandoen en die ging schoenen aandoen.'</i>	
22	Stelt uw kind vragen met 'waarom'?	
23	Maakt uw kind goede meervoudsvormen? Bijvoorbeeld: <i>'tafels'</i> i.p.v. <i>'tafelen'</i> .	
24	Kan uw kind de volgende zinnen afmaken: Niet zwart maar.... Niet hoog maar.....	
25	Maakt uw kind zinnen met 'omdat'?	
26	Maakt uw kind goede verleden tijd vormen? Bijvoorbeeld: <i>'gedronken'</i> i.p.v. <i>'gedrinken'</i> of <i>'liep'</i> i.p.v. <i>'loopte'</i> .	
	Totaalscore:	

Normtabel Early Language Scale (ELS)

Interpretatie

Minimale totaalscore of hoger: taalontwikkeling voldoende

Lager dan de minimale totaalscore: twijfel over taalontwikkeling

Leeftijd kind		Minimale Totaalscore:
1 jaar	12 – 14 maanden	3
	15 – 17 maanden	7
	18 – 20 maanden	8
	21 – 23 maanden	10
2 jaar	24 – 26 maanden	15
	27 – 29 maanden	16
	30 – 32 maanden	17
	33 – 35 maanden	19
3 jaar	36 – 38 maanden	21
	39 – 41 maanden	22
	42 – 44 maanden	23
	45 – 48 maanden	23
4 jaar	48 – 53 maanden	24
	54 – 59 maanden	24
5 jaar	60 – 71 maanden	25

Samenvatting

Taal is een basisvaardigheid die essentieel is voor bijna alles wat we in het leven doen. Deze vaardigheid ontwikkelt zich niet bij alle kinderen vanzelf. Taalontwikkelingsproblemen zijn helaas de meest voorkomende ontwikkelingsproblemen bij kinderen. Taalontwikkelingsproblemen kunnen grote gevolgen hebben voor een kind, het gezin en de samenleving als geheel. Vroegtijdige signalering van deze problemen en doorverwijzing naar een passende behandeling kunnen de nadelige gevolgen aanzienlijk beperken. Daarom moeten alle kinderen met taalontwikkelingsproblemen zo vroeg mogelijk worden opgespoord. Vroegtijdige opsporing van ontwikkelingsproblemen, waaronder taalontwikkelingsstoornis, is een belangrijke, wettelijk vastgelegde, taak van de Jeugdgezondheidszorg (JGZ). Om vroegtijdige op te kunnen sporen, hanteren organisaties voor jeugdgezondheidszorg een gestandaardiseerde aanpak voor het systematisch volgen van de ontwikkeling van kinderen en het screening van kinderen op ontwikkelingsstoornissen. Dit is een integrale functie en verantwoordelijkheid van alle professionals in de jeugdgezondheidszorg. Tot op vandaag worden niet alle kinderen die baat zouden kunnen hebben bij vroegtijdige interventie op het gebied van taalontwikkeling gesignaleerd. Daarom was het doel van dit proefschrift om bij te dragen aan het verbeteren van het signaleren van atypische taalontwikkeling bij kinderen. Dit doel is vertaald in zes onderzoeksvragen. In deze samenvatting zullen de grondgedachte, de belangrijkste methoden en de resultaten van deze onderzoeken kort besproken worden.

In hoofdstuk 2 beschrijven we ons onderzoek naar de eerste onderzoeksvraag: Wat is de consensus van logopedisten over rode vlaggen voor een taalontwikkelingsstoornis? We hebben een Delphi studie uitgevoerd om nationaal een valide consensus te bereiken over klinische signalen en rode vlaggen (d.w.z. meest urgente klinische signalen) voor atypische spraaktaalontwikkeling bij kinderen van één tot zes jaar oud. In twee Delphi rondes stemden spraaktaalprofessionals (logopedisten en linguïsten) in met 124 klinische signalen van atypische spraak- en taalontwikkeling bij kinderen van één tot zes jaar oud. De 34 meest urgente klinische signalen, die we ‘rode vlaggen’ noemden, zijn symptomen van ernstige moeilijkheden die kunnen wijzen op een reeks onderliggende problemen, waaronder taalontwikkelingsstoornis, dyslexie, aandachtstekortstoornis met hyperactiviteit (ADHD), of

autismespectrumstoornis. Deze bevindingen dragen bij tot de identificatie van atypische taalontwikkeling, omdat nu duidelijk is welke taalkenmerken als rode vlaggen werden beschouwd die onmiddellijke doorverwijzing voor verdere evaluatie vereisen. De rode vlaggen werden geïdentificeerd voor kinderen van één tot zes jaar, en kunnen dus worden gebruikt door professionals in zorg en onderwijs.

In hoofdstuk 3 bespreken we onze systematische review van de literatuur met betrekking tot onderzoeksvraag twee: welke kenmerken van auditieve verwerkingsstoornis overlappen met kenmerken van taalstoornis? Het is bekend dat verschillende neurologische ontwikkelingsstoornissen overlappende kenmerken hebben, waardoor het voor klinici en onderzoekers moeilijk is om stoornissen te onderscheiden. Inzicht in overlap en in het onderscheiden van kenmerken zou nuttig zijn voor klinici uit verschillende beroepsgroepen die kinderen met ontwikkelingsstoornissen moeten doorverwijzen, diagnosticeren of behandelen. Wij vonden 13 studies die voldeden aan de inclusiecriteria, allen van matige methodologische kwaliteit. Deze studies toonden aan dat kinderen gediagnosticeerd met verschillende neurologische ontwikkelingsstoornissen (taalontwikkelingsstoornis, auditieve verwerkingsstoornis, dyslexie, aandachtstekortstoornis met hyperactiviteit (ADHD), en leerstoornis) gelijk presteren op tests van intelligentie, geheugen of aandacht, en taal. Er waren slechts minimale verschillen in prestaties voor zintuiglijke en perceptuele functietaken tussen groepen kinderen met deze verschillende neurologische ontwikkelingsstoornissen. Er werden geen significante verschillen gevonden met betrekking tot IQ, werkgeheugen, taal, en leestesten tussen kinderen met taalstoornissen en kinderen met auditieve verwerkingsstoornissen. Deze studie toonde aan dat er een aanzienlijke overlap bestaat tussen kinderen met verschillende ontwikkelingsstoornissen. Daarom is het belangrijk om een interprofessionele focus te hebben op de prestaties van een kind om te bepalen welke behandeling geschikt is. Deze bevindingen waren gebaseerd op een klein aantal studies van matige methodologische kwaliteit. Meer hoogwaardig onderzoek is nodig om de verschillen en overeenkomsten van kinderen met verschillende neurologische ontwikkelingsstoornissen beter te begrijpen.

In hoofdstuk 4 beschrijven we onze cohortstudie die tot doel had de volgende onderzoeksvraag te beantwoorden: Wat is de validiteit van het huidige

preventieve kindergezondheidszorg taalscreeningsprotocol op de leeftijd van twee jaar, en heeft de klinische beslissing van de professional invloed op de validiteit? We rekruteerden een gematchte steekproef van kinderen die niet slaagden en kinderen die wel slaagden voor de taalscreening op 2-jarige leeftijd. Dit resulteerde in een steekproef van 124 kinderen. We beoordeelden de taalontwikkeling van de kinderen op de leeftijd van twee jaar en een jaar later op de leeftijd van drie jaar. De taalscores van de taaltest en van de screening op de leeftijd van twee jaar (met en zonder afwijkingen van het protocol door de professionals) werden gebruikt om de concurrente en predictieve validiteit van de screening te beoordelen. We vonden dat het huidige protocol voor taalscreening in staat was om huidige en latere taalproblemen op te sporen, met vrij weinig protocolafwijkingen door de professionals. De uitkomsten van de taalscreening op 2-jarige leeftijd vergeleken met die van de referentiestandaard op twee jaar (concurrente validiteit) hadden een sensitiviteit van 0,79, en een specificiteit van 0,86. In vergelijking met de referentietests één jaar later (predictieve validiteit) waren deze waarden respectievelijk 0,82 en 0,74. Klinische beslissingen door de professional die afweken van het protocol waren zeldzaam (7%) en hadden geen significante invloed op de validiteit van de screening. Deze studie toonde aan dat de huidige taalscreening op de leeftijd van twee jaar voldoende presteert. Voor screening op bevolkingsniveau is echter een hogere specificiteit wenselijk om oververwijzing te minimaliseren. Bovendien heeft dit screeningsprotocol betrekking op kinderen van twee jaar oud. Problemen in de taalontwikkeling kunnen zich voordoen na de leeftijd van twee jaar, aangezien taal zich ontwikkelt tot zes jaar. Deze kinderen zullen met de huidige taalscreening niet worden geïdentificeerd.

In de volgende twee hoofdstukken (5 en 6) beschrijven we de constructie en validatie van een nieuw taalscreeningsinstrument voor kinderen in de leeftijd van één tot zes jaar. In de eerste studie hebben we het instrument geconstrueerd aan de hand van de volgende onderzoeksvraag: Welke set van duidelijke en onderscheidende mijlpalen weerspiegelt empirisch de taalontwikkeling bij kinderen van één tot zes jaar? In totaal 1.381 ouders uit heel Nederland rapporteerde over mijlpalen in de taalontwikkeling van hun kind. De mijlpalen waren afgeleid van klinische signalen voor atypische taalontwikkeling die we vonden in de Delphi-studie beschreven in hoofdstuk 2, en van momenteel

gebruikte screeningsinstrumenten. Alle mijlpalen waren geformuleerd als ja/nee vragen over taalvaardigheid van kinderen van één tot zes jaar. De begrijpelijkheid van deze mijlpalen voor ouders werd ook gecontroleerd. We gebruikten niet-parametrische itemresponstheorie analyse om de meest kenmerkende mijlpalen te identificeren die de taalontwikkeling van kinderen van één tot zes jaar beschreven. De analyse resulteerde in een set van 26 mijlpalen in taalontwikkeling die we de Early Language Scale (ELS) noemden.

De tweede studie betrof de validiteit van de ELS. De onderzoeksvraag was: Wat is de validiteit van de nieuw ontwikkelde Early Language Scale? Om de ELS te valideren, beoordeelden we de taalontwikkeling van 265 kinderen van één tot zes jaar met uitgebreide taaltesten. Deze groep kinderen was een op leeftijd en geslacht gestratificeerde steekproef uit de vorige studie. We vonden een sensitiviteit van 0,62 en een specificiteit van 0,93. Dit is geschikt voor screening in een gezonde populatie, aangezien de hoge specificiteit leidt tot een relatief lager aantal kinderen dat wordt doorverwezen. Deze studie toonde aan dat de ELS valide is voor het signaleren van atypische taal bij kinderen van één tot zes jaar in een ‘community-based’ setting, zoals binnen de jeugdgezondheidszorg. De ELS zou ondersteuning kunnen bieden aan professionals werkzaam in de kinderopvang en vroegschoolse educatieve settingen om kinderen (van één tot zes jaar) met een atypische taalontwikkeling snel te signaleren. Een beperking van de steekproef in beide studies was dat er een zekere oververtegenwoordiging was van moeders met een hoog opleidingsniveau. Dit resulteerde in normen die hogere percentages kinderen van moeders met een laag opleidingsniveau zouden identificeren dan van kinderen van moeders met een hoog opleidingsniveau, aangezien bewijs suggereert dat kinderen van hoog opgeleide moeders (en vaders) een betere taalontwikkeling hebben. De normen zijn echter van toepassing op onze steekproef, met inbegrip van de volledige breedte van onderwijsniveaus van de ouders. Bovendien wordt een betere taalontwikkeling geassocieerd met meer schoolsucces en een betere kwaliteit van leven. Daarom verdienen alle kinderen met een atypische taalontwikkeling identificatie van deze cruciale tekortkoming.

In hoofdstuk 7 beschrijven we een onderzoek waarin thema’s voor verbetering van de structurele taalscreening vanuit het gezichtspunt van de ouders werd verkend. De onderzoeksvraag was: Wat zijn de percepties,

opvattingen en wensen van verzorgers van kinderen met en zonder atypische taalontwikkeling met betrekking tot taalscreening in de preventieve jeugdgezondheidszorg? We interviewden een steekproef van 38 ouders van eentalige en meertalige kinderen met en zonder atypische taalontwikkeling. De thema's die uit de interviews naar voren kwamen om de structurele screening te verbeteren waren: een meer gezinsgerichte aanpak implementeren, interprofessionele samenwerking implementeren in de structurele screening, en ouders meer specifieke kennis verschaffen over taalontwikkeling en mogelijke verwijsroutes voor kinderen met atypische taalontwikkeling.

Conclusies

Wij concluderen dat screening op taalachterstand in de Nederlandse Jeugdgezondheidszorg al goed verloopt, maar verder verbeterd kan worden door de ELS te implementeren. Dit geldt mogelijk ook voor andere landen en settingen. Het gebruik van de resultaten uit dit proefschrift in de praktijk kan het aantal capabele en zelfverzekerde kinderen vanaf de vroege schoolleeftijd vergroten en kinderen een positieve start bieden door taalachterstanden grotendeels te verminderen en de vroege taalontwikkeling te versnellen.

Dankwoord

Een proefschrift schrijf je niet alleen. Zonder hulp of begeleiding zou ik dit proefschrift niet hebben kunnen voltooien en daarom wil ik allen die mij geholpen of ondersteund hebben van harte bedanken:

Allereerst wil ik mijn begeleiders bedanken. Mijn eerste promotor Cees van der Schans: **C**onsequent leerde jij mij kritisch denken, **E**fficiënt schrijven met 'landmarkzinnen', **E**thisch denken en handelen. Dank voor jouw altijd **S**cherpe en analytische bijdrage! Mijn tweede promotor Menno Reijneveld: **M**eesterlijk in positieve feedback leerde jij mij **E**fficiënt schrijven, **N**adenken over analyses, uitkomsten en presentatie, jouw enorme **N**etwerk en kennis van volksgezondheidszorg was onmisbaar. Dank voor jouw **O**ntwikkelingsgerichte begeleiding! Mijn promotor Margreet Luinge: **M**otiverend, van dag één tot nu: altijd **A**andacht, **R**espectvol straalde jij **G**eloof uit in mijn kunnen, je bracht **R**ust wanneer nodig en in gesprekken vinden wij makkelijk **E**nergie. Dank voor jouw **E**chtheid en **T**oegewijde begeleiding! Ook een woord van dank voor alle co-auteurs en specifiek Wim Krijnen voor alle hulp en advies bij de statistische analyses.

Ellen en Kris, lieve paranimfen, dank voor jullie steun in de afgelopen jaren. Ellen, met jou zette ik de eerste schreden in het onderzoek. We presenteerden samen op internationale congressen. Onze eerste gezamenlijke lezing is een mijlpaal om nooit te vergeten. Ik heb genoten van alle keren die volgden, met jou als steun op de achtergrond aan mijn zij kon het niet misgaan. Dank voor al die mooie ervaringen. Kris, ook jij laat me iedere keer voelen dat je me onvoorwaardelijk steunt. Mede dankzij jouw energie ligt hier een proefschrift. Ik hoop met jullie beiden nog veel nieuwe mooie ervaringen te maken.

Zonder kinderen was er geen Early Language Scale. Veel dank aan alle ouders en kinderen die hebben meegewerkt aan de studie. Ook dank aan alle professionals die hebben geholpen met de werving en dan specifiek Betty Carmiggelt voor het enthousiasmeren van GGD's. Dank aan Judith van der Boom en Ineke van der Hoek voor het faciliteren van de werving. Ook een woord van dank voor Trudy en Hillie van de administratie van het lectoraat Healthy Ageing, Allied Health Care and Nursing.

Zonder onderzoeksgroep(en) geen inspiratie; dank aan alle (oud-) medewerkers van de lectoraten: Mechteld, Jan-Jaap, Thea, Priya, Martine, Lies,

Willemke, Rik, Linet, Baudina, Marlies, en alle anderen van het lectoraat Healthy Ageing, Allied Health Care and Nursing; Jeanette Doomenbal en collega's van lectoraat Jeugd Educatie en Samenleving en collega's van Research Instituut SHARE van de RUG .

Ook een dank aan ZonMw, Simea, de Hanzehogeschool Groningen en de Rijksuniversiteit Groningen voor het faciliteren van deze promotieplaats. In het bijzonder wil ik mijn voormalige teamleider van de opleiding logopedie, Sake van der Ploeg bedanken voor het faciliteren en in laten plannen van lesdagen en onderzoeksdagen, dat gaf me de ruimte dit onderzoek te starten en af te maken.

Dank ook aan alle collega's van logopedie, logo-vriendinnen van 'het Kookclubje', vrienden en vriendinnen, zangers van Collage, die me regelmatig vroegen hoe het ging met mijn onderzoek.

Natuurlijk ook dank aan mijn ouders en mijn broer Marty, jullie stonden (en staan) altijd voor me klaar. Lieve mam, je mocht de afronding niet meer meemaken. In gedachten ben je er wel altijd bij geweest. Lieve pap, jij bent trots voor twee en dat is een grote steun. Ook dank aan José voor je positieve blik vol energie. Tenslotte, mijn grote dank aan Jan, je hebt me altijd gestimuleerd en ruimte geboden om dit traject te doen.

Dank ook aan de logopedisten en studenten van de Hanzehogeschool Groningen, Fontys Hogeschool, Universiteit Utrecht en Rijksuniversiteit Groningen voor hun kritisch en creatief denken en inspirerende samenwerking en inzet die onmisbaar waren tijdens de dataverzameling:

About the author

Margot Bochane was born on the 27th of July 1973 in Leiderdorp, The Netherlands. She completed her secondary school at Visser 't Hooft Lyceum in Leiden. From 1991-1995, she studied speech pathology at the University of Applied Sciences Leiden. After completing her studies, she started working as a speech language pathologist in a private practice in The Hague, followed by three years at a special needs school for primary education for children with communication needs 'the Spreekhoorn' in Breda succeeded by almost twelve years at special school for primary education 'the Aquamarijn' in Leeuwarden. Additionally, she started teaching speech language development at NHL-Stenden University of Applied Sciences. Since 2008, Margot is working at Hanze University of Applied Sciences, where she is still teaching in the program of speech and language pathology in the School of Health Care Studies. She started her PhD research in 2012 at the research group Healthy Ageing, Allied Health Care and Nursing at the Hanze University of Applied Sciences and Department of Health Sciences at the University Medical Centre Groningen. She was a member of the Graduate School of Health Research (SHARE) at the University Medical Centre Groningen.

Research Institute SHARE previous dissertations

This thesis is published within the Research Institute SHARE (Science in Healthy Ageing and healthcaRE) of the University Medical Center Groningen / University of Groningen. Further information regarding the institute and its research can be obtained from our internet site: <http://www.share.umcg.nl/>

More recent theses can be found in the list below.

(supervisors are between brackets)

2021

Zijlstra E

The impact of the hospital environment; understanding the experience of the patient journey

(prof CP van der Schans, prof M Hagedoorn, dr MP Mobach)

Feenstra M

Trajectories of health status in older people

(prof SEJA de Rooij, prod BC van Munster, dr N Smidt)

Roossien C

Sensors@work; towards monitoring physical workload for sustainable employability

(prof MF Reneman, prof GJ Verkerke)

Klerk R de

Development of a wheelchair propulsion laboratory

(prof LHV van der Woude, dr HEJ Veeger, dr RJK Vegter)

Zhou Y

Machine learning for identifying patterns in human gait; classification of age and clinical groups

(prof CJC Lamoth, prof T Hortobayi)

Brink MJ van den

Management of heavy menstrual bleeding

(prof MY berger, prof MY Bongers, dr JH Dekker)

Brouwer MA

At the boundaries of life; suffering and decision-making in children with life-threatening conditions (1-12 years)

(prof AAE Verhagen, dr E Maevkelberghe)

Beune IM

Standardization in fetal growth restriction; progression by consensus

(prof JJ Erwich, dr SJ Gordijn, dr JW Ganzevoort)

Keizer MNJ

Neuromuscular control of knee laxity after an anterior cruciate ligament reconstruction

(prof E Otten, dr JM Hijmans, dr RW Brouwer)

Habtewold TD

Data-driven subphenotypic dissection of the clinical heterogeneity of schizophrenia spectrum disorders

(prof HM Boezen, prof R Bruggeman, dr BZ Alizadeh, dr EJ Liemburg)

Luchtenberg ML

A network of exchange; towards the empowerment of children in medical research

(prof AAE Verhagen, dr ELM Maeckelberghe)

Ferdiana A

Quality of life, work, and social participation among individuals with spinal cord injury

(prof U Bültmann, prof JJL van der Klink, prof MWM Post)

Tuvdenjorj A

Economic evaluation of tobacco control in Asia; dynamic population health impact assessment in Mongolia

(prof E Buskens, prof TL Feenstra)

Kouwijzer I

HandbikeBattle A challenging handcycling event

(prof LHV van der Woude, prof MWM Post, de LJM Valent)

Nanninga CS

Home & place making after stroke; exploring the gap between rehabilitation and living environment

(prof K Postema, prof LB Meijering, dr AT Lettinga, dr MC Schönherr)

2020

Wassink-Vossen S

Physical activity and functional recovery in late-life depression

(prof RC Oude Voshaar, dr P Naarding, dr R Collard)

Hovenkamp-Hermelink A

The long-term course of anxiety disorders; an epidemiological perspective

(prof RA Schoevers, dr H Riese, dr B Jeronimus)

Moazzen S

Nutrients and diet quality in gastrointestinal cancers

(prof GH de Bock, dr BZ Alizadeh)

Poonsiri J

Exploring cycling and sports in people with a lower limb amputation: prosthetic aspects

(prof JHB Geertzen, prof PU Dijkstra, prof R Dekker, dr JM Hijmans)

For earlier theses visit our website