Language development in children with psychiatric impairment.
Blankenstijn, C.; Scheper, A.R.

Citation for published version (APA):
3 Research methods

Claudia Blankenstijn and Annette Scheper

3.1 Introduction
In our project we examine the morphological/syntactic and semantic/pragmatic development of 120 children with a psychiatric disorder. These data are compared with those from a group of 75 normally developing children (Roelofs, 1998) and 240 normally developing children (STAP; Van den Dungen and Verbeek, 1999). First, the selection procedures with regard to children with a psychiatric disorder are explained (3.2), including specific selection criteria (3.2.1), some background information and a description of the psychiatric disorders (3.2.2), and language test data (3.2.3). Second, the selection procedures, criteria and background information of the normally developing children are discussed (3.3). Third, we describe some details of the spontaneous language analysis (3.4) with respect to narrative (3.4.1) and conversation (3.4.2). Fourth, we present the morphological/syntactic and semantic/pragmatic models of analysis (3.5) and the classification of language impairment based on such models with respect to morphology/syntax (3.5.1) and semantics/pragmatics (3.5.2). Finally, we describe the different statistical analyses procedures used (3.6) and we finish with the formulation of the specific research questions (3.7).

3.2 Description of subjects: children with psychiatric disorders
Most of the Dutch-speaking PI-children (n=110) were attending psychiatric diagnosis and treatment in an Academic Clinic for Youth and Child Psychiatry. In the clinic children were seen for the first time in nearly all cases and participated in the language research project, temporarily appended to the standard psychiatric diagnostic procedure. The small inflow of four-year-olds at the clinic made it necessary to recruit half of the group four-year-olds from a Medical Day-care Centre (n=10), already diagnosed as having a psychiatric disorder and receiving special education.

The PI-children recruited were selected regardless of their Socio-Economic Status (SES), including PI-children with relatively high to low social backgrounds. Beitchman and colleagues (1990) and Cohen and colleagues (1998a) did not find that a lower SES increases the risk for the emergence of LI and/or PI (see 2.2). Both the Medical Day-care Centre and clinic are situated in the west urban region of the Netherlands (South-Holland).

3.2.1 Selection criteria
All the PI-children who finally took part in the project had to meet the following criteria:

1. age range 4;0 to 10;0 years
2. monolingual Dutch
3. no hearing loss, IQ > 70
4. no other severe disabilities
5) able to carry out research tasks  
6) diagnosed as psychiatric disordered, not autistic or schizophrenic

During the data collection (1993 - 1999) 180 children with a psychiatric problem within the age range of 4;0 to 10;0 years in the clinic and Medical Day-care Centre participated in the psychiatric investigation, and were language tested for this study. Of these 180 PI-children 60 had to be excluded for various reasons: 30 because they did not meet the selection criteria (9 drop-outs were caused by missing test data, 4 drop-outs had an IQ below 70, 1 drop-out had unintelligible language, 3 children were bilingual and 13 were diagnosed as autistic), 10 because of no parental permission; 20 children fulfilled the criteria, but illness of children, defect of audio-visual material, or no availability of a research room made it not possible to include them.

The population of 120 psychiatric children that met the selection criteria, was limited to 4;0 to 10;0 years. Before age 4;0 a psychiatric diagnosis is hard to establish and referral to the clinic younger than four years of age rarely occurs. No reliable linguistic diagnostic instruments are available for children older than 10;0 years of age. The PI-children were selected in such way that there were 20 children per age group (Table 3.1). The group PI-children consists of 33 girls and 87 boys, with a preponderance of boys as is typical for impaired populations. The sex ratio found in the population with a psychiatric disorder equals reported ratios in the literature: 25% girls and 75% boys. This is indicative for the population it represents (Cantwell and Baker, 1987; Cohen et al., 1989) (see 2.2). Although girls below age six are said to develop language qualitatively differently from boys (Van Alphen, 1999:32), sex of child will not be studied as a variable.

Table 3.1 Age groups, total number of PI-children per age group, mean age per age group and total numbers of girls and boys per age group

<table>
<thead>
<tr>
<th>Age groups</th>
<th>n=120</th>
<th>mean age*</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 yrs</td>
<td>20</td>
<td>4;7</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>5 yrs</td>
<td>20</td>
<td>5;7</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>6 yrs</td>
<td>20</td>
<td>6;9</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>7 yrs</td>
<td>20</td>
<td>7;7</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>8 yrs</td>
<td>20</td>
<td>8;6</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>9 yrs</td>
<td>20</td>
<td>9;6</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

* The first number indicates the number of years and the second the number of months.

We only included children from monolingual Dutch-speaking homes, as multilingual language development may differ from monolingual language
Research methods

development (e.g. De Houwer, 1995). We excluded children with a moderate or severe hearing loss, because of its possible negative influence on language development (Bishop and Mogford, 1988; Bol and Kuiken, 1988). Children with an IQ lower than 70 diagnosed as 'mentally retarded' were excluded. These children have severe information processing difficulties that may negatively influence the development of language and social-cognition (see 2.2 and 2.3.1). The standard diagnostic test used at the clinic was the Wechsler Intelligence Scale for Children (WISC) (Wechsler, 1949) and at the Medical Day-care Centre the Raven’s intelligence test (Raven, Court and Raven, 1986).

As can be seen from Table 3.2, the IQ scores of 115 PI-children were available and 5 were missing. The mean total IQ scores per age group indicate normal intellectual functioning. However, we observed some variation: more than half (n=61; 53%) of the 115 PI-children are of normal intelligence; the rest of the PI-children could be classified as low normal functioning (IQ 85 to 99) (n=44; 38%) or borderline intellectual functioning (IQ 71 to 84) (n= 7 IQ < 80; n=3 IQ 80 to 84; 9%) (see 2.2 and 2.3.1). Of the 7 PI-children with an IQ below 80 5 PI-children received the diagnosis PDD-NOS. It is still disputable where the dividing line between normal and abnormal intellectual functioning should be situated. Some researchers take a total IQ of 80 as opposed to the total IQ of 70 (e.g. Maassen, Poppelaars, Pasman and Rotteveel, 2001). The language and psychiatric impairments of 47% of 115 PI-children might very well be based on moderate to mild executive dysfunctioning and information processing difficulties (see 2.3.1).

Table 3.2  Age groups, means scores of IQ Total, IQ Performat and IQ Verbal of 115 PI-children

<table>
<thead>
<tr>
<th>Age groups</th>
<th>IQ Total</th>
<th>IQ Performat mean</th>
<th>IQ Verbal mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 yrs</td>
<td>107.1</td>
<td>109.1</td>
<td>101.8</td>
</tr>
<tr>
<td>5 yrs</td>
<td>107.7</td>
<td>109.1</td>
<td>104.6</td>
</tr>
<tr>
<td>6 yrs</td>
<td>100.8</td>
<td>103.6</td>
<td>98.1</td>
</tr>
<tr>
<td>7 yrs</td>
<td>94.5</td>
<td>97.5</td>
<td>93.9</td>
</tr>
<tr>
<td>8 yrs</td>
<td>103.7</td>
<td>109.0</td>
<td>99.1</td>
</tr>
<tr>
<td>9 yrs</td>
<td>98.1</td>
<td>99.6</td>
<td>99.3</td>
</tr>
</tbody>
</table>

PI-children with a clear clinical syndrome, such as Down's Syndrome or Williams' Syndrome, or any neurological, sensomotorical and/or visual problems were excluded. Neurological disabilities (e.g. Tannock, 1998), sensomotorical disabilities (e.g. Ayres, 1994) and visual disabilities (Wegener Sleeswijk, 1986; Baker, 1993) can have a negative effect on the language and social-emotional learning process.
3.2.2 Psychiatric diagnosis

The psychiatric disorder of all PI-children at the clinic was classified according to the DSM-III-R (APA, 1987) and at the Medical Day-care Centre the ICD-10 classification for mental and behavioural disorders (WHO, 1992) was used. Autistic and schizophrenic children were excluded, because inclusion would have increased the comorbidity rates. These children form separate groups of the most severely language impaired PI-children, in which the LI is a necessary — although not a sufficient — condition for PI (Cantwell and Baker, 1987; Prizant et al., 1990).

The medical and psychiatric information on all PI-children was collected by the psychologists and psychiatrists of the diagnostic teams. The medical records were used as documented in a computerized database of medical and psychiatric information (Treffers, Goedhart, Waltz and Koudijs, 1989). Part of this complete patient file consists of the information of a child, such as sex, age, life history and socio-economic status of the parents. The PI-children themselves were further tested. At the clinic the Diagnostic Interview Schedule for Children-Child version (DISC-C) was used; this is a standardized diagnostic interview developed for use in epidemiological studies of children and adolescents (Costello, Edelbrock, Dulcan, Kalas and Klaric, 1984).

It must be clear, in advance, that the research population is not representative for the general clinic and Day-care centre population, because of the specific selection criteria used. The 10 Day-care PI-children were all diagnosed as psychiatric disordered before age 4;0, which probably means that they are more severely disordered than the 110 PI-children recruited from the clinic. There is no documentation on the type of PI-children that are referred to the Medical Day-care Centre or the clinic and none on the medical situation, so that it is impossible to judge the representativity of the PI-population studied.

The following main classifications of internalizing and externalizing psychiatric disorders were used (see 1.2): Depression and Anxiety Disorder (internalizing disorders) as opposed to Oppositional Behavioural Disorder and Attention Deficit Disorder with Hyperactivity Disorder (ADHD) (externalizing disorders). Both Pervasive Developmental Disorder Not Otherwise Subscribed (PDD-NOS), characterised by a delay in development on all levels, and the classification 'No Diagnosis' implies that several psychiatric problems exist without one main classification being obvious (both symptoms of internalizing or externalizing disorders might be present) (Table 3.3). At the Day-care Centre the same main classifications were used after transforming the ICD-10 classifications into DSM-III-R classifications (APA, 1987).
Table 3.3: Number of PI-children with specific disorders per age group

<table>
<thead>
<tr>
<th>PI-children</th>
<th>Internalizing</th>
<th>Externalizing</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Depression</td>
<td>Anxiety</td>
<td>Oppositional behaviour</td>
</tr>
<tr>
<td>4 yrs</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>5 yrs</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>6 yrs</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>7 yrs</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8 yrs</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>9 yrs</td>
<td>4</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>33</td>
<td>13</td>
</tr>
</tbody>
</table>

| Total 120   | 41            | 32            | 25    | 22   |

A significant correlation between age and a specific psychiatric disorder (Mantel-Haenszel = 11.66183, df 1; p<.0064) was found; the older PI-children have relatively more internalizing disorders, especially Anxiety Disorder compared to the other categories. It was not possible to achieve an equal distribution of type of psychiatric disorder, since the project was dependent on the intake in the clinic and Day-care Centre.

When we correlate the different types of psychiatric disorder with the IQ-scores (WISC IQ Total, IQ Perwormal and IQ Verbal) we can observe a trend for significance (p<.10): the PI-children with PDD-NOS relatively have the lowest Perwormal IQ scores compared to the PI-children with other specific psychiatric impairments.

The Children’s Global Assessment Scale (CGAS) (APA, 1994) was used at the clinic in order to measure the global functioning of PI-children at home and at school. In the Netherlands, the CGAS score is the only assessment tool in every day practice that can give a reliable indication of the severity of the PI. Usually, further treatment procedures are based on the different scores: CGAS 1 (score 0 to 50) indicates serious symptoms of PI; CGAS 2 (score 50 to 60) indicates a moderate PI and CGAS 3 (score 61 to 70) indicates a mild PI. A CGAS score above 70 indicates normal functioning. The CGAS-scores of 105 PI-children were available and 15 were missing.

With respect to the 120 PI-children, 35 (33%) out of 105 PI-children with severe symptoms (CGAS 1) needed immediate treatment and some of them even needed to be taken into the clinic, since they showed serious impairments in social and school functioning. The 41 (39%) of the 105 PI-children with moderate symptoms (CGAS 2), showing impairments in social and school functioning, needed ambulatory treatment. This forms the group of out-patients. Only 29 (28%) of 105 PI-children were diagnosed as psychiatrically impaired, needing no direct treatment (CGAS 3).
They had some difficulty in social and school functioning. Most parents of these mildly disordered PI-children received some help in the form of an evaluation of their PI-children's functioning with one of the clinicians in order to learn to cope better with the abnormal social-cognitive behaviour of their child.

We found a significant correlation of CGAS outcome with age. We observed an almost linear decrease with age in CGAS 1/2 (severe/moderate symptoms) and an almost linear increase of CGAS 3 (mild symptoms) in 105 PI-children (p<.04). This means that the younger the PI-children, the more severe the symptoms of their psychiatric impairment were judged to be.

When we correlate CGAS scores with the three PI-categories (1) internalizing PI, (2) externalizing PI and (3) both internalizing and externalizing symptoms (see Table 3.3), no correlations were found. The following division proved to be more informative, since a significant correlation was observed between the CGAS and the following four different types of psychiatric impairment: (1) internalizing PI, (2) externalizing PI, and both (3) PDD-NOS and (4) 'No Diagnosis'.

Most PI-children with externalizing PI, especially those with ADHD (60%) and most PI-children with PDD-NOS (67%) had the most severe symptoms (CGAS 1). Most PI-children with internalizing PI, especially those with Anxiety Disorder (82%) had moderate/mild symptoms (CGAS 2/3). Most PI-children with 'No Diagnosis' (75%) had mild symptoms (CGAS 3). In sum, 76 (72%) of 105 PI-children need some form of therapy, especially the PI-children with ADHD and PDD-NOS, followed by the PI-children with Anxiety disorders.

3.2.3 Parental checklist and language test data

Language (dis)abilities can be detected in many different ways. We used a Parental Checklist (Blankenstijn and Scheper, 1993) and a language test (TvK, Van Bon and Hoekstra, 1982) in order to get insight in the language history of each individual PI-child and to collect information about the possible existence of receptive and productive LI.

Parental checklist

We developed a parental checklist, called Language Development Checklist (Gehoor-Communicatie-Spraak-Taal Oudervragenlijst; Blankenstijn and Scheper, 1993) (see Appendix 3a). This checklist contains 46 questions about the PI-children's development in the area of auditory abilities (8 questions), communication (15 questions), language (17 questions) and articulation (5) and 1 open question in order to elicit additional personal information about the PI-children's development in general. Unfortunately, no questions were asked about early language intervention or therapy. We only know that 10 four-year-old PI-children recruited from the Day-care Clinic receive language therapy. Each parent was asked to fill in this checklist at home, before the PI-child was language tested. Of the 120 PI-children the parents of 18 (15%) did not fill in this checklist, resulting in available etiologic information for 102 PI-children. On the basis of the checklist,

---

1 We thank Daniela Polišenská for developing a database of the etiologic information in the parental checklists.
the parental respondents judged that 18 (18%) of all 102 PI-children had articulation problems, i.e. a language disorder in the area of phonology (see 1.2). In total, 49 (48%) of all 102 PI-children were judged as being language impaired (LI) in the area of morphology/syntax and semantics/pragmatics by their parents before referral to the psychiatric clinic (see for more details 15.2.1).

Language test data
The results of a language test were used as an index of language (dis)ability. As part of the standard and intake procedure of the clinic, all 120 PI-children were language tested with the *Taaltests voor Kinderen* (TvK; Van Bon and Hoekstra, 1982). This is one of the most frequently used standardized language tests in the Netherlands with norms based on over 1900 children from 4;0 to 10;0 years of age with appropriate reliability ranges. The children of the Medical Day-care Centre were also tested to make their data complete.

The number of PI-children that are deviant according to the five language subtests (TvK) are presented in Table 3.4. These subtests are differentiated into receptive (1, 2 and 3) and expressive language subtests (4 and 5). Within this differentiation of the subtests we ordered them from word level (1 and 4) to sentence level (2, 3 and 5). They cover the areas of morphosyntax (2, 5), morphosyntax/semantics (1, 4) and semantics/pragmatics (3). Pragmatics is underrepresented with only one receptive subtest. However, the test protocol is not only unclear about how to judge children as being language-impaired on the basis of low scores on one or more receptive and/or expressive subtests, but also about to what extent low subtest scores in any combination should contribute to the diagnosis 'language disorder'. Despite this classification problem, we will use the criterion that PI-children are diagnosed as language impaired (LI) when they have too low scores\(^2\) on one or more subtests.

![Table 3.4 Number and percentage of PI-children diagnosed as LI based on the Taaltests voor Kinderen (Van Bon and Hoekstra, 1982)](table3_4.png)

* the abbreviation between brackets indicate the Dutch subtest labels
** 14 calculated over 77 PI-children

\(^2\) A standard score of 3.0 and below indicates abnormal language performance on the specific language subtests (TvK; Van Bon and Hoekstra, 1982).
The results on the receptive language subtests on word level indicate that 17 (14%) of the 120 PI-children have a disorder in making receptive lexical choices (matching one word meaning with one picture out of four). They have many failures (match a word with a wrong picture) and seem not to understand the meaning of words that should have been acquired according to their age. Thus, their lexicon probably is restricted.

The results on the receptive language subtests on sentence level show that comparably 17 (14%) of all PI-children have a disorder in their ability to judge what is a correct morphological/grammatical sentence in Dutch. These PI-children thus also have difficulties in identifying morphological/syntactic errors. When we look at pragmatics, we first have to mention that for reasons of time, age and attention deficits, the receptive subtest (3) that tests the ability to infer implicit meaning in single sentences was not performed by 43 PI-children. We see that 14 (18%) of 77 PI-children have a disorder in inferring implicit semantic/pragmatic meanings within sentences.

When we look at the combination of receptive difficulties, we observe that no PI-children have deviant scores on all receptive subtests (1, 2 and 3); 10 (8%) PI-children had deviant scores on two out of three subtests in any combination; 24 (20%) PI-children had deviant scores on only one subtest (1, 2 or 3). These last two groups thus contain 34 PI-children, i.e. 28% of the 120 PI-children, that have a receptive language disorder on word and sentence level in the area of semantics (1), morphology/syntax (2) or pragmatics (3).

The results on the productive language subtests on word level indicate that a very small number (8; 7%) of all PI-children have difficulties in naming pictures, indicating a productive semantic language disorder related to a limited vocabulary. The results on the productive language subtests on sentence level show that 18 (16%) of all PI-children have difficulties in sentence production. Only 6 (5%) PI-children had deviant scores on all productive subtests (4 and 5), whereas 13 (11%) PI-children had deviant scores on one out of the two productive subtests (4 or 5). We computed that 19 PI-children, i.e. 16% of the 120 PI-children, fall out on one or more productive subtests and are detected as having a productive language disorder on word and sentence level in the area of semantics and morphology/syntax.

3.3 Description of subjects: normally developing children
In 1993 Roelofs (1998) collected language data from 75 Dutch normally developing children (42 girls and 33 boys), equally divided over the age groups 4:0 to 8:0 years. The N-children were recruited from three different schools (two school with mainly children with middle Socio-Economic Status (SES); one school with mainly children with lower SES; Roelofs, 1998:63) in the west urban region of the Netherlands (North-Holland), including pupils from a range of social backgrounds. In consultation with the teacher the N-children were selected according to the following criteria:

(1) age range 4:0-4:6 to 8:0-8:6 years
(2) monolingual Dutch
(3) no overt educational, psychological and social problems
(4) able to carry out research tasks

Children were not included if they had a known history of language delay, or social-cognitive problems, although other additional language/psychological test data were not available (Table 3.5).

Table 3.5  
Age groups, total number of N-children per age group, mean age per age group and total number of girls and boys per age group (Roelofs, 1998)

<table>
<thead>
<tr>
<th>Age groups</th>
<th>n=75</th>
<th>mean age</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 yrs</td>
<td>15</td>
<td>4.4</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>5 yrs</td>
<td>15</td>
<td>5.3</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>6 yrs</td>
<td>15</td>
<td>6.2</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>7 yrs</td>
<td>15</td>
<td>7.2</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>8 yrs</td>
<td>15</td>
<td>8.3</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

The 75 Dutch-speaking N-children from the Roelofs-population are judged to be representative for the group normally developing children in general since they were in regular educational settings. Parental permission was necessary to let the child participate in the project. In total 8 N-children had to be excluded: 6 dropouts were caused by illness, non-co-operative behaviour, an inability to carry out research tasks or by a defect of audio-visual equipment; 1 drop-out was judged as language disordered; and 1 had hearing problems (Roelofs, 1998).

The 240 Dutch-speaking N-children from STAP-population were recruited from different schools in Amsterdam, also located in the west urban region of the Netherlands (North-Holland), including pupils from a range of social backgrounds and according to similar selection criteria as the Roelofs-population. The N-children from the STAP-population were also equally divided over the age groups 4;0 to 8;0 years (Van den Dungen and Verbeek, 1994, 1999; see also 4.3).

3.4 Language Assessment

Spontaneous language data were gathered in two settings: a narrative (narrative genre) and a semi-structured interview (conversational genre). In general the language problems in children are more obvious in such complex settings compared to language testing (McTear and Conti-Ramsden, 1992; Smith and Leinonen, 1992). The two settings are not only exploring more complex language behaviour than elicited by language tests, but also give more insight in problems with verbal communication in everyday life. It is difficult to decide whether the interview is a more complex language task than the narrative. In the conversational genre PI-children might be guided by the interviewers, but they have to plan language information in a creative way and therefore can experience some psychosocial
stress. Conversely, in the narrative genre PI-children can be helped by the pictures that trigger the language information, despite the fact that they have to tell the whole story on their own.

3.4.1 Narrative

In order to elicit the narrative, the child was invited to look at the picture book 'Frog, where are you?' (Mayer, 1969), frequently used in cross-linguistic research (Berman and Slobin, 1994). A short summary of the 24 picture book follows (see also Trabasso and Rodkin, 1994:85-86; Appendix 3b):

'This is a story about a little boy who had a pet dog and frog in a jar. One evening, the boy, dog, and frog were in the boy's bedroom. The boy and the dog went to sleep and while they were sleeping, the frog escaped from its jar. The next morning the boy and the dog woke up and found the empty jar. The boy was upset and tried together with the dog to find the frog. First, they were searching inside the room. Accidentally the dog fell out of the window. Thereafter they were searching for the frog outside. But in each place the boy and dog searched, they found nothing or encountered a different animal. Finally, a deer accidentally carried the boy to a pond being challenged by the dog. They found the frog, being in love with a female frog, accompanied by little frogs. The story ended with the boy taking home the frog or a baby frog, being followed by the dog.'

Children were first asked to look through the entire book, and then to tell the story while looking at the pictures. We gave them the following instructions:

'Here is a picture book. This book tells a story about a boy [point to picture on cover], a dog [point], and a frog [point]. First, I want you to look at all the pictures. Afterwards you look at the pictures again and tell the story to me. I do not know the story.'

The influence of the investigator was minimized. Younger children were sometimes helped with turning the pages. The following neutral prompts were used: silence, nod of head, 'okay' and 'yes'. Less neutral prompts were occasionally necessary, like 'and then?' and 'what happened next?'. The elicitation procedure used is comparable to Berman and Slobin (1994) and Roelofs (1998).

3.4.2 Conversation

In order to collect spontaneous language data, we interviewed the children in a semi-structured interview, according to a Dutch language analysis procedure STAP (Vanden Dungen and Verbeek, 1999). This is a standardized diagnostic instrument for children from 4:0 to 8:0 years. The STAP procedure is based on spontaneous conversation with a researcher unfamiliar to the child and involves topics outside the here-and-now. The interviewer has to ask open questions as much as possible about every-day events. Typical subjects are pets, brothers and sisters, school, friends, sport, holiday and home activities. Television programs, films, computer programs
or books are avoided, as it is known that these topics are relatively difficult to talk about, especially for younger children (see also Roelofs, 1998).

The ST AP is suitable to diagnose children as having expressive language disorders in the areas of morphology/syntax and semantics/pragmatics (see Appendix 3c). The focus is on grammaticality and morphosyntactic complexity; only 2 variables involve semantics and 3 pragmatics. Global language measures, such as mean length of utterance (MLU), mean total number of elliptical answers or the mean number of unintelligible utterances are also included. We used part of this linguistic model of analysis to describe the morphological/syntactic disorders of the 120 PI-children (see Chapter 4).

Both narrative and conversational data were collected in a one-hour session (see also Roelofs, 1998:64), each child being audio- and video-taped in a special testing room. One of the investigators sat face-to-face and opposite to the child, while the other was running the camera.

We used the Child Language Data Exchange System (CHILDES) to transcribe and analyse the (non)linguistic data from a time-coded video. A selection of the transcription notation conventions in the Codes for the Human Analysis of Transcripts (CHAT) were used (MacWhinney, 1995). The transcription and segmentation conventions were mainly based on STAP (Van den Dungen and Verbeek, 1999) together with additional conventions related to nonverbal behaviour (Blankenstijn, Roelofs and Scheper, to appear). Transcription was in standard orthography, although non-standard forms of words were also transcribed, such as clitics, phonological, dialectal- or sociolectic variants.

3.5 Morphological/syntactic and semantic/pragmatic models of analysis

We developed different models of analysis to investigate the morphological/syntactic abilities (Scheper, 1996) and the semantic/pragmatic abilities (Roelofs, 1996; Blankenstijn, 1996) of the PI-children in comparison to N-children (Roelofs, 1998) in both conversation and narrative. Approximately, 25,000 T-units (see 4.1) (divided in 30,000 clauses) and 30,000 communicative contributions of both N- and PI-children were collected and scored in both the conversational and narrative genre. In the interview genre, approximately 10,000 communicative contributions of the N- and PI-interviewers were also analysed. We did not exclusively explore morphological or semantic abilities, but only those morphological abilities most closely connected to syntax and only those semantic abilities most closely connected to pragmatics in both genres.

We divided the morphological/syntactic component of language in two parts, the grammatical form and the complexity of morphological/syntactic categories in order to evaluate the type of morphological/syntactic impairment. Within these categories a differentiation is made into lexical (nouns, verb and adverbs) and functional categories (e.g. determiners) to observe the type of morphological/syntactic problem that might be presented in the PI-children (e.g. Chomsky, 1986; Clahsen, 1989, 1992) (see 4.1).
Following Roelofs (1998), we divided the semantic/pragmatic component of language used in conversation into three parts with respect to the structure (turn taking abilities; adjacency pairs and episodes), function (speech acts and their form), content (topic management; coherence and cohesion). The semantic/pragmatic component of language used in narrative was predominantly related to the content (plot management; cohesion) in order to evaluate the type of semantic/pragmatic impairment (see 10.1).

A morphological/syntactic error or marked unit with respect to the morphological/syntactic complexity has a negative impact on the information transmission at the semantic/pragmatic level. If no morphological/syntactic problems exist, there still might be problems in the area of semantics/pragmatics. Therefore, we integrated the two models of analysis at different points, for example with respect to the semantic/pragmatic effect of missing arguments. In order to gain insight in possible differences in language capacity related to a specific genre, we compared PI-children's capacities in the two genres: most morphological/syntactic results on the level of grammatical form and complexity are compared in both genres, whereas only the results of conjunctive cohesion and referential cohesion are compared in both genres on the semantic/pragmatic level of analysis.

In general, language errors were analysed in four steps. These steps are worked out by making morphological/syntactic and semantic/pragmatic paraphrases following the flow-chart developed by Roelofs (1996, 1998:202) (Appendix 3d). Making a paraphrase is adding, leaving out or changing words or sentences in order to explicate the 'ideal' message the child aims at, while staying as close as possible to the actual erroneous communicative contribution. The four steps are:

(1) identification of the error

(2) comparison between the error and the 'ideal' language unit(s) according to the spoken adult target language

(3) classification of the error. Each error was categorized as an error in the area of morphology/syntax or semantic/pragmatics, although some errors at the interface between language areas could be counted as both (Bloom and Lahey, 1978:22)

(4) evaluation of the error. Each error was evaluated with reference to its morphological/syntactic and semantic/pragmatic consequences in the communicative setting in which it occurs

The classification of language behaviour as morphological/syntactic or semantic/pragmatic LI on the basis of the spontaneous language analysis is discussed in more detail in the following subsections (see 3.5.1 and 3.5.2).
3.5.1 The classification of morphological/syntactic LI based on spontaneous language analysis

With respect to the analysis of morphology/syntax a group comparison on the most general level was possible. The instrument used, STAP (Van den Dungen and Verbeek, 1999), makes it possible to classify children as morphologically/syntactically disordered related to based on deviant z-scores on two general production variables: ungrammatical sentences and grammatical errors (see 4.2). On the basis of a deviant score on one of these general variables a child is classified as LI.

There are further 23 subvariables (see Appendix 3c) that specify ungrammaticality. The STAP spontaneous language protocol is also clear on how to judge children as being language-impaired on the basis of deviant scores on one or more of these 23 subvariables. The subvariables are not hierarchically ordered in terms of their negative effect on communication. Errors in one area of morphology/syntax are judged as being equally disturbing as errors in other areas of morphology/syntax. According to STAP, all subvariables thus equally contribute to ungrammaticality, although this is highly debatable. A relatively infrequent morphological/syntactic error could be more serious than relatively many morphological/syntactic errors of another type. The diagnosis LI cannot be based at this point in time on a deviation in one single subvariable. In the future, more research is needed to classify deviancies in different areas of morphology/syntax on a scale of severity related to age.

As pointed out above, in addition to the STAP variables an even more finely-grained model of morphological/syntactic analysis (Scheper, 1996) was used to investigate the morphological/syntactic abilities of the PI-children. This model contains even more detailed subvariables that mostly fall under the subvariables of the STAP. However, since the STAP did not examine such variables, a comparison with the 240 N-children from the STAP-population was not possible. A comparison was therefore made with a selection of the 75 N-children from the Roelofs-population. This more detailed analysis may contribute to the proposed development of the LI scale mentioned above that is needed to detect the morphological/syntactic symptoms in the PI-population.

3.5.2 The classification of semantic/pragmatic LI based on spontaneous language analysis

Since a comparison between the PI-children with the N-children from the STAP-populations following the STAP-norms (Van den Dungen and Verbeek, 1994, 1999) results in 100% semantic/pragmatic disorder in the 120 PI-children, this analysis procedure proved to differentiate insufficiently.

The STAP-procedure has five general measures to classify children as semantic/pragmatic disordered. These variables are moderate and severe semantically marked utterances, implicit reference and moderate and severe pragmatically marked utterances (Van den Dungen and Verbeek, 1994, 1999) (Appendix 3c). These variables, however, were found to be unclear and incomplete. Firstly, the different types of semantic and pragmatic errors mentioned in STAP under the general variables form only a very small part of the types of errors that can be classified. For
example, interruptions (see 10.6) are not included. Secondly, with respect to the variable implicit reference, some implicit referents are scored under this heading, whereas other types, for instance, the use of definite NP’s or the use of proper names in the conversational interview genre (see 13.5), are scored as incorrect under the heading of ‘pragmatically marked utterance’. The boundaries between different instances of unclear reference assignment therefore proved to be too unclear (Hickmann, 2003). Furthermore, it is debatable whether the use of implicit reference should be considered as a general pragmatic measure. Thirdly, the boundary between semantically/pragmatically impaired and normal language behaviour is very thin. For instance, one severe semantic/pragmatic error that causes unintelligible information in a STAP interview is enough to diagnose a four- to eight-year-old child as semantically/pragmatically impaired (Van den Dungen and Verbeek, 1994, 1999). Therefore, the STAP procedure proved to be insufficient suited to assess semantic/pragmatic LI.

A new, more finely-grained semantic/pragmatic analysis procedure was needed. Before such an assessment tool could be developed, however, more information about the semantic/pragmatic development of normally developing Dutch-speaking children was needed. Roelofs (1996, 1998) created such a semantic/pragmatic model of analysis in order to describe the pragmatic development of Dutch-speaking normally developing children. We used this model to detect semantic/pragmatic LI in the 120 Dutch-speaking PI-children. The subvariables used in this semantic/pragmatic model (Roelofs, 1996, 1998) are not the same as the five general semantic/pragmatic measures used by STAP. Roelofs (1998) did not include a general measure for semantic/pragmatic LI, such as the number of semantically/pragmatically incorrect sentences or the total number of semantically/pragmatically errors. Individual variables are used to determine a problem in this area. In the future more research is needed to classify deviancies in different areas of semantics/pragmatics based on these more specific Roelofs-(sub)variables. This can be done in a similar way to that was used for the STAP procedure, where only the 23 most informative subvariables were finally selected to determine the morphological/syntactic LI.

The different LI-symptoms out of all the Roelofs-(sub)-variables that proved to contribute to the semantic/pragmatic LI-symptoms in PI-children can be a first step towards the development of a new assessment tool to identify semantic/pragmatic disorder. A general measure for semantic/pragmatic LI should be included, for instance, the ‘total number of semantically/pragmatically marked/incorrect utterances’. Since most researchers only investigate either morphological/syntactic abilities or semantic/pragmatic abilities, no model of analysis exists that integrates these two areas. In the future, such a model should be developed.

A description of both morphological/syntactic (dis)abilities and semantic/pragmatic (dis)abilities in 120 PI-children enables us to compare the language abilities of our research population with children that are Specific Language Impaired (grammatical SLI) (Van der Lely, 1993, 1994; Leonard, 1998) on the one hand and with children who are Semantic-Pragmatic Impaired (Bishop and Adams, 1989) or Pragmatic Language Impaired (PLI) (Bishop, Chan, Adams, Hartley and Weir, 2000) on the
other hand. On the basis of this comparison, we might find evidence for the existence of specific morphological/syntactic and/or semantic/pragmatic (dis)abilities exclusively observed in PI-children or not.

3.6 Statistical analysis
The models of analysis were developed first and the interrater reliability of the main variables (and subvariables) were tested on small subgroups of the total population of PI-children (Appendix 3e). Blankenstijn, Roelofs and Scheper established sufficient interrater reliability for the main language transcription and segmentation variables. Roelofs and Blankenstijn established the interrater reliability for the semantic/pragmatic variables in both genres. Scheper established sufficient percentages of agreement (see Baker, De Geus, Van Amstel, Gosselaar, Schuijt, Ursem, Verkoeijen and De Wijckerslooth, 1998) or interrater reliability for the main morphological/syntactic variables (Schuijt, 1999)\(^3\) in both genres. We used the corrected Cohen's Kappa (see Van den Brink and Koele, 1985, 1986, 1987; Brennan and Prediger, 1981) most frequently as the measure of interjudge agreement. For the segmentation and some of the morphological/syntactic variables with regard to the conversational genre percentages of agreement were used. For the MLU the Pearson's product-moment correlation coefficient (PMC) was calculated with the use of SPSS/PC+.

According to the Explanatory Criterion (Burisch, 1984) that uses significant group-effects to classify deviant behaviour from normal behaviour, the selected Morphological/Syntactic (MS) and Semantic/Pragmatic (SP) variables are used to explore whether there exist significant differences between the PI and the N-children.

As statistical methods used in order to measure predominantly group- and age-effects in the linguistic performance of 120 PI-children compared to 75 N-children, we mainly used analysis of variance (ANOVA's, ANCOVA's, MANOVA), Chi-square, Pearson's product-moment-correlation (pmc) coefficient, Fisher's Exact Test or a Binomial test (Van den Brink and Koele, 1985, 1986, 1987). Generally, the ANOVA is designed to be sensitive to differences among the means of the groups under study. In order to avoid spuriously inflating the number of significant results, we developed where possible a hierarchical coding system of subordinate coding categories for the general variables that represent differences between appropriate versus inappropriate morphological/syntactic and semantic/pragmatic language abilities. We tried to define mutually exclusive general variables.

If a significant main effect for age or a group*age interaction effect was found, post hoc trend analyses using one-way ANOVA or ANCOVA were executed to examine the linearity of the age-effect in both populations. Additionally, we used a Factor analysis (Van den Brink and Koele, 1985, 1986, 1987) in order to find specific types

---

\(^3\) We would like to thank our colleague Marja Roelofs, the students of the research group 'First Language Acquisition': Krista de Geus, Mariët van Amstel, David Gosselaar, Janet Ursem, Jesse Verkoeijen and Gertje de Wijckerslooth and especially Stephanie Schuijt for their time and effort in establishing the interrater reliability.
of 'language impairment profiles' in relation to specific types of psychiatric disorder, divided into internalizing and externalizing disorders.

We frequently checked significant F-values using the non-parametric Kruskal-Wallis test for independent groups, as suggested by Craig and Evans (1993), but since this always confirmed our findings with other statistical measures, we will not report the results.

We carried out a common Factor analysis followed by Varimax rotation (Gorsuch, 1983; Van den Brink and Koele, 1987:50) was used in order to explore structural relations between the variables representing differences between appropriate versus inappropriate morphological/syntactic and semantic/pragmatic language abilities. The result of the Factor analysis, i.e. more or less separate, specific dimensions of LI, were used to identify extreme LI in specific types of psychiatric disorder.

3.7 Specific research questions

We addressed the following specific research questions based on the main research questions (see 2.5):

Subquestions under main research question 1a:

Subquestion 1a1
To what extent do Dutch-speaking PI-children have problems with morphology/syntax related to the grammatical form and complexity of lexical and functional categories in both genres? (Chapter 4)

Subquestion 1a2
Which morphological/syntactic disabilities related to the grammatical form and complexity of lexical categories (lexical verbs, argument structure, prepositions and adverbs) and functional categories (tense and agreement marking) are characteristic of Dutch-speaking PI-children in both genres? (Chapters 5 to 8)

Subquestions under main research question 1b:

Subquestion 1b1
To what extent do Dutch-speaking PI-children have problems with semantics/pragmatics with respect to the structure (turn taking ability), function (responsiveness) and content (topic management, coherence and cohesion) in the conversational genre and content (plot management and cohesion) in the narrative genre? (Chapters 10 to 13)

Subquestion 1b2
Which semantic-pragmatic (dis)abilities are characteristic of Dutch-speaking PI-children with respect to the structure (turn taking ability), function (responsiveness) and content (topic management, coherence and cohesion) in the conversational genre and content (plot management and cohesion) in the narrative genre? (Chapters 10 to 13)
Subquestions under main research question 2

Subquestion 2a
Is there a difference in performance at the level of morphology/syntax related to the grammatical form, temporality, transitivity, agreement marking and morphosyntactic packaging comparing the two genres, conversation and narrative? (Chapter 9)

Subquestion 2b
Is there a difference in co-referential cohesion at the level of semantics/pragmatics comparing the two genres, conversation and narrative? (Chapter 14)

Subquestion under main research question 3
Is there a relationship between specific profiles of 'morphological/syntactic and semantic/pragmatic impairments' with specific psychiatric disorders, such as internalizing and externalizing disorders? (Chapter 15)