International Journal of Pediatric Otorhinolaryngology 77 (2013) 1469–1473

Contents lists available at SciVerse ScienceDirect

International Journal of Pediatric Otorhinolaryngology

journal homepage: www.elsevier.com/locate/ijporl

Measuring speech sound development: An item response model approach

Gertrude H. Priester [a,*], Wendy J. Post [b], Sieneke M. Goorhuis-Brouwer [c]

[a] Faculty of Health & Welfare, Speech Therapy, Windesheim University of Applied Sciences, Zwolle, The Netherlands
[b] Department of Educational Sciences, University of Groningen, The Netherlands
[c] Department of Otorhinolaryngology, University Medical Center Groningen, University of Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 4 November 2012
Received in revised form 10 June 2013
Accepted 11 June 2013
Available online 6 July 2013

Keywords:
Speech sound development
Screening
Mokken model
Item response theory

ABSTRACT

Research aim: The primary aim of our study is to investigate if there is an ordering in the speech sound development of children aged 3–6, similar to the ordering in general language development.
Method: The speech sound development of 1035 children was tested with a revised version of Logo-Articulation Assessment. The data were analyzed with the Mokken Scale Program (MSP) in order to construct scales with satisfactory scalability (*H*-coefficient) and sufficient reliability (rho).
Results: The majority of children over 4.3 years of age turned out to have mastered most speech sounds. An ordering was only found in the youngest age group (3.8–4.3 years of age), for the sounds of /r/ in initial and final position and /s/ in initial position. This resulted in a set of scales. The scales developed for /r/ (in initial and final position) and /s/ were moderately scalable ($H > 0.43$) and reliable (rho > 0.83), and independent of gender. Moreover, we found variation in the judgment of speech sound development, which may perhaps have been due to where exactly the examiner was positioned during the assessment procedure: in front of the child, or sitting beside the child.
Conclusions: We could not detect an ordering for all speech sounds. We only found an ordering for /r/ in initial and final position and /s/ in initial position. In the Mokken analysis we conducted, these scales turned out to be moderately strong and reliable. Our research also underlines that speech sound development is judged not only in an auditory sense, but judgment also depends on the visual interpretation of the listener.

1. Introduction

Several Dutch studies [1–6] have shown that, in the group of children between 3 and 7 years of age, speech therapists often identify far more problems with speech sound development (SSD) than with language development. In connection with this, Van der Ploeg et al. [3,4] have raised the question whether the standards for SDD presently employed are truly appropriate for evaluating SDD in the Dutch language.

To get a clear idea of the clinical practice of speech therapists diagnosing SSD problems, we began our study by conducting a preliminary survey [7,8]. This survey showed that speech therapists in the Netherlands mainly judge SSD by assessing how children name (the images presented by) pictures or imitate certain words. However, the methods used in clinical practice provide no answer to the question which speech sounds exactly are mastered at particular ages. The interpretation of delayed or incorrect SDD is therefore mostly an intuitive process, and not based on normative data. This insight underlines the importance and relevance of the question raised by Van der Ploeg et al. [3,4].

As a second step, we reviewed the British English and Dutch literature on normative data in SSD [9]. The most important studies [10–13] on this subject agree more or less on the average ages at which vowels and consonants are acquired and mastered. As a rule, all vowels have been mastered around the age of 3; most single consonants have been mastered by the age of 4, with the exception of /s/ and /r/. The studies in our literature review disagreed on the age of acquisition of consonant clusters. Dodd et al. [10] have found that most consonant clusters have been mastered by the age of 5, whereas Stes [11,12] concludes that consonant clusters are only mastered between the ages of 6–10. Both studies also demonstrate that, when it comes to the development and mastering of consonant clusters, the data show substantial individual variation.

The analytical studies of English and Dutch normative data on SSD suggest that an ordering in SSD can be said to exist: vowels are mastered at an earlier age than single consonants, and consonant

* Corresponding author at: Speech Therapy, Windesheim University of Applied Sciences, P.O. Box 10900, 8000 GB, Zwolle, The Netherlands.
Tel.: +31 88 469 6244/649 626768.
E-mail addresses: hg.priester@windesheim.nl, gertrude.priester@upcmail.nl (G.H. Priester).

0165-5876/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijporl.2013.06.011

clusters are the last to develop. Views differ, however, on the ordering in the development of single consonants and consonant clusters. Based on the theory of Jakobson [14], it is often suggested that certain single consonants are mastered earlier than others, and that this trend is universal across languages [10–13]. Other studies, on the contrary, suggest that there is no strong evidence for a universal ordering. Pearson et al. [15], for instance, have found variation in the order in which speech sounds are acquired by children learning different dialects of the same language. Likewise, Goldstein et al. [16] argue that bilingual children learn the same speech sounds in a different order in their respective native languages.

Building on earlier research, in which a certain ordering in the acquisition of language milestones [17] has been found, we raise the question whether a universal ordering in SSD can also be said to exist. Such an ordering could be used as a solid basis for developing screening procedures.

In order to study this question, we assessed children's SDD by asking them to name pictures, and judged whether certain target speech sounds were pronounced correctly or not. Analogous to the research on the ordering of language milestones, we applied the nonparametric Mokken model [18,19] to analyze our data. In this non-parametric item response model, the relation between the acquisition of speech sounds and SSD must satisfy certain measurement properties, such as monotonicity and non-intersection [20].

This article forms part of a study which was set up to gain a clearer understanding of the discrepancy between the number of children who are treated for speech sound development (SDD) problems and the lack of reliable standards for SDD in the Dutch language.

2. Method

2.1. Participants

Our sample consisted of children between the ages of 3.8 and 6.3. For children under the age of 4, a number of kindergartens were approached; to sample children older than 4, we contacted primary schools. All participants in the study were native Dutch speakers.

We chose the age span of 3.8–6.3, because two pilot studies had indicated that children younger than 3.8 years of age cannot respond adequately to the instructions that would have to be followed [7,8,21]. According to the teachers and the school medical service we consulted, all children in our sample displayed normal cognitive development, and had normal vision and hearing abilities. Moreover, Dutch was their first language. To ensure that the outcomes on SSD were not influenced by any possible linguistic problems, the language development of all participants was checked first with the language screening instrument SNEL [17]. Only children with normal SNEL scores for their age (i.e. language development within the normal range) were included in the study. The parents of all participants were asked for their informed consent.

A well-balanced distribution of the participants was created by drawing a random sample from the total population of the kindergartens and primary schools in the study. We made sure to include different cities and villages from each of the twelve provinces of the Netherlands.

3. Materials and procedure

3.1. Assessment procedure

The children in the sample were examined with the revised version of LOGO-art [7,8]. This revised assessment procedure consists of a set of 84 pictures whose images normally can be named spontaneously by at least 75% of all children. For 41 words (pictures) that were more difficult to label (less than 75% of the children respond adequately), an imitation procedure was followed. In this manner, the most common Dutch consonants and consonant clusters, both in initial and final positions, could be assessed. Some speech sounds occurred more than once, being associated with several different pictures. The number of sounds and the number of pictures per type of sound are presented in Table 1.

Table 1
Number of sounds and pictures per type of sound in the revised version of LOGO-Art.

Type of sound	Number of sounds	Number of pictures
Single consonant initial	17	34
Single consonant final	12	23
Consonant cluster initial	29	35
Consonant cluster final	27	33
Total	85	125

A group of speech therapy students received special training for assisting in the project. The students performed the actual speech assessment, under the supervision of the first author and a research assistant in speech and language pathology. We divided the students into 14 test teams [22]. Each child in the sample was scored by 2 students, a test leader and an observer. The test leader was seated in front of the child, whilst the observer was seated beside the child. Both the test leader and the observer noted down whether the child pronounced a certain speech sound (the target sound in the named word) correctly (score = 1) or incorrectly (score = 0). The correct pronunciation was defined as being consistent with the speech pattern of an adult Dutch speaker [23]. Speech sounds that were wrongly pronounced (e.g. /s/ pronounced interdentally, omission of a certain sound, substitution, or phonological errors such as devoicing or fronting) were scored as incorrect.

The research assistant carried out quality checks after the assessment of the first 10 children tested by each team. A digital voice recorder (Olympus VN3100PC) registered the assessment. These sound files were subsequently digitally archived, and used to check on the quality of the assessments that had been made. The quality checks consisted of establishing whether pronunciation had been scored correctly, and the entry made in the database was accurate. Students who had failed to meet the required standards received additional training; quality checks were performed again in a follow-up evaluation. Inter-observer reliability was analyzed by Cohen's Kappa; also, the percentage of agreement between observers was calculated.

3.2. Mokken analysis

Before discussing the results of the Mokken analysis, we will first explain some of the terminology involved. In the Mokken analysis, the pictures used to assess the children are referred to as 'items'. A positive or negative response of a child to a certain item means that the child pronounced the intended speech sound (represented by the item) in either a correct or an incorrect manner. Items that satisfied the assumptions of the Mokken model could be used to form a Mokken scale.

The Mokken model is based on a number of measurement properties, which can be formulated in terms of four different assumptions: (1) the existence of a uni-dimensional latent trait SSD; (2) local independence (the responses to different items are independent); (3) monotonicity (there is a monotone positive relationship between the probability of a positive response to a certain item and SSD); and (4) non-intersection (the ordering of the