SUMMARY

This thesis renders an account of the research, conducted for of the construction, standardisation and validation of two tests for the measurement of language development in children between the ages of fifteen and seventy-eight months. The tests that were developed are the Reynell Test voor Taalbegrip (RTB) and the Schlichting Test voor Taalproductie (STP). The tests are designed to diagnose children’s language problems. The starting point for the design was formed by the Reynell Developmental Language Scales (RDLS; Reynell, 1977, 1985). The RTB can be seen as an elaborate revision (and Dutch translation) of the RDLS. The STP was newly designed during the research project. The central question in this thesis concerns the instrumental utility of the tests: do the tests answer their purpose? To answer this question a description is given of the construction of the tests, the guidelines for this construction and the standardisation research. Also an account is given of the research conducted in order to establish the reliability and the validity of both tests.

In Chapter 2 the concepts ‘language’, ‘language development’ and ‘language problems’ are elaborated upon in order to establish the suitability of the RTB and the STP in diagnosing and treating children with language problems. Because the much used term ‘developmental language deficit’ does not sufficiently describe the group of children where language diagnostics are relevant, the term ‘language problems’ is used in this thesis as a general term. By means of a scheme, based on Bloom and Lahey (1978), in which aspects of language are combined with language modalities comprehension and production the measuring claims of the tests can be outlined.

The research design is presented in Chapter 3. The research for the development of the tests has two stages: the construction stage and the standardisation stage. The construction research is conducted in two subsequent stages. This was necessary because a part of the STP, ‘Zinsontwikkeling’ was developed in stages, pending the results of a longitudinal study of grammatical development between the ages of four and six years. The longitudinal investigation was especially designed for this and took place within the present the research project. In the first construction research 208 children were tested with the experimental material; in the second 245 children were involved. In the standardisation research the final instruments were tested on 1049 children. The (national) standardisation sample can be described as a stratified stepwise sample that consists of eleven age groups, with ages between 1;3 and 6;3 years and half-year age intervals. In view of the validity of the Dutch standards demands are put on the representativity and the size of the sample. The size of the standardisation...
The representativity was examined by investigating the level of education of the subjects’ parents. The conclusion was drawn that the level of representativity for this variable is acceptable.

In Chapter 4 details are given about the construction of the RTB and the STP. Both content oriented and methodological considerations were at the basis of this construction. Methodological considerations involve the optimization of the reliability and validity of the tests. The most important consideration with regard to content was to measure language comprehension and language production in a separate and mutually independent way. This offers possibilities for a differential diagnosis. The original goals of Reynell and recommendations from RDLS users formed other important considerations. Furthermore the cooperation with the university of Louvain was of importance in constructing the RTB. Regarding the construction of the STP, at the beginning of the project it had been decided to develop a test part that would measure the grammatical development by means of elicitation procedures in a structured test situation. The construction of this part was based on the results of the previously mentioned longitudinal study on the development of grammar between the ages of four and six. Data from the research for of the development of the TARSP (Schlichting, 1993) were another main source for the construction. During the project a new part for measuring vocabulary was also developed. This was done because the reliability of the original RDLS design for this part proved to be low, while in the higher age span insufficient differentiation between levels of proficiency was obtained.

The reliability of the RTB and the STP (Chapter 5) has been investigated at every research fase by establishing internal consistency measures of all test parts per age-group. The main conclusion which can be drawn from the final results is that the internal consistency of the RTB and the different parts of the STP are sufficiently high for use in clinical situations. Also, the test-retest reliability has been studied. Retests have been conducted, six months after the first testing in the standardisation research, on 134 subjects from different age groups. The test-retest reliability of the RTB and the STP proved to be fairly substantial, especially considering the long interval between the measurements. Finally examiner influence has been investigated, using the data from the standardisation research. A oneway analysis of variance was performed on the standard scores of the children that did not speak a dialect, examiner being the independent variable. Only one examiner effect, in ‘Auditief Geheugen’, proved to be substantial.

The procedure, followed to develop norms for the tests and the results of this standardisation procedure are presented in Chapter 6. For all test parts a standardisation model could be found which met the requirements. However, for the STP part ‘Auditief Geheugen’ it was necessary to limit the age range in order to meet the requirements. This decision was also based on reliability results. In Chapter 7 the validity of the RTB and the STP is discussed. The content validity of these tests has been investigated thoroughly. A test for non-verbal intelligence
(SON 2½-7; Tellegen, Winkel & Wijnberg, 1996) was administered to a large number (574) of subtests from the standardisation sample. Also a ‘third’ test was administered besides the combination RTB/STP and the SON 2½-7. The third test was either a second language test or a second test for cognitive development. Non-cognitive test parts (motor development) were also administered to some of the subjects. The tests have good validity when correlations between language tests are high (convergent validity), correlations between language tests and cognitive tests are lower and correlations between language tests and non-cognitive tests are very low (discriminant validity). The overall conclusion which can be drawn is that convergent and discriminant validity are good: the pattern of correlations is consistent with the desirable pattern. The predictive validity of the RTB and the STP has not yet been investigated. Regarding the RTB a good predictive validity can be expected on the grounds of favourable results found with the RDLS and the RTOS.

In Chapter 8 the results are discussed and recommendations are made for further research as well as for the practical use of the tests. The final conclusion is that with the completion of the RTB and the STP good, workable and child-friendly tests have become available for the measurement of language comprehension and language production in children who are suspected of having language problems.