This section includes 154 letters of provided by members of the CHILDES consortium documenting their usage of the CHILDES system. These letters are provided by groups representing a total of 232 researchers. On February 4, the following letter was sent out to the info-childes@googlegroups.com mailing list, as well as the IASCL (International Association for the Study of Child Language) mailing list. All letters we received were responses to this request. The collection of published papers listed in these letters is now also included in the cumulative bibliography found in section 7. That bibliography includes a total 3104 publications, of which 1500 appeared during the last 5-year period, many of them cited in these letters.

Dear Colleague,

The NICHD grant for CHILDES is up for renewal this year and I am busy soliciting letters documenting its usage over the last five years. So, if you, your colleagues, or your students have indeed been making use of data or programs in CHILDES, could you please ask them to send me a letter documenting their usage. The new proposal is due March 1. However, I would appreciate receiving your letter as soon as possible. There is no need to send the letter in paper form; email is fine.

The new proposal will propose work on:
- continued expansion of the database,
- web-distributed user tutorials,
- creation of a new, user-friendly interface for CLAN,
- more powerful search programs that work with the new interface,
- fully automatic computation of IPSyn and DSS,
- extension of automatic morphosyntactic analysis to more languages,
- fuller support for coding of sign language with links to ELAN,
- fuller support for Conversation Analysis (CA) coding,
- computer-assisted methods for very rapid first-pass transcription,
- extended support for video-based transcription,
- extended interoperability with other major computer programs, and
- a first generation system for web-based collaborative commentary.

Letters should include citations for articles, conference presentations, talks, teaching activities, or student papers based on the use of CHILDES programs or data. I will then include these citations in the CHILDES-BIB that is on the web, as well as in the proposal. Also, they should mention any new features that you would like to see included in the programs and the database.

You can just email your letter to macw@cmu.edu. I will acknowledge each email, as I receive it. Many thanks for your support for the continuation of this important component of the scientific infrastructure for the study of child language development.

Brian MacWhinney
Professor of Psychology, CMU
macw@cmu.edu
<table>
<thead>
<tr>
<th></th>
<th>Support Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Aguilar-Mediavilla, Eva</td>
</tr>
<tr>
<td>2.</td>
<td>Albalá, María José</td>
</tr>
<tr>
<td>3.</td>
<td>Argus, Reili</td>
</tr>
<tr>
<td>4.</td>
<td>Aukrust, Vibeke Grøver</td>
</tr>
<tr>
<td>5.</td>
<td>Andrén, Mats</td>
</tr>
<tr>
<td>6.</td>
<td>Arkenberg, Marnie</td>
</tr>
<tr>
<td>7.</td>
<td>Avram, Larisa</td>
</tr>
<tr>
<td>8.</td>
<td>Baker, Anne</td>
</tr>
<tr>
<td>9.</td>
<td>Bassano, Dominique</td>
</tr>
<tr>
<td>10.</td>
<td>Bartsch, Karen</td>
</tr>
<tr>
<td>11.</td>
<td>Babarczy, Anna</td>
</tr>
<tr>
<td>12.</td>
<td>Bazzanella, Carla</td>
</tr>
<tr>
<td>13.</td>
<td>Beeghly, Marjorie</td>
</tr>
<tr>
<td>14.</td>
<td>Behrens, Heike</td>
</tr>
<tr>
<td>15.</td>
<td>Berko-Gleason, Jean</td>
</tr>
<tr>
<td>16.</td>
<td>Berman, Ruth</td>
</tr>
<tr>
<td>17.</td>
<td>Bernstein-Ratner, Nan</td>
</tr>
<tr>
<td>18.</td>
<td>Bertinetto, Pier Marco</td>
</tr>
<tr>
<td>20.</td>
<td>Bok-Benneman, Reineke</td>
</tr>
<tr>
<td>21.</td>
<td>Bol, Gerard</td>
</tr>
<tr>
<td>22.</td>
<td>Borensztajn, Gideon</td>
</tr>
<tr>
<td>23.</td>
<td>Bosco, Cristina</td>
</tr>
<tr>
<td>24.</td>
<td>Bryant, Judith</td>
</tr>
<tr>
<td>25.</td>
<td>Bunta, Ferenc</td>
</tr>
<tr>
<td>26.</td>
<td>Chaney, Carolyn</td>
</tr>
<tr>
<td>27.</td>
<td>Chang, Chien-ju</td>
</tr>
<tr>
<td>28.</td>
<td>Chang, Franklin</td>
</tr>
<tr>
<td>29.</td>
<td>Channell, Ron</td>
</tr>
<tr>
<td>30.</td>
<td>Clark, Eve</td>
</tr>
<tr>
<td>31.</td>
<td>Conti-Ramsden, Gina</td>
</tr>
<tr>
<td>32.</td>
<td>Codesido-García, Ana</td>
</tr>
<tr>
<td>33.</td>
<td>Christofidou, Anastasia</td>
</tr>
<tr>
<td>34.</td>
<td>Corrigan, Roberta</td>
</tr>
<tr>
<td>35.</td>
<td>Crutchley, Alison</td>
</tr>
<tr>
<td>36.</td>
<td>Cruz-Ferreira, Madalena</td>
</tr>
<tr>
<td>37.</td>
<td>De Houwer, Annick</td>
</tr>
<tr>
<td>38.</td>
<td>Demuth, Katherine</td>
</tr>
<tr>
<td>39.</td>
<td>Dickinson, David</td>
</tr>
<tr>
<td>40.</td>
<td>Dressler, Wolfgang</td>
</tr>
<tr>
<td>41.</td>
<td>Edelman, Shimon</td>
</tr>
<tr>
<td>42.</td>
<td>Ely, Richard</td>
</tr>
<tr>
<td>43.</td>
<td>Erkelens, Marian</td>
</tr>
<tr>
<td>No.</td>
<td>Name</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>90</td>
<td>Naigles, Letitia</td>
</tr>
<tr>
<td>91</td>
<td>Nakayama, Mineharu</td>
</tr>
<tr>
<td>92</td>
<td>Narasimhan, Bhubana</td>
</tr>
<tr>
<td>93</td>
<td>Nicholas, Johanna</td>
</tr>
<tr>
<td>94</td>
<td>Ninio, Anat</td>
</tr>
<tr>
<td>95</td>
<td>Noccetti, Sabrina</td>
</tr>
<tr>
<td>96</td>
<td>Ogura, Tamiko</td>
</tr>
<tr>
<td>97</td>
<td>Pan, Barbara</td>
</tr>
<tr>
<td>98</td>
<td>Pearson, Barbara</td>
</tr>
<tr>
<td>99</td>
<td>Perfors, Amy</td>
</tr>
<tr>
<td>100</td>
<td>Peterson, Shira</td>
</tr>
<tr>
<td>101</td>
<td>Petrič, Teodor</td>
</tr>
<tr>
<td>102</td>
<td>Pfaff, Carol</td>
</tr>
<tr>
<td>103</td>
<td>Pine, Julian</td>
</tr>
<tr>
<td>104</td>
<td>Quay, Susan</td>
</tr>
<tr>
<td>105</td>
<td>Radford, Andrew</td>
</tr>
<tr>
<td>106</td>
<td>Richard, Brian</td>
</tr>
<tr>
<td>107</td>
<td>Roberts, Joanne</td>
</tr>
<tr>
<td>108</td>
<td>Roepner, Tom</td>
</tr>
<tr>
<td>109</td>
<td>Rogers, Vivienne</td>
</tr>
<tr>
<td>110</td>
<td>Rose, Yvan</td>
</tr>
<tr>
<td>111</td>
<td>Roy, Deb</td>
</tr>
<tr>
<td>112</td>
<td>Rowe, Meredith</td>
</tr>
<tr>
<td>113</td>
<td>Rowland, Caroline</td>
</tr>
<tr>
<td>114</td>
<td>Rozendaal, Margot</td>
</tr>
<tr>
<td>115</td>
<td>Santos, Ana Lúcia</td>
</tr>
<tr>
<td>116</td>
<td>Sanz-Torrent, Monica</td>
</tr>
<tr>
<td>117</td>
<td>Savickiene, Ineta</td>
</tr>
<tr>
<td>118</td>
<td>Seliar-Cabral, Leonor</td>
</tr>
<tr>
<td>119</td>
<td>Selimis, Stathis</td>
</tr>
<tr>
<td>120</td>
<td>Shirai, Yasuhiro</td>
</tr>
<tr>
<td>121</td>
<td>Silver, Rita</td>
</tr>
<tr>
<td>122</td>
<td>Sinigaglia, Jociele</td>
</tr>
<tr>
<td>123</td>
<td>Slancova, Daniela</td>
</tr>
<tr>
<td>124</td>
<td>Slater, Carol</td>
</tr>
<tr>
<td>125</td>
<td>Slobin, Dan</td>
</tr>
<tr>
<td>126</td>
<td>Soderstrom, Melanie</td>
</tr>
<tr>
<td>127</td>
<td>Sosa, Anna</td>
</tr>
<tr>
<td>128</td>
<td>Stephany, Ursula</td>
</tr>
<tr>
<td>129</td>
<td>Syret, Kristen</td>
</tr>
<tr>
<td>130</td>
<td>Szagun, Gisela</td>
</tr>
<tr>
<td>131</td>
<td>Tang, Gladys</td>
</tr>
<tr>
<td>132</td>
<td>Tardif, Twila</td>
</tr>
<tr>
<td>133</td>
<td>Tare, Medha</td>
</tr>
<tr>
<td>134</td>
<td>Theakston, Anna</td>
</tr>
<tr>
<td>135</td>
<td>Thomadak, Evangelia</td>
</tr>
</tbody>
</table>
136. Tonelli, Livia .......................................................... 312
137. Tran, Jennie............................................................ 313
138. Tremblay, Anne ....................................................... 314
139. Tsay, Jane ............................................................. 316
140. Tsou, Chi-Zong........................................................ 321
141. Ullman, Michael ..................................................... 322
142. Uziel-Karl, Sigal ..................................................... 323
143. Vaccaro, Santo ....................................................... 326
144. Van der Lely, Heather ............................................. 328
145. Veneziano, Édy ....................................................... 329
146. Viau, Joshua .......................................................... 335
147. Voeikova, Maria ..................................................... 337
148. Weirather, Randy .................................................... 338
149. Weist, Richard ....................................................... 339
150. Wellman, Henry ..................................................... 341
151. Winskel, Heather .................................................... 342
152. Wong, Anita .......................................................... 343
153. Wong, Richard ....................................................... 345
154. Yip, Virginia and Matthews, Steven ......................... 347
155. Zhou, Jing ............................................................. 349
Dear Dr. MacWhinney,

As a researcher who has used the CHILDES I’m glad to describe my work. I’m working on phonology and morphosyntax in SLI, so I have used the CHILDES to do the phonetic transcription and the phonological and morphosyntax analyses. For this proposes I have used the UNIBET alphabet and developed a special codification system based in the dependent tiers called: %pho, %sil, %pal, %num and %erp. Then we used the FREQ and PHONFREQ CL AN programs to extract the data for the analyses.

We have published this research in these articles:


I hope this information would be useful for the CHILDES Project.

Palma de Mallorca, 18th February of 2008

Eva Aguilar Mediavilla
2. Albalá, Maria José

From: Maria Jose Albala <malbala@ile.csic.es>
Date: February 18, 2008 12:21:22 PM GMT+01:00
To: Brian MacWhinney <macw@cmu.edu>
Subject: Re: CHILDES renewal request

Dear Professor MacWhinney: thank you very much for your dedication to CHILDES, a Project so useful for all of us.

In my Center at Spanish High Council of Scientific Research, I am teaching CHILDES methodology to students. As a result, in recent years several master dissertations using CHILDES data-base and CLAN programs were presented. I have presented the results of my work with CHILDES tools at several workshops:


I am working also in the international Project on Pre and Proto Morphology in Language Acquisition, directed at the Austrian Academy of Sciences by professor Wolfgang U. Dressler. CHAT transcriptions from the corpus Marrero-Albalá and CLAN programs have allowed us to publish, in the framework of this Project, two articles:


Other publications using our CHAT transcriptions and CLAN programs are:


Aguirre, Carmen, María José Albalá & Victoria Marrero (2005), “Mami, te quierito: la adquisición del diminutivo en español”. In Mayor Cinca, María Ángeles, Begoña Zubiauz de Pedro & Emiliano Díez-Villoria (eds.), Estudios sobre la adquisición del lenguaje. Salamanca: Ediciones Universidad de Salamanca, pp. 120-144.

Thank you very much and best regards,

María José Albalá
Dear Prof. MacWhinney,

I have used CHILDES project services already almost 15 years and am using it now in my everyday work. I am very pleased by the information that you propose to NIH to continue to expand the database of the CHILDES project, to create a new, user-friendly interface for CLAN, to extend automatic morphosyntactic analysis to more languages.

The services that the CHILDES project provides is essential for work on child language in my university in Estonian language department. I have introduced it to my students working on child language and they have used it several times for their bachelor and master thesis. I am working on the new database for CHILDES this year, 14 hours of total 44 hours of videofilms (recording of spontaneous speech of triplets aged 1;0-4;0) is already transcribed and partially coded. I have also made some preparations for constructing the MOR-program for Estonian language data. The CHILDES programs have allowed me to publish, among others the following books and articles:

Argus, Reili 2008. Eesti keele muutemorfoloogia omandamine [Acquisition of Morphology in Estonian]. Tallinn University Dissertations on Humanities.

Argus, Reili, in print. Acquisition of Estonian: some typologically relevant features. Sprachtypologie und Universalenforschung.


Cordially,
Reili Argus
Tallinn University, Institute of Estonian Language and Culture, Estonia
Dear professor Brian MacWhinney,

I write this letter in support of the CHILDES system. At the University of Oslo we make extensive use of the CLAN and CHAT facilities in our child language and classroom discourse research. Attached are some publications and conference presentations since 2003. I do hope you will succeed in securing funding for the CHILDES system for the years to come.

Oslo, February 8, 2008

Sincerely,

Vibeke Grøver Aukrust
Professor/ Head of department

Published papers making use of CHILDES since 2003:


**Aukrust, Vibeke Grøver.** Learning to talk about talk: A study of dinner table conversations in two communities in Norway and the United States. . I: *BUCLD 27:*

Strømsø, Helge Ivar; Aukrust, Vibeke Grøver. Lesing og kognitiv utvikling: er det noen sammenhenger?. Nordisk Pedagogik 2003;23(2):65-78


Rydland, Veslemøy; Aukrust, Vibeke Grøver. Second language learners' out-of-frame talk in peer pretend play. EARLI 12th Biennial Conference for Research on Learning and Instruction; 29.08.2007 - 29.08.2007

In press:

Aukrust, Vibeke Grøver: Classroom conversations as immediate and long-term resources for oral language and literacy acquisition in young immigrant second language speakers. Accepted for publication in Journal of Applied Developmental Psychology.


Aukrust, Vibeke Grøver & Rydland, Veslemøy: Does it matter? Talking about ethnic diversity in preschool and first-grade classrooms. Accepted for publication in Journal of Pragmatics

Aukrust, Vibeke Grøver: Girls’ and boys’ conversational participation across four grade levels in Norwegian classrooms: taking the floor or being given the floor? Accepted for publication in Gender and Education

Submitted:

Aukrust, V. G. (submitted) Classroom discourse across four grade levels – a study of participant structures and conversational moves in teacher-led classroom conversations. Submitted to Linguistics and Education.

Hello!

I am a PhD student at the department of Languages and Literature, Lund University, Sweden. My thesis is about the development of child gestures (between 18-36 months) and I use CLAN a lot to analyze a longitudinal Thai/Swedish corpus of mother-child interaction that we have created here in Lund, using the CHAT fileformat. The main reason is that it saves its files as text files which allows me to easily write all kinds of scripts to analyze the transcriptions and the coding I've done on my data. CLAN does a fantastic job with that. Other "competing" tools like ELAN have much more complicated file formats and is therefore less easy to use when it comes to creating your own automatized scripts for analysis. The simple file format also has the benefit that it is fast to produce in comparision to other somewhat similar tools. I have tried a lot of different programs for transcribing video data (like Transana, Anvil, ELAN and others), and in the end CLAN has proven to be most useful to me. I have also been using both the PC and Mac version and needless to say, the fact that CLAN exists for both these platforms is also great, since it makes collaboration between different researchers working on different platforms easy.

You mention potential improvement of sign language coding in CLAN and all improvements in that direction is of course direct improvements of my abilities to produce a good PhD thesis. Even though I find CLAN very useful, there ARE improvements to be done when it comes to sign language/gesture analysis, and therefore I recommend anyone to grant all the money there is to further development of CLAN! :)

I whish you luck.

Mats Andrén
PhD student in General Linguistics,
Department of Languages and Literature,
Lund University,
Sweden
6. Arkenberg, Marnie

Dear Professor MacWhinney,

I am writing in support of your renewal effort for NIH funding for the CHILDES project. The CHILDES database has played a significant role in my program of research on the investigation of children’s vocabulary acquisition and the components of the child’s environment that support lexical learning.

I have used the CHILDES database for two previous projects, and I am currently using it for a third. In the first project I investigated the reciprocal role of questioning by parents and children on children’s vocabulary knowledge. In the second I supervised an undergraduate senior’s honors thesis that investigated the differences in question-asking by parents and children from low- versus middle-income families and the effect of questioning on more general conceptual knowledge.

In my upcoming work I plan to look at differences in the pattern of question asking by parents in different cultures—work I could not do if it were not for the international collaborative effort of the CHILDES community and the work you have done and continue to do to promote that collaboration. As a junior scientist I have found the CHILDES database to be not just immensely helpful, but imperative; I was able to take advantage of longitudinal data that would have been impractical for me to collect myself at this point in my career. Furthermore, as an Assistant Professor at a teaching college that leaves little time for research, the use of the CHILDES database and the tools associated with it, allows me to still remain a productive investigator despite the little time I have to collect language data myself. I heartily support a new proposal that would enable the continuation of the use of quality child language data and the tools used to best investigate those data. Below is a list of works that have utilized the CHILDES system.


Yours sincerely,

Marnie E. Arkenberg
Dear Professor MacWhinney,

In support of your application for the continuation of a NIH research grant I would like to give a brief description of how I have been using CHILDES at the University of Bucharest, Romania, over the last six years.

With your support, I have collected a small data base of Romanian child language: weekly 60 minute recordings of two monolingual Romanian children. The data were collected for a period of 20 months at the children's home on a weekly basis, with a small number of breaks caused by holidays (B. 1.03 – 3;01, A. 1;09 – 3;07). 16 recordings from the Bianca corpus were transcribed with CHILDES and they are in the CHILDES data base. 10 recordings from the Antonio corpus and 6 from the Bianca corpus were transcribed (a first draft) and they will be sent to the CHILDES data base as soon as the transcribed files have been checked. Work on the corpus is ongoing. Using the CHILDES programmes has made the transcriptions possible, offering the first longitudinal corpus in electronic format which can be used for the study of the acquisition of morphology and syntax in child Romanian.

The data have been used by my colleague Martine Coene (Antwerpen) and myself for several publications, conference talks and invited talks (see the list below). Our focus is on the acquisition of morpho-syntax. Using the CHILDES programmes has made our analyses possible. We are also using the data on work for a sub-project (focussing on the acquisition of adjectives) of the Pre- and Protomorphology project coordinated by Professor Wolfgang Dressler, The Austrian Academy. The data have also been used by some of my MA students for their MA graduation paper. I also used the CHILDES data base in my acquisition classes.

I believe that research on the CHILDES programmes is extremely important especially for those who, like myself, study languages other than English, for which no other longitudinal corpora of child language are available. The absence of such corpora negatively affects not only the study of typical language development but also the study of language impairments, whose description is practically impossible without a description of typical development.

Sincerely yours,
Larisa Avram
List of publications based on data transcribed and analyzed using CHILDES (2002-2007)


Coene, M. L. Avram (submitted) 'Can children tell us anything we did not know
about parameter clustering?". Amsterdam: John Benjamins.


**Invited talks**


Avram, L. 'On finiteness and the complementizer phrase in child Romanian', Ca' Foscari University, Venice (Erasmus exchange programme), September 2005.

**Papers presented at conferences/ workshops**

Avram, L., M. Coene' Why is it difficult to reach Agreement?' *LSRL 32* Toronto, 2002.


Avram, L. 'The No Functional Projection Hypothesis Revisited', The Annual Conference of the English Department, Transilvania University, Brašov, March 2003.


Coene, M., L. Avram 'What early clitics can tell us about early subjects', GALA 2003, Utrecht (poster).


Avram, L. M. Coene 'On the acquisition of finiteness. The case of child Romanian', The Annual Conference of the English Department, University
of Bucharest, June 2005.
Coene, M., H. Taelman, L. Avram, S. Gillis: 'Early bare infinitives are universally non-finite...but not always infinitives!', Congress of the International Association for the Study of Child Language (IASCL), Berlin July 2005.
Avram, L., M. Coene: 'The Complementizer Phrase in child Romanian: an early discourse-anchor', Generative Approaches to Language Acquisition (Gala 2005), Siena, 7-10 September (poster).
Avram, L. 'On the acquisition of direct object clitics.' The Annual Conference of the English Department, University of Bucharest, 23-25 June 2006.

Avram, L. 'Object clitics as Last Resort: implications for language acquisition', The Romance Turn II, Utrecht, 7-9 September 2006 (invited speaker).
8. Baker, Anne

From: "Baker, A.E." <A.E.Baker@uva.nl>
Date: February 5, 2008 8:47:23 AM GMT+01:00
To: <macw@cmu.edu>
Subject: support letter for CHILDES: if you need more, let me know

Dear Professor MacWhinney,

As director of the Amsterdam Centre for Language and Communication and chair of Psycholinguistics at the University of Amsterdam, The Netherlands, I wish to express my wholehearted support for the continuation and expansion of the CHILDES facilities. We use CHILDES in both our undergraduate and graduate teaching programs teaching the students to use the database and analysis tools, and requiring them to carry out studies using both the Dutch and data from other languages. The continual improvement of the tools has meant that different types of studies have become possible and this we have greatly appreciated.

For our Ph.D. students CHILDES means that they can carry out quite ambitious projects using available data rather than necessarily having to collect their own data in their brief three year projects. Many of my PhD students have made use of the CHILDES database. For example a thorough cross-linguistic search with regard to the acquisition of aspect was carried out by Annerieke Boland (2006): Aspect, tense and modality: theory, typology and acquisition. LOT, Utrecht. Also the dissertation of Erica Thrift would not have been possible without the database: 2003, Object drop in the acquisition of Dutch, LOT Utrecht. Currently Margot Rozendaal and I are publishing work on the acquisition of reference based on CHILDES cross-linguistic data to appear in the Journal of Child Language this year. Esther Parigger working on disorders has also used data from non-typical populations from the database.

As is evident, much work would not have been possible in research and teaching without the database and therefore it is essential that it continues and also is extended to include current technical possibilities.

Yours sincerely,

Anne Baker

prof. dr. Anne E. Baker Psycholinguistics, Language Pathology and Sign Linguistics Universiteit van Amsterdam Spuistraat 210, NL 1012 VT Amsterdam. Tel. +31-20-5253853, 3864 (secr.) email:a.e.baker@uva.nl email
With this letter, I want to give warm support to the CHILDES new proposal, which I consider to be one of the most central and helpful projects in the area of child language research.

During last years, the work I carried out on early productive child language with students, collaborators and colleagues have made great use of CHILDES data and programs. In particular, we relied on the CHAT coding system to collect and transcribe a large amount of data from French children. We have contributed to the CHILDES database with the corpus of a French child from 14 to 30 months (Pauline). Moreover, although we usually elaborate specific coding corresponding to our research purposes, we also use the CHILDES programs for general or initial analyses.

Among the articles recently published in which I referred to CHILDES, I should like to mention the following ones:


Reference to the CHILDES system was also made in various conference presentations, such as:


Sincerely yours

Dominique Bassano
10. **Bartsch, Karen**

February 11, 2008

Brian MacWhinney  
Department of Psychology  
Carnegie Mellon University  
5000 Forbes Ave.  
Pittsburgh PA 15213

Dear Professor MacWhinney:

I write to support your continuing efforts to maintain and improve the CHILDES database and programs. Roughly a quarter of my current research program, which is devoted to investigating children’s developing understanding of mind and other aspects of social cognition, depends on my access to the CHILDES data and programs. My students, undergraduate and graduate, are involved in research using the database. Finally, I also use transcripts from the database in several classes that I teach. I’ll elaborate briefly.

As you know, I have used the CHILDES database and programs for over twenty years to investigate conceptual development in non-laboratory contexts. Within the past five years, I have used the database to investigate children’s developing appreciation of persuasion, as it relates to concepts of mental states, and also to explore children's earliest moral sensibilities. The former project, which was supported by a 3-year grant from the National Science Foundation (grant # BCS0131496), produced evidence that young children are surprisingly engaged in conversation that resembles attempts to change the beliefs of those with whom they interact. On the other hand, we also learned that mind-oriented persuasion techniques were rare in the at-home conversations of these young children. Two graduate students and ten undergraduates had the opportunity to receive training on the use of such data as they were involved in one or more aspects of this project. A revised manuscript "...." reporting the primary results of this study is currently under review at *Social Development* [Bartsch, K., Wright, J.C., & Estes, D. Young children’s persuasion in everyday conversation: Tactics and attunement to others’ mental states].

In a separate line of research, I have collaborated by my graduate student, Jen Wright, in an exploration of children's earliest references to moral issues. In this study, the conversations of two children represented in the CHILDES database were subjected to intensive scrutiny with an eye to examining every conversation that could possibly be construed as having moral import. Among other results, we found that young children talked with adults about some topics (e.g., emotion) more than others (e.g., fairness), not always in the patterns that would be predicted from previous literature. A manuscript reporting these results was recently published (Merrill-Palmer Quarterly [Wright, J.C., & Bartsch, K. (2008). Portraits of early moral sensibility in two children's everyday conversation. *Merrill-Palmer Quarterly, 54*(1), 56-85].
Finally, I continue to use the CHILDES data in several of my classes. I regularly use language transcripts as data for my undergraduate Experimental Psychology course (PSYC 4050), in which I want students to acquire experience with naturalistic methodologies. Students come up with research questions, design coding schemes, code and analyze language data, and report on their findings. I use the data in similar but more limited ways in my Developmental Psychology (PSYC 2300) and graduate-level Advanced Developmental Psychology courses. Finally, as noted above, my graduate students have used CHILDES data in their dissertation projects and numerous undergraduates have benefited from training through participation in my research utilizing this critical resource.

In sum, the CHILDES program is essential to my professional activities and products. Thank you for continuing to maintain and improve this unique resource.

Sincerely,

Karen Bartsch, Ph. D.
Professor
Dear Prof. Brian MacWhinney,

Child language researchers, computer scientists and linguists at the University of Technology and Economics in Budapest have used the CHILDES corpus and tools of analysis in several research projects in recent years. While child language is a major focus of research at our department and use of CHILDES is part of our postgraduate curriculum, the corpora available from CHILDES have also been an essential resource in natural language engineering projects at other departments in our university. The most recent projects and related publications are listed below.

We would very much like to be able to rely on the data and tools of CHILDES in the future and would especially welcome further developments in automatic morphosyntactic tagging in languages with highly complex morphological systems and in computer-assisted methods for rapid first-pass transcription.

Yours sincerely,

Anna Babarczy,
István Fekete,
Péter Halácsy,
Ágnes Lukács,
Csaba Pléh,
Eszter Simon,
Vera Szamarasz


When talking about time, people frequently resort to spatial expressions (e.g. *Her life is still in front of her.* / *They play tennis between 7:00 and 8:00.*) Such space-time parallelisms (Bennet, 1975; Clark, 1973; Traugott, 1978) are present in all of the 53
languages studied thus far (Haspelmath, 1997) and have triggered much research in the field (e.g. Boroditsky, 2000; Gentner et al. 2002; Kemmerer, 2005).

The aim of the present study is to assess whether children use the spatial sense of spatiotemporal expressions prior to their temporal senses in early childhood. This question is of importance because it is still unclear how exactly the relationship between the two distinct domains of space and time comes about. The Theory of Conceptual Metaphor (Lakoff & Johnson, 1980) implies that this relationship develops during childhood through recurrent bodily experience. Others, like the Career of Metaphor Theory (Gentner et al., 2001) take a more flexible stance and emphasise that metaphoric representation changes as the metaphor evolves through frequent and widespread use from novel to conventional. So while the interpretation of novel metaphors involves sense creation, the interpretation of conventional metaphors involves meaning retrieval.

To test whether children use the spatial sense of spatiotemporal expressions prior to their temporal meanings, we examined Hungarian target words having both a spatial and a temporal sense in ordinary language use. These included postpositions (e.g. előtt “before”), adverbs (e.g. közel “near”), and inflections (-tól/-től “from”). Hungarian is a language with very rich morphology: these target inflections and particular postpositions can be affixed to nouns as well as pronouns.

We used the Hungarian data of the CHILDES (Child Language Data Exchange System) database for testing. We analyzed 99 files comprising 64771 utterances. The ages of target children varied between 18 and 39 months. All utterances of the Hungarian CHILDES corpus were morphologically annotated.

Partitioning target phrases into semantic classes is not trivial. In addition to spatial and temporal uses, we identified a third class of phrases containing occurrences of target words in an obligatory argument position or idiom. For instance, the sublative inflection –ra/-re (“onto”) can appear in a temporal (déle “by noon”) or a spatial sense (az asztalra “onto the table”), but can also be the inflection of an obligatory argument (darabokra esik “fall apart”).

Results show that the spatial meaning of spatiotemporal terms does not in fact precede their temporal meaning on all occasions and may even occur exclusively. Of the instances of the 22 spatiotemporal expressions found in the data, 16 appeared exclusively in a spatial sense, while 3 appeared solely with a temporal meaning. The three remaining target expressions were found with both senses in the corpus. The determination of the first appearance of these latter 3 spatiotemporal expressions in chronological terms, varies according to the matching of subjects. If subjects are matched according to age, 2 of the 3 spatiotemporal expressions appear first in the spatial, and only later in the temporal sense. However, if we match subjects by MLU, 2 of the 3 spatiotemporal expressions appear first in the temporal, and only later in the spatial sense. Nevertheless, results generally show that the majority of spatiotemporal expressions included in the analysis appear earlier in a spatial than in a temporal sense. More importantly, however, there are at least 4 (5 if we match subjects by MLU) spatiotemporal expressions that occur either exclusively or primarily in the temporal sense in the data.

Development of abstract language, Data collection and narrative corpus study of differences between typically developing children and children with Williams
syndrome in the use of abstract language – Part of EU NEST Grant Number 028714, "The Origins, Representation and Use of Abstract Concepts" 2006-2008

Narrative data of Hungarian children telling the picture story “Frog, where are you” have been collected from 17 typically developing children (age range 4;4-7;6) and 8 children with Williams syndrome (age range 9;5-15;5). The children have also been subjected to a series of standard tests (Raven, Trog, Peabody) measuring linguistic and non-linguistic cognitive abilities. Each content word used by the children in telling the story has been assigned an abstractness rating using the experimental methodology of WP1. Mean abstractness scores are computed for each child based on the word ratings.

A corpus of English language narratives of the same picture story has been compiled from the CHILDES database (http://childes.psy.cmu.edu/data/Frogs/). The above method has been applied to the English corpus for cross-linguistic comparison.

**Analogue learning in child language development**, Supported by the Hungarian Science Research Council. Grant F No. 61735. 2006-2008

The project investigates statistical patterns in linguistic errors in the spontaneous speech of English-speaking and Hungarian-speaking preschool children with an emphasis on the development of argument structures. The error patterns are correlated with the patterns of use of analogous target-like constructions in the children’s language. A computational model is built in an attempt to simulate the learning process leading to the observed empirical results.

**The development of negation in Hungarian child language**

The main concern of the project was the development of word order restrictions triggered by negation and certain other linguistic features such as focussing and subjunctive mood. It was part of a multidisciplinary research project on negation. The study involved a detailed quantitative analysis of child language corpora from the CHILDES database, using the tools of analysis provided by the CHILDES system.

**Disambiguating automatic PoS tagging in short sentences**

Our research group has been working on a Hungarian morphological analyzer since 2003. First we extended the codebase of MySpell, a reimplementation of the well-known Ispell spellchecker, yielding a generic word analysis library. At this point the development of the library has forked. Now the extended MySpell, called HunSpell, is part of the OpenOffice.org multilingual office suite. Hunmorph is the program tuned to morphological analysis. Hunmorph is supplemented by a statistical disambiguator. As short sentence length presents special difficulties in statistical disambiguation, the programme was trained on Hungarian child language data available in CHILDES.

**Publications and conference presentations**


12. **Bazzanella, Carla**

*To the National Institutes of Health (NIH)*

Since many years I have been using both the data and programs of the Child Language Data Exchange System (CHILDES), as I have already documented in 1998 and 2003. Recently I (with some colleagues of mine, first of all Cristina Bosco) used it for:

- **articles**


  Bazzanella, Bosco C., Gili Fivela B., Miecznikowski J., Tini Brunozzi F. forth.: *Polifunzionalità dei segnali discorsivi, sviluppo conversazionale e ruolo dei tratti fonetici e fonologici*, in atti del convegno *La comunicazione parlata* (Napoli febbraio 2006).


- **conference presentations:**

  Bazzanella C., Bosco C., Gili Fivela B., Miecznikowski J., Tini Brunozzi F. forth.: *Polifunzionalità dei segnali discorsivi, sviluppo conversazionale e ruolo dei tratti fonetici e fonologici, Congresso La comunicazione parlata* (Naples, February 2006).

  Bazzanella: Using corpora in pragmatic research: data and theory, *3rd Freiburg Workshop on Romance Corpus Linguistics* (September 2006)

Bazzanella, Miecznikowski Johanna: Central/peripheral functions of *allora* and ‘overall pragmatic configuration’, *ICHL* (Montreal; July 2007).

Hereby I confirm my high opinion of this most significant and useful research tool in our scientific field; I think CHILDES is a very important and fruitful component of the scientific infrastructure for the study of child language development, and I warmly hope it can be further supported by NIH funding.

Best regards,

Prof. Carla Bazzanella

Department of Philosophy
University of Turin
v. S. Ottavio 20
10124 Torino
Italy

carla.bazzanella@unito.it
fax +39 011 8124543
http://hal9000.cisi.unito.it/wf/DIPARTIMEN/Discipline1/Professori/Carla-Bazz/

Turin, February 17, 2008
Dear Brian,

I am very happy to write a note of support for your NIH continuation grant for the further development and refinement of the CHILDES system. I have two R01 grants from NICHD in which we propose to use the CHILDES to transcribe and evaluate child language from videotapes of mother-child social interaction in multiple contexts. One evaluates a longitudinal sample of children born prematurely below 32 gestational weeks (with and without white matter brain injury) and demographically similar term children from birth to 30 months, and the other evaluates a healthy, term sample of African American children and their mothers from heterogeneous socio-economic backgrounds from birth to age four. A primary goal in each project is to evaluate whether mother-child mutual regulatory processes observed during social interaction mediate the association between risk factors such as white matter brain injury or low SES and children's outcomes in various domains. One important outcome domain to be evaluated in each project is children's language development. We are assessing children's language using age-normed instruments and measures derived from transcriptions of children's spontaneous language. We also intend to submit the language transcripts derived from each of these understudied child populations to the CHILDES data base for further analysis by interested students and other researchers in the child language field.

I strongly believe that your CHILDES continuation grant deserves to be funded. It is a very important project with far-reaching and international applications.

Sincerely,
Marjorie Beeghly

Marjorie Beeghly, Ph.D.
Department of Pediatrics, Harvard Medical School
Child Development Unit, Children's Hospital
1295 Boylston Street, Suite 320
Boston, MA 02215
phone: 857-218-4340
fax: 617-730-0074
e-mail: marjorie.beeghly@childrens.harvard.edu

Contact information at WSU:
phone: 313-577-4061
e-mail: beeghly@wayne.edu
Dear Madam/Sir

That CHILDES is an indispensible resource and tool for Child Language Acquisition Research is more than obvious from the many thousand articles and tens of thousands of presentations that use the corpora and/or the programs for analysis.

In fact, CHILDES is one of the few instruments in corpus linguistics and computational linguistics that is actively used by many people all over the world – as opposed to many projects which never made it beyond the test phase. So is it time for Brian MacWhinney and NSF to relax and say "fine, we achieved our goal, let's move on to something else"? Nothing could be more damaging to the field of first language acquisition research! I have just finished editing a book on the use of corpora in acquisition research (REF), and point out that corpora are dynamic entities that require continued and intensive maintenance. A data collection that is not maintained will become unusable in just a couple of years due to changes in technology and storage media. Moreover, a large data collection is of little value or perhaps even counterproductive if there is no quality control on transcription and standardization. Here, the work of Brian MacWhinney and his team is truly admirable: not only did they spend years and years on labour-intensive and cumbersome fixing of spelling and formatting inconsistencies, but most importantly, they developed a vision as to where the new possibilities in computational linguistics can contribute to the study of child language. The morphosyntactic annotation now available for English will enable us to run quicker, more reliable and more precise analyses for almost all domains of grammatical knowledge. And they will allow us to critically examine and hopefully validate claims from the previous literature that were often based on just a few examples from small case studies.

My own use of CHILDES over the past 6 years included work on building up a German database of about 5 million words, coding part of these data regarding the inflectional properties of German child and adult language, and analyzing them regarding plural marking, verb tense, finiteness and word order (see list of publications and presentations). For these activities I made intense use of the CHILDES facilities, and was always helped by Brian MacWhinney and his team who are ready to adopt the system to its users' needs.

In addition, I just finished editing a volume for Benjamins Publishers on the use of "Corpora in Language Acquisition Research" in which CHILDES is the basis of all the articles included. It becomes clear that the future of child language research lies in the integration of the audio- and video signal to allow for new types of analyses, the standardization of transcriptions to allow for automatic coding and analyses. These will make not only make new types of analyses possible, but also entail more...
scientific rigor since analyses can be run over a larger database and more individuals, bringing out the general and the particular in acquisition.

I therefore fully support the application for further funding of the CHILDES-activities.

Yours sincerely,

Heike Behrens

Heike Behrens (U Basel): Publications based on CHILDES, 2001-2008

Behrens, Heike (in press). Usage-based and emergentist approaches to language acquisition. *Linguistics*


Behrens, Heike & Wittek, Angelika (2001). *Wer macht was im Spracherwerb? Fragebogenerhebung zu Spracherwerbsaktivitäten im deutschsprachigen Raum*. PDF-Manuskript: MPI für evolutionäre Anthropologie, Leipzig


Behrens, Heike (2006). *Discovering the middle field: The acquisition of discontinuous word order in German*. IV International Conference on Construction Grammar [ICCG 4]. Tokyo, September 1-3 (angenommen)


Behrens, Heike (2006). The social basis of language acquisition. JASGIL, Basel, May [invited]


Behrens, Heike (2005). *Discovering the rules of the German plural system.* Colloquium of the Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin, June 10 [invited].


Behrens, Heike (2005). *Goed onderzoek heeft een "Vertrauensvorschuss" nodig.* 2nd Meeting of the Academic Senate, University of Groningen, May 9 [invited].


Behrens, Heike (2004). *A Cognitive Linguistics perspective on language acquisition.* English and German Department, University of Basel, October 26 [invited].


Behrens, Heike (2004). Input - output relationship in the acquisition of parts-of-speech and constituency. Poster at Child Language Seminar, University of Bristol, July 12-14


Behrens, Heike (2003). *The empirical basis for acquiring irregular morphology.* Amsterdam Colloquium on Language, University of Amsterdam, November [invited].


Behrens, Heike (2003) *Discovering the middle field: the acquisition of discontinuous word order in German.* Cognitive Linguistics Conference. La Rioja (Spain). July


Behrens, Heike (2002). *Complex Morphological Systems and Language Acquisition.* Center for Research in Language (CRL), University of California at San Diego, April. [invited]


Behrens, Heike (2001). *Der Erwerb der Verbflexion.* Englisches Seminar, Universität Hannover, December. [invited]


Behrens, Heike (accepted). *Processes of generalizing German plural morphology.* 7th International Cognitive Linguistics Conference, Santa Barbara, CA, July.


15. **Berko-Gleason, Jean**

18 February 2008

Professor Brian MacWhinney  
Department of Psychology  
Carnegie Mellon University

Dear Brian,

I am writing to express my support for your research efforts on behalf of the child language community, and in the hope that your funding from NICHD for this invaluable work will be continued. Your development of CHILDES, including the database and the programs, has had a major impact on our discipline and it continues to do so as you move forward with new methods. I am particularly looking forward to the new CLAN interface you plan, and to creative use of the Internet. As I was thinking of my own wish list for CHILDES it seemed to me that one of the new things I would like to see developed is something akin to instant Internet access to each CLAN manual set of instructions (or even the program itself). For instance, one might Google CLAN COMBO and find just the instructions for how to perform a COMBO search, or even a COMBO applet.

We continue to rely on CHILDES in our research, and Richard Ely has sent you references to a paper we published in the *Journal of Child Language* as well as to a book chapter. I have also included a section on CHILDES in the newest edition of our textbook, *The Development of Language, 7th edition*. This section begins: “One of the most significant events in language development research has been the creation of the Child Language Data Exchange System (CHILDES).” My colleagues and I are currently involved in a project that draws on the database and programs to investigate parents’ use of language to direct children’s attention to things in the world. We are presenting our work on this topic at the meeting of the International Association for the Study of Child Language in Edinburgh this summer. The textbook and IASCL references are:


With best wishes,

Jean Berko Gleason  
Professor Emerita  
Department of Psychology  
Boston University
Dear Brian,

This is to express my total and unreserved support for continuation of the funding for CHILDES. I do not know how my lab, my research group, the informal weekly research seminar on language acquisition, development, and use that I have conducted over the past ten years or so -- and most particularly the graduate students whose dissertations I supervise in these domains -- how all these would be able to survive without your ready access to your programs and support system.

A short letter will not suffice to detail all the manifold uses that I and others in my group have made of CHILDES over the past six years. Let me mention only a few:

(1) indepth and detailed analyses of a range of corpora on early child language in Hebrew -- part of which has already been entered on the CHILDES data base, and ongoing samples now being collected and analyzed in Hebrew and also in Arabic by students at Tel Aviv University.

(2) analyses of the data base of narrative and expository texts, both written and spoken, in different languages (at Tel Aviv we have worked mainly on the English, Hebrew, French, and Spanish samples of the large-scale project on developing text construction abilities by schoolchildren, adolescents, and adults funded by a major Spencer Foundation grant to Ruth Berman as PI) -- with resulting publications in a range of journals including Journal of Pragmatics, Journal of Child Language, Linguistics, Written Languages & Literacy as well as chapters in edited books.

(3) extension of the mor tier analyses to written materials produced by native speakers and advanced learners of English in the framework of a crosslinguistic project funded by the German Israel Foundation for Research and Development to Chrstiane von Stutterheim of Heidelberg University and myself -- yielding findings that are currently being written up

(4) analysis of comparable narrative and expository texts produced by schoolchildren in English and Hebrew from different backgrounds -- high and low SES as well as children suffering from SLI in the framework of a grant from the United States Israel Binational Science Foundation to Ruth Berman, Dorit Ravid, and Marilyn Nippold -- currently under way.

(5) M.A. theses and doctoral dissertations by a score or more of students in a range of domains analyzing Hebrew language data -- early child grammar, narratives produced by children from different backgrounds, lexicon, clause-level syntax, clause linkage, and global text construction.

(6) CHILDES is used in a grant titled “The Impact of L1 on Advanced Learner Language: A cross-linguistic study of spoken and written usage” from the German-Israel Foundation for Research and Development [GIF Research Grant No. 1-789-109.4/2003]
to Ruth A. Berman, Tel Aviv U and Christiane von Stutterheim, Heidelberg U, 2005 to 2007. A database of over 500 texts (and over 120,000 words) produced in 5 languages -- English, Hebrew, German, Russian, and Arabic -- was transcribed and coded based on CHAT and CLAN conventions. The transcription conventions for Hebrew were refined, as well as those for Arabic and for Russian. Texts included both spoken and written versions of the same film-based narrative and written expository discussions.

In all these, the data have all been transcribed following CLAN conventions and subsequently coded using CHILDES programs -- in many cases elaborated for our special purposes, under the supervision of my doctoral supervisee Bracha Nir-Sagiv, who is today the leading expert on CHILDES programs in Israel and who has trained dozens of researchers and students -- not only in my lab, but at Bar Ilan, Hebrew, and Haifa Universities as well -- in use and applications of these programs.

PUBLICATIONS, TALKS & SYMPOSIA, DISSERTATIONS USING CHILDES

Ruth A. Berman, Department of Linguistics, Tel Aviv University

PUBLICATIONS (since 2002)

2002 R. A. Berman & L. Verhoeven., eds. Cross-linguistic perspectives on the development of text production abilities in speech and writing. Written Languages and Literacy, Volume 5, Parts 1 and 2 [Special Issue – 4 + 4 articles].

Crosslinguistic comparisons in later language development. In S. Strömqvist, ed. The diversity of languages and language learning. Lund: Center for Languages and Literature, 25-44.


2005 R. A. Berman. Editor: Journal of Pragmatics, 37,2. Developing Discourse Stance across Adolescence. [Special Issue – 7 articles on different languages]
D. Ravid & R. A. Berman. Information density in the development of spoken and written narratives in English and Hebrew. Discourse Processes 41, 117-149.


In press:


B. Nir-Sagiv, M. Sternau, R. Berman, & D. Ravid. School-age development of linguistic register as a distinguishing characteristic of text types across genre and modality. Israeli Journal of Literacy and Language. [In Hebrew]


R.A. Berman & B. Nir-Sagiv. Spoken and written language across the school years. In D. Aram & O. Korat, eds. The Development, learning, and promotion of literacy in the
family, at kindergarten, and at school: Festschrift in Honor of Iris Levin, Volume I.
Jerusalem: Magnes Press [in Hebrew]

Submitted:
D. Ravid & R. A. Berman. Developing linguistic register across text types: The case of
Modern Hebrew

B. Nir-Sagiv & R. A. Berman. Complex syntax as a window on contrastive rhetoric.

TALKS & SYMPOSIA (since 2002)

2002  Developmental perspectives on genre differentiation across adolescence
Tel Aviv University Linguistics Department colloquium, January 10
The role of discourse stance in developing text production abilities
Unit on Human Development, Tel Aviv University, January 29
Talking and writing about conflict situations across adolescence – with D. Ravid
Conference on Language and Society, Tel Aviv University, April
Development of middle voice as a means for agency downgrading in French and
Hebrew. International Association for the Study of Child Language, Madison, WS,
July, with Harriet Jisa Symposium on “The notion of register in sociolinguistic and
psycholinguistic research”, Israel Association for the Study of Language and Society,
Hebrew University, Jerusalem -- June, with Dorit Ravid, Bracha Nir-Sagiv, Marit
Sternau

2003  Information packaging in written and spoken narratives: developmental and
cross-language perspectives, Georgetown University Roundtable on Language and
Linguistics [GURT], February, with Dorit Ravid.

Rethinking the development of text production abilities [in Hebrew] Tel Aviv University
Unit on Human Development and Education, May

Tense/aspect shifting in narrative development across adolescence. Child Language
Seminar Annual Conference, Newcastle on Tyne, July

On modeling information density in narrative texts. Workshop on Multidisciplinary
Approaches to Discourse 2003 (MAD’03): Determination of Information and Tenor
in Texts, Utrecht, October, with Dorit Ravid.

Developing linguistic temporality across text types and languages. Script, Israel
Association of Literacy, Tel Aviv University, December [in Hebrew]

2004  Developing text production abilities across adolescence: genres, modalities, and
ages, Université Laval, Québec, February
Maturely proficient text production: means and measures. Graduate Center, City University of New York, April

Developing narrative and expository text construction: Same and/or different? Keynote Address, Society for Text and Discourse, Chicago, August

2005 The elusive notion of text quality: Discourse construction across adolescence. Keynote Address, Conference on the Phonology-Morphology Interface, Haifa University, March

Symposium on Noun Phrase structure and content in later language development: Text-based cross-linguistic analyses. International Association for the Study of Child Language (IASCL), Berlin, July, with Harriet Jisa, Dorit Ravid, Liliana Tolchinsky


Symposium on Acquiring Literacy across Genres: A multi-lingual perspective. International Association for the Study of Child Language (IASCL), Berlin, July, with Dorit Ravid and Judy Reilly

Text-embedded measures of linguistic expression in L1 English: Data from later language development. Plenary lecture, Conference on Research on Language Testing (ACROLT), Jerusalem, June

2006 The psycholinguistics of text construction. Plenary lecture, The Forum Research on Brain and Language, Bar Ilan University, January

Modality-driven versus modality-neutral features of narrative text construction, 2nd Annual Cognitive Science Conference, St. Petersburg, Russia, June, with Bracha Nir-Sagiv

Markedness in comparing oral and spoken language, Russia State University for the Humanities, Linguistics Colloquium, June

Developing N-N constructions in the language and the individual Plenary lecture, Israel Association for Language and Literacy, Levinsky College, July

Cognitive and linguistic factors in developing text quality: Global versus local? Conference on New Directions in Cognitive Linguistics, University of Sussex, Brighton, October, with Bracha Nir-Sagiv

Complex syntax in narratives: A window on contrastive rhetoric, International Conference on Language, Culture, and Mind, Paris, July
(with Bracha Nir-Sagiv)

2007  The lexicon-syntax interface in developing binominal constructions in Hebrew. L’Association Française de Linguistique Cognitive (AFLiCo), 2nd annual conference Université Villeneuve, Lille, May

The pragmatics of text construction: Comparing L1 and advanced L2 learner language Panel at International Pragmatics Association (IprA) 10th Annual Conference, Göteborg, July, with Christiane von Stutterheim

Between lexicon and syntax in developing Hebrew nominal constructions. Formal Approaches to Language Acquisition, Haifa University, October

2008  Language development beyond age five and even age ten. English Department Colloquium, Bar Ilan University, January Linguistic and cognitive facets of language development across childhood and adolescence, School of Education, Haifa University, February

THESES AND DISSERTATIONS [supervised by Ruth Berman]


Sandbank, Ana -- Writing narrative texts: A crosslinguistic and developmental study. 2004


Kupersmitt, Judy -- Temporality in texts: A crosslinguistic study of form-function relations in narrative and expository discourse, 2006


Nir-Sagiv, Bracha - Clause packages as constructions in developing narrative discourse, 2008.


Kotzer, Ayala – Effect of conditions of elicitation in the ability of children aged 3 to 7 years to reconstruct narrative text [in Hebrew], 2008
To: Professor Brian MacWhinney  
From: Nan Bernstein Ratner  
Re: NIH Competing Continuation support of CHILDES Project  
Date: February 6, 2008

For use in documenting the importance of continued NIH support of the CHILDES Project, I would like to take this opportunity to update you on use of the Project resources by my program over the past six years. We have completed, ongoing and projected research that extensively relies on the programs and databases provided by the Project. Below, I briefly sketch the projects, associated support when applicable, project team members and relevant publications and presentations that have been enabled, in large part, by the CHILDES Project resources.

- Characteristics of stuttering at onset in childhood/Comparing normal and stuttered disfluency in the speech of preschool-aged children (major co-investigator, Stacy (Silverman) Wagovich, University of Missouri). This project has utilized virtually every utility currently hosted by the CHILDES Project, including Sonic transcription, and CLAN analysis programs (MLU, VOCD, Freq, etc.). To date, it has generated the following publications:


- Infant predictors of later language development (with Rochelle Newman, University of Maryland). This project has been supported by the Bamford-Lahey Children’s Foundation and now has pending support (2008-11, Newman, PI) from the National Science Foundation. Its major publication to date is:


This project has used, and will continue to use, the full array of CHILDES Project resources and utilities, including but not limited to Sonic CHAT coding, CLAN programs for language analysis (DSS, Freq, KWAL, RELY, etc.). It will also be performing acoustic analysis of mothers’ infant-directed speech, and so we welcome anticipated changes to CLAN that will provide it with extended interoperability with other major computer programs (such as PRAAT).

- Plasticity of Language in Epilepsy Research (POLER; 1RO1 NS44280-01 NINDS); William Gaillard, PI (Children’s Hospital National Medical Center). This project has utilized the same array of CHILDES Project resources as the others in its ongoing analysis of the narrative abilities of children with chronic or recent-onset epilepsy. The first publication from this project is anticipated to appear in 2008; numerous conference presentations have already appeared (see below).

The following list details conference presentations reliant on CHILDES Project resources since 2001; students are listed in italics. I would like to call your attention to those marked with an asterisk (*). These particular national conference presentations were made by students participating in a University of Maryland Summer Research Initiative that educates students from historically under-represented groups in research techniques. Over the past decade, we have found the CHILDES Project, its archive and its programs to be a virtually unparalleled teaching device that allows students to develop a unique research question and answer it using Project resources within the short summer session.


American Speech-Language and Hearing Association Annual Convention, San Diego.


Finally, I am pleased to hear that you plan to continue to improve methods for very rapid first-pass transcription, which is important not only to our research, but to many of our clinical trainees, who make extensive use of CLAN to compute values for spontaneous language samples in their assessment of a variety of communication disorders. In this regard, planned further improvements to DSS and the addition of IPSyn will be invaluable as well.

Please let me know if I can provide any further information that will be helpful in your application for continued support for the CHILDES Project. In my opinion, it is quite clearly one of the most important initiatives in the social sciences today.
18. Bertinetto, Pier Marco

From: Pier Marco Bertinetto <Pier.Marco@Bertinetto.eu>
Date: February 14, 2008 10:27:21 AM GMT
To: macw@cmu.edu
Subject: CHILDES use

Dear Brian,
I am happy to state that I and my co-workers have been lately engaged in a project concerning the acquisition of tense and aspect by Italian children. The data of two of the three children analyzed belong to the Italian CHILDES corpus, which is indeed a very useful source of data.
Good luck with your application.
Best regards

pm
--

==================================================================================================

= Pier Marco Bertinetto
via Matteotti 197, I-55049 Viareggio (tel: ++39/0584/32215)

Ufficio:
Scuola Normale Superiore, p.za dei Cavalieri 7, I-56126 PISA
tel.: ++39/050/509111 fax: ...563513 <Bertinetto@sns.it>
Laboratorio di Linguistica: http://linguistica.sns.it

==================================================================================================

=}
19. **Bittner, Dagmar**

**From:** Dagmar Bittner <dabitt@zas.gwz-berlin.de>
**Date:** February 21, 2008 6:26:42 PM GMT+01:00
**To:** Brian MacWhinney <macw@cmu.edu>
**Subject:** Re: CHILDES renewal request

Dear Brian,

my colleagues and I are very pleased to hear that you plan to continue your work on the CHILDES project. The services that the CHILDES project provides is essential for work on child language and psycholinguistic research in general. I have been working with the CHILDES-system, and especially the German data, for many years now. First with respect to the analysis of morphological and syntactic development in first language acquisition, then for cross-linguistic comparisons in noun and verb acquisition. A great part of this work would have been impossible without the existence of CHILDES. In our research group at the ZAS in Berlin we exclusively use the CLAN analysis tool for transcribing, coding, and analyzing our data. I would like to express you and your group our gratitude for founding, maintaining, adjusting and expanding the database and the CLAN analysis tool. Below, I list recent publications and talks for which I used the CHILDES database and/or CLAN.

Publications:


Talks
2002 together with Klaus-Michael Köpcke: Grammatical complexity and the acquisition of case morphology in German. 10th International Morphology Meeting, Budapest.
2002 Grammaticalization of the input: Overgeneralizations and system pressure. IX. IASCL Congress, Madison.
2003 Aspectual interpretation of early verb forms in German; Conference on the Acquisition of Aspect, ZAS Berlin.
2004 Case before gender in L1- and L2-acquisition of German. 37th meeting of the SLE, Kristiansand.
2005 (zusammen mit N. Gagarina, M. Kühnast, V. Popova), Aspect before tense in first language acquisition of Russian, Bulgarian, and German, VIII. International Congress on Cognitive Modelling in Linguistics, Varna.
2006 The acquisition of verb inflection and finiteness in German under SLI, 12th International Morphology Meeting, Budapest.
2006 (zusammen mit Julia Siegmüller), Relational words and utterance complexity in normally developing and SLI children learning German, 2. Konferenz der DGKL, München.
2007 Infinite contrasts in the acquisition of verb inflection - evidence from normally developing and SLI children in German, 29. DGfS Jahrestagung, Siegen.

with my best wishes
Dagmar

--
Dagmar Bittner

Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung

Schützenstr. 18

10117 Berlin

Tel. 030 20192562
Fax 030 20192402
Dear Brian,
In the past few years I have been teaching two master seminars on language acquisition in Romance, together with my colleague Brigitte Kampers-Manhe. I also had my students utilize CHILDES in a seminar on 'reduced language use'. I, as well as my students, considered CHILDES a fantastic tool in these contexts. Of course project must go on. If not, this would be a great loss for the research on language acquisition and for linguistics in general. Most students of the seminars used CHILDES for their final papers. Also some students wrote their Master thesis using the corpus and its devices. I include concrete reference to the latter below.
Beste wishes,
Reineke

Reineke Bok-Bennema, associate professor in Spanish and Romance linguistics, Dptm.of Romance Languages and Cultures, Faculty of Arts, University of Groningen.

Seminars:
Master Werkgroep Taalkunde Romaans: eerste en tweede taalverwerving - february - june, 2004 (R. Bok-Bennema, B. Kampers-Manhe)
Master Werkgroep Taalkunde Romaans: eerste en tweede taalverwerving - february - june, 2006 (R. Bok-Bennema, B. Kampers-Manhe)
Master Werkgroep Taalkunde Spaans: beknopt taalgebruik, february-june 2007 (R. Bok-Bennema)

Master Theses:
Oosterwoud, Simone (2003) /La adquisición de la proyección funcional DP en español. /
Oosterkamp, Mark (2007) /La subida de los clíticos en la adquisición del español como primera y segunda lengua./
Dear Sir or Madam,

I am writing this letter to inform you about the use of the CHILDES database and the CLAN programme that I have made during the last years. I am an associate professor at the Department of Linguistics of the University of Groningen in the Netherlands. In my teaching I am using the data of Dutch children with Specific Language Impairment (the Bol and Kuiken corpus) every year in two or more advanced courses. In order to compare the linguistic characteristics of these children to typically developing children, I use the data of Dutch children with normal development (the Groningen corpus). I have been training my Dutch students to make analyses with the CLAN programme.

I am teaching as well international students that visit the University of Groningen to follow the courses of the European Master in Clinical Linguistics (Erasmus Mundus Programme). These students come from all over the world and in the CHILDES class I teach them to work with the CLAN-programme in order to analyse data coming from their own mother tongue that are to be found in the CHILDES database.

In the following publications I used the Dutch data of the CHILDES database mentioned above.


Yours sincerely,

Dr. Gerard W. Bol
22. **Borensztajn, Gideon**

**From:** "Gideon Borensztajn" <gideonbor@gmail.com>
**Date:** February 7, 2008 9:24:46 AM GMT+01:00
**To:** alavie+@cs.cmu.edu
**Cc:** "Brian MacWhinney" <macw@cmu.edu>, "Eric Davis" <dhdavis+@cs.cmu.edu>, "Jelle Zuidema" <jzuidema@science.uva.nl>, "Rens Bod" <rens@science.uva.nl>
**Subject:** Our experimental results using your syntactic annotation of the Brown corpus

Dear Alon,

I am delighted to offer you and your colleagues the privilege to be among the first to read about our experimental results based on the syntactically annotated Brown corpus. The attached article was submitted to the cognitive science conference, and is being reviewed at the moment. Rens, Jelle and I want to thank you again for all the kind help you offered, and we hope to continue our fruitful cooperation in the future!

Many friendly regards,

Gideon Borensztajn.
To whom it may concern,

even if there is an increasing availability of linguistic resources, CHILDES remains among the more valuable projects in the area of corpus linguistics, and this is attested by the fact that its tools and data are useful and interesting for a variety of researches.

Carla and I used the Italian data in our recent works with a group of young Italian researchers, and in particular, in our talks and papers at:

- the VIIth Congress of the Associazione Italiana di Linguistica Applicata (AltLA), held in Milan at the Università Cattolica, 22-23th February, 2007: Carla Bazzanella, Cristina Bosco, Barbara Gili Fivela, Johanna Miecznikowski, Francesca Tini Brunozzi; “Segnali discorsivi e tipi di interazione”, which is published in Cristina Bosisio, Bona Cambiagli, Emanuela Piemontese, Francesca Santulli (eds.) Aspetti linguistici della comunicazione pubblica e istituzionale, Guerra, Perugia.

- the Congress Comunicazione parlata, held in Naples at the Università Federico II, 24th February, 2006: Carla Bazzanella, Cristina Bosco, Barbara Gili Rivela, Johanna Miecznikowski, Francesca Tini Brunozzi “Polifunzionalità dei segnali discorsivi, sviluppo conversazionale e ruolo dei tratti fonetici e fonologici”, which will be published in the Proceedings of the conference.

Moreover, Carla cited CHILDES in its invited plenary talk at the 3rd Freiburg Workshop on Romance Corpus Linguistics "Corpora and Pragmatics - Speech acts and context-bound interaction in the light of corpora and data-bases", held at the Albert-Ludwigs University Freiburg i.Br. (Germany) in September 14th-17th, 2006 (Carla Bazzanella: “Using corpora in pragmatic research: data and theory”).

CHILDES is also used in our publications Cristina Bosco, Carla Bazzanella

And I've cited the CHILDES-Talkbank as an example of system that integrates collections of video and audio, in the paper that I've presented in the same 3rd Freiburg Workshop on Romance Corpus Linguistics, currently in press (Cristina Bosco: "Multimodality and multimedia in the pragmatic annotation of dialogue").

Cristina Bosco and Carla Bazzanella

Cristina Bosco
Assistant Professor
Dipartimento di Informatica, Università di Torino
Corso Svizzera 195, 10149, Torino
011 - 6706845
bosco@di.unito.it
February 12, 2008

Professor Brian MacWhinney
Department of Psychology
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890

Dear Professor MacWhinney,

It is my pleasure to write this letter in support of continued NIH funding for the Child Language Data Exchange System. I have used CHILDES both in my research and my teaching for many years.

Most recently, CHILDES has figured into my teaching in several ways. First, many undergraduate research assistants gained experience analyzing transcripts and gotten a sense of what the everyday language of preschoolers and their parents is like. Second, I use excerpts from transcripts to illustrate points in lectures in my Developmental Psychology and Language Development classes. Finally, I designed several student projects using CHILDES. Some of them accompany my chapter on communicative competence in the preschool years in Jean Berko Gleason and Nan Bernstein Ratner’s textbook *The development of language*. The seventh edition of this popular textbook is coming out soon.

CHILDES is enormously important for child language researchers. It enables scholars in small or poorly-funded labs to work with data they might not otherwise be able to collect. Furthermore, the availability of videos of conversations is invaluable for helping to disambiguate the function of some language, which is particularly critical in the domain of pragmatics. The range of languages, child ages, and contexts afford the possibility of analyses of multiple domains of language. In short, CHILDES is widely seen as a central tool. Please let me know if I can provide any further information about my use of this system.

Sincerely,
Judith Becker Bryant, Ph.D.
Professor and Area Director, Program in Cognition, Neuroscience, and Social Psychology
This letter is written in support of the Child Language Data Exchange System (CHILDES) Project. The CHILDES database and the Computerized Language Analysis (CLAN) are extremely valuable resources for child language researchers, linguists, and other professionals who are interested in researching child language. I have used the CHILDES database and CLAN in my research and instruction.

My primary area of research is bilingual language acquisition and phonological acquisition, and I find all of the components of the CHILDES database very useful. Due to the nature of my research, I find it especially helpful that the database has bilingual and cross-linguistic child language samples. I have worked mostly with the typed transcripts and the audio files performing linguistic and acoustic analyses on the data. This resource allows access to an abundance of shared data that may otherwise not be accessible to child language researchers. Moreover, I believe that the data available via the CHILDES database can also be a valuable instructional resource in linguistics and communication disorders courses.

The utility and popularity of CHILDES and CLAN are extremely widespread, which became evident to me when I was asked to give workshops on the database and its uses by researchers outside my immediate field. For example, I was invited to give numerous workshops on the CHILDES database, CLAN, and CHAT to audiences varying from linguists to ESL teachers, speech scientists, and communication disorders majors. The audience is always broad enough to include interested individuals from undergraduate students to senior professors. Continued expansion of the database would be extremely valuable for those of us who do research on child language. Having more tutorials that address a wide audience from students to researchers would be very beneficial. Along the same lines, funding could also be allocated to improve the interface of CLAN so as to enhance its instructional value and encourage even more widespread use. The new interface could also incorporate novel, more powerful search functions to maximize the capabilities of the program. A more complete support for Conversation Analysis coding as well as extended support for video-based transcription would also be useful enhancements. A final suggestion would be to develop a functional system for collaborative commentary (i.e., blogging) on transcripts over the web.

In my opinion, the CHILDES Project is an extremely valuable resource that has moved the field of child language research forward. Without it, our understanding of child language acquisition would not be as advanced as it currently is, and I think that its...
continued improvement is essential not only for understanding child language development, but also for communication disorders. The benefits of CHILDES have translated into real results that moved a variety of fields forward from linguistics to speech science and communication disorders. I am convinced that maintaining and enhancing CHILDES will continue to lead to important discoveries that would otherwise be difficult if not impossible to achieve. Therefore, I support the maintenance and expansion of the Child Language Data Exchange System wholeheartedly.

Sincerely,

Ferenc Bunta, Ph.D.

Department of Communication Sciences and Disorders Temple University

E-mail: Ferenc.Bunta@temple.edu
To whom it may concern:

As a teacher of undergraduates in Child and Adolescent Development, Communicative Disorders, and Liberal Studies (pre-teacher training), I am very grateful for the CHILDES system. My students make extensive use of the bibliography that is available at childes/bib; although it has not been systematically updated since 2002, it is an excellent resource for articles about children's communication development. My students also use phonology and fonts and Jean Berko Gleason's "wug" fles, available at childes. The most important childes resources for me are the discussion groups, especially info-childes. No matter what little problem may arise in my teaching or research, I have access to 1600 other minds, some of whom are always willing to lend a hand. And I encourage my students to enlist in info-childes and follow discussion strands as a way of bringing classroom lecture/discussion to life.

Carolyn Chaney
cchaney@sfsu.edu
27. Chang, Chien-ju

A Support Letter for Renewal of CHILDES Grant

Chien-ju Chang
Department of Human Development and Family Studies, National Taiwan Normal University
162 Ho-ping East Road, Section 1
Taipei 106, Taiwan
changch2@ntnu.edu.tw or wu2305@ms25.hinet.net
886-2-23696023

I have been using CHILDES since I was a student at Harvard Graduate School of Education in 1992. Over the past 16 years, I found the function of CHILDES is getting more powerful and more helpful for Chinese speakers. Being a child language researcher, I really appreciate the creation and development of the CHILDES system. I am also very grateful for the efforts Dr. Brian MacWhinney and his research team put on developing and promoting this system.

In light of the importance of the CHILDES system on research on language development of Taiwanese children, my colleagues and I proposed a three-year project to National Science Council in Taiwan to support the development of Taiwan CHILDES last October. The proposal was passed and we just started the project from this winter break. There are three main tasks we proposed to accomplish in the project: (1) to set up a Chinese child language database open to the general public; (2) to standardize the format of child language samples collected in Taiwan; (3) to develop computational tools that address the language specific properties of Chinese.

In the past five years, my students and I made a great use of the CLAN programs and we completed a number of studies based on CHILDES in Taiwan. In the following, I will list (1) reference of these publications/presentations, (2) my research projects using the CLAN programs for analysis, and (3) the CHILDES workshops held in Taiwan from 2002 to 2007:

● Publications/presentations

A. Journal papers


B. Book chapters/conference proceedings


C. Conference presentation


D. Student master’s thesis


Yeh, Y. (2004). *Teacher-pupil talk and interaction in a bilingual kindergarten English class.* Graduate School of Children English Education. National Taipei University of Education.


**CHILDES Workshop**


December 15-16, 2002. Workshop on Child Language Data Exchange System. National Taiwan University, National Taipei University of Education, National Taipei University

October 20, 2007. Workshop on computing in linguistic analysis. National Taiwan University.

**Research projects**


Co-principal investigators: Chien-ju Chang, Jane S. Tsay, and Hwa-wei Ko.

28. Chang, Franklin

From: "franklin chang" <chang.franklin@gmail.com>
Date: February 13, 2008 4:11:28 AM GMT
To: macw@cmu.edu
Subject: support letter

Brian,

I have been using 12 of the corpora from CHILDES to test unsupervised approaches to syntax acquisition (Cantonese, Croatian, English, Estonian, French, German, Hebrew, Hungarian, Japanese, Sesotho, Tamil, Welsh). The main thing that I did was to come up with a method for evaluating the accuracy of a syntax acquisition algorithm that could work in these typologically different languages. I also tested several n-gram methods and also an algorithm based on my connectionist model of syntax acquisition and sentence production (Chang, Dell, & Bock, 2006). I attached both papers in case you are interested. Below are the paper and posters that have used the corpora.


Thanks a lot for making these corpora available.

Franklin
http://www.kecl.ntt.co.jp/clip/member/chang/index.html phone: 81+774-93-5273
29. Channell, Ron

From: ron channell <rwc5@byu.edu>
Date: February 6, 2008 4:50:23 PM GMT+01:00
To: Brian MacWhinney <macw@cmu.edu>
Subject: CHILDES renewal

6 February 2006
Dear Dr. MacWhinney,

I have been using the CHILDES project resources in teaching and research here at Brigham Young University. My department offers bachelor's and master's degree training in speech-language pathology. As part of this training, our students learn to analyze samples of language both from children who are language impaired and children who are typically developing. The CHILDES samples have allowed our students to study language samples from children representing a much broader cultural perspective than could otherwise be obtained here in Provo, Utah, and to study a much greater number of samples than they would otherwise study.

The CHILDES resources have also allowed me to test a greater number hypotheses by having data readily at hand; again, these data are samples from children of various racial, ethnic, cultural, and language groups that I could not otherwise readily obtain. For example, some of the probability information that I used in my contribution to the Computerized Profiling software was extracted from CHILDES samples. A resulting article was:


Spanish CHILDES samples were also used in two recent master's theses:


I look forward to the expanded analysis tools in the CLAN part of CHILDES, so that some of my home-brewed efforts can be gracefully retired. I am grateful for the continuation of this product as a national (as well as international) resource which is of great benefit in teaching and research.

Sincerely,
Ron W. Channell, PhD
Associate Professor and Graduate Coordinator
Department of Communication Disorders
Dear Brian,

The CHILDES Archive has proved to be an invaluable resource in both my research and teaching. The best way for me to attest to the immense research value of the Archive is to append a list of recent papers and dissertations that have drawn extensively on Archive data (see below).

I also use the Archive in teaching, and have the students in my classes on first language acquisition use data from a range of languages to do mini-research projects. It's an excellent tool for this purpose too.

You have done an immense service to the field of developmental psycholinguistics with this Archive, and we now all depend on it in many different kinds of research on language in young children.

With all good wishes,

Eve V. Clark
Richard W. Lyman Professor
& Professor of Linguistics
Stanford University

Publications and projects (2002-2008) using CHILDES data:
Clark, E. V., & Nikitina, T. (2008/in press) One vs. more than one: Antecedents to plurality in early language acquisition. Linguistics 00, 000-000. [ms. 58pp.]
31. Conti-Ramsden, Gina

Dr. Brian MacWhinney
Professor of Psychology
Carnegie Mellon University

Dear Brian,

Re: Child Language Data Exchange System (CHILDES)

I am writing to let you know that we continue to find the Child Language Data Exchange System (CHILDES) immensely helpful. In the past few years we have made use of its variety of programs as well as have donated narrative data on typically developing children and children with specific language impairment (SLI). Below find a list of four recent publications which used the system extensively. In addition, having data available is most helpful for students who are in need of refining their research ideas. The data available in CHILDES is now extensive an provides a unique opportunity for students to try their ideas. We wanted to provide you with this feedback. We intend to continue using the CHILDES system in our work in the future.

Yours sincerely,

Gina Conti-Ramsden, Ph.D., FRCSLT, FBPS, AcSS

Publications using CHILDES:


Dear Brian MacWhinney,

My name is Ana Isabel Codesido García and I am currently working as a lecturer on General Linguistics at the Universidade de Santiago de Compostela (Galice-Spain). One of my research topics deals with Clinical Linguistics, and related to that I’m also interested on the typical language development in bilingual children in our speech community (Galice). That’s why, through this letter, I would like to manifest my support for the renewal of the NICHD grant for CHILDES, in order to make the continuation of this important component of the scientific infrastructure for the study of child language development possible.

Over the last five years I’ve been using CHILDES in my current work, not only as a researcher work (*cfr.* References below), but also as a teacher. In particular, in three of the subjects whose teaching I’m involved with –i.e., Computational Linguistics, Clinical Linguistics, and Applied Linguistics- my studentes have been making use of CHILDES system in order to check by themselves how Corpus Linguistics can be applied to child language studies. In other cases my students have had the chance to compare several child language studies based on different criteria such as clinical, bilingual, or typologically different languages around the world. Also they have had the chance to practice their abilities as future child language researchers transcribing by themselves, checking transcriptions from the database and even extracting some data using a selection of CLAN programs. In all the referred activities we have used the useful CHILDES webtool. Unfortunately, I do not have a copy of the results got by my students. Actually, all of them were just kind of short exercises developed during the lecture sessions.

Finally, I would like to thank you and your team for your enormous effort in order to have the “CHILDES’s world” always ready for researchers studying with maximum efficacy the exciting field of child language.

Yours sincerely,

ANA ISABEL CODESIDO-GARCÍA (MA, Doctor in Linguistics)

Facultade de Filoloxía – USC

Campus Norte, s/n

15782 Santiago de Compostela (Spain)

E-mail: fekuna@usc.es

Tel.: 00 34 981 563 100 – ext.11776
REFERENCES


CONFERENCE PRESENTATIONS:

33. Christofidou, Anastasia

From: "Christofidou Anastasia" <christo@academyofathens.gr>

Date: February 20, 2008 4:48:40 PM GMT+01:00
To: <macw@cmu.edu>
Subject: CHILDES

Research Professor Dr. A. Christofidou
Head of the Research Center for Scientific Terms
and Neologisms, Academy of Sciences, Athens
84, Solonos str., Greece
christo@academyofathens.gr

Prof. Dr. Brian MacWhinney

Dear colleague,

I have continuously been using CHILDES for 15 years in my research on Greek first language acquisition and I cannot imagine what I would do without this wonderful tool. For many years, I have been teaching postgraduate seminars on the acquisition of Greek at the Pedagogical Faculty (School of Primary Education) of the University of Athens. For the empirical work my students have to do for their term papers, I introduce them to the use of the CHILDES tools (data transcription in CHAT format, lexicon-based grammatical coding of the learners’ utterances, data analysis with the help of the CLAN programs).

CHILDES has also been used and continues to be used by more than twenty international researchers united in the Crosslinguistic Project on Pre- and Protomorphology directed by Wolfgang U. Dressler from the University of Vienna and Austrian Academy of Sciences, who explore the development of early grammar in typologically and genetically diverse languages, such as Turkish, Finnish, Estonian, Croatian, Russian, Greek, (Austrian) German, Italian, Spanish, French, Palestinian Arabic, Yucatec Maya. The latest publications and papers in which I am involved are the following:


Christofidou, Anastasia, and Ursula Stephany


Xanthos, Aris, Laaha, Sabine, Dressler, Wolfgang U., Stephany, Ursula, Voeikova, Maria, Christofidou, Anastasia et al. (submitted). On the role of morphological richness in the early development of noun and verb inflection. First Language.

I do hope that the NIH will grant you the possibility of continuing your work on the CHILDES Data Base and the CLAN programs so that your plans from which the scientific community of linguists and psycholinguists in the whole world can profit will materialize.

Thanking you for the help you have offered us,
with best wishes for the future,

Anastasia Christofidou
Dear Brian,

I am writing in support of your continuing application to NIH for CHILDES. The database and accompanying computer programs have provided an exceptional service to researchers in Child Language. As a researcher who was actually conducting research in the days of typewritten lists and hand counting of frequencies, I am always grateful for the enormous advances that CHILDES has provided. CHILDES allows those of us who are at institutions without much research support and without large cadres of graduate students to continue to make contributions in the field. I look forward eagerly to the new developments outlined in your current proposal, particularly the new CLAN interface and expansion of its search capabilities. I strongly urge NIH to support your proposal.

In the last 6 years, I have used CHILDES to support the following presentations and publications:


Corrigan, R. (August, 2004) Information provided in adult-child discourse about the meaning of adjectives. Paper presented at the Fourteenth Annual Meeting of the

Corrigan, R. (October, 2003). A case study of the construction of evaluative meaning in
adult-child conversations. Poster presented at the Cognitive Development
Society, Park City, Utah.

Corrigan, R. (June, 2003) Learning word connotations from spoken discourse: The case
of “happen.” Paper presented at the Society for Text and Discourse, Madrid,
Spain.

Corrigan, R. (July, 2002) “The company words keep” in adult conversations with
children: The case of happen. Paper presented at the Society for Text and
Discourse, Chicago.

Sincerely,

Roberta Corrigan
Professor, Educational Psychology
University of Wisconsin-Milwaukee
Milwaukee, WI 53211

e-mail: corrigan@csd.uwm.edu
35. Crutchley, Alison
Division of English
School of Music, Humanities & Media
University of Huddersfield
Queensgate
Huddersfield HD1 3DH
UK

Tuesday, 19 February 2008

To whom it may concern:

Letter in support of the renewal of funding for the CHILDES project

I have made use of data from CHILDES in my teaching over the last four years on two modules: an undergraduate English Language module, ‘Child Language Acquisition’, and a module on the MA in Modern English Language, ‘Language Development and Diversity’. For both of these modules transcripts from CHILDES are used both to exemplify linguistic behaviours at different ages, to give students practice in analysing specific behaviours across a number of children (e.g. turn-taking, topic maintenance) and to give students practice in analysing a range of behaviours within a single child (e.g. calculating MLU (Mean Length of Utterance), estimating stage of syntactic development, describing lexical development, looking for indications of comprehension or non-comprehension, etc.)

Furthermore, the MA module requires students to download and analyse data from CHILDES as part of an assessment on later language development.

For both of these modules, access to real language data from such a variety of children has been invaluable. Module feedback repeatedly indicates how useful students find opportunities to analyse this kind of data. I therefore strongly support the proposal for renewal of funding to enable the continuation and development of this very valuable resource.

Yours faithfully

Dr Alison Crutchley
Course Leader, English Language
36. Cruz-Ferreira, Madalena

Prof Brian MacWhinney
Department of Psychology
Carnegie Mellon University

17th February 2008

Letter in Support of the CHILDES Database

Dear Professor MacWhinney,

This is a comment and a report on my use of CHILDES data and programmes in the past five years, in support of the NICHD continued grant to the CHILDES Project.

Comments on CHILDES

The CHILDES database is unique in the extent and quality of resources it offers on child language. The choice of allowing free access to these resources results in a win-win situation, in that researchers in turn naturally feel motivated to contribute their own data, language corpora and/or insights free of charge.

One other reason for the steady growth of the database is the flexibility of the tools it offers for language coding and analysis. Coding and analysis tools have steadily improved over the years to contemplate features of a) languages other than English, including orthographic conventions, and b) phonetic and intonational transcription, particularly Phon resources.

My own research interests lie in child multilingualism, particularly phonology and prosody. Suggestions for further improvement of CHILDES resources from this perspective are:

- A more user-friendly interface between software in common computer operating systems and CHILDES programmes. For example, the ability to directly copy phonetically and intonationally transcribed material from word processing applications.

- The incorporation into coding facilities of phonetic symbols used in transcription of impaired speech (e.g. ExtIPA).

- A fully integrated interface between CLAN and Phon programmes.

- Regular updates to the CHILDES-Bib, particularly in view of the recent surge of research interest in child multilingualism and child prosody.
Contributions to the CHILDES database


Contributions to the CHILDES database, in preparation

- Completion of the Mcf Multilingual corpus (first edition 2000), which concerns the longitudinal and cross-sectional acquisition of Portuguese, Swedish and English of three siblings up to age 10.
- Audio and video linkage to transcripts of the whole Mcf corpus.
- Yearly update of the *Bibliography of Singapore Child Language*.

Publications drawing on the CHILDES database


This thesis was awarded the prize for Best Honours Thesis of its cohort.


Use of the CHILDES database in research and teaching

This involves familiarisation of undergraduate and graduate students with web-based and computer-based resources provided by the CHILDES database, coding and analytical tools. The related modules I chair and/or coordinate are:

- **Language Development** and **Child Language**, both undergraduate, since 2001/2002;
- **Normal Functioning 1** and **Normal Functioning 2**, both part of the Linguistics Framework for the MSc in Speech-Language Therapy launched in 2006-2007 at the National University of Singapore.

CHILDES-dedicated teaching sessions include an interactive survey of CHILDES resources (databases, CHAT, CLAN, bibliographies, tutorials, links to child language resources), which special emphasis on material which students can use for their immediate research purposes.

Research students deal with English child data, including the two varieties of the language spoken in Singapore, and with data in any of the other three official languages in Singapore, Mandarin, Malay and Tamil. Students make use of CHILDES corpora in English and, for example, in other Chinese languages, in order to make overall sense of child language acquisition.

For all students, this is also their first encounter with electronic treatment of child and/or language data. Students are encouraged to use CHAT coding conventions in their own project work for each module, which involves collection and analysis of language data from typically and/or atypically developing Singaporean children. This practice also encourages them to consider future submission of their own data to the CHILDES database.

**CHILDES-related initiatives**

In 2005, I founded a network, the SCLSIG (Singapore Child Language Special Interest Group), online at http://groups.yahoo.com/group/sclsig/

The founding of this SIG, which I chair, was inspired by CHILDES networks, namely, info-childes, on my realisation of their invaluable contributions to sharing knowledge, finding resources and fostering engagement in collaborative work across countries, languages and cultures.

Cooperation with SIG members has allowed me to put together several editions of a Singapore Child Language Bibliography, the first such publication dedicated to an officially multilingual and multicultural country, contributed to the CHILDES database.
I have now convened three SCLSIG conferences at the National University of Singapore, whose announcements and reports are disseminated to the CHILDES community, and which I intend to continue promoting on a yearly basis. Their dates and topics are:

2005. SCLSIG1 - Child Language Research in Singapore: The State of the Art and Where Do We Go from Here? (First Meeting of the Singapore Child Language Special Interest Group).

2006. SCLSIG2 - A Survey of Current Research on Language Norming in Singapore (Second Meeting of the Singapore Child Language Special Interest Group).

2007. SCLSIG3 - Language models at home and in school (Third Meeting of the Singapore Child Language Special Interest Group).

The CHILDES databases and programmes continue to enable cooperation among scholars and students worldwide, proving beyond doubt its fruitfulness and desirability. Our endeavour is to understand children through their linguistic production and environment, and thereby their linguistic and educational needs. It is my sincere hope that the NICHD will renew the CHILDES grant, so that progress in this endeavour may be sustained.

With best regards
Madalena Cruz-Ferreira

Madalena Cruz-Ferreira, BA Hons., MA, PhD
Senior Lecturer
Department of English Language and Literature
National University of Singapore
Email: ellmcf@nus.edu.sg
Academic website: http://profile.nus.edu.sg/fass/ellmcf/
Dear Professor MacWhinney,

I would like to thank you for your continuing efforts in coordinating CHILDES. CHILDES has been an invaluable tool for research in child language development all over the world, and in the analysis of spoken discourse in general. Research in child language development has taken a huge step forward since CHILDES became available to language researchers everywhere.

I have been a data-contributing member of CHILDES since 1989, and helped to publicize the CHILDES mirror site CHILDES-in-Antwerp when it was first started up by Steven Gillis in 1992. In my own more recent empirical work I have continued to rely on the tools made available through CHILDES. Please find an overview below of my more recent work that has in some way or another involved CHILDES.

I hope that you will be able to continue your much appreciated support of CHILDES, and once again I would like to express my gratitude to you for all the great work you have done as a major service to the field.

Very best regards,

Annick De Houwer, PhD
Research Professor, University of Antwerp, Belgium
Director, Research Center on Language, Media and Socialization, University of Antwerp, Belgium
Collaborative Investigator, National Institute for Child Health and Human Development (NIH), Bethesda, United States
2002-2008: Overview of my work involving CHILDES
(Annick De Houwer)

I. Monolingual Dutch project 4-year-olds

2002-2004 Finalization of the phonological and orthographic transcription and coding using CHAT of a large database of audio-recordings of Dutch-speaking four-year-olds in interaction with family members and with peers at school and teachers

2002-2003 Analysis of these transcriptions using CLAN

2002 Master's thesis supervised by A. De Houwer partially based on these transcriptions and analyses:

Kuppens, An, 2002. 'Codewisseling van dialect richting Standaardnederlands in kindgerichte spraak: een onderzoek in het Antwerps' (Code alternation from dialect to standard Dutch in child directed speech: an investigation in Antwerp), Dept. of Communication Sciences, University of Antwerp

2002 Conference presentation relying on these transcriptions and analyses:


2003 Conference presentations relying on these transcriptions and analyses:


2003 Publications of articles relying on these transcriptions and analyses:

conferentie (Articles of the fourth sociolinguistic conference), Delft: Eburon, pp. 268-276.


2004 Contribution of this corpus to CHILDES and publication through Talkbank:


II. Bilingual Dutch-English: Kate data
2004 Publication through Talkbank of my previously contributed corpus of English-Dutch bilingual data:


III. Large longitudinal (child age 5 months - 48 months) project following 30 monolingual Dutch-speaking, 31 bilingual Dutch-French-speaking and 3 bilingual Dutch-English speaking families and their firstborn children (in collaboration with Dr. Marc H. Bornstein, NICHD)

2005-now Orthographic transcription and coding using CHAT of maternal speech in video-recordings of mother-child interactions at child ages 5, 13, 20 and 48 months

2006-2007 Phonetic transcription and coding using CHAT of child speech in video-recordings of mother-child interactions at child age 20 months

2007-now Orthographic transcription and coding using CHAT of child speech in video-recordings of mother-child and investigator-child interactions at child age 48 months

2005-now Ongoing analyses of the above transcriptions and codings using CLAN

2005 Several master's theses supervised by A. De Houwer based on these transcriptions and analyses:

Decaluwé, Lien, 2005. De invloed van het geslacht van het kind op de interactiestijlen van moeders: verbale en non-verbale interactie met peuters op de leeftijd van 13 maanden (The influence of child gender on maternal interaction styles: verbal and non-verbal
interaction with 13-month-old children), Dept. of Logopedics and Audiology, University of Ghent.


Lamote, Mieke, 2005. Taalverwerving en tweetaligheid: een onderzoek naar de soorten verzoeken van moeders ten opzichte van hun 13 maanden oude jongens en meisjes in eentalige en tweetalige situaties (Language acquisition and bilingualism: an investigation into the kinds of requests by mothers of their 13-month-old sons and daughters in monolingual and bilingual situations), Dept. of Logopedics and Audiology, University of Ghent.

Rahier, Nathalie, 2005. Lexicale variatie in kindgerichte spraak: een studie van moeders in interactie met hun peuters op de leeftijd van 13 maanden (Lexical variation in child directed speech: a study of mothers in interaction with their 13-month-old children), Dept. of Logopedics and Audiology, University of Ghent.

2006 Invited plenary address in which I presented some of these analyses based on the 20 month data.


2006 Several master's theses supervised by A. De Houwer based on these transcriptions and analyses:

Cauwe, Hanne, 2006. Hoeveelheid taalaanbod van moeders in het 2e levensjaar van hun kinderen: variatie, stijlen en consistentie over de tijd (Amount of maternal language input to children in the second year of life: variation, styles and consistency over time), Dept. of Logopedics and Audiology, University of Ghent.

Flama, Alexandra, 2006. Taalkeuze in conversaties tussen moeders en hun tweetalige peuters (Language choice in conversations between mothers and their bilingual toddlers). Dept. of Logopedics and Audiology, University of Ghent.

Peeters, Lynn, 2006. De dynamiek in de non-verbale communicatie in de moeder-kind-interactie (The dynamics of non-verbal...
communication in mother-child interaction), Dept. of Communication Sciences, University of Antwerp

2007 Master's thesis supervised by A. De Houwer partially based on these transcriptions and analyses:

2008 Two master's theses supervised by A. De Houwer are in progress that rely on CHAT transcriptions and codings and on CLAN analyses of the 48 month material

IV. Other
2005 In this overview chapter, I refer to many studies of bilingual development that have relied on CHILDES:

2008 in my capacity of General Editor with Dr. Steven Gillis of the official book series of the International Association for the Study of Child Language, 'Trends in Language Acquisition Research' (John Benjamins Publishing Company) I co-directed the publication of the following volume, which refers a great deal to work carried out using CHILDES:

in progress I am preparing a textbook on Bilingual First Language Acquisition for Multilingual Matters. In it, I give overviews of the CHILDES corpora that contain data on BFLA children. I also describe CHILDES in a fair amount of detail, and advocate use of its transcription and coding methods for new studies.
in progress  Supervision of a doctoral dissertation by An Kuppens (University of Antwerp) that uses CHAT and CLAN to transcribe and analyse radio programs and television advertisements

in progress  Supervision of a doctoral dissertation by Natascha Ringblom (University of Stockholm) that uses CHAT and CLAN to transcribe and analyse recordings of family interactions (in a bilingual setting)
Dear Brian,

I am very pleased to write in support of continued funding for CHILDES. As you know, we have recently contributed 2 large longitudinal, phonetically transcribed corpora of mother-child speech for 6 children each for both English and French (the Providence Corpus (365 hrs.) and the Lyon Corpus (185 hrs.)), complete with acoustically-linked files. We have used these in many of our recent publications (see below), and they are already being used in several other labs both at Brown, and at the University of Washington, MIT, University of Lyon 2, etc. We have also contributed 98 hours of parent-child longitudinal speech interactions from 4 Sesotho-speaking children (the Demuth Sesotho Corpus), and this is being used by other researchers as well. My students and I (graduates and undergrads) have also made extensive use of other corpora on the CHILDES database as our research evolves. This has included the use of Spanish and SLI corpora, as well as many other English corpora. I routinely use all of these sources in teaching my classes on language acquisition and language disorders, providing students with abundant opportunity to explore many aspects of what both parents and children actually say, and how this changes over time. In addition, it serves as an extremely important source of ‘pilot’ data for designing experiments. Finally, it provides a means for testing certain hypotheses and/or theories about how language is acquired, thereby enhancing our understanding of this critical important domain of scientific inquiry, and the implications for both typically developing children and those with language delay.

In short, the CHILDES database, and its many research tools, provide an invaluable resource for both my teaching and research. That said, it could also be

• more user-friendly
• support other types of searchers
• improve the noun/verb categorizer with more contextual information for better accuracy (part of the search program function)
• improve computer-assisted methods for more rapid first-pass transcription – at least of
parental and older-child speech.
• increase support for the collection, annotation, and donation of new corpora.
All of these improvements would ensure maximizing the use of outstanding resource.

Sincerely,
Katherine Demuth, Professor
Department of Cognitive & Linguistic Sciences
Center for Human Development
Brown University, Box 1978
Providence, RI 02906

References where my lab has used CHILDES corpora – since 2002
Song, J.Y. & Demuth, K. In press. Compensatory vowel lengthening for omitted coda
consonants: A phonetic investigation of children’s early representations of prosodic
words. Language & Speech.

Demuth, K. In press. Exploiting corpora for language acquisition research. In H. Behrens
(ed.), Trends in Corpus Research: finding structure in data. TILAR Series, John
Benjamins.

Demuth, K. In press. Prosodic development: syllables, words and morphemes. In E.
University Press.

Demuth, K. In press. The acquisition of phonology. In J. Goldsmith, Jason Riggle, &

Demuth, K. & Ellis, D. In press. Revisiting the acquisition of Sesotho noun class
prefixes. In Crosslinguistic approaches to the psychology of language: Festschrift
for Dan Slobin, J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, Seydazçalikan, K.
Nakamura (eds.), Lawrence Erlbaum.

onsets and rising diphthongs: Evidence from the acquisition of French and
Spanish. Language Acquisition, 15, 5-57.

Cognition, 103, 163-179.

Demuth, K. 2007. The role of frequency in language acquisition. In Gü lizow, I. &
Gagarina, N. (Eds.), Frequency effects in language acquisition. Studies on

Demuth, K. 2007. Acquisition at the prosody-morphology interface. In Proceedings of


AHRC Project

In September 2005 we were awarded a 5-year grant by the Arts and Humanities Research Council (AHRC) for a project called “Code-switching and convergence in Welsh: a universal versus a typological approach” which aimed to evaluate alternative competing models of bilingual conversations in which speakers switch back and forth between two languages. Our methodology involved recording 40 hours of audio recordings of informal conversation by Welsh-English bilinguals and transcribing the data following LIDES conventions.

This project was developed with consideration of lessons learnt from our pilot project, funded by the British Academy, which was submitted to Talkbank in 2005. Transcripts include a main tier with sound link to the relevant recording, a gloss tier for Welsh data (using our own glossing conventions) and a free translation tier. Recordings are transcribed manually. Transcripts also include speaker extralinguistic variables, namely age, sex, broad areas lived in, situational information and relationship between speakers. I should note that our youngest speaker is 10 years old, and that the vast majority of our speakers (96%) are 16 or over.

At present, we have collected all recordings (c. 40h15m) and have transcribed 31 hours using CHAT. Transcripts are not yet finalised, due to our current development of an editing/proofing procedure. At present, we intend to submit to Talkbank before Christmas 2008. However, we do aim to have ready a small number of transcripts soon, to send to you for approval. We will also send an updated and much extended readme file describing the transcription conventions which are particular to our corpus.

Some analysis has already been conducted on the transcripts currently available, most notably theses that two PhD students working on the project are hoping to submit this
year. Peredur Davies is working on convergence in the speech of Welsh-English bilinguals and Jonathan Stammers is studying English verbs in Welsh and the issue of code-switching versus borrowing. Other publications and presentations using our data have been listed below.

**Publications**


**Presentations**

Bury, D. What is a bilingual clause? Colloquium on Language Contact, Language Change, Gregynog, 18th March 2007


Bilingualism Centre Corpus-Based Group

The ESRC Centre for research into Bilingualism in theory and Practice was awarded a 5 million pound grant by the ESRC in 2007. There are four main groups within the centre and these include the Neuropsychological Group, Experimental Group, Corpus Group and Education and Ethnography Group. Based at Bangor University, the new research centre aims to create greater understanding of bilingualism across the world. Research will focus on the nature of the relationship between the two languages of individual bilingual speakers as well as amongst the wider community.

The corpus group is currently at the early stages of data collection. We are hoping to achieve a large amount of recordings that will include Spanish-English data from Miami, Spanish-Welsh data from Patagonia and Welsh-English data from Wales in conjunction with the AHRC project (above). These recordings will be transcribed using the CHAT transcribing system. Please note that we are primarily interested in adult language. Our
researchers are currently being trained to use the system whilst following conventions already used by the AHRC project, as well developing our own that will be of specific use to our research. We already have 40 hours of Welsh-English recordings at our disposal (from the AHRC project) which are in the process of being transcribed using the CHAT system and we also hope to achieve around 30 hours of Spanish-English data and 30 hours of Spanish-Welsh data which will also be transcribed using the CHAT system.

Data transcribed in the CHAT system has already been used in a presentation on the centre’s official launch on the 16th of November of 2007 by the Corpus Group. The presentation was entitled ‘Sgwrsio Dwyieithog/Bilingual Conversation’

Data transcribed using CHAT was also used in the following presentation,

Dear Brian,

Since coming to Vanderbilt University I have made a special effort to introduce PhD students to the amazing repository of data and analytic tools that the CHILDES represents. Just yesterday in my doctoral class I made reference to the recent SRCD monograph that drew on the CHILDES database to examine the role of children's questions in driving development. I have a doctoral student now using the rare word list created by Catherine Snow and others as part of an analysis of her own transcript data. In the future I intend to use CHILDES software to analyze my own data,

The CHILDES project is an invaluable resource that must continue to be maintained and developed.

Sincerely,

David Dickinson
Professor and Interim Chair, Department of Teaching and Learning
#330 Peabody, 230 Appleton
Vanderbilt University
Nashville, TN 37203-5721
Dear Professor MacWhinney: I am very pleased by the information that you propose to NIH to continue to expand the database of the CHILDES project, to create a new, user-friendly interface for CLAN as well as more powerful search programs that work with the new interface, to extend automatic morphosyntactic analysis to more languages, to elaborate computer-assisted methods for very rapid first-pass transcription, as well as several other innovations that you have in mind. The services that the CHILDES project provides are essential for work on child language and other psycholinguistic research. In my university department I introduced a regular teaching of CHILDES and CLAN methodology and there it is generally used by students. Without these programs the work on child language in the Department of Linguistics and Communication Research of the Austrian Academy of Sciences would be deprived of one of its bases. And the same holds for the international Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition that investigates 15 languages. The CHILDES programs have allowed us to publish, among others the following books and articles:


Professor Wolfgang U. Dressler
Head of the Dept. of Linguistics, University of Vienna
Head of the Dept. of Linguistics and Communication Research of the Austrian Academy of Sciences
Coordinator of the Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition
Dear Brian,

I am writing to inform you of the recent use that my lab has been making of the CHILDES database.

As you may know, the work in my lab focuses on computational and behavioral aspects of language acquisition. In particular, we develop algorithms for learning syntax from raw corpus data. The algorithms are subsequently tested on a variety of corpora. In this connection, the CHILDES corpora collection constitutes an invaluable resource. I will cite three examples of its use in our work. First, when trained on child-directed speech data from CHILDES, our first large-scale algorithm, ADIOS (Solan, Horn, Ruppin, and Edelman, 2005), performed better than on other corpora, highlighting the importance of realistic input. Second, more recently, we have been able to identify some of the characteristics of child-directed speech that make it easier to learn from (Brodsky, Waterfall, and Edelman, 2007). Third, by mimicking some of these characteristics, we improved the learnability of a mini-grammar in artificial language acquisition in adults (Onnis, Waterfall, and Edelman, 2008).

We plan to keep relying on CHILDES in our future research. If I were to single out one feature that would be most useful for our purposes, it would be the amount of longitudinal data (such as Lara). We are now working on integrating extra-linguistic cues into our algorithms, so it would be nice to see synchronized information such as prosody, gesture, and attentional shifts coded alongside speech.

Thank you, and best regards.

- Shimon

p.s. The relevant references are included below.

Shimon Edelman
Professor, Department of Psychology, 232 Uris Hall
Cornell University, Ithaca, NY 14853-7601
http://kybele.psych.cornell.edu/~edelman


February 16, 2008

Dear Brian,

Thank you for your recent email requesting information regarding our use of CHILDES. CHILDES continues to be a valuable research tool for our work. The CHILDES database and analytic tools were the basis for the following study:


This project examined the use of apologies terms in parent-child discourse. Because these terms occur infrequently, approximately once every 1000 utterances, having an existing database of parent-child discourse was essential to the success of the project.

CHILDES was also the basis of the following study:


In addition, CLAN programs were used in another project that looked at the use of reported speech in the autobiographical memory reports of young adults. This work will appear in the following publication:


Finally, my Ph. D. student, Ann MacGibbon, is using CLAN programs in her dissertation research. She is looking at genre differences in evaluation in children's narratives. We have gathered a large corpus of narratives from 40 children including personal narratives, frog stories, emotion narratives, and projective narratives that we plan to contribute to the CHILDES database.

In summary, our work could not proceed without CHILDES. We are grateful for
all the past support CHILDES has received and hope that it continues to be funded.

If you have any questions or need further documentation of the value of CHILDES in our research, please feel free to contact me.

With best regards,

Richard Ely

--
Richard Ely
Department of Psychology
Boston University
64 Cummington Street
Boston, MA 02215
617-353-3204
43. Erkelens, Marian

6 February 2008
Dear Professor MacWhinney:

Over the last four years I have been conducting my PhD research focusing on the acquisition of verbs and nouns by children acquiring Dutch as their first language. During my research, I have repeatedly used CHILDES for several different goals. In fact, I could not have done the things I did if it were not for CHILDES. With this letter, I wholeheartedly would like to support your proposal for renewal of funding.

Starting my research, I wanted to get an impression of how Dutch children use their content words across sentences in early language production. Therefore, I have encoded the data available in CHILDES of four Dutch children between 2;0 and 3;6 by means of the coder option. Unfortunately, the Dutch corpora cannot yet be tagged with an automatic morpho-syntactic tagger, which would definitely have speeded up this process. After encoding, I have analyzed the data by means of the CLAN programs ‘combo’ and ‘freq’. The results of these analyses have formed the basis of my thinking about the proper way to investigate the categorization of content words in Dutch. A clear picture of proper and erroneous uses of content words across syntactic functions appeared.

Since I wanted to find out what role specific distributional cues play in category acquisition, I had to investigate the distributional cues from the input speech. Again, CHILDES was my very rich source of input data to Dutch children. I was able to replicate the major part of an input study done earlier in English, for Dutch. Unfortunately, the dense databases used for English are not (yet) available for Dutch. It would be great if people got the opportunity to provide more Dutch data over the same period of time per child. Based on the outcomes of my input study, I was able to design and conduct perception experiments testing whether Dutch infants use the same cues in early categorization as English infants.

In sum, every part of my PhD thesis relies heavily on information from Dutch child language corpora available from CHILDES. Without CHILDES it would have been impossible to do as much work as I did now. However, with certain improvements of the system I could even do much more work in less time. These improvements include expansion of the Dutch part of the database, especially with denser corpora, and the development of an automatic morpho-syntactic tagger for Dutch. I truly hope that you will get the opportunity to continue the very important work of maintaining and further developing CHILDES in the future.

Sincerely,
Marian Erkelens.
Presentations reporting on my work with CHILDES:


Dear Brian,

This letter is in response to your request for updates on uses of the CHILDES database to include in your renewal application to the NIH. In my lab, we have recently begun to include computational modeling in our research, and have many plans for such work in the future. Given these goals, I depend on the CHILDES resources you have worked to create, and I certainly want to help ensure that this essential database is maintained and expanded.

First, a graduate student, Rose Scott, is working on a project exploring the acquisition of transitivity-alternating verbs in early childhood. Previous findings in the computational linguistics literature had suggested that surface features of sentences that reflect verb argument structure can be used to divide verbs into meaningful classes (e.g., Merlo & Stevenson, 2001). Scott examined the distribution of the relevant surface features in a large corpus of child-directed speech taken from CHILDES, and used an unsupervised clustering algorithm to determine whether the features could be used to differentiate two classes of transitivity-alternating verbs. Her results suggested that animacy cues, linked to particular grammatical positions, play an important role in verb classification. Next, she conducted experiments which showed that 2-year-olds use some of the cues uncovered in her corpus analyses to assign distinct meanings to two classes of transitivity-alternating verbs. Preliminary data from this project have been published in the Cognitive Science Society proceedings (Scott & Fisher, 2006), and reported at SRCD (Scott & Fisher, 2007). A paper that describes this work is now in preparation.

Second, with a colleague in Computer Science, Professor Dan Roth, a post-doc, Yael Gertner, and a graduate student in Computer Science, Michael Connor, I have embarked on a collaborative project combining computational and experimental work. The experiments with children investigate young children's detection and use of a set of proposed simple structural cues to sentence interpretation and verb learning. The computational work involves experiments with a system for automatic semantic role labeling (SRL). Computational models of semantic role labeling face a learning problem very similar to the one children face in early sentence comprehension: The system learns to identify, for each verb in a sentence, all constituents that fill a semantic role, and to determine their roles, such as agent, patient or goal. The system is presented with a set of lexical and structural features for each input sentence, and trained on a corpus of child-directed speech. Results from experiments with children constrain the set of features presented to the model. The goal is to test the main claims of our account of early sentence comprehension by explicitly modeling learning based on the proposed shallow description of sentence structure, and to assess the consequences of these
simple representations in a substantial sample of natural child-directed speech.

We are now annotating adult-to-child utterances in a subset of American English CHILDES corpora, following the Propbank annotation scheme proposed by Palmer and colleagues (Palmer et al., 2005). We first automatically annotate the sentences using a currently available semantic role labeling system (Punykanok et al., 2005), and then hand-correct the automatic annotation. As a byproduct, the program also produces a syntactic analysis of the input sentences. We hope that, as the semantic role annotation is completed, these data will be useful to other users of CHILDES, and that we could work with you to add them to the CHILDES database.

While the hand-annotation is under way, we have begun to experiment with the learning consequences of our proposed simple structural features, using already annotated text corpora data from Propbank as training data. This work has yielded interesting results, which we described in a paper submitted to the annual meeting of the Association for Computational Linguistics (ACL). This project is supported by funding from the NIH (R01 HD054448) and the NSF (BCS 0620257).

CHILDES provides essential raw material and analytic tools for computational work on language acquisition. The field can't do without it. My collaborators and I thank you for working so hard to create this resource and keep it growing.

Sincerely,

Cynthia Fisher

References:


45. Fernández Pérez, Milagros

Over the last years, the CHILDES system has undoubtedly proved to be a most helpful tool for the group “Koiné”, which has already a decade of researching experience in the field of child language study. First of all, our attachment to CHILDES has provided us with an extensive set of guidelines for the treatment of child speech raw data, precise enough to be significant, but also supple enough to be handy and adaptable. Secondly, it has afforded us the opportunity to carry out some really enlightening crosslinguistic comparisons, owing to the huge number of corpora stored in the CHILDES web page (they compose, in fact, a reasonably large sample of the world’s main languages). Last, but not least, it has also made quite easier the task of making our attainments known both among the scientific community and among the general public. Yet we must not finish this letter without putting on record how much we are in debt with CHILDES, by making a list of the works and deeds which show signs of its influence in one way or another.

Work in progress on the “Koiné” Corpus:

At the moment, we are tackling a small number of tasks related to the encoding of relevant information in our corpus, namely:

1. As regards those special features of children speech which concern verb inflection and syntax, we are devising a more comprehensive set of labels.
2. As regards those special features of children speech which concern phonemics, we are trying to improve of the existing set of labels.
3. We are also creating a new set of labels, aimed at marking the increase in syntactical complexity of children utterances.

Selected academic works:

- Pablo Cano López & Isabel Fernández López, "Dos lingüistas pioneros en el estudio del lenguaje infantil: Marcel Cohen y Oscar Bloch", Jornadas sobre eficacia comunicativa y evaluación del habla infantil y afásica (Santiago de Compostela, noviembre y diciembre de 2005).
- Isabel Fernández López, "El incremento de constituyentes en las emisiones lingüísticas infantiles: del período telegráfico a la expansión ilimitada", Jornadas sobre eficacia comunicativa y evaluación del habla infantil y afásica (Santiago de Compostela, noviembre y diciembre de 2005).
- Isabel Fernández López "El sistema CHILDES: coficiación y búsquedas en el corpus 'Koiné'", Jornadas sobre eficacia comunicativa y evaluación del habla infantil y afásica (Santiago de Compostela, noviembre y diciembre de 2005).
- Gabriela Prego Vázquez, “Competencia social y evaluación de la afectividad comunicativa en el habla infantil”, Jornadas sobre eficacia comunicativa y evaluación del habla infantil y afásica (Santiago de Compostela, noviembre y diciembre de 2005).


**Selected conferences and reports:**

**Ph.D. thesis**

— **Title:** “La adquisición del sistema fonológico en niños castellanohablantes: perfiles y condicionamientos”
— **Ph. D. Student:** Isabel Fernández López
— **Defense date:** 20 July 2007
— **Mark:** Excellent cum laude

**Ph.D. thesis (work in progress)**

— **Title:** “El papel de las predicaciones en la emergencia de gramáticas en el habla infantil”
— **Ph. D. Student:** Ruth M. Martínez Barbosa

**Master thesis (work in progress)**

— **Title:** “La emergencia de la deixis con valor adverbial y temporal en el habla infantil”
— **Predoctoral Student:** Soraya Cortiñas Ansoar

**Selected scientific publications:**


— Milagros Fernández Pérez, "Lingüística clínica y habla infantil", M. Villayandre Llamazares (ed.), Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística (León, diciembre

Dear Prof. MacWhinney,

First, I would like to express my gratitude for the continuous effort and energy that you put into the invaluable CHILDES database system. Without this database the scientific progress that we see in the field of language acquisition would not have been possible. It is of vital importance that analyses based on child language data can be repeated and refined by other researchers to provide new insights into the same data, and leading to important discussions on fundamental issues of both linguistics and developmental psychology. Without any doubt this contribution will signify an important change to the field of child language studies. I am also very pleased with the recent progress in the development of phonetically transcribed corpora in Phon, which include the sound files and/or videos. This is an extremely valuable addition, as current discussions in phonology and in phonological acquisition focus on the amount of phonetic detail that is stored in the mental lexicon. Without inspection of detailed transcriptions and sound files research into this matter will not be possible.

There are two new research projects that I am carrying out and are currently being made available in the new Phon format, although these data themselves are not yet all available, but will be in the future, if CHILDES remains supported. One is related to a number of issues in Portuguese phonological acquisition, and is done in collaboration with Joao Freitas and Susana Correira. This concerns the acquisition of stress in Portuguese, and the fact that in the current format the actual sound files are available for analysis is a gigantic step forward in investigating different realizations of word and sentence prominence. The other project is investigating manner of articulation features in early words in Dutch, German and French, which I am carrying out with Nicole Altvater-Mackenson and Christophe dos Santos. Manner of articulation, the phenomena of lenition, strengthening and harmony can only be studied on the basis of detailed transcription and the study of the sound files. This year, our work will among others be presented at a number of major conferences on child language development such as ICIS at UBC in Vancouver, IASCL at the University of Edinburgh and GALANA at the University of Connecticut.
The two projects (studying top-down and bottom-up stress assignment and the acquisition of vowel contrast) that I mentioned in my previous letter are continuing. Although the Phon format has made tremendous progress in the past few years, and many databases are now compiled using that structure, the data themselves still not all available. For the acquisition of phonology it is very important that this is further developed.

In addition, teaching a language acquisition course without using Childes is unthinkable, and students usually use the wealth of Dutch data that are available in the database for their final projects.

I am therefore happy that Clara Levelt and myself have been able to contribute to making the CLPF database (Levelt 1994, Fikkert 1994) available in the new format, including the sound files. We moreover stimulate our colleagues and students to contribute their data as well. I think that the Phon project will have far reaching consequences for the study of child language and beyond. This is an outstanding project that has changed the language acquisition world for good, and I am delighted to be able to contribute to such a system. I am without hesitation interested in participating in the consortium organizing this new phonological database and I am very much looking forward to using it.

Sincerely yours,

Prof. Paula Fikkert

Professor first language acquisition and phonology
Radboud University Nijmegen
PO Box 9103
NL 6500 HD Nijmegen
The Netherlands
p.fikkert@let.ru.nl

Selected references


47. Franceschina, Florencia

From: "Florencia Franceschina" <Florencia.Franceschina@campus.jyu.fi>
Date: February 4, 2008 7:32:06 PM GMT+01:00
To: <macw@cmu.edu>
Subject: Re: CHILDES renewal support letter

Dear Brian

I would like to express my complete support for the continuation and proposed extensions to the CHILDES project. The CHILDES database and software tools have been extremely useful to my work as a university teacher and researcher at Lancaster University, UK, and the University of Jyväskylä, Finland. I have used the database (CHILDES and TalkBank) for teaching courses on research methodology and language acquisition. I have also used the software tools for coding and analysing data for research purposes (e.g., as part of the CEFLING research project http://www.jyu.fi/hum/laitokset/kielet/cefling at the Department of Languages, University of Jyväskylä, Finland).

CHILDES is a very efficient system for the coding and analysis of learner data in different languages. It is also very valuable that all the project's resources are made freely available on the Internet. In my view the data and software sharing model that CHILDES has adopted is an example of best practice in research. I sincerely hope that research funding can be found to continue to support this invaluable resource for students, teachers, researchers and other language professionals.

Yours faithfully

Florencia Franceschina, PhD
Department of Languages/English
University of Jyväskylä
PO Box 35, 40014, Jyväskylä, Finland
Telephone: +358 14 260 1219
Email: florencia.franceschina@campus.jyu.fi
48. Freudenthal, Daniel

Dr. T.D. Freudenthal
School of Psychology
University of Liverpool
Liverpool, United Kingdom

Prof. B. MacWhinney
School of Psychology
Carnegie Mellon University

Re: Use of CHILDES data base
Liverpool, 8 Februari 2008

Dear professor MacWhinney

I would like to take this opportunity to state my support for the continued expansion of the CHILDES data base and associated utilities. As a researcher working on the cross-linguistic computational simulation of Language Acquisition I use the corpora and analysis tools provided through the data base on a daily basis. The availability of these data and tools has been crucial in this work which has resulted in 4 journal and 10 conference publications including 2 papers in the journal Cognitive Science as well as 6 papers in the prestigious proceedings of the Annual Conference of the Cognitive Science Society.

As we seek to apply our model to a wider range of languages and phenomena I fully expect the CHILDES data base to remain the crucial resource that it has proved to be over the past years.

Best wishes,

Daniel Freudenthal

Papers in Peer Reviewed Journals:


Papers in Peer Reviewed Conference Proceedings:
49. Gagarina, Natalia

1. The CHILDES system has been successfully used in the following projects sponsored by the German Science Foundation (DFG):

- 2008-2013: Projektbereich „Discourse cohesive means in language acquisition: Intersentential anaphoric relations“
- 2003-2005: DFG-Projekt „Acquisition of the marking of +/- Telicity“
- 2000-2002: DFG-Projekt „Syntaktic Consequences of the acquisition of Morphology“

2. The CHILDES system is used successfully used in the project, sponsored by the German Ministry of Education and Research (Dr. Gagarina is a PI):

- Language Acquisition as a Window to Social Integration among Russian Language Minority Children in Germany and Israel

3. The special tagging programm for the analyses of the Russian data compatible with The CHILDES system had been created:


4. The CHILDES system was used to prepared the following publications and Talks:

Publikations (since 2001)

Book

Acquisition of the verb categories in child Russian (forthcoming 2008)
Reviewed publications *(journal articles, chapters in books and selected proceeding)*


Gagarina, N. 2004. Does the acquisition of aspect have anything to do with aspectual pairs? *ZAS Papers in Linguistics* 33, 39-61. → the article is revised for the publication in *Slavic Linguistics.*


**Publications in preparations**

Gagarina, N. Verbal errors and their reformulations in the acquisition of Russian.

Gagarina, N. The development of verbs in monolingual and bilingual context of the acquisition of Russian.

Gagarina, N. Did uncle Peter drink up the juice? On the comprehension of resultativity in Russian-speaking monolingual toddlers.

**Edited books**


**Selected invited talks**

2007, Juni The pronominal reference by mono- and bilingual German-Russian children, Universität Hamburg


**Selected conference presentations**

2008
<table>
<thead>
<tr>
<th>Monat</th>
<th>Veranstaltung</th>
</tr>
</thead>
<tbody>
<tr>
<td>Februar</td>
<td>Acquisition of the verb categories in child Russian, 27.02. 2008 (Similar to Tenure track)</td>
</tr>
<tr>
<td>Februar</td>
<td>Verb inflection beyond Protomorphology. Workshop Pre- and Protomorphology in language acquisition, Universität Wien, (mit Sigal Uziel-Karl)</td>
</tr>
<tr>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>Verb categories in a bilingual German-Russian child. 2nd Congress of the Slavic Linguistic Society, 22-26 August, Berlin</td>
</tr>
<tr>
<td>Juni</td>
<td>Deutsch-Englischer Bilinguismus: allgemeine Punkte und praktische Empfehlungen. Eine Vorlesung für die Erzieher und Eltern des Deutsch-Englisch bilingualen Kindergarten</td>
</tr>
<tr>
<td>April</td>
<td>Entwicklung des Diskurs im Spracherwerb. Staatliche Pädagogische Herzen-Universität, St. Petersburg</td>
</tr>
<tr>
<td>Februar</td>
<td>Functional and modal elements in child Russian. DGfS Tagung, Februar 28 – March 2, Siegen</td>
</tr>
<tr>
<td>Februar</td>
<td>Comprehension and production of anaphoric demonstratives by German, Bulgarian and Russian speaking children (mit D. Bittner, I. Guelzow, M. Kuehnast), DGfS Tagung, Februar 28 – March 2, Siegen</td>
</tr>
<tr>
<td>Februar</td>
<td>From protomorphology to morphology, Workshop Pre- and Protomorphology in language acquisition, Universität Wien, (mit Sigal Uziel-Karl)</td>
</tr>
<tr>
<td>Februar</td>
<td>Comprehending Russian aspectual pairs with Uncle Peter, COST Action 33 Conference, Februar 15-17, Berlin</td>
</tr>
<tr>
<td>Januar</td>
<td>Annotation und Analyse von deutschen und russischen Spracherwerbskorpora, Uni Tübingen</td>
</tr>
<tr>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Dezember</td>
<td>Natalia Gagarina &amp; Insa Gülzow (ZAS) NPs, pronouns and anaphoric reference in young children's narratives, ZAS, Berlin</td>
</tr>
<tr>
<td>November</td>
<td>Developmental variation in the acquisition of L1 Russian verb inflection by monolinguals and bilinguals. Boston University Child Development Conference, November, 3-5. (With S. Armon-Lotem and O. Gupol)</td>
</tr>
<tr>
<td>Mai</td>
<td>Adult’s reformulations of errors in child Russian. Budapest Morphology Conference.</td>
</tr>
<tr>
<td>Mai</td>
<td>Inflection verb errors in the acquisition of Russian by bilingual and monolingual children (with Armon-Lotem, Sharon, N. Gagarina and Olga Gupol). Toronto, Conference on Bilingualism, 5-7. May</td>
</tr>
<tr>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>Oktober</td>
<td>Ball oder m’atsch? The choice of early words by German-Russian bilingual children: a case study. St. Peterburg University, Conference on bilingualism. – published in the proceedings</td>
</tr>
</tbody>
</table>

Salience as a factor in models of language acquisition X. IASCL Congress, Berlin, Juli 2005

Frequency effects in language acquisition. Introduction to the Workshop on frequency in language acquisition. The annual conference of the German linguistic society, Cologne.

Emergence of verbal morphology in Russian from a typological perspective. Workshop on the acquisition of pre- and protomorphology, Vienna University.

Aspectual interpretation of early verb forms in Russian and German (with D. Bittner). International conference on Language Acquisition, Russian Academy of Sciences, St. Petersburg (in Russian)

Development of verb inflection in Russian: beyond protomorphology. Workshop on the acquisition of pre- and protomorphology, Vienna University

Do children need aspectual pairs for language acquisition? MPI for Psycholinguistics, Nijmegen.

Development of finiteness in L1 (the case of Russian), MPI for Psycholinguistics Nijmegen.

Verbal aspect – a universal category in language acquisition? A typology of the development of aspectual oppositions. Humboldt University, Berlin (in Russian)

Emergence of verb classes in Russian. Workshop on the acquisition of pre- and protomorphology, Vienna University

Development of infinitives in child Russian. Academy of Sciences, St. Peters burg (in Russian)

The acquisition of inflectional verb morphology by German and Russian children International Association for the Study of Child Language Conference, Madison, USA.

The non-target infinitives in child language. Conference on Language acquisition, the University of Samara (in Russian) – published in the proceedings

Reduplication in language acquisition (with Wolfgang Dressler, Dziubalska-Kolaczyk and Kilani-Schoch), Conference on Reduplications, Graz.

Project presentation. ZAS, Berlin

New version of morphological coding for the speech production of Russian children (in the framework of CHILDES), (with M. Voelkova). 4th European Conference on Formal Description of Slavic Languages (FDSL-4), University of Potsdam

Aspectuality and its acquisition in Russian. Conference ‘Perspectives on Aspect’, OTS/Utrecht, Netherlands, poster
Dear Brian,

The CHILDES database has been extremely valuable in my studies of generic language, and of children’s requests for causal information. Papers making use of these data include the following:


Talks making use of these data include the following:


Furthermore, the CHILDES database has been an important teaching tool. Every year that I teach the U. Michigan undergraduate Language Acquisition class, one of the assignments involves having the students analyze transcripts from CHILDES. Students very much enjoy the project, and uniformly learning a tremendous amount about language development from this “hands-on” experience.

Sincerely,

Susan A. Gelman
Frederick G. L. Huetwell Professor of Psychology
51. Gillis, Steven

Dear Brian:

In support of the new NIH proposal for CHILDES, I hereby provide the information that you asked for.

First of all, the University of Antwerp provides the necessary ‘cyberspace’ for establishing a mirror site of the CHILDES website in Antwerp. That also includes the necessary hardware for backup, personal assistance for writing updates of scripts for automatic synchronization, and the like. My university is very cooperative given the importance of the CHILDES enterprise.

As to teaching: I am devoting 5 2-hour lectures every year to CHILDES in a course on Corpus Linguistics. Moreover I am including 3 2-hour lectures to CHAT transcription issues in a course on Language Acquisition. These courses are taught on a yearly basis in Masters program in linguistics, and in the Advanced Masters in linguistics.

As to research, in what follows I present an overview of (1) the funded research projects in which we collected / are collecting corpora using the CHILDES CHAT standard, or used CHILDES corpora, (2) the publications presenting analyses of those corpora, (3) PhD dissertations using CHILDES tools and/or corpora, and (4) presentations at scientific meetings.

I hope that you will be successful with your NIH grant proposal.

Best,
Steven Gillis, PhD
Center for Dutch Language and Speech - CNTS
Department of Linguistics
University of Antwerp - City Campus
Prinsstraat 13
2000 Antwerpen
Funded research projects in which CHILDES corpora are used or in which corpora are collected and transcribed / coded using the CHILDES CHAT standard

Title: Dual electric-acoustic speech processor with linguistic assessment tools for deaf individuals with residual low frequency hearing.
Funding Agency: EU, FP7
Period: 2008-2010
Budget: €892.880

Title: Mind your syntax. Oral language development and development of Theory of Mind in deaf children with a cochlear implant.
Funding Agency: Special Research Fund (BOF) University of Antwerp
Period: 2007-2010
Budget: €210.000

Title: Speech and language acquisition in Dutch speaking children with different degrees of hearing: Hearing children and deaf children with a cochlear implant.
Funding Agency: Special Research Fund (BOF) University of Antwerp
Period: 2006-2009
Budget: €440.000

Title: Acoustic phonetic analysis of the speech of very young children with a cochlear implant.
Funding Agency: Special Research Fund (BOF) University of Antwerp
Period: 2005-2009
Budget: €130.000

Title: Lexical and morphosyntactic development in young children with a cochlear implant: A crosslinguistic study of Dutch and Hebrew.
Funding Agency: Research Foundation – Flanders
Period: 2005-2008
Budget: €285.000

Title: A constructivist analysis of 'fillers' in Dutch child language.
Funding Agency: Research Foundation – Flanders
Budget: €202.000

Title: Language acquisition by children with cochlear implants: A longitudinal investigation
Funding Agency: Research Foundation – Flanders
Period: 2001-2004
Budget: €238.000

Publications in which the CHILDES instrumentarium is used and/or CHILDES corpora are used:


Helena Taelman (2004) "Syllable omissions and additions in Dutch child language"

TO WHOM IT MAY CONCERN:

Prof. Brian MacWhinney has indicated to me that the grant application for continuing the CHILDES data base is forthcoming. I would like to support his application most enthusiastically.

I am a general linguist, and have come to the CHILDES data base and work on child language development only two years ago, when I started working on the acquisition of complex (embedded) clauses by children at two age ranges, 2;0-2;8 and 3;0-4;6. The extensiveness, fidelity, great care and wide age-and-language (and subjects) range of the CHILDES data base is an absolute Godsend for the kind of research project I have been pursuing. The CHILDES data base is by far the most conscientiously assembled data base I know of in linguistics. The carefully standardized format allows a researcher like myself, who was not there at the time of data-collection, to study and evaluate not only the utterances but, in particular, the adaptive-communicative context within which they were produced. The data base is rich enough to allow fine contextual discriminations.

The two papers I have just finished writing, based on the CHILDES data-base, are destined for a Symposium on the Genesis of Complex Syntax, to be hosted at Rice University, Houston this coming March, with M. Shibatani & myself acting as co-organizers. It is an international affair bringing together experts on diachronic syntax & grammaticalization, child language development, cognitive neuro-scientists and evolutionary biologists. All three people in the child language section have depended on the CHILDES data base for their papers in this symposium. I cannot imagine being able to do what we have done without the CHILDES data base.

I urge you, in the strongest terms I can marshal, to fund Prof. MacWhinney's application.

Best regards,

T. Givón
Distinguished Professor b(Emeritus) of Linguistics
and Cognitive Science
University of Oregon
<tgivon@uoregon.edu>
53. Granfeldt, Jonas

From: Jonas Granfeldt <Jonas.Granfeldt@rom.lu.se>
Date: February 24, 2008 5:27:35 PM GMT+01:00
To: Brian MacWhinney <macw@cmu.edu>
Subject: Re: CHILDES usage

Published or submitted research papers where I have used CHILDES (2005-present)


Presentations


Dear Brian

During the past years Childes was of great support for my research and teaching activity. In the articles cited below and in the conference presentation, I used data from CHILDES. I described some aspect of childes in my textbook written in Italian and used materials for illustration and for exercises. In my courses, I discuss some aspect of the Childes system.

Best whishes

Maria Teresa Guasti


Haman, Eva

Warsaw 10th February 2008

Dear Professor MacWhinney,

The CHILDES System which you are developing is an exceptionally useful tool for language data transcription and analysis for all language researchers, in particular those involved in studies on language development. I am happy to be among them and here I would like to acknowledge my latest research in which I have used CHILDES to transcribe and analyze the data. The project “Cross-Cultural Examination of Parent-Child Narratives and Conversations” (2005-2007) was awarded a Twinning Grant from the National Academy of Sciences to Dr. Andrea Zevenbergen (SUNY, Fredonia) and I was the Polish partner in this project. We have done a Polish – American English comparison of mother-child everyday conversations and co-narratives about past events. The CHILDES system enabled us to do the transcription, coding and analysis of the data. We are planning to share our data with the child language community in the future. Our first results are going to be presented during the IASCL Congress in Edinburgh (July, 2008). Currently one of my M.A. students, Aleksandra Olszaska, is preparing her M.A. thesis using these data and CLAN software. I strongly support your continuous efforts to improve the system and to help researchers all over the world in child language data gathering, storing, analyzing and sharing with our community.

Sincerely yours,

dr. Eva Haman
Psycholinguistics Lab
Faculty of Psychology
University of Warsaw
Dear Brian

I have received your request concerning CHILDES. First I have to apologize for answering so late, but I have been abroad during February. We are using CHILDES in our current work. We have been making great use of your data concerning Kim Plunkett’s Danish speech corpora as well as programs in CHILDES.

We have transcribed the Odense Twin Corpus (a longitudinal corpus of spontaneous child speech from six Danish twin pairs and their families) as well as the Script Study (a semi-structured experiment with 80 Danish children in the age of 3-10 years) in CLAN using the CHILDES-conventions. We have subsequently coded the CLAN-transcriptions in our own data linguistic coding and search system for Danish, OLAM (in FileMaker 9, developed together with Claus Lambertzsen, Berlin) (phonetic, phonological, morphological). When we obtain the resources we shall prepare The Odense Twin Corpus to the CHILDES-database.

Furthermore we have OLAM-coded Kim Plunkett’s Danish speech corpora as well.

**Publications:**

**Presentations:**
“The Classical Task: From singular to plural form in Dutch, Danish, Austrian German, and Hebrew”. Authors: Steven Gillis, Agnita Souman, Inge Molemans, Sim Dhollander (University of Antwerp); Katja Rehfeldt, Laila Kjærbæk Hansen, Hans Basbøll (University of Southern Denmark); Sabine Laaha, Johannes Bertl, Wolfgang U. Dressler (Austrian Academy of Sciences, Vienna); Naama Lavie, Ronit
Levie, Dorit Ravid (University of Tel Aviv). Accepted for presentation at the IASCL, Edinburgh 2008 as part of the symposium Why are noun plurals hard to acquire? A multi-task approach.

“Cross-sectional Naturalistic Elicitations: Scripts and Conversations in Hebrew, Austrian German, and Danish”. Authors: Bracha Nir-Sagiv, Helli Zwilling, Netta Abugov, Dorit Ravid (University of Tel Aviv); Sabine Laaha, Katharina Korecky-Kröll (Austrian Academy of Sciences, Vienna); Katja Rehfeldt, Laila Kjærbæk Hansen, Hans Basbøll (University of Southern Denmark). Accepted for presentation at the IASCL, Edinburgh 2008 as part of the symposium Why are noun plurals hard to acquire? A multi-task approach.

“Longitudinal Child Speech and Child Directed Speech: Distributional analyses of noun plurals in Danish, Dutch, Austrian German, and Hebrew”. Authors: Hans Basbøll, Laila Kjærbæk Hansen, Katja Rehfeldt (University of Southern Denmark); Agnita Souman, Steven Gillis (University of Antwerp); Katharina Korecky-Kröll, Laura E. Lettner, Wolfgang U. Dressler (Austrian Academy of Sciences, Wien); Ronit Levie, Ephratt Raz, Dorit Ravid (University of Tel Aviv). Accepted for presentation at the IASCL, Edinburgh 2008 as part of the symposium Why are noun plurals hard to acquire? A multi-task approach.


**Thesis etc.:**


Best regards
Laila Kjærbæk Hansen
Research Assistant
Center for Child Language
University of Southern Denmark
Dear Brian,

Thank you for your letter of February 9th regarding my use of CHILDES. I continue to find this an invaluable tool in several ways. First, in my research, I have often used it to explore and/or check on the emergence of particular concepts. For example, as reported in my book (The Work of the Imagination, Blackwell, 2000, p. 125 -6) I carried out studies of 2- and 3-year-olds appreciation of events that are imminent but do not actually come to pass. Having watched, for example, a toy horse gallop across the table and come to an abrupt halt at the edge, 2- and 3-year-olds refer to the near-fall. In addition, with the help of the CHILDES system, we established that children start to produce the word ‘almost’ before their third birthday, and as might be expected, they use it to refer to near misses or threatened mishaps that do not materialize.

More recently, my students and I (Corriveau, K.H., Pasquini, E. S. & Harris, P.L.) used CHILDES to check for children’s emerging references to the mind as compared to the brain and for the contribution of these two concepts to children’s developing understanding of enduring personal identity. This was an important complement to the experimental work that we conducted. The findings were published under the title “‘If it’s in your mind, it’s in your knowledge’: Children’s developing anatomy of identity’ in Cognitive Development (2005), 20, 321-340.
I also had the privilege of serving as a commentator on a landmark SRCD monograph by M. Chouinard which included a comprehensive CHILDES-based analysis of the questions that children ask, and the functions that questions serve between approximately 2 and 5 years. (Chouinard, M. (2007). Children’s Questions: A mechanism for cognitive development. SRCD Monographs, 72, Serial No. 286). Studies of children’s questions have been reported for more than 100 years. Chouinard’s monograph was the first to offer a comprehensive analysis of their emergence and cognitive role. Without CHILDES, the project would have been next to impossible.

Finally, I have used CHILDES in my laboratory class teaching. I ask individual students to analyze the data set from one or two children and then the results obtained by all students are pooled into a composite picture. Students benefit enormously from moving back and forth between the particulars of the individual children that they analyze and the larger picture that emerges when age changes are plotted across larger groups of children.

I wish you well in your request for further funding. In my judgment, it has been and will continue to be a vital tool for developmental psycholinguistics and the study of cognitive development.

Yours sincerely,

Paul Harris, M.A., D. Phil, F.B.A.
Victor s. Thomas Professor of Education
To Whom It May Concern:
I have made extensive use of the CLAN software that is provided within the CHILDES database for the last three years. In my PhD research project, I collect my own (digital) video data, which are transcribed using the CHAT editor in order to create my own corpus of mother-child and teacher-child language interactions. All transcripts are subsequently coded using the CLAN software. I created my own coding scheme with the help of the CHILDES manual, enabling me to identify features of academic language in the language interactions of very young children in interaction with their parents and teachers. I make extensive use of the analysis features such as FREQ, MOR and VOCD. The CLAN program is without any doubt essential for my work.

An overview of the papers and conference presentations I have given over the last year can be found below.

Sincerely,

Lotte Henrichs
University of Amsterdam


Dear Dr. MacWhinney:

The PAROLE corpus team (Equipe Langages, laboratoire de recherche LLS, Université de Savoie, in Chambéry, France) is honored to express its very strong support for continued development of the CHILDES language transcription and analysis suite of software. We have been using your indispensable tools over the past two years to constitute the PAROLE corpus of oral productions by speakers of three foreign languages – English, French, and Italian. Despite the fact that CHILDES was developed for the transcription and analysis of child language, we find the tools adequate and very useful for the analysis of learner language. Our subjects are young adults, having learned their L2 at school (not bilingual subjects).

During the course of our transcription process, we have submitted several questions to CHILDES, and have been truly amazed at the immediate, personal contact by either Leonid Spektor or yourself (frequently the same day!), and the efficient solution to any problem. This really is a phenomenal investment on the part of the developers and “keepers” of the system, and something that should be preserved (if possible, although we realize the unique nature of the investment on your part), recognized, and held up as a model of caring support in the scientific community. We feel that the CHILDES team is a driving force at a global level, gently and gradually obliging researchers in language acquisition to adopt coherent methods for language analysis, which enables us in turn to share data and tools, and therefore speeds and deepens our scientific understanding.

Attached, please find a brief description of the PAROLE corpus (which we hope to submit to TalkBank in May), as well as brief suggestions for certain types of analyses to take into account in future versions of CLAN. We have become extremely interested in temporal aspects of spoken fluency, and would be interested in automating some of the steps involved in calculating fluency indicators (total hesitation time, when all hesitation times are entered in the main transcription line, for example). We also include a list of the talks and publications to date on our PAROLE data.

Our sincere apologies for getting this letter of support off to you so late, and our deepest thanks for your unstinting and effective support of projects investigating the acquisition of language at all levels.

Most sincerely,

Heather E. Hilton, PAROLE project director
The PAROLE corpus (description)

The PAROLE corpus (PARallèle, Oral, en Langue Etrangère) is presently being compiled at the Université de Savoie (France). This corpus consists of oral productions by learners of L2 French (L1 Spanish, English, Swedish, German, Japanese, Korean and Chinese), of L2 Italian (L1 French) and of L2 English (L1 French and L1 German). The samples are collected at various levels of proficiency, ranging from A1 to C1 of the Common European Framework, and are obtained from guided and semi-guided oral production tasks based on picture and video prompts, and from free narration of personal experience. It is thus possible to compare productions on identical tasks across languages and across proficiency levels. These data are complemented by a battery of tests chosen to measure other aspects of each learner’s L2 knowledge and skill (vocabulary, grammar, listening) and language-learning aptitude (phonological memory and grammatical analysis skill), and by questionnaires on the subjects’ language learning background and motivation. In addition to the L2 corpus (85 subjects to date), we have compiled a corpus of native speakers of the three project languages performing the same tasks, as a source of “benchmark” data for comparison with the L2 productions (26 NS subjects).

Talks based on the PAROLE corpus:


Publications based on the PAROLE corpus:

HILTON, H. E. (submitted) Connaissances, procédures et productions orales en L2. AILE.


Suggestions for functions in CHILDES:

We find it user-friendly already! It really isn’t complicated to use.

We have been carefully timing all hesitations, pauses, etc. (anything over 200 milliseconds), using the “sonic Mode” waveform. This works beautifully. All times are entered on the main transcription line. We have therefore gone through a relatively complex process of running a keyword search on the pause symbol, saving the list as text file, and importing it into Excel, in order to calculate hesitation times for each subject. It might be fun to play with the timing aspects of a new tool – making calculations of hesitation time a bit simpler. Would it be possible to “mark” the hesitations in the sound file, and calculate the hesitation times from this, for example? (I spoke to the EXMERALDA people in Holland about this, and they seem to think it’s not too difficult.)

Time is everything; we are having great fun analysing where the most disfluent hesitations occur in our L2 productions, and interpreting why. The temporal aspects of spoken data strike us as extremely important to take into account in future generations of transcription software.

In L2 production, we’re a bit sceptical of the usefulness of first-pass transcription assistance, since the productions can be so far from a NS norm...
February 14, 2008

Dr. Brian MacWhinney
Professor of Psychology, CMU

Dear Brian,

CHILDES has been a model for inter-laboratory collaboration both in terms of the available database and in terms of the list serve that reaches our colleagues worldwide. My lab has made limited use of the data base feature in the last 5 years, but we did examine a set of transcripts for the use of grammatical ellipsis in the speech of English speaking 3 and 4-year olds (Mary is cooking, now Bill wants to). We also used some of the Brown corpus to study the use of motion verbs in parental speech (e.g., run, hop).

This kind of normative data is extremely costly to collect and becomes impossible for those who want to collect cross culturally data. I hope that you will continue to support the language research community by funding CHILDS for another 5 years.

Sincerely,

Kathy Hirsh-Pasek
Lefkowitz Professor of Psychology
Dear Professor MacWhinney

I am writing this letter to explain the projects which my colleague and I extensively use CHILDES for data analysis at The University of Queensland, Australia.

The first project (The measurement of primary traits in second language oral proficiency in second language acquisition research) is a collaborative research with Professor Ortega at University of Hawai‘i at Manoa. This project is to investigate how four primary second language traits (i.e. syntactic complexity, lexical range, accuracy and fluency) each contribute to advanced second language proficiency. An in-depth analysis of 300 speech samples produced by learners of English, Japanese and Spanish will be compared with results taken from independent proficiency measures. The project has been funded by the Australian Research Council Discovery Grant Scheme for 2005-2008. To analyse the speech data of the three languages, we have used CHILDES extensively. Several research assistants learnt the program as well. The results have been reported, and written up for journal publication.

The second project (Investigation of syntactic complexity measures in the oral and written discourse in Japanese) is related to the project above, but in this project I compared syntactic complexity of speech data of learners of Japanese with that of native speakers of Japanese. The results were presented at the conference in New Zealand in April 2007. The third project (Effects of learner background on the acquisition of writing skills in Japanese as a second language) is to investigate various aspects of written Japanese. Investigating syntactic complexity and lexical range, we have used CHILDES. The results have been presented at a few conferences and reported in a journal. The fourth project (The patterns of interaction in adult EFL classroom in China: the interrelationship between noticing and types of input) is to analyse classroom interaction in EFL classes in China. This project is The results have been reported at a conference in New Zealand last year.

Apart from the projects explained above, I have used CHILDES for data analysis for a few projects. The results are reported in various publications as shown below.

Finally one of my PhD students has been using CHILDES for her data analysis. The working title of her project is Interactional feedback and EFL learning in
China and is going to present at an international conference in Beijing in March 2008.

My colleagues and I continue to appreciate the opportunity for us to use CHILDES for data analysis and also your support. If you need further information regarding the projects, please don’t hesitate to contact me.

Yours sincerely

Noriko Iwashita

-----------------------------------------------------------------------

Noriko Iwashita, PhD
Lecturer in Applied Linguistics
Discipline Coordinator of Applied Linguistics Program
School of Languages and Comparative Cultural Studies
The University of Queensland
Brisbane, 4072
Queensland AUSTRALIA
ph. 61-7-3365 2381
fax 61-7-3365 6799
e-mail: n.iwashita@uq.edu.au

Publications

Conference presentations

Iwashita, N. & Li, F.H. (2007). “The patterns of interaction and students’ perceptions of feedback in adult EFL classroom in China” at the Investigating interaction inside and outside the classroom: A focus on research methodology colloquium, Social & Cognitive Aspects of Second Language Learning and Teaching Conference, 12-14 April, Auckland


Dear Dr. MacWhinney,

I am writing to you to manifest how important the CHILDES project has been for me in the development of several research projects I have directed over the last years at the Universidad Autónoma de Querétaro in Mexico. I first used CHILDES in a language and cognition study of Mexican, Spanish-speaking children between the ages of 10 and 36 months. Both monolingual and bilingual children were seen and the natural language samples were transcribed using CHILDES. Gestural, vocabulary and early syntactic data was obtained. That data base is currently part of the CHILDES database. Several CLAN programs were used to analyze the data. In the development of the study, not only was the system itself of great use, all queries, problems and doubts were solved almost immediately by you. Publications and presentations from this project after 2003 include:

Conference presentations:

Publications


After that research project, I carried out a study with Spanish-speaking late-talkers. These children were between 18 and 30 months of age. Data from spontaneous language samples of both late-talkers and age controls were gathered and transcribed using CHILDES. Once again, gestural, vocabulary and morphosyntactic information was obtained and analyzed using several CLAN programs. This data base and transcriptions are still not finished, but will be sent to CHILDES once it is clean and complete. Publications and presentations after 2003 include:

Conference presentations
10/07 Sustantivos y predicados en niños con retraso inicial de lenguaje (Nouns and Predicates in Late-Talking Children). IX Congreso Nacional de Lingüística (Nacional Linguistics Conference). Zacatecas, México
I am currently working on narrative samples of children with Specific Language Impairment (SLI) in a larger study. A preliminary study compares 5 children with SLI to data taken from the CHILDES database Spanish language samples. A publication of this data will be coming out soon:


We have also used CHILDES in many Masters thesis in both the Linguistics and the Educational Psychology programs at my university. CHILDES is presented, as well, as part of the coursework in Language Development courses in both programs.

I have found CHILDES to be most useful as it is a very time efficient means for obtaining vocabulary and MLU counts and to find the contexts of particular language structures. My students and I have also used the database to contrast our data with that of other countries or researchers. One part of the whole CHILDES system that we find a necessary component for anyone involved in child language development, for communicating with colleagues within the field, for updates of current research, for discussion of current topics and for exchanging information is info-childes.

For future development of CHILDES I would suggest that more information and coding be done on the main lines rather than dependent tiers. Also, a more user friendly mor line would be useful. What is good is that the program now works for Spanish. Another facet that may be helpful is to have the program compare data bases. I would also suggest that more workshops for less advanced and advanced users be given at the conferences the way it was done many years ago. As I recently acquired a computer with Windows Vista,
I am not sure if the program will work well. Hopefully it will be developed to function well with this new system.

I truly hope that the CHILDES complete system continue for many years. It is a necessary asset for anyone involved in child language. Most of all, I thank you for all of your hard work and unconditional help.

Sincerely,

Dr. Donna Jackson-Maldonado
Facultad de Lenguas y Letras
Universidad Autónoma de Querétaro
e-mail: djackson@uaq.mx
Dear Brian,

During the last ten years I have used the Child Language Data Exchange System (CHILDES) for many different projects.

1. Spoken and written texts produced by French monolingual children and adults.

There are two large corpus studies that could not have been completed without the CHILDES system. One corpus is part of an international project financed by a Major Grant from the Spencer Foundation (Chicago, Illinois, USA) « Developing literacy in different contexts and in different languages » Ruth Berman, Tel Aviv University, Principal Investigator, 1997-2003. The second is a replication of this first study using low SES monolingual French children and adolescents and was financed by the Agence National de la Recherche (ANR-05-BLAN-0313-01, Harriet Jisa, Principal Investigator, 2003-2009).

Publications using the French texts included in this corpus:


2. LONGITUDINAL STUDY OF CONVERSATIONS BETWEEN MONOLINGUAL FRENCH-SPEAKING MOTHERS AND CHILDREN WITH COCHLEAR IMPLANTS.

Fifteen children were recorded in natural conversational situations with their mothers at 6 month post-implant intervals, beginning at 12 months post-implant and finishing at 50 months post-implant. This research was supported by a grant from the Programme Cognitique (Ministère de la Recherche et de l'Education Nationale) "Plasticity and aspects of cognitive organisation in prelingual deaf children with cochlear implants: Spoken language development and cognitive abilities", Josiane Bertoncini, Principal Investigator.

Publications using this corpus:


3. LONGITUDINAL STUDY OF EIGHT NORMALLY DEVELOPING FRENCH-SPEAKING CHILDREN IN CONVERSATIONAL SITUATIONS WITH THEIR MOTHERS.

Eight children were recorded every two weeks at home in normal everyday interaction with their mothers. The data are audio- and video-recorded. Transcriptions include video linkage. This project is financed by a grant from the National Institute of Mental Health (Grant #1R01 MH60922-01A2) awarded to Katherine Demuth, Brown University (2001-
2006) and by a grant from the French government “Action Concertée Incitative, Terrains, techniques, théories, Travail interdisciplinaire en Sciences humaines et sociales” (2003-2006, Harriet Jisa, Principal Investigator. All of the data from these two studies will be given to the CHILDES data bank. At present the transcriptions for four children have been donated.

Publications using this corpus:


Rose, Yvan, Dos Santos, Christophe. 2006. Influences multiples dans l'harmonie consonantique et la métathèse en acquisition du français. Recherches linguistiques de Vincennes, n°35, p. 77-102


4. Frog stories.

A number of my graduate students are studying the development of narrative competence by using the textless picture book, Frog, Where are you? The languages included are Amharic, Arabic, French and Hungarian.

Publications using this corpus:


Finally, during the first semester of each academic year I teach a seminar in crosslinguistic perspectives on language acquisition. This seminar is attended by French-speakers as well as a number of foreign students. For the last three years I have required that the students use the data available in their native language from the CHILDES data
base for the seminar work and the final paper. Without CHILDES this seminar would be impossible.

The Child Language Data Exchange System is an essential tool for my research and teaching and constitutes an invaluable resource for the field.

Harriet Jisa
Professor of Linguistics
In the ongoing studies of the acquisition of Danish as a second language by Turkish-speaking minority students (the Køge Project) as well as in other projects carried out at this institution, we have been using the CHILDES package with great pleasure, and we intend to continue doing so. The Køge Project has used the CHILDES package since 1987, and we have a range of Danish, Turkish, and international studies which have benefited from the data collected 1989-1998 and thoroughly transcribed following the CHILDES conventions, altogether more than 140 group conversations among grade school children and several hundred conversations involving the same students and adults in face-to-face interaction. This material has been used as the basis for more than 100 books and papers written in English, Turkish, German, Swedish, and Danish (cf. the attached list). During all these years the Childes team has provided courses, administrative and practical help as well as a stream of useful communications. For this reason our project is set to continue working in the CHILDES format, and with the CHILDES programs. We heartily endorse the CHILDES team and its extraordinary work, and we strongly support that it be continuously funded in order to execute the existing plans for amending and expanding the CHILDES package.

Sincerely,

J. Normann Jørgensen
Professor
Danish as a Second Language

normann@hum.ku.dk

+45 35 32 83 64
Bibliography of the Køge Project by 2007

This list includes works which use (some of) the Køge data or describe (parts of) the Køge Project: Excluded are works that refer to the works here, or works that only (comparatively) briefly include aspects or data from the Køge project. Likewise excluded are popular articles in newspapers, etc.


Huls, Erica (2000): Politeness in Group Conversations of Turkish Students in Denmark. In Holmen & Jørgensen (eds), 101-120.


Jørgensen, J. N. (ed) (2001c): Multilingual Behavior in Youth Groups. Scandinavian studies in the simultaneous use of two or more languages in group conversations


Kirilova, Marta (2006): “Han er en fra Amager, men også lidt fra Afrika”. Dansk med accent. En empirisk baseret analyse af indfødte danskers holdninger til dansk


Pfaff, Carol (2000): Bilingual verbal Repertoires Represented in the Speech of Turkish-Danish and Turkish-German Bilingual Children in the Køge and KITA Projects. In Holmen & Jørgensen (eds), 195-229.


65. Kamandulyte, Laura

I would like inform you where we have used the program CHILDES: 1. Project „The Development of Spoken Lithuanian Corpus“ (2007–2008, prof. I. Savickiene)

The aim of the project is to collect about 80 hours (150 000 – 200 000 words) of spoken language records and to develop a grammatically annotated spoken language corpus available for users on the Internet. As collecting natural spontaneous language data is very complicated and requires special preparation, a lot of time, human effort and financial means, it consists of several consequent stages: composing a corpus, preparing methodology for grammatical annotation; data collection, transcription of records (using CHAT format), digitalisation, coding of the grammatical digitalised corpus (using CHILDES), lexicon collection, disambiguation and developing an access on the Internet.

The stages of corpus development using program CHILDES were described in:


The stages of Lithuanian corpus development using program CHILDES were presented at conference 3rd Baltic Conference on Human Language Technologies on 4-5 October 2007 (http://conference.vdu.lt/index.php?cf=7)

I work on my PhD Thesis „The acquisition of modifiers in Lithuanian“. Corpora of 5 children are created using CHILDES


Best regards
Laura Kamandulyte
Lithuania
66. Katis, Demetra

University of Athens
Faculty of Early Childhood Education
13a Navarinou str.
106 80 Athens – Greece
Fax: 0030210 368 8510
Associate Professor Demetra Katis
dimkati@ecd.uoa.gr

To
Prof. Dr. Brian MacWhinney
Dept. of Psychology
Carnegie Mellon University
Pittsburgh, PE 15213-3890, USA
macw@cmu.edu

Dear Dr. MacWhinney,

I am writing in order to express my gratitude for the chances given to researchers at the University of Athens to use the Childes Data Base for the past 12 years. My students – especially those at a graduate level – would not be able to carry out their projects, without the availability of data both from Greek (the Stephany corpus) as well as other languages. In fact, as cross-linguistic comparisons are often necessary in developmental psycholinguistics, we would face a serious obstacle otherwise. I wish to especially note a recent Ph.D. on motion verbs which compared Greek data with the Brown corpus in English (also resulting in publications).

I am also taking this chance to note that I have not unfortunately been able to contribute to the data base yet, even though I have an extensive range of corpora on the acquisition of Greek mostly as a first but also as a second language. More specifically, they include a large range of semi-spontaneous narratives (roughly 500 children), semi-experimentally elicited narratives and descriptions of pictures (L2 data here as well) (roughly 200 children), limited cross-sectional data (30 children) as well as in-depth longitudinal studies (4 children). At least one of the longitudinal corpora (of a girl aged 1;8-4;6) has
been transcribed in Chat form, but still needs a final checking of the transcriptions. In addition, the 200 narratives and descriptions of pictures from L1 and L2 speakers of Greek have also been transcribed in Chat, but with some revisions imposed by a cross-linguistic project on the acquisition of discourse in L1 and L2 (under the auspices of M. Watorek, University of Paris 8). It is lack of funding and personal time that has kept me back from preparing this data for contribution. I do certainly intend to soon contribute however the longitudinal corpus—noted above- for which work is nearly complete.

I can not even think of how research on language acquisition could continue without the data base and tools of transcription and analyses which you have made possible.

With best wishes for the continuation of the project

Demetra Katis
Associate Professor of Psycholinguistics
As a researcher in early language development in several languages, I am used to analyze spontaneous productions from children and their parents. The data are audio and video recorded in informal settings before being transcribed with the International Phonetic Alphabet. In these days, I am working with colleagues (national and international) on 20 children acquiring French, Romanian, Turkish, Tachelhit, Tunisian and Dutch. So far, more than 600 hours were transcribed and entered into a program called the LIPP program. This program allows automatic calculations of sound and sound association frequencies. Unfortunately, this software is limited to phonetic analyses, oriented to the analysis of English language and not very user friendly. Thus, we have decided to transfer our data from the LIPP program into the CHILDES Database. This transfer which is under progress with the help of Brian MacWhinney and Yvan Rose will give us a lot of opportunities in terms of data sharing and data analysis. We will not only concentrate on phonetics and phonology with the new Phon program but also on the interface between phonetic and phonological and lexical development. Moreover, not only the children’s utterances will be considered but also the mothers’ productions. Phon provides functionality for multimedia data linkage and segmentation, multiple-blind transcriptions as well as flexible analytical functions accessible through a user-friendly graphical interface. Last but not least, I am regularly using the CHILDES database in my teachings. Normal developing children but also atypically developing children are represented in the database. In the corpora one can find nice examples to illustrate the theoretical part of the lessons and make the students used to children’s speech at different ages and in different languages.
Dear Professor MacWhinney,

I am very happy to hear that you propose to NIH to continue to expand the database of the CHILDES project, to create a new, user-friendly interface for CLAN as well as more powerful search programs that work with the new interface, to extend automatic morphosyntactic analysis to more languages, to elaborate computer-assisted methods for very rapid first-pass transcription, as well as several other innovations that you have in mind. The services that the CHILDES project provides is essential for work on child language and other psycholinguistic research. Without these programs the work on child language would be deprived of one of its bases. In the international Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition that investigates 15 languages to which I belong, the CHILDES programs have allowed us to publish, among others the following books and articles:


Marianne Kilani-Schoch
MER1
Université de Lausanne
EFLE
Anthropole
1015 Lausanne
SWITZERLAND
EFLE
Dear Professor MacWhinney,

Here is a short blurb about my usage of CLAN in response to the NICHD grant solicitation letter. I sincerely hope that the grant is renewed because the program is proving to be immensely useful for my research.

As a final year doctoral candidate at Boston University’s School of Education, I chose to analyze the Sinhala and English language acquisition patterns of Sri Lankan schoolchildren in two contrasting schooling systems for my dissertation. I’m currently using the VOCD program within CLAN to measure lexical diversity in the pupils’ oral narrations replicating the following study:


Overall, CLAN was easy to learn and provides detailed reports – it is indeed wonderful to have access to such an invaluable tool free of charge. Please keep up the great work!

Best wishes,
Samantha Kulatilake
Hi Brian,

Here is my statement for the renewal; I hope it's OK. Thanks for all your great work and good luck!

Evan

***************************************************

Over the last 6 years I have used the CHILDES database and the CLAN programme extensively in my research. It is an essential part of my research programme and I am hugely indebted to everyone who has worked on the project and people who have submitted their data. In the last 6 years I have produced the following publications that have made use of CHILDES.


Cameron-Faulkner, T., & Kidd, E. (2007). I'm are what I'm are: the acquisition of 1PS-present BE. Cognitive Linguistics, 18, 1 - 22.


I will continue use the database in the future. I am currently in the initial stages of preparing a grant application to collect naturalistic data from children learning Fijian as a first language. If I am successful I will require the support of the CHILDES-using community so that my data can be transcribed to the high standards expected by the child language community (e.g., morphological coding, audio/video linking). This process will be facilitated by continued funding support for the project.

Yours sincerely,

Dr Evan Kidd (The University of Manchester, UK).

Dr Evan Kidd
Lecturer in Psychology
School of Psychological Sciences
University of Manchester
Oxford Road M13 9PL
Manchester, UK
Ph: +44 (0) 161 275 2578
Fax: +44 (0) 161 275 8587
http://www.psych-sci.manchester.ac.uk/staff/108727
Dear Brian,

I am a graduate student working towards a PhD who has used CHILDES extensively for my MA thesis (which also turned into an article and several conference presentations), a joint research project, and preparatory research for my dissertation (in progress). I will also be using CHILDES in the coming weeks for a tutorial for the students in my undergraduate Language Acquisition class, and it is likely that a number of them will make use of the database for their final projects in this course. I am including the references for the work just listed in hopes that this will help indicate how vital a resource the CHILDES database is for L1A researchers and students like myself.

PUBLICATIONS


THESES


PRESENTATIONS


= 

Thank you for all the work you do on this project, and best of luck in getting continued funding. It will certainly benefit us all!

best,
susannah kirby
unc-ch
72. Korecky-Kröll, Katharina
From: Katharina Korecky-Kröll <Katharina.Korecky-Kroell@oeaw.ac.at>
Date: February 14, 2008 11:45:05 AM GMT
To: <macw@cmu.edu>
Subject: CHILDES renewal request

Dear professor MacWhinney,

I am a PhD student at professor Dressler's department. I am currently working on a PhD thesis on the acquisition of German noun morphology by two Austrian children, and I am doing all my basic analyses in CHILDES. I have been using CHILDES for each of my publications and talks (see the list below).

Best regards,
Katharina Korecky-Kroell

Publications:


Talks:


Klampfer, Sabine, Dorit Ravid, Katharina Korecky-Kröll, Gregor Laaha & Wolfgang U. Dressler: Early noun plurals in German: regularity, productivity or default? 9th International Congress for the Study of Child Language, Madison, Juli 2002

Korecky-Kröll, Katharina. Wimmbad gehma? Der Erwerb von Subjekt- und Objektpronominen im Wiener Deutschen. 31. Österreichische Linguistiktageung,


Katharina Korecky-Kröll: Austrian longitudinal data. Francqui Colloquium on the Acquisition of Noun Plurals, Antwerpen, 3.12.2005

Dressler, Wolfgang U., Steven Gillis, Johannes Bertl, Katharina Korecky-Kröll, Sabine Laaha, Hans Basboll, Katja Rehfeldt: Core morphology and its acquisition: Plural formation in Dutch, German and Danish. Vortrag: 12th International Morphology Meeting, Budapest, 27.5.2006.


Dear Professor MacWhinney,

Thank you for informing that you propose to NIH to continue the CHILDES project. I feel that the CHILDES project is very important for child language studies especially from the comparative perspective. I have used the programs and the database in my university teaching for several years and I hope that in the future we could add also material about Finnish to it. I have supervised several MA thesis which have benefited from the CHILDES-system, and my participation in the international research project "Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition" has also relied on CHILDES.

Let me also mention a few of my publications which have benefited from the CHILDES:


With best wishes
Klaus Laalo
professor of Finnish language
University of Tampere
Dear Brian,

I hereby want to give the CHILDES project my full support. Lately a very exciting development has taken place within CHILDES, namely the Phon project. Finally phonologists and phoneticians too can fully profit of the CHILDES database. The Dutch CLPF database with data collected by Prof. Paula Fikkert of the Radboud University and myself formed the pilot data for the project, so I have followed the developments closely. It is wonderful that phonetic transcriptions (in IPA) can be viewed and searched. Sound and/or video files that go with the transcriptions can be opened with a single click for immediate inspection. The search engine that is being set up is extremely useful for phonologists. For example, one can search for specific features in specific syllabic positions, in stressed or unstressed syllables.

Phon, and therefore CHILDES plays and will play a major role in three externally funded projects I am currently (co-)supervising. Two of these projects are funded by the Netherlands Organization for Scientific Research (NWO). One five-year project (€600,000), which started last year, studies the development of the speech production mechanism, specifically the phonological encoding part of the mechanism. In another five-year project (€550,000), which has just started, sound inventories in child language, aphasia, and languages of the world are studied. The third project (€200,000, four years) started September 2007 and is funded by the Leiden Institute for Brain and Cognition (LIBC) and compares the ways young infants and young zebrafinches acquire their respective phonological systems.

In all of these projects CHILDES (Phon) is either used to study phonological patterns in available data, or it is used to store data that is currently being collected. Phon is still being developed and more (user-friendly) search devices are constantly being added to the program in close collaboration between you, Prof. Yvan Rose and researchers and users of Phon like me. The collaboration and good communication has been very pleasant and it is great to see Phon developing into a very powerful tool.

On behalf of my co-PIs Prof. Marc van Oostendorp (Meertens Institute, Amsterdam) and Prof. Carel ten Cate (Dept. of Biology, Leiden University) and my PhD students I take the opportunity to thank you for creating the possibility of making child language data available to the research community in an ever-improving way. This certainly helps to bring child language research to a higher scientific level and to keep it there.

Best wishes,

Prof. Clara Levelt
Dept. of Linguistics
Leiden University
Dear Brian,

I am very happy to have this opportunity to describe the role that the CHILDES system and your team in Pittsburgh have played in my work in recent years. My work focuses on the acquisition of Hebrew in normal and pathological populations of children. We analyze spontaneous samples of children and do comparative studies between children with different pathologies. In the last 10 years my work has focused on children with Williams syndrome, Downs syndrome, Fragile X and mental retardation of unknown etiologies. Our Hebrew coding system is fully compatible with the CHAT format and we continuously use CLAN programs for analyses. Furthermore, my research group and I have often made direct contact with you and your team and your on-line help with problems we encountered has been invaluable.

Let me take this opportunity to thank you and your team for years of work and dedication that have benefited us all. In wishing you luck in the renewal of the grant for the CHILDES project I wish luck to all of us, who are using the database ever so often and for sure will continue to use it in the future!

All the best!!

Yonata Levy,
Psychology Department,
The Hebrew University
Jerusalem,
Israel.
Dear Brian,

I am writing to express my sincere gratitude to you for your creation of the CHILDES database. In the past decade I have used the CHILDES database extensively (since my graduate years), and found that it has benefited my research in very significant ways. Looking back, I cannot imagine how much of my research could have been done if I had not had access to this important database. I believe that many researchers in the field have exactly the same feeling as I do.

There are roughly three ways in which the CHILDES database has benefited my research.

First, because CHILDES contains electronic databases not only for English, but also for other languages (especially Chinese and Cantonese in my case), it is extremely helpful to researchers like me who have a strong interest in cross-linguistic studies of language acquisition. In my recent research, I have made use of the Bilingual Chinese-English Corpus contributed by Steven Matthew and Virginia Yip in connectionist modeling of bilingual language processing (see the reference below; Li & Farkas, 2002; Zhao & Li, 2007). My students and I have also used the Tardif corpus and other Mandarin Chinese corpora for corpus-based analyses in my research (see Liu, Zhao, & Li, 2008; Zhao & Li, 2008).

Second, because CHILDES includes not only transcribed speech of children, but also parental speech directed to children, it is extremely important to researchers like me who are interested in identifying the role of parental input in determining the lexical-morphological patterns of children’s speech. For example, we have recently published an extensive analysis of the parental speech in CHILDES in order to identify the role of frequency in the input (see Goodman, Dale, and Li, 2008), and have developed the so-called CHILDES parental corpus. The CHILDES parental corpus has been useful to

Feb. 7, 2008
many researchers in developmental psycholinguistics (see Goodman, et al. 2002, 2003), and hence you agreed to put some of the frequency analysis on the official CHILDES homepage.

Third, because CHILDES contains massive amount of empirical child language data, covering all aspects of children’s lexical, morphological, sentential, and discourse acquisition, it is very important to researchers like me who would like to connect computational modeling results with empirical data from children. CHILDES aids me in at least three ways in this regard. (1) I have evaluated my modeling results against data available in CHILDES. (2) I have derived a comprehensive list of the early nouns and verbs that appear in the CHILDES English corpus, with information on the co-occurring morphological markers and frequency of use. This information serves as a basis for us to construct a realistic learning schedule for simulation. (3) In collaboration with Burgess and Lund, I have analyzed both children’s and adults’ speech in CHILDES according to the HAL methods of Burgess and Lund (1997), in order to derive semantic representations that correspond more closely to the child’s developing lexical representations at different stages of development. Moreover, this co-occurrence based method has been incorporated into our connectionist model, DevLex, which extracts semantic representations from CHILDES parental speech and projects these representations to a self-organizing map (see Li, Farkas, & MacWhinney, 2004; Li, Zhao, & MacWhinney, 2007).

Although all of these aspects are concerned with research only, CHILDES has proven to be an important educational tool. In my cognitive science (PSY333) and psycholinguistics classes (PSY409), I have been teaching my students how to use CHILDES as a tool to study language acquisition. Given the highly electronic nature of CHILDES (e.g., being available on the world-wide-web 24 hours a day), students can access the database, retrieve the CLAN program, analyze individual children’s speech, and write short reports. Several students have also gone beyond this level, and have conducted independent research (PSY361) on the database. For example, they have derived comprehensive lists of all the verbs that appear between the 10th and 48th months in the CHILDES English corpus, describing the emergence of lexical categories (see Liu, 2007).

Over the years your CHILDES has become part of my research, and you have also been my closest collaborator, especially with regard to computational modeling using information from CHILDES. I want to go back to my earlier comment: researchers like me cannot imagine how it would have been if we had not had CHILDES. I want to thank you and congratulate you on making such an important contribution to the field.

Sincerely

Ping Li, PhD
Professor of Psychology and Cognitive Science
Appendix: Published articles or book (since 2003) that have used CHILDES:

Book:


Articles:


Dear Brian,

It would be impossible for us to document all the ways in which we have continued to use the data and programs of the CHILDES system in the last six years but we ourselves, our postdocs, PhD students and Masters students all use CHILDES pretty much continuously. We list below the major features of our usage and where your plans for development would be of assistance to us:

1. Transcription: We are continuing to use CHAT to transcribe our English and German dense databases. These will, of course, be donated to CHILDES in the long-term. Recently we have also collected three English-German bilingual corpora and two more corpora from British-English children between the ages of 3;0 -4;0. From the point of view of continuing with this work, the creation of rapid first-pass transcription programs and more extended support for video transcription would be of enormous benefit.

2. Analysis: In designing almost every one of our experiments, we first use CLAN with data from CHILDES to assess the relative frequencies of particular lexemes and constructions. Secondly, one of the major strategies in our research is to use the Competition model to analyse the validity of cues available to children in the input to make predictions that can be experimentally tested. So far we have done this for English, German, Polish and Cantonese. Clearly the development of automated morphosyntactic analysis to languages other than English is of central importance to this enterprise. In addition a fully automated method of assessing children’s linguistic level that goes beyond MLU, such as IPSYN, is also very important – particularly if this could be done crosslinguistically – a major challenge as we realise.

3. Finally, the development of a more user-friendly interface for CLAN would be great, since it would reduce the amount of time required for teaching and explanation!

Again it is really not possible for us to list all the papers, talks and taught courses in which this work has been presented. We both give large numbers of talks and intensive courses on our research in language development every year. So we have confined ourselves to listing the major publications from 2003 in which our use of the marvellous CHILDES system has been of major importance.

With best wishes
Michael Tomasello and Elena Lieven
Max Planck Institute for Evolutionary Anthropology, Leipzig
Max Planck Child Study Centre, University of Manchester
Relevant publications


Lieven, E. (in press). Language development in a cross-linguistic context. In M. Kail, M. Fayol & M. Hickmann (Eds.) First and second language acquisition


Dear Brian

I have used CHILDES regularly during the last six years, but my use of the programs and corpuses has increased quite considerably in the last few months. My doctoral research comprised of three studies that examined verbal communication in children with impaired hearing, and I used CHAT and CLAN for all of the transcription and most of the analyses. More recently, I have turned my attention to linguistic socialisation in typically developing children, and I am examining the ways in which children and their interlocutors talk about language and language use in everyday conversations. In a series of related studies I am examining language-focused talk in child-child conversations (Gathercole, Garvey and Evans corpuses), comparing mothers’ and fathers’ use of language-focused talk (Warren-Leubecker corpus), and examining the development of language-focused talk and it’s relationship to parental input (Forrester and Providence corpuses; possibly MacWhinney and other corpuses). Furthermore, I am interested in other aspects of pragmatic development, such as the types of feedback that children give during conversations (i.e., backchanelling). I hope to examine the development of feedback strategies in children, focusing on video data from CHILDES (such as the Providence corpus).

Regarding the use of CHILDES in teaching, I started a new post in January 2008, and I am hoping to develop a third year undergraduate module on the study of children’s language acquisition. The module would include mastery of CHAT transcription and use of some of the CLAN programmes in its learning outcomes, and the main method of assessment would be a project that utilises analyses of the CHILDES corpuses.

Good luck with the renewal of the support grant for CHILDES. The existence of the CHILDES project is extremely important for an academic in my current context, which is predominantly a teaching role. I would find it extremely difficult to carry out the research I have described without the availability of the CHILDES corpuses. The developments you are proposing will improve the utility of CHILDES and make it an even more valuable tool to language researchers. The availability of video and audio corpuses also makes CHILDES more useful to researchers interested in pragmatics, which can be difficult to study reliably from transcripts alone. Your efforts are greatly appreciated.

Many thanks,

Julian
References


To Whom It May Concern:

I am an assistant professor in La Salle University’s Speech-Language-Hearing Sciences Program. I currently teach a variety of undergraduate and graduate courses, including Language Development, Diagnostic Procedures and Introduction to Communication Disorders. I have found the CHILDES database to be extremely helpful in guiding students through the process of locating, evaluating and critically reviewing research information as it applies to the field of speech-language pathology. It is imperative that students in our field develop the skills to analyze and use information grounded in research. Questions often arise in relation to the normal developmental process, as well as the differential between “normal” and “abnormal”.

The CHILDES database affords my students and me access to a large variety of research data, including older databases that provided foundational understanding of children’s language development, such as the work of Roger Brown and Jean Berko. It further allows us to compare and contrast the work of such pioneers to that of current researchers. In addition, the database is extremely useful for students to view and review language data from diverse groups.

As useful as the data itself is, the CHILDES also provides referential links to other sites involved with research of language development for a variety of populations, such as the Cochlear Implants Mailing List, the Stanford Child Language Research Forum and Computerized Profiling, to name a few.

While I realize that many use the CHILDES to assist in active research, I believe that the kind of benefits I’ve mentioned above are just as important. As students progress through the undergraduate and graduate programs to become fledgling professionals, they must understand the importance of research to informed practice. The contributions of the CHILDES and its resources are an excellent resource to support this learning.

Sincerely,

Joan A. Luckhurst
Joan A. Luckhurst, Ph.D., CCC-SLP
Assistant Professor
Dear Brian,

I am writing in support of your renewal application for the CHILDES database.

I lecture in language development at University College London and City University, and frequently turn to the CHILDES database for language data that I can use with my students in class. This is an extremely valuable teaching resource, and I am excited that you have plans to expand the amount of data available.

I am also starting to research on sign language, so your plans to support coding of sign language and links with ELAN are particularly welcome.

I wish you the best of luck with your funding application. CHILDES has revolutionised child language research, and I'm sure it will continue to support researchers and teachers for many years to come.

Best wishes,

Chloe

Dr Chloe Marshall, Research Scientist

(please note new temporary address to Spring 2008)

Centre for Developmental Language Disorders & Cognitive Neuroscience
Department of Human Communication Science
University College London
4th Floor
Eastman Dental Institute
123-126 Grays Inn Road
London WC1X 8WD

tel: 020 79051214
e-mail: c.marshall@ucl.ac.uk
web: www.ucl.ac.uk/DLDCN
Dear Brian,

Following your email, I'm writing to document my use of CHILDES over the past 6 years.

I use CHILDES and other available corpora to check the properties of the stimuli in almost every experiment I run with preschool children. This has been recorded in the publications listed below.

I also run tutorials teaching undergraduate linguistics and psychology students how to use the service. I have found it to be an invaluable teaching tool. Indeed we make reference to CHILDES as a 'relevant website' in the following encyclopaedia entry:


The main improvement that I would look forward to is more accurate MORing - we often have to post process the mor line for accurate counts, e.g. of different verb forms. Also further computational functions would be great (we still sometimes write our own scripts for complex frequency counts and extraction of other distributional properties).

I very much hope CHILDES gets further funding for maintenance and expansion. This is probably a good time to say thanks very much for setting it up in the first place!

All the best,
-Danielle
Publications using CHILDES (or, for some 'in press' articles, using corpora that will hopefully soon be donated to CHILDES):


Danielle Matthews  
Max Planck Child Study Centre  
School of Psychological Sciences  
University of Manchester  
Manchester, M13 9PL  
U.K.  
Tel: 00 44 (0) 161 275 2594  
Web: www.psych-sci.manchester.ac.uk/staff/84838
Dear Brian,

I write to tell you how important the CHILDES project has been for my research, and how greatly I value its contributions to the field. Several publications from my lab have made use of detailed analyses of children's speech and speech to children, relying exclusively on the CHILDES corpus. Within the past five years we drew heavily on CHILDES to provide the basis for understanding the distribution of exposure young children have to language about concrete objects and the subordinate and superordinate categories to which they belong. This work featured prominently in analyses and simulations in Chapter 5 of the book cited below.


More generally, there can be no doubt that the openness of this resource has vastly increased the extent to which our understanding of child language is based on facts about what children heard and what children say.

Best wishes,

James L. McClelland
Professor of Psychology
Center for Mind Brain and Computation
Stanford University
83. McKee, Cecile

15-Feb-08

Dear Brian,

In response to your query regarding usage of CHILDES in the last five years, I will emphasize how it enhances my teaching activities.

I use CHILDES in courses at both graduate and undergraduate levels at the University of Arizona. Every semester during the last five years, I have taught a psycholinguistics course at one or both levels. Because my university has an unusual commitment to interdisciplinary links, such courses are listed across departments and thus gather students from several majors. These are usually Psychology, Linguistics, Philosophy, and Speech/Hearing/Language Sciences. In such courses, my webliography always includes a link to CHILDES. I use it in a variety of ways at several points during the semester.

As an example, I used the section on Louis XIII and some of the files under Snapshots to introduce the study of child language in LING/PSYC/PHEL 432 last semester. Our textbook in that course was Berko Gleason and Bernstein Ratner’s Psycholinguistics. Its chapter on child language includes a section on calculating MLU. Supporting materials for that chapter include audio samples from Adam at 1;9 through 4;6, as well as transcriptions of the audio. I integrated a demonstration of some data and tools on CHILDES with an in-class exercise in which students calculated the MLU for one of Adam’s tapes.

Best,
Cecile McKee, Ph.D.
Professor of Linguistics
University of Arizona

U. of Arizona courses
- Spring 2008
  - PSYC/LING 543 – Advanced Language Development, 7 students enrolled
- Fall 2007
  - LING/PSYC/PHEL 432 – Psychology of Language, 57 students enrolled
  - LING/PSYC 533 – Theories of Language Development, 4 students enrolled
- Spring 2007
  - LING/PSYC/SPLH 341 – Language Development, 23 students enrolled
February 18, 2008

Professor Brian MacWhinney  
Department of Psychology  
Carnegie Mellon University  5000 Forbes Ave.  Pittsburgh, PA 15213  

Dear Professor MacWhinney,

I would like to support the new proposal to the NIH in order to continue and improve the CHILDES project. Below, I will describe my use of the CHILDES system and my publications that have relied on the use of the CHILDES programs. One of the courses that I created and have taught for the past six years at San Francisco State University (SFSU) is “Second Language Acquisition,” which is an introductory comprehensive course on second language acquisition and learning for graduate students. In this class, I introduce the CLAN programs because I believe that learning the CLAN programs, which have played a significant role in spurring the pace of research in the field of language acquisition and language development, would provide a tremendous opportunity to students who are studying at SFSU. Although the programs primarily deal with children’s data, they are also applicable to adults’ and second-language data. Specifically, in the course I have explored areas of research impinging on the problems of language acquisition and learning (particularly second language acquisition and learning). To accomplish this end, in addition to the regular classroom, I have students work regularly in the computer laboratory.

Prior coming to San Francisco State University, I participated in a great number of CHILDES related projects at Harvard University. I also worked with Japanese researchers on developing CHILDES formatting applicable to Japanese data (JCHAT). Furthermore, in the past at Harvard, I also taught two courses: (1) Child Language and Education and (2) Discourse Analysis. Using these experiences, here at SFSU, I developed materials that could introduce the CLAN programs to students efficiently and effectively, and satisfy students’ specific needs (i.e., analyzing Japanese and English data). I created two manuals explaining how to use this language-analysis program: one in a conventional handout format, and the other in a demonstration format using PowerPoint. Up until today, I have continued to improve these two manuals.

I have published a number of articles, testifying to the breadth of my professional interests and the quality of my research. Many of them have relied on the use of the
CHILDES system (mainly transcribing the data that my research team collected). At SFSU, I have sought funding for my research on bilingual children’s literacy skills development. I specifically choose the storytelling task as a method of inquiry because research must meet the demands of evidence based and scientifically based inquiry. The participants in the study are Japanese/English bilingual, elementary school children in the multilingual San Francisco area. Bilingual children ages 6 to 12 narrated a picture story book in both English (L2) and their native/home language (L1). For the past six years, my research team has evaluated the stories produced by the bilingual children in terms of grammar, tense marking, flow, and other factors. The material used is the so-called frog story consisting of series of pictures representing a number of dynamic interactions among a boy, a dog, and a frog over time and in different physical settings. We chose this method because the use of a single story source keeps the content constant throughout and allows comparison among children not only across age groups but also across the two languages.

My representative (edited) books and articles since 2002 that used the CHILDES system will appear below, on pages 3 and 4. I hope that the information I have included in this letter will be useful for the NIH CHILDES grant, which is for renewal. Sincerely,

Masahiko Minami, Ph.D.
Professor and Japanese Program Coordinator
San Francisco State University

PUBLICATIONS


Dear Brian,

You were asking for a description of our use of CHILDES data or programs over the last 5 years. I will answer in describing the activities we conducted in the frame of our JCHAT project (a project promoting the collection of Japanese data for CHILDES) and the Index Development Project (a government-funded project developing an index for grammar development of Japanese).

As you know, the JCHAT Project, which started out in 1993, concentrates on
1) the development of a transcription format for Japanese (a task which is completed with the appearance of Miyata, S. 2002. *Wakachi 2002 v.2.0 JCHAT nyuryoku foomatتو oyobi JMOR keitaiso bunsekiyo wakachigaki gaidorain* [Word separation scheme for Japanese developed for the use in JCHAT and JMOR]. Nagoya: Aichi Shukutoku U.)
2) the development of grammar and lexicon files for a Japanese version of the CLAN morphemicization program MOR (completed in 2002, available at CHILDES)
3) conducting annual workshops on data entry and analysis of Japanese data.
5) The compilation of a bibliography of research on language acquisition in and/or on Japanese (the current version of JBIB by Seiko Ono is available on the CHILDES server).
6) The extension of the Japanese database. Besides data currently collected and transcribed by individual researchers, we are also working on the conversion of older corpora published in print. Specifically we are working on a) the Miipro corpus (4 children, observational study, from 1;2 – 5;0), b) Yokoyama corpus (1 child, diary study), and c) Okayama corpus (observation study, cross-sectional, 130 children).
7) Furthermore we have a mirror site of CHILDES at Chukyo University, Nagoya, managed by Hidetosi Sirai.
The work in relation with CHILDES data-collecting and transcription format development has inspired a considerable amount of research. In the first years we performed lively research sessions at the JCHAT workshops, which finally lead in 1999 to the formation of the Japanese Society for Language Sciences (JSLS), a bilingual international research society concerned with language acquisition.

Also the “Index Project” (“A Crosslinguistic Study for the Universal Developmental Index” 1999-2000; Head Investigator H. Sirai, and “Nihongo Kakutoku oyobi Dainigengo shutoku ni okeru gengo hattatsu shihyo no kaihatsu to nicheitaisho gengohattatsu kenkyu [The development of a developmental index for first and second language acquisition in Japanese], 2002-2004”, Head investigator K. Otomo, "Development of a Developmental Index for Japanese and its Application to Speech Developmental Disorders". 2006-2008. Head Investigator S. Miyata) was inspired by the work in the frame of the JCHAT/CHILDES Project. In this project we are developing an index for grammar development for Japanese, inspired by the Developmental Sentence Score (DSS) by Lee, L. (1974). This research is conducted on the basis of data collected for CHILDES, and the analysis heavily relies on the CLAN programs. The final index will be adapted to a program based on the CLAN program DSS. A by-product will be an overview over general tendencies in Japanese language acquisition until 5, a badly needed compilation (since Clancy’s epoch-making summary on Japanese child language research in Slobin (1985), no synopsis of acquisition studies has appeared).

All the research mentioned above is inspired by the CHILDES database, and relies on tools of the CHILDES database. The introduction of CHILDES to Japan has had a strong impact on child language research, which is reflected best in the formation of the Japanese Society for Language Sciences (JSLS). Without the support of CHILDES and your (often very concrete) personal help and assistance, the situation of language acquisition research in Japanese would be very different. In this sense, I strongly recommend a renewal of the grant from the National Institutes of Health (NIH).

With kind regards,

Susanne Miyata
Communication Disorders and Science Course
Department of Health Science, Faculty of Medical Welfare,
Aichi Shukutoku University
23 Sakuragaoka Chikusa-ku
Nagoya, 464-8671 Japan
smiyata@asu.aasa.ac.jp
Dear Brian,

I would like to lend my support to you in renewing the NIH funding for the CHILDES project. I was first introduced to the concept of CHILDES database as an undergraduate in linguistics at University College London, and the immense scope of the project struck me. I still remember using CHILDES for the first time as being pretty amazing and awesome - so much data! and free! I don't think the whole field of language acquisition would be anywhere near where it is now, if we did not have CHILDES. It is one of the best tools.

I can only boast one published work (with Ken Wexler) that uses CHILDES, but countless searches were performed by me and my undergraduate research assistants. We've been primarily focusing on the role of adult input to children, especially of articles/determiners ('the', 'that', etc). Much of this is work in progress.

I am very much looking forward to an easier search interface!!

Sincerely,

Nadya


--

Nadezhda (Nadya) Modyanova
Ph.D. Candidate in Neuroscience
the Wexler ab/Normal Language Lab
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge MA 02139
Office: 617 253 7626
Málaga, Spain, February 23rd, 2008

Dear Brian,

in support of your application for the continuation of your research grant from the NIH I would like to give an outline of how we have been using CHILDES for our research at the University of Málaga (Spain)

In the last three years we have collected a data base of Spanish child language. This has done in a longitudinal study with weekly samples of a deaf child with cochlear implant. This study will serve as the basis for a large group study that we have started just 3 weeks ago (Feb 2008). Our research has been founded by our regional Government (Proyecto de Excelencia de La Junta de Andalucía. P7-SEJ-3119). All the samples of our case study (156) were transcribed with CHILDES. They are in the CHILDES data base as the MOC Corpora. A large number of publications, invited talks and conference publications are based on this data set (see attached list), and work with the data is ongoing.

Yours sincerely,

Ignacio Moreno-Torres

In my view, it is essential that research on the CHILDES programmes continues which enables the programmes to be continually improved and become more useful to researchers studying child language.

Universidad de Málaga (Spain)
Publications based on data transcribed and analyzed using CHILDES – in English


Publications based on data transcribed and analyzed using CHILDES – in Spanish


Invited Talks


Moreno-Torres, I. Language acquisition with cochlear implant. Seminaire de Psicologüística de l'Université Libre de Bruxelles. Université Libre De Bruxelles (Belgium). March 22, 2007

Presentations at conferences (English)


Presentations at conferences (Spanish)

Dear Professor MacWhinney,

I would like to support the new proposal you are making to the NIH in order to continue and improve the CHILDES project. Indeed, my research team (Leonard project http://anr-leonard.ens-lsh.fr/) and myself in my personal work have been making extensive use of the CHILDES data and programs. The existence of the CHILDES project has motivated us to contribute data and meta-data to the database (recordings and transcriptions in progress) and serves as a wonderful example of data-sharing in the entire French community working on child language.

Without CHILDES and the generosity, the talent, the patience you have had for years in building the project, in uniting researchers, I don’t think there would be such an international community in child language able to share data, analyses, ideas, and their passion for language acquisition. Although diversity is necessary in Science, we also need a certain degree of standardization in order for the community to attain a certain level of inter-comprehension and communication. I think you and the CHILDES project have served as a link between various communities in the field of language acquisition and without your mediation, the research would not have become as rich and in keeping with the progress of technology.

I must admit that I have been aware of the existence of CHILDES since the beginning of my PhD in the 1990s and only used during that period to make hypotheses on English data. There were two drawbacks at the time: I needed the transcriptions to include phonetics, I needed to work extensively on my video-recordings. But CHILDES has made astounding progress since those two weaknesses have disappeared a few years ago and we now have the possibility to link the transcriptions and the video, and we can transcribe in IPA in the PHO line.

Therefore, CHILDES has served as a model to launch our own project in France (Projet Léonard) and we have decided to adopt the CHAT format and
to use the CLAN tools for all the data we use. Indeed the CHAT format is flexible enough for us to add our own codings (for example, a specific coding for fillers, for points, for prosody…) which we will be able to share with the community once they are stabilized and tested).

Not only does CHILDES provide data and tools for child language analyses, but the CHILDES website helps us with specific issues and problems we encounter in our work:
- video recording, digitalizing and compression
- bibliography
- pedagogical material
...

The CHILDES project is therefore essential for the community working on child language and needs to be continued and expanded for the development of research in the language acquisition field on the international level.

Specific requests
We would eagerly welcome the improvements that could be made thanks to a new NIH grant.

1) We do think it would be important if the database could be more user-friendly, if the contents of the data could be described better. It would be a big help if it were possible to make requests and easily (and quickly) find specific situations. Here are a few examples:
- father-child interaction
- parent-child in a bath/dinner/book reading situation
- all data involving two year olds speaking English, French or Spanish making a (dialogic) narrative of their day away from home

2) It would also be more pleasant if the data was presented in a type of interactive table of contents linked to the data itself with language, age, number of recordings, existence of the audio or video.

3) For more precise work on child phonology, we have used PHON (thanks to our collaboration with Yvan Rose) rather than CLAN and we are very impatient for the two programs to be easily compatible (CHAT transcriptions transformed into PHON and PHON transcriptions to CHAT in order to use the CLAN programs).

4) We could welcome further support for coding for more pragmatic and discursive analyses as well as further support for coding of child language
5) We do have a problem to find and train students to transcribe data and link it to the video recordings and would need help to organize training sessions in order to motivate them.

6) For our type of analyses, we often invert the main tier (in orthography) and the PHO line. In further developments of CLAN, we wonder if it would be possible NOT to have a main tier…and NOT to have a hierarchy between orthography (which is often an interpretation) and the phonetic transcription.

7) We would appreciate having the possibility to use colors in the editing of the data.

8) We would appreciate having a possibility to SEE corrections. Indeed we require the data to be transcribed by two coders, one after the other and it would be a great help to be able to confront alternative transcriptions (as in PHON). It would also be nice if different layers of coding, comments etc… could be visible in the transcripts… Each time a different researcher uses the data, he/she might want to add layers of coding that could be helpful to others.

**USE of CHILDES for our scientific production**

Almost all the papers we have presented or written since the beginning of the project in 2005 make use of the data or tools developed in the framework of the CHILDES project. The oral presentations are in the course of being submitted as papers (International journals).

**Papers**


**Presentations**


6- Morgenstern, A., Mathiot, E., Leroy, M. Pointing gestures and demonstrative words : Deixis between the ages of one and three. First SALC Conference. LUND, Suède. 29 voembre au 1er décembre 2007.


The symposium I have submitted has been accepted for the IASCL congress in Edinburgh end of July and all the authors will make use of the CHAT format and the CLAN programs for their analyses.

Personal work

I have myself used the CHILDES database to make hypotheses on the acquisition of Selfwords in English for my book on Self words and have then worked on French and French Sign language (Morgenstern (2006) Un JE en construction: genèse de l’auto-désignation chez le jeune enfant, Paris, Ophrys).

Along my older papers, two make extensive use of the Bloom data (Peter):

Thank you for having launched and developed the CHILDES project, for your constant presence on the international scene, for your generosity and PLEASE continue!
Hi Brian,

I am an early career researcher/lecturer at the University of Sydney and have just submitted my PhD thesis for examination. I am responding to your email regarding CHILDES. I have used basic commands (MLU, freq etc) for child language sampling analysis and have used CLAN to code parent-child interactive book reading sessions. The results of which include conference presentations and a manuscript on the way (see below). I do hope that this program and the proposed improvements come to fruition. I am yet to use CHILDES as part of my teaching but I would like to work towards that. To this end, improvements such as web-based tutorials with transcripts coded/analysed from existing data would be most helpful.

All the best with your proposal.

Natalie

Natalie Munro
Lecturer, Speech Pathology
Faculty of Health Sciences
The University of Sydney
Cumberland Campus C42PO Box 170 Lidcombe NSW 1825 AUSTRALIA
Phone (Mon-Wed): 61 2 93519880
Fax: 61 2 9351 9173
Email: n.munro@usyd.edu.au
Webpage: http://www2.fhs.usyd.edu.au/csd/natalie_munro.htm
Papers/presentations where CHILDES programs have been used.
Dear Brian,

The purpose of this letter is to enthusiastically support your NIH proposal to continue and improve the CHILDES system for the study of young children’s early language development.

Since 2003, I have used the CHILDES databases for two research projects: One compared English- and Spanish-learning children’s early use of different types of motion verbs (Hohenstein, Naigles & Eisenberg, 2004) and the other investigated the maternal input of children learning Mandarin (Lee & Naigles, 2005). For both of these projects, the existence of corpora in languages other than English was absolutely essential, as we cannot come to understand how children acquire language without detailed examination of the way(s) children and adults use all the different languages of the world. The existence of these databases within the CHILDES system made the projects tractable and able to be efficiently conducted. I am currently involved in a third project for which I will be, for the first time, intensively using the CLAN tools. We are collecting mother-child interactions in a longitudinal study of the acquisition of English by children with autism. My students and I have been training ourselves to use CHAT and CLAN, and I am happy to say that Brian has been immensely helpful whenever we have had questions. We especially look forward to using the IPSyn calculations, the Conversation Analysis coding, and the rapid first-pass transcription methods that are forthcoming in this new proposal. In sum, the CHILDES system has been very useful already to my research, with expectations that it will only become more essential. Thanks for the good work you have done and are planning to do!

Sincerely,

Letitia R. Naigles
Professor, Psychology
University of Connecticut

CHILDES-related publications since 2003:
91. Nakayama, Mineharu

February 13, 2008

National Institute of Health

RE: Supporting the CHILDES project

Dear members of the NIH Grant Committee,

I am writing this letter to strongly support Brian MacWhinney's grant proposal to enhance the CHILDES system.

The Ohio State University’s Japanese language and linguistics program is one of the largest in the Continental USA and the program is very strong in both undergraduate and graduate education. It offers different courses pertaining to Japanese linguistics. In particular, we have had several graduate students specializing Japanese language acquisition, and these students used the CHILDES system. When I offer a Japanese language acquisition course, I discuss the system and some students use it for their projects. The system is vital to enhance our education in the US. It is important that it is available and maintained so that everyone can share the data through the system. I am confident that the system will be used by our students very frequently and serve as a vital research reference material.

Given where our program strength lies, and the necessity of the CHILDES system in the US, I strongly endorse Brian MacWhinney’s proposal to maintain and enhance the CHILDES system. If you have any questions, please do not hesitate to contact me at nakayama.1@osu.edu. Thank you very much for your consideration.

Sincerely yours,

Mineharu Nakayama
Associate Professor of Japanese Linguistics, Department of East Asian Languages and Literatures
Adjunct Associate Professor, Department of Linguistics
The Ohio State University
Dear Brian,

I am writing to let you know about my recent research activities involving the use of the CHILDES database. My collaborator, Marianne Gullberg (Max Planck Institute for Psycholinguistics), and I have been investigating the development of verb meaning in children acquiring Tamil and Dutch, targeting the use of "caused posture" expressions ("set", "stand", "lay") in children acquiring the two languages. Our study employed an elicited production task to collect data in experimental contexts from children and adults. But as we are also interested in examining the influence of patterns of use in the input to the children, we explored longitudinal, spontaneous production data using corpus data available on CHILDES. Specifically, data were analyzed from the Tamil "Vanita" corpus (R.Narasimhan, 1981) and the Dutch "Groningen" corpus (Wijnen & Bol, 1993). The results of our research have been reported in a paper submitted for publication and currently under revision:

Narasimhan, B. and Gullberg, M. (under revision). "Caused posture expressions in Dutch and Tamil child language: The role of input frequency and verb meaning."

In addition, the data on CHILDES are an invaluable resource in teaching activities. In my classes, I regularly draw on examples of different linguistic phenomena that are culled from transcripts of spontaneous conversation (from various corpora e.g. the Brown corpus, the Gleason corpus, etc.). Further, many of the student term projects in my language acquisition class will be based on the CHILDES data using the CLAN tools that are freely available on the web. These projects will involve topics as different as relative clause acquisition, the use of metaphor, verb-particle constructions, and argument structure, among others.

The availability of computerized data and tools for analysis on CHILDES is an invaluable resource to the child language research community. If you need any additional information, please do not hesitate to let me know.

Sincerely,

Bhuvana Narasimhan
Assistant Professor
Linguistics Department
University of Colorado, Boulder
93. Nicholas, Johanna

February 11, 2008

Brian MacWhinney
Department of Psychology
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213

Dear Prof. MacWhinney:

This letter is written in support of the continued funding for the CHILDES project that you direct. The resources that have become available to us as a result of this project have been tremendously important for the work of my lab, allowing us to be much more productive than would otherwise be possible. For about 15 years we have used CLAN programs and tried to keep up with the interesting new analytic tools that you and your staff have created for investigations in all aspects of language study.

In my laboratory we study the language development of very young deaf children who are learning spoken language. Our work focuses on natural language sample analysis of preschoolers who received a cochlear implant at various ages in infancy.

In addition to providing a vehicle for language analysis, the CHILDES database has provided a means for us to share data from our studies with other researchers who may have an interest in studying the emerging language of this population. Since deafness is a low-incidence disability, data collection from a homogeneous sample is a very difficult task. Therefore, a repository of language samples that are archived in a conventional format could potentially save others much time, effort, and expense when seeking to answer questions that require language samples from these children.

Attached is a list of the publications and presentations from our laboratory that have used CHILDES resources as an integral part of the methodology. I dare say that much of this work would not have been possible with the tools your project has provided. We are very grateful and give our unqualified support for the continuation of this project. We are especially looking forward to the development of fully automated IPSYN analysis and computer assisted methods for first-pass transcription.

Sincerely,

Johanna G. Nicholas, Ph.D.
Research Associate Professor
Department of Otolaryngology
Washington University School of Medicine
Publications that include CHILDES tools in the methods:


**Presentations**


Geers, A., Nicholas, J., Tye-Murray, N., and Uchanski, R. “Rehabilitation factors contributing to auditory speech, and language skills in children implanted before age 5” Presentation to meeting of the British Association for Pediatric Otorhinolaryngology and the British Association of Teachers of the Deaf, Nottingham, United Kingdom. April, 2000.


Dear Brian,

This letter is in support of your application to the National Institutes of Health (NIH) for continuing support of the Child Language Data Exchange System (CHILDES). This is an appropriate moment for me to explain the centrality of CHILDES archive for my research projects, and to thank you for making this wonderful resource available for all of us.

In the last year, I have taken on an ambitious project all of which is built on the availability of transcribed observations on the CHILDES archive. I am in the process of hand-coding large speech samples of all the English-speaking parents currently on CHILDES -- of which there are over 500 -- for the syntax of the verbs they use in their utterances. The method of analysis I am employing is a combination of Chomsky's Minimalist program and Dependency Grammar, and I am looking for each verb's subject and objects, and maybe for some other syntactic contexts. Ultimately, I hope to be able to describe the syntactic characteristics of English-language parental speech addressed to children under 3 years of age, focusing on the special and sometimes anomalous features of verbs used by them. The result of this study will be, I hope, the thorough characterization of the linguistic input to the learning process, based on a truly large sample of parents-speakers. In such a large sample, it is possible to see phenomena that escape the small samples single researchers are able to master on their own. The results promise to be surprising and mostly counter-intuitive, based on the work already done.

Concurrently, I will also analyze English-speaking children's use of verbs in a parallel large sample, to see what is the output of the acquisition process given its complex input. The results will be written up in a book-length monograph currently considered by one of the major academic publishers.

It is obvious that without the availability of the CHILDES archive, this research project would be impossible to carry out. Without CHILDES, it is simply impossible to arrive at the kind of data that the large pooled sample based on the CHILDES archive reveals. I urge the NIH in the strongest possible language to continue the support of the CHILDES project. This international data-exchange system is invaluable for the field and we all hope that it will continue to exist and thrive for all of us.

The best,

Anat Ninio, Joseph and Belle Braun Professor of Psychology
Department of Psychology, The Hebrew University
Jerusalem, Israel 91905, Email: Anat@Ninio.org
95. Noccetti, Sabrina

From: Sabrina Noccetti <noccetti@hotmail.com>
Date: February 14, 2008 2:09:35 PM GMT
To: <macw@cmu.edu>
Subject: CHILDES renewal request

Dear Professor MacWhinney,

I am very pleased to hear that you are planning to expand the database of the CHILDES project, to create a new, user-friendly interface for CLAN. I believe that the CHILDES project is an essential tool for those who work on child language and other psycholinguistic research. I am one of the participants to the international Crosslinguistic Project on Pre- and Protomorphology in Language Acquisition that investigates 15 languages, coordinated by Prof. Wolfgang Dressler (Head of the Dept. of Linguistics, University of Vienna) and I used CHILDES to carry my research on the Italian Language (L1). The CHILDES (CLAN) programs have allowed me to publish the following book and articles:


Noccetti, Sabrina 2003 Acquisition of Verb Morphology in Italian: a Case Study. In Bittner Dagmar, Kilani-Schoch Marianne, Dressler Wolfgang (Eds.), Development of Verb Inflection in First Language Acquisition. A Cross-linguistic Perspective


My best regards,
Sabrina Noccetti
Dipartimento di Anglistica
Università di Pisa
Dear Professor Brain MacWhinney,

Thank you for your great works to the Programs of the CHILD Language Data Exchange System. The programs were very useful and we can progress the research about lexical development in the Japanese children, using the Clan program. I could publish the following papers in the Journal.


I am eager to hope that National Institute of Health continues the funding for CHILDES project. Thanks to NIH funding, the researchers all over the world will receive benefits from the progress of CHILDES project.

Thank you for your dedication to the progress of CHILDES.

Tamiko Ogura

Professor of Psychology, Ph.D.
Brian,

Off the top of my head, students I advise doing work using CHILDES:

Shaher Banu Vagh, doc student, has an article under review on nouns in maternal input, using our Early Head Start data (same data Meredith used for her dissertation articles).

Maria Fusaro, doc student, is looking at mother and child use of head nods/shakes in our NE data. Just starting dissertation; no article yet.

Julia Hayden, doc student, is looking at instances where mothers challenge accuracy of elements of children's personal narratives in the Early Head Start data.

We're also using CHILDES to look at child lexical production in videotaped bilingual mother-child interaction for my Head Start Measurement study. We are using word types to validate parent and teacher CDI reports for 2-3 year olds in Early Head Start/Head Start.

Finally, about half the students in my course From Language to Literacy do research projects using data from the CHILDES database, or use CLAN to analyze other data every year, as part of course requirements.

Hope this helps.

Barbara
February 18, 2008

To the Review Committee,

I took my first training workshop in the CHILDES system at Harvard in 1989, nearly 20 years ago. In the interval since then, CHILDES has become deeply embedded in the practice of our field. If I’m going to look up a word in the dictionary, I just have to say I’m going to “look it up.” You know I mean in the dictionary. Likewise, in our weekly research meeting today, in the space of an hour and a half, I counted eight times when someone said something like, “Let’s check when children begin to use that expression.” “Do they say ‘each other’ before they say ‘one another’?” “Is ‘ein ander,’ (the phrase used in a cross-linguistic experiment), common in German child-directed speech? Our colleague Tanja doesn’t think so, but we can check the German transcripts.” We never mentioned CHILDES, but the database has reached the status that everyone knew that’s what was meant. “We can look it up in CHILDES.” What’s more, we can do it on our own computers, easily linked to the web interface maintained for us at Carnegie Mellon.

These simple functions, that we teach students in the course of a couple of sessions in the lab, are perhaps the most pervasive contribution of the CHILDES system, but they are far from its only contribution to our field.

Brian MacWhinney and the CHILDES lab are continually moving us in the direction of more and more sophisticated functions and greater functionality. There is automatic coding that we would never have dreamed of 20 years ago; video links, and tools to aid first pass transcription. I have to admit that I’m still struggling with getting my transcriptions to line up with the video links, but when I get closer to my deadline on the project, I know that I can count on personalized help by email and/ or phone. In the past when my commands haven’t seemed to work right for me, I have been able to send my files to the CHILDES lab and it is 2 for 2 in finding my error. There is also a separate ListServ connected with CHILDES called ChiBolts to let us bring our questions to other colleagues. MacWhinney monitors it so if no one else offers a solution, he will work on it. A check in the archive—also maintained by CHILDES—will verify his consistent attention to issues large and small.
The main Listserv, InfoChildes, and the very informative website at http://childes.psy.cmu.edu/ are further perks of the system. They may not be cutting edge themselves, but they help us know where the field is, and they let those of us who are not early adopters catch up when we need to, without having to be experts and follow every new technology ourselves. I think it’s also a great boon for less developed countries, such as in Eastern Europe, that the system is so comprehensive, and it is free. I believe our field is ahead of other NSF fields in terms of international cooperation, and much of the credit can go to the availability and versatility of CHILDES tools and the relatively close-knit community the system engenders. Two years ago, when I was looking for international collaborators, my link with CHILDES put me in contact with researchers and educators in rural India, Sardinia, Malawi, and Croatia, to name the ones that are most inaccessible through other channels.

So many computerized systems that started when CHILDES did in the days of DOS never made the transition to Windows. Very few can be used with OSX. Those other systems didn’t have the resources and the personnel to keep up with the pace of developments in computing, much less linguistics, and so now their usefulness is limited. Not so CHILDES. It went to Windows; it can be used by MACs; and it has extended well beyond child-language. I cannot say enough in support of the CHILDES system and how well it has kept pace with and led the technology transformations in that time.

I put my data on the CHILDES system more than ten years ago. It was a project that I and five research assistants spent four years collecting, coding, and error-checking. The analysis itself took about two weeks, and then it was published in a few papers, most notably in the 2002 monograph, Language and Literacy in Bilingual Children.* Since then, I have heard from many people that they have used those data, mostly for student summer projects or for several types of class projects in education courses. Its most extensive use—that I know of—was as the pilot data to develop new analysis techniques in Spanish. It is so gratifying that all that effort has borne more fruit than just my own publications, and it continues to be useful. Even when I want to consult them, I download my own stories from CHILDES, because my archive CDs are not always accessible to me in my different offices.

In addition to giving me access to other people’s data, I am very thankful to CHILDES for giving my data new life. I cannot imagine being in this field without it.

Sincerely,

Barbara Zurer Pearson, Ph.D.
Research Associate

Hi Brian,

I am writing because I have used the CHILDES database in several research projects over the past three years. Both primarily used the Adam corpus (Brown). In one, I investigated a Bayesian model for the learnability of hierarchical phrase structure and used the child-directed speech in CHILDES as the corpus of data given to the model. This has yielded work leading to several journal articles and talks (see below). In the most current work I use the Adam corpus to extract the statistics of construction usage for many different English verbs in child-directed speech. This work is still in preparation, and is also a part of my PhD thesis, which will be completed this year.

sincerely
Amy Perfors

PAPERS


TALKS

5th Biennial Meeting of the Child Development Society; symposium on computational approaches to language acquisition, Santa Fe, NM. October 2007

40th Annual Meeting of the Society for Mathematical Psychology; symposium on computational models of language, Irvine, CA. July 2007


Recursion in Human Languages Conference, Normal, IL. April 2007

32nd Conference of the Society for Philosophy and Psychology, St. Louis, MO. June 2006
100. Peterson, Shira

Brian MacWhinney
Professor of Psychology
Carnegie Mellon University

February 20, 2008

Dear Dr. MacWhinney,

In 2004, while I was a doctoral student in the Warner School of Education at the University of Rochester, I had the opportunity to attend a training on the CHILDES system at Carnegie Mellon University. For my dissertation, I used the CHILDES system of transcription and coding (CHAT and CLAN program) for my analysis of large group (Circle Time) discourse in preschool classrooms. The following is a list of awards, publications, and presentations resulting from this work:

**Awards:**
- 2003 Spencer Dissertation Fellowship for Research Related to Education
- 2002 AERA/Spencer Pre-Dissertation Fellowship, Mentor: Dr. Elinor Ochs, UCLA
- 2002 Doctoral Award for Research and Travel, University of Rochester

**Publications:**

**Presentations:**


I would like to offer my wholehearted support for the continuation of the CHILDES project. The CHILDES software has been, and continues to be, an extremely valuable tool in my research on the role of adult-child discourse in children’s language development. This project is an important resource for the community of scholars working in the field of child language.

Sincerely,

Shira May Peterson, Ph.D.
Research Associate
Children's Institute
274 N. Goodman Street, Suite D103
Rochester, NY 14607
(585) 295-1000, ext. 233 (phone)
(585) 295-1090 (fax)
speterson@childrensinstitute.net
www.childrensinstitute.net
I am using the CLANWIN program since 2003 to produce transcripts according to the CHAT format, to produce lexicons of Slovenian and German lexical items and to calculate statistical relationships between language categories. Since 2006, I use CLANWIN in combination with ELAN. I have been using CLANWIN for first and second language acquisition research. At my home university in Maribor (Slovenia) I have been using it in the years 2003/04 to 2005/06 in Language acquisition classes with undergraduate students of German as a foreign language. I have published data calculated with CLANWIN in several articles and contributions in conference proceedings in the international scientific press, presentations and posters in Slovenian and international conferences.

**Articles and Contributions in Conference Proceedings**

*on 1st or 2nd Language acquisition*

*where the Clanwin program was used for statistical purposes*

*on transcripts written in the CHAT format*


**Conference Presentations and Posters on 1st Language acquisition where the Clanwin program was used for statistical purposes on transcripts written in the CHAT format**


PETRIČ, Teodor. Acquisition of Slovene verb paradigms compared to German. V: International Morphology Meeting, Abstracts. [Vienna: University, 2004, s. a.], str. 116-118. [COBISS.SI-ID 15078920]

PETRIČ, Teodor. Acquisition of Slovenian morphological paradigms compared to corresponding developments in German verb morphology. V: Poznan Linguistic Meeting, May 18 - 20, 2004 [Poznan: University, 2004].


Dear Brian,

Actually, just this very morning I was teaching a student about CHILDES, CHAT, CLAN and I was planning to write to you to let you know what we are doing here in a couple of days when I get to meet my colleague in France. I am now currently heading a project on "later language development" of language minority adolescents in Berlin, part of a collaborative project with Dr. Mehmet-Ali Akinci at the University of Rouen in France. So far we have completed a pilot study of 12 graders at a gymnasium in Berlin. (This is the highest type of secondary school, which ends in 13th grade with the "Abitur", needed for entrance into university), planning to get data from kids in "lower" types of schools, Hauptschule, Realschule, Gesamtschule -- and, if possible in some bilingual programs; and lower ages too)

In June 2007, we collected monologic written and oral texts in three languages using the video devised by Ruth Berman and used in the cross-linguistics L1 project reported in Berman/Verhoeven 2005 and subsequent publications. We're doing the same, except that we have not only first language Turkish (+2Kurdish +1Arabic) (possibly undergoing attrition for some) but also second language German and first foreign language English. These are to be compared with the results from Akinci’s study of Turkish/French bilingual children and adolescents with his monolingual Turkish control data. We got language background, family and afterschool language practices data too, from an extensive questionnaire.

In addition to the monologic texts, we just last Tuesday Feb 12, 2008, recorded these these kids in small group discussions of various aspects of controversial or problematic issues in migration / integration / assimilation in their 3 languages (separate topics for each language). Several of the class members had been among those 300 German and Turkish kids invited to participate on the panel or in the audience of a "forum" event in which prime minister Erdogan and chancellor Merkel addressed these issues. It turned out to be pretty controversial, I guess it may even have made the news in the USA.

Best wishes,
Carol
Dear Brian,

In support of your application for the continuation of your research grant from the NIH, I would like to describe how my colleagues and I have used CHILDES over the past 6 years.

Our research involves the use of a variety of methods, including naturalistic data analysis, elicitation experiments, grammaticality judgment experiments and computational modelling, to study first language acquisition in a number of different languages.

Our naturalistic work has involved the collection of two very rich corpora of early child Spanish, and the analysis of previously collected corpora of early child English from typically developing children and children with SLI. CHILDES was critical in providing a transcription system and a suite of analysis programs that allowed us to code and analyse large enough samples of child speech to properly test current models of early grammatical development.

In our experimental work (which has focussed mainly on English), CHILDES has been important in allowing us to construct experimental materials (e.g. sentences in grammaticality judgment experiments) that reflect the statistical properties of child-directed speech as opposed to speech between adults.

In our computational work, which involves using a single computational model (MOSAIC) to simulate developmental data across several different languages, CHILDES has been critical in providing both a cross-linguistic database and a set of analytic tools that has allowed us to manipulate and analyse the cross-linguistic data in a systematic way.

Much of this research would have been impossible and all of it would have been much more difficult to carry out without the database, transcription methods and analysis programmes provided within CHILDES.

A list of publications, conference presentations, research grants, and doctoral theses is attached, all of which have used the data and/or analysis tools available in CHILDES in some way.
In my view, it is essential that work on CHILDES continues, including further expansion of the database to include an even wider range of languages, and further development of the supporting programmes to facilitate the transcription and analysis of corpora that are large enough to allow proper quantitative analyses to be conducted.

Yours sincerely,

Julian

Refereed Journal Papers


Conference Presentations:


Freudenthal, D., Pine, J. M. & Gobet, F. (2002). Modelling the Optional Infinitive phenomenon in English and Dutch. 9th Congress of the International Association for the Study of Child Language, Madison, Wisconsin, USA.


Funded Research Proposals

Leverhulme Trust (2008-2010). Examining the cognitive mechanisms that underpin Specific Language Impairment.


Economic and Social Research Council (2006-2008). Modelling the cross-linguistic pattern of finiteness marking in declaratives and questions.


Economic and Social Research Council (2003-2006). Title: Modelling the development of finiteness marking.

Economic and Social Research Council (2002-2003). Title: Modelling the dynamics of children’s syntax acquisition.

Economic and Social Research Council (2001-2003). Title: Contrasting syntactic and limited scope accounts of early grammar in children with SLI.

February 17, 2008

Dear Prof. MacWhinney,

I am writing in support of the CHILDES renewal application. CHILDES has been invaluable not only for my own research in the past six years but also for teaching new generations of students how to conduct research requiring transcription and analyses of discourse.

With regard to my own research, the following articles have been published where I cite the usage of the CHILDES programs for the transcription and analysis of data from multilingual children:


Conference papers have also been presented citing the usage of CHILDES:


I have also introduced the use of CHILDES in the supervision of undergraduate and graduate theses at my university that are now in progress. In Autumn 2008, I will teach a new course called Child Language Development where I will introduce a unit to upper-level undergraduate students on using CHILDES to study child language development. I have also introduced MA students in our Graduate School of Education’s English teaching program to use CHILDES in their research on second language acquisition in the Japanese classrooms as proposed in:


As outlined above, the CHILDES Project is important as a research and teaching tool for the future of scientific study in the field of child language development and multilingual language acquisition. I urge the National Institutes of Health (NIH) to continue supporting the further development of CHILDES as a bridge and link between researchers around the world.

Sincerely,

Suzanne Quay, Ph.D.
Associate Professor of Linguistics
International Christian University, Tokyo
Dear Brian,

I know that the NICHD grant for CHILDES is up for renewal this year, and I would just like to say what an invaluable research tool it has proved to be for me, my colleagues and my department’s graduate and undergraduate students over the years. The data-base has not only proved to be an essential tool for research (by staff and students alike), but also for teaching (I designed a whole set of course materials around practical analysis of data taken from CHILDES).

It is of the utmost importance for the scientific community that CHILDES continues to be supported, so that it can continue to grow, e.g. by expanding the data-base, including more video material, developing more powerful search engines and so on.

Good luck with your application, which has my unreserved support.

Best wishes,

Andrew

Prof. Andrew Radford

Head of Department of Language and Linguistics

University of Essex
Since 2002, CLAN software has been used to continue our work on lexical diversity and developmental trends in vocabulary acquisition. Analyses using the vocd and freq programs have been carried out on samples of first language speech and writing and have enabled us to develop new language measures for studying the development of inflectional morphology and more valid type/type measures such as rare word diversity and noun/verb ratios. This work has been funded by the Economic and Social Research Council (ESRC) in the UK and by the University of Reading’s Research Endowment Trust Fund.

More recently, during work on foreign language learning and assessment, also funded by the ESRC, my colleagues and I have been using the CHILDES tools in CLAN and the CHAT editor to study the development of writing in French by Year 12 students attending schools in England. We have found the following programs particularly useful in studying lexical diversity, use of words in different frequency bands (rare word usage) and collocations: freq, vocd, kwal and combo. Work on this extensive data base from nearly 200 students in 20 schools is still continuing.

The following publications report research that has been supported by the CHILDES project:


I wish you every success in your application for further funding.

With best wishes

[Signature]

Brian Richards
Professor of Education
107. Roberts, Joanne

February 25, 2008

Dr. Brian MacWhinney
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

Dear Brian:

I have been using the CHILDES program for data analysis of language interactions of children, including at-risk children, children with fragile X syndrome, Down syndrome, or autism, for the past 14 years. The program is very flexible, allowing me to do basic analyses such as MLU and more sophisticated pragmatic analyses. CHILDES is user friendly, time effective, and computes the needed linguistic measures. More importantly, whenever I have a question, I have found Brian MacWhinney and the CHILDES support team to be invaluable. The program is constantly updated so that the most current software can be used.

I strongly support the continued funding of CHILDES. This is a very cost effective mechanism to ensure that research projects use common analysis procedures. It saves research projects hours of time. The quality of the CHILDES program and the technical support provided is outstanding. I highly support continued funding of this excellent project.

Recent articles based on the use of CHILDES programs or data:


Sincerely,

Joanne E. Roberts, Ph.D.
Senior Scientist &
Professor of Speech and Hearing Sciences & Research Professor of Pediatrics
Dear Brian,

This letter is intended to provide enthusiastic support for continued funding of CHILDES. I have been associated with the project from the beginning and it has more amply fulfilled its expectations than almost anyone could have expected. In a word, CHILDES has become essential to modern research in language acquisition. I require all undergraduates and graduates at UMass in my courses to learn how to use it. And in many cases the undergraduates say that the opportunity to do really original research as undergraduates is one of the highlights of their careers. Every dissertation that I have directed or been a member of over the last 15 years has had searches as an important part---usually the start. The continued expansion of the database, both for English and other languages, is extremely important. Here's why: as the field progresses more and more refined questions are asked. They call for a larger database due to the rarity of some of the crucial phenomena. If one wants to get an overview of how a word like "promise" is acquired, there are insufficient examples in the combined childes databases of ten years ago to answer the question. We have developed a diagnostic program, DELV (Diagnostic Evaluation of Language Variation---Harcourt Brace) which focusses on the fact that "who bought what" is a crucial diagnostic of language ability. Children without problems can do it at 3yrs and those with problems cannot do it at 9yrs. CHILDES was instrumental in providing an important negative result: it is extremely rare in adult input to children. Therefore increasing exposure is important in helping children with disorders. The database proved that it was very rare.

Tom Roeper
Professor of Linguistics
University of Massachusetts
Dear Prof MacWhinney,

I am a third year PhD student at Newcastle University, UK. I have been using the CHILDES programs (namely CLAN with the CHAT conventions) extensively for my research. I am collected L2 French data as part of the FLLOC project (www.flloc.soton.ac.uk) and transcribing it using your programs, which have been invaluable. I believe that this database is also available on the CHILDES website and that we will be sharing our new data with you when the project is complete - my data will be included in that. Using the walker controller within the program has made it much easier to transcribe although for some of my colleagues they would prefer if it also supported wma files. I convert mine to mp3 so it doesn't matter. I haven't published anything but I have given some papers at conference - albeit mainly in the doctoral workshop.

2006 - Syntactic development of L2 French, Newcastle Postgraduate conference in Theoretical and Applied Linguistics, Newcastle UK
2006 - The syntactic development of L2 French, EUROSLA, Turkey and AFLS, Bristol. Both in doctoral workshop.
2007 - Verb raising in L2 learners of French, EUROSLA, Newcastle, UK
2007 - Développement du placement des verbes chez les anglophones apprenant le français, AFLS, Boulogne sur mer, France
2007 - David, A., Rogers, V. and Myles F. Développement morphosyntaxique et lexical chez les apprenants du français L2, AFLS, Boulogne sur mer, France

I hope this letter is helpful to you in your funding bid as CHILDES is a fantastic resource and it would be a catastrophe for acquisition researchers if it were not to continue.

Regards

Vivienne Rogers
vivienne.rogers@newcastle.ac.uk
www.students.ncl.ac.uk/vivienne.rogers
110. Rose, Yvan
Monday, February 18, 2008

Dr. Brian MacWhinney
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213-3890

Re: Support to CHILDES renewal further development

Dear Brian,
I would like to express my most absolute support to your current project aiming at the renewal of funding for CHILDES and the further expansion of this extremely valuable project.

CHILDES has been at the centre of my research and training activities ever since I entered the profession. In both my undergraduate language acquisition courses (LING 3155 - Introduction to Language Acquisition; LING 4150 - Language Acquisition II), I regularly use CHILDES data and analysis tools to provide my students with direct access to examples of developmental patterns in language acquisition. My students also spontaneously research the CHILDES resources for their term papers at both undergraduate and graduate levels. One former student of mine, Christophe dos Santos, recently completed his Ph.D. thesis on a longitudinal corpus documenting phonological development in a young French learner from Lyon. I also published a number of papers with Christophe based on these data. (See references below.) A second student, Laetitia Almeida, is currently working on bilingual French-Portuguese data coded using CHILDES tools (CLAN and Phon). She has already made plans to contribute these data to the CHILDES database at the end of her studies. I am myself actively working on the formatting of two corpora, which I will contribute to the database this coming spring.

CHILDES is also at the centre of several areas of development for systematic methods for empirical investigations, which themselves provide a stronger foundation for theoretical and applied research. In addition to providing data and tools, CHILDES provides traction to existing research projects worldwide. This in turns adds a significant degree of empirical and methodological credibility to several subfields in important areas of theoretical research in linguistics, psychology, cognitive science, computer science, the majority of which have significant spin-offs in important areas of applied and health research and education. I have first-hand experience in this area through our collaboration in the development of the Phon and PhonBank software and database, one that has already had a significant impact to the field of phonological development (e.g. the thesis work described above; on-going contributions of babbling and phonological development data to PhonBank; development of several new corpora in Catalan, Spanish, European Portuguese, Québec and European French, German, Czech and Saami). CHILDES is thus more than a research project or service to the community. It is at the basis of an entire, and still growing, research community, on which countless researchers and research
organizations can, and do, rely. One simple but significant example of this relates to the uniform formats that CHILDES helps developing and deploying; these standards enable large-scale investigations that provide researchers with fine-grained characterizations of the issues addressed in their research. This is particularly significant if one considers that the data and tools that are at the centre of the CHILDES community would otherwise be very hard, if not impossible, to access, in any format at all.

For all these reasons, I do believe that the proposed improvements to the CHILDES system should be considered not as an expense but rather as an extremely valuable investment, one that will further facilitate and streamline several areas of research, in addition to adding to the quality of it. In closing, I want to express my sincerest gratitude for your excellent work at developing and maintaining your brilliant research project.

Kindest regards,

[Signature]

Dr. Yvan Rose
Department of Linguistics
Memorial University of Newfoundland
St. John's NL A1B 3X9
CANADA

******************************************************************************************

111. Roy, Deb

From: Deb Roy <dkroy@media.mit.edu>
Date: February 18, 2008 1:50:15 PM GMT+01:00
To: Brian MacWhinney <macw@cmu.edu>
Subject: letter

Dear Brian,

I am writing to express my support for the CHILDES effort. Observational data and analysis tools form the bedrock of any scientific endeavor. CHILDES plays a vital role for child language research around the world. I believe there are many improvements that can be made to how we collect and analyze data that will be worthy of future research and development. In this regard, I will be happy to advise your team on making improvements to your CLAN system to increase speed of speech transcription. Furthermore, future efforts in our lab may lead to new corpora that we can contribute to your database.

Regards,
Deb Roy

-----------

Deb Roy
Director, Cognitive Machines Group • Chair, Academic Program • MIT Media Laboratory
AT&T Career Development Professor of Media Arts & Sciences • Massachusetts Institute of Technology
www.media.mit.edu/~dkroy
Dear Brian,

I have several CHILDES-related journal articles and presentations from the past six years. These studies were collaborations with Catherine Snow and Barbara Pan while I was a graduate student at Harvard University. For all of these papers we used data from low-income families participating in Early Head Start. Parents and children were videotaped interacting at 14, 24 and 36 months and videos were transcribed and analyzed using the CHAT and CLAN components of CHILDES. Here are the citations:

**Articles in Peer Reviewed Journals:**

**Presentations:**


Teaching:
I taught an undergraduate Introduction to Language Development course at the University of Chicago in the Winter quarter of 2007. Thirty-five undergraduates took this course, mostly juniors and seniors majoring in psychology, human development or linguistics. This was a lab course where students learned about research methods related to child language development by learning the CHILDES system. I gave 2 hands-on tutorials describing the system and how it is used and then students had 3 homework assignments which supplemented what they were learning in class. The assignments focused on 1) transcription 2) CLAN analysis related to lexicon (freq, ttr), and 3) CLAN analysis related to syntax (MLU). For assignments 2 and 3 they analyzed data from several corpora in CHILDES and discussed their findings based on what they were learning in class about language development and individual differences.

I feel very strongly that CHILDES is an incredible resource to people in our field. Thanks so much for everything you do to keep it going and improving all the time!

Best,
Meredith Rowe
Postdoctoral Fellow
University of Chicago
rowemer@uchicago.edu
Dear Brian,

We have continued to use CHILDES extensively in both research and teaching work. It is an invaluable resource that has transformed the way in which we do research. We use it a) to analyse existing datasets donated by other researchers b) to analyse transcripts we have collected and c) to guide the way in which we transcribe and d) to create materials for experimental studies (e.g. to identify frequent and non-frequent words in child directed speech).

We have two datasets that we intend to donate to CHILDES over the next year or so. Here is a list of some work (since 2001) that relied in some way on access to the CHILDES system:

**Journal publications**


Conference presentations and talks


Grants awarded


Dr. Ben Ambridge. Amount awarded: £26,500.


Good luck with this

Dr. Caroline Rowland
School of Psychology
University of Liverpool Bedford Street South
LIVERPOOL, L69 7ZA, UK Tel: +44 151 794 1120 Fax: +44 151 794 2945
Child Language Study Centre: http://www.liv.ac.uk/psychology/clrc/clrg.html
Dear Brian,

As a frequent user of CHILDES I would like to express my gratitude for founding, maintaining, adjusting and expanding the CHILDES-database and the CLAN analysis tool. I have been working with the CHILDES-system for nearly eight years now. First when I was writing my MA-thesis, for which I used data from the Groningen-corpus. As a student-assistent I have expanded my knowledge of Childes and for my PhD-thesis, I use CHILDES again. My project focuses on the interaction between morphosyntax and pragmatics in the acquisition of reference from a cross-linguistic perspective (Dutch, English and French). It would have been impossible to carry out this research without CHILDES. Collecting, transcribing, coding and analyzing longitudinal data in three languages for three children each within a period of four years is impossible. That is why I am very glad with CHILDEs from which all data were taken. Furthermore, I used the CLAN analysis tool for coding and analyzing my data. I also use MOR automatic coding to get a fuller picture of the morphosyntactic forms that children use.

I think that it is very important that the CHILDES-project is continued. The database must be expanded by new corpora, also from languages that are not well represented yet, for example the Scandinavian languages. The integration of dense databases into CHILDES is also crucial, given the new trends in language acquisition research.

Publications
Rozendaal, M.I. The acquisition of the morphosyntax-pragmatics interface in French L1: Evidence from reference with articles and pronouns. In S. Baauw, J. van Kampen & M. Pinto (eds.), The Acquisition of Romance Languages. Selected Papers from The Romance Turn II. LOT: Utrecht, pp 145-164. 2007
To whom it may concern

My experience as a linguist working with language acquisition made me understand the importance of a project such as CHILDES, which allows sharing large collection of data and makes tools available for transcription and coding of spontaneous speech.

During the research for my PhD, I must say that the Clan tools, as well as the manuals providing indications concerning transcription, were essential for the transcription and the analysis of the data. The CHILDES database also made crosslinguistic comparisons possible, such as a comparison between my data (child European Portuguese) and data concerning the acquisition of French (York corpus) or the acquisition of English (Brown corpus). Recognizing the importance of CHILDES, I am preparing to contribute data to this important database.

The following papers and dissertation could not have their current form without CHILDES:


Santos, Ana Lúcia (2006) A compreensão de só em posição pré-verbal por crianças entre os 2;9 e os 4;11. In F. Oliveira e J. Barbosa (eds.) Textos seleccionados. XXI Encontro Nacional da APL. APL.


Lisbon, February 27th 2008

Ana Lúcia Santos
(Assistant Professor, Universidade de Lisboa)
To Whom It May Concern,

Monica Sanz Torrent, Ph.D in Psychology by the University of Barcelona and associate at the Basic Psychology Department in the Universitat de Barcelona, ensures that:

The CHILDES Project is a very useful tool for studying language. It allows a detailed analysis of children and adult’s spoken language’s samples, both with its transcription format (CHAT), as with its analysis programs (CLAN). Its new applications and those resources that are currently under development are central to the research in this area. Within its multiples applications, infrastructure and resources CHILDES allows the study of language development in children with communication disorders and/or language pathology, and its large database allows cross-linguistic comparisons. In Barcelona’s University, we are currently running different training courses for diverse groups of Spanish investigators that are also using CHILDES's system tools and some master's works (I. Badia, Cl. Machado, LL. Andreu,) and PhD. thesis (M. Sanz y E. Aguilar) of Basic Psychology’s Department have been based on this method of analysis.

CHILDES System is one of the main methodologies used in the following projects financed by the Spanish and the Catalan governments:

PARTICIPATION IN FUNDED R+D PROJECTS FROM PUBLIC CALLS:

Title of the project / contract: Interacción entre factores cognitivos i lingüísticos en la adquisición normal y patológica.
Kind of contract/Program: Programa Nacional de Promoción General del Conocimiento
Financing Firm/administration: Ministerio de Ciencia y Tecnología.
Number of the project / contract: BSO2003-02200 Amount: 16.400,00 Duration, since: 2003 Until: 2006
Main researcher: Miquel Serra Raventos
Number of researchers participating: 08
Title of the project / contract: Avaluació de la competència en comunicació lingüística: Anàlisi del llenguatge, aspectes morfo-sintàctics i semàntics.
Kind of contract/Program: Ajuts a la Recerca
Financing Firm/administration: AGAUR. Agència de Gestió d'Ajuts Universitaris i de Recerca de la Generalitat de Catalunya
Number of the project / contract: 2006ARIE10048 Amount: 4.800,00 Duration, since: 2007 Until: 2008
Main researcher: Monica Sanz Torrent
Number of researchers participating: 11

Title of the project / contract: Procesamiento del lenguaje y comprensión del evento en niños con diferente competencia lingüística.
Kind of contract/Program: Programa Nacional de Información para la Investigación Científica y el Desarrollo Tecnológico
Financing Firm/administration: Ministerio de Educación y Ciencia
Number of the project / contract: SEJ2007-62743 Amount: 35.332,00 Duration, since: 2007 Until: 2010
Main researcher: Monica Sanz Torrent
Number of researchers participating: 6

And it also has being used for the following published studies:

JOURNAL PUBLICATIONS

Authors (signature): Aguilar, E.; Sanz-Torrent, M.; Serra, M.
Title: A comparative study of the phonology of pre-school children with Specific Language Impairment (SLI), Language Delay (LD) and Normal Acquisition.
Journal: Clinical Linguistics & Phonetics
Volume: 16 Number: 8 Pages, Initial: 573 final: 596 Year: 2002 Place of publication: (ENGLAND) ISSN: 0269-9206

Authors (signature): Sanz-Torrent, M.
Title: Los verbos en niños con Trastorno del Lenguaje.
Journal: Revista de Logopedía, Foniatria y Audiología
Volume: 2 Number: 2 Pages, Initial: 100 final: 110 Year: 2002 Place of publication: (SPAIN) ISSN: 0269-9206

Authors (signature): Serra, M.; Aguilar, E.; Sanz-Torrent, M.
Title: Evolución del perfil productivo en el Trastorno del Lenguaje.
Journal: Revista de Logopedia, Foniatria y Audiología
Volume: 2 Number: 2 Pages, Initial: 77 final: 89 Year: 2002 Place of publication: Barcelona (SPAIN) ISSN: 0269-9206

Authors (signature): Serra, M.; Sanz-Torrent, M.
Title: Acotando la sobreextensión de 'gramática' y de otros términos lingúisticos
Journal: Cognitiva
Volume: 15 Number: 2 Pages, Initial: 221 final: 233 Year: 2003 Place of publication: (SPAIN) ISSN: 0214-3550

Authors (signature): Aguilar, E.; Sanz-Torrent, M.
Title: Errors d'omissió en nens amb trastorn especific del llenguatge (TELL): Aspectes morfològics i fonològics
Journal: Revista del Col·legi de Logopedes de Catalunya
Authors (signature): Serra, M.; Sanz-Torrent, M.
Title: Las 400 palabras: La interacción léxico-sintáctica en la adquisición del lenguaje
Journal: Anuario de Psicología.
Volume: 35 Number: --- Pages, Initial: 183 final: 206 Year: 2004 Place of publication: Barcelona (SPAIN) ISSN: 0066-5126

Authors (signature): Serrat, E.; Sanz-Torrent, M.; Bel, A.
Title: Aprendizaje del lenguaje y desarrollo gramatical: vocabulario verbal, aceleración léxica y complejidad sintáctica.
Journal: Anuario de Psicología.
Volume: 35 Number: --- Pages, Initial: 221 final: 2234 Year: 2004 Place of publication: Barcelona (SPAIN) ISSN: 0066-5126

Authors (signature): Andreu, LL.; Sanz-Torrent, M.; Serra, M.
Title: Dificultades y trastornos del lenguaje en la escuela
Journal: Quaderns digitals
Volume: 2 Number: 2 Pages, Initial: 1 final: 26 Year: 2007 Place of publication: Valencia (SPAIN) ISSN: 1575-9393

Authors (signature): Aguilar, E.; Sanz-Torrent, M.; Serra, M.
Title: The influence of phonology on morpho-syntax in Romance languages in children with Specific Language Impairment (SLI)
Journal: International Journal of Language & Communication Disorders
Volume: 42 Number: 3 Pages, Initial: 325 final: 347 Year: 2007 Place of publication: (ENGLAND) ISSN: 0269-9206

Authors (signature): Sanz-Torrent, M.
Title: Specific Language Impairment: lexical, semantic and morpho-syntactic aspects of verb acquisition
Journal: Language Acquisition
Volume: 15 Number: 1 Pages, Initial: --- final: --- Year: 2008 Place of publication: (UNITED STATES) ISSN: 1048-9223

Authors (signature): Sanz-Torrent, M.; Serrat, M.; Serra, M.
Title: Verb morphology in Catalan and Spanish in children with Specific Language Impairment: a developmental study.
Journal: Clinical Linguistics & Phonetics
Volume: --- Number: --- Pages, Initial: --- final: --- Year: in press Place of publication: (UNITED KINGDOM) ISSN: 0269-9206

BOOK PUBLICATIONS

Authors (signature): Serra, M.; Matínez, I.; Hidalgo, M.; Sanz Torrent, M.
Title: La evolución inicial del Trastorno del Lenguaje (TL): Datos y modelos para una mejor caracterización diagnóstica y para el pronóstico.
Book: Homenaje al Dr. Jorge Perelló Gilberga.
Publisher: Publicaciones Universidad Pontificia de Salamanca
Volume: --- Number: --- Pages, Initial: --- final: --- Year: 2002 Place of publication: Salamanca (SPAIN) 004

Authors (signature): Sanz-Torrent, M.; Serrat, E.
Title: Pragmàtica i discurs
Authors (signature): Sanz-Torrent, M.; Badia, I.; Serra, M.
Title: Contributions from the bilingual Specific Language Impairment in Catalan and Spanish to the understanding of the normal and pathological language acquisition.
Book: En C. Pérez-Vidal, C.; M. Juan-Garau & A. Bel (Eds.) A portrait of the young in the new multilingual spain.
Publisher: Multilingual Matters

Authors (signature): Andreu Barrachina, LL., Sanz-Torrent, M.; Aznar, J.A.; Cámara, E.
Title: La comprensión y la producción del lenguaje en niños con TEL: un estudio mediante los movimientos oculares.
Book: Estudios de desarrollo del lenguaje y educación.
Publisher: Aula abierta
Volume: --- Number: --- Pages, Initial: --- final: --- Year: in press Place of publication: Oviedo (SPAIN)

Authors (signature): Sanz-Torrent, M.; Andreu-Barrachina, Badia, I.; Mayoral, X.; Serra, M.
Title: 'Mirar para hablar': movimientos oculares y producción en niños con TEL
Book: Estudios de desarrollo del lenguaje y educación.
Publisher: Aula abierta
Volume: --- Number: --- Pages, Initial: --- final: --- Year: in press Place of publication: Oviedo (SPAIN)

And finally for the following national and international congress presentations:

CONFERENCE CONTRIBUTIONS

Authors: Aguilar Mediavilla, E.; Sanz-Torrent, M.; Serra Raventós, M.
Title: Evolution of phonological simplification processes of children with specific language impairment.
Kind of participation: Poster
Conference: IX IASCL Congress & SRCLD
Place of celebration: Madison (Wisconsin) (UNITED STATES) Year: 2002

Authors: Sanz Torrent, M.; Aguilar Mediavilla, E.; Serra Raventós, M.
Title: SLI deictic, semantic and morphosyntactic difficulties in Catalan/ Spanish verb production
Kind of participation: Poster
Conference: IX IASCL Congress & SRCLD
Place of celebration: Madison (Wisconsin) (UNITED STATES) Year: 2002

Authors: Serra Raventós, M.; Aguilar Mediavilla, E.; Sanz Torrent, M.; Hidalgo, M. C.; Martinez, I.
Title: Psycholinguistic Profiles Of Spanish Specific Language Impaired (SLI) Children: The Evolution of Cognitive And Linguistic Difficulties From 3 To 5 Years
Kind of participation: Poster
Conference: IX IASCL Congress & SRCLD
Place of celebration: Madison (Wisconsin) (UNITED STATES) Year: 2002

Authors: Serrat, M., Sanz-Torrent, M. y Bel, A.
Title: Verb acquisition in Catalan and Spanish speaking children: lexical, morphological and syntactic aspects.
Kind of participation: Poster
Conference: IX International Congress for the Study of Child Language y al Simposium on Research in Child Language Disorders
Place of celebration: Madison (UNITED STATES) Year: 2002

Authors: Serra, M.; Sanz-Torrent, M.
Title: La evolución inicial del Trastorno del Lenguaje: Datos y modelos para una mejor caracterización diagnóstica y para el pronóstico.
Kind of participation: Conferència invitada
Conference: II Jornadas Luso Espanholas de Linguagem infantil
Place of celebration: Maia (PORTUGAL) Year: 2002

Authors: Sanz-Torrent, M.; Hidalgo, M.; Badia, I; Serra, M.
Title: Extended Verb Island Period in SLI children?.
Kind of participation: Poster
Conference: VIII Meeting of the ‘European Child Language Disorders Group’
Place of celebration: Gregynog (ENGLAND) Year: 2003

Authors: Sanz-Torrent, M.; Aguilar, E.; Badia, I.; Serra, M.
Title: Event representation, argument structure and verb use in Specific Language Impairment Children
Kind of participation: Poster
Conference: 6th Symposium on Psycholinguistics
Place of celebration: Barcelona (SPAIN) Year: 2003

Authors: Sanz-Torrent, M.; Badia Llenas, I.; Serra Raventos, M.
Title: La lectura en los niños con Trastorno del Habla y del Lenguaje
Kind of participation: Conferència invitada
Conference: XIV Congreso de la Sociedad Española de Logopedia, Foniatria y Audiologia
Publication: Actas del XIV Congreso de AELFA
Place of celebration: Madrid (SPAIN) Year: 2004

Authors: Sanz-Torrent. M.; Badia, I.; Serra, M.
Title: Aspectos léxicos y morfo-sintácticos asociados a la adquisición del verbo en niños con TEL.
Kind of participation: Presentació comunicació
Conference: IV Congreso Internacional sobre Adquisición de las Lenguas del Estado.
Place of celebration: Salamanca (SPAIN) Year: 2004

Authors: Sanz-Torrent, M.; Badia Llenas, I.; Serra Raventos, M.
Title: Grammaticalization processes in SLI children
Kind of participation: Poster
Conference: Second Lisbon Meeting on Language Acquisition
Place of celebration: Lisboa (PORTUGAL) Year: 2004

Authors: Sanz-Torrent, M.
Title: Habilitats i coneixements implicats en el desenvolupament del llenguatge
Kind of participation: Conferència invitada
Conference: Col·legi de Logopedes de Catalunya
Place of celebration: Barcelona (SPAIN) Year: 2004

Authors: Sanz-Torrent. M.; Badia, I.; Serra, M.
Title: El trastorno del Lenguaje y el inicio de la lectura
Kind of participation: Conferència invitada
Conference: XXI Congreso Internacional de Logopedia, Foniatria y Audiología
Place of celebration: Madrid (SPAIN) Year: 2004

Authors: Sanz-Torrent. M.; Badia, I.; Serra, M.
Title: Aspectos léxicos y morfo-sintácticos asociados a la adquisición del verbo en niños con TEL.
Kind of participation: Presentació comunicació
Conference: IV Congreso Internacional sobre Adquisición de las Lenguas del Estado.
Place of celebration: Salamanca (SPAIN) Year: 2004

Authors: Sanz-Torrent, M.; Serrat, E.; Olmo, R.
Title: Verb vocabulary and progress on verb morphology: a case study
Kind of participation: Poster
Conference: X International Congress for the Study of Child Language
Place of celebration: Berlin (GERMANY) Year: 2005

Authors: Serra, M.; Sanz-Torrent. M.; Badia, I.; Hernández, A.
Title: 'The 400 words': relations between the lexicon and the morphosyntax.
Kind of participation: Poster
Conference: X International Congress for the Study of Child Language
Place of celebration: Berlin (GERMANY) Year: 2005

Authors: Serrat, E.; Olmo, R.; Badia, I.; Sanz-Torrent. M.; Aguilar, E.; Lara, M.F.; Serra, M.
Title: Catalan versión of the MacArthur-Bates CDI-II: words and sentences.
Kind of participation: Poster
Conference: X International Congress for the Study of Child Language
Place of celebration: Berlin (GERMANY) Year: 2005

Authors: Badia, I; Sanz-Torrent, M.; Serra, M.
Title: Verb use patterns of Children with Specific Language Impairment: a comparative study of event representations
Kind of participation: Poster
Conference: IX Meeting of the 'European Child Language Disorders Group'
Place of celebration: Abbaye de Royaumont (FRANCE) Year: 2005

Authors: Badia, I; Sanz-Torrent, M.; Serra, M.
Title: Patrones de uso de verbos en niños con Trastorno Específico del Lenguaje
Kind of participation: Poster
Conference: VII Symposium of Psycholinguistics.
Publication: Actas del XIV Congreso de AELFA
Place of celebration: Valencia (SPAIN) Year: 2005

Authors: Andreu, LL.; Sanz-Torrent, M.; Badia, I.; Machado, C.; Serra, M.
Title: Procesamiento del lenguaje y movimientos oculares en niños con TEL
Kind of participation: Poster
Conference: VIII Simposio de Psicolingüística
Place of celebration: Palma de Mallorca (SPAIN) Year: 2007

Authors: Sanz-Torrent, M.; Badia, I.; Andreu, LL.; Serra, M.
Title: Verbs, argument structure and visual attention in children with SLI
Kind of participation: Poster
Conference: AFASIC 4th International Symposium
Place of celebration: University of Warwick (ENGLAND) Year: 2007
For all those reasons exposed, we believe it is entirely necessary the continuity and evolution of this key component for the improvement of the study of language acquisition.

En Barcelona, 19th February 2008

Monica Sanz Torrent
Laboratori de Psicolinguistica I Patologia del Llenguatge
Departament de Psicologia Basica
Universitat de Barcelona
117. Savickiene, Ineta

Dear Brian,

in order to support all the activities related to CHILDES, CLAN I would like to express my strong support; therefore I am sending you the list of my papers.
Mu Phd student Laura Kamandulyte already sent you the info concerning Spoken Lithuanian corpus where we use CHILDES.

I work with CHILDES all the time for investigating Lithuanian children speech as well as CDS. Now we have 3 girls and 3 boys longitudinal data recorded, transcribed and partially coded by using CHILDES.

With my best regards,

Ineta Savickiene

p.s. we’ve met in Pittsburgh in 2003 at the seminar you organised for CHILDES users. I hope that you will come to Lithuania one day

Prof. dr. Ineta Savickienë

Vytautas Magnus University, Kaunas, Lithuania


Brian MacWhinney
Department of Psychology
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213 USA

Dear Colleague,

I have been using the Child Language Data Exchange System (CHILDES) for many years. Besides the general goals to be attained by any researcher who works in the field of cross-culture child language acquisition, my specific ones were the following:

1 – adapting the codes to the phonologic/phonetic systems of the Brazilian Portuguese variants;
2 – feeding the CHILDES database with the corpora of my doctoral thesis which on the occasion deserved from Professor Roger Brown the following quotation: “It is a great thrill for me whenever I learn that someone in another country has done a study testing the generality of the results in A First Language. Not many have done it so carefully, intelligently, and comprehensively as yourself” (letter of March 15, 1977);
3 – discussing theoretical and methodological issues, which are crucial for the field of language acquisition, e.g., how children build up their mental lexicon and MLU;
4 – helping my students in their research projects for master and doctoral degrees with a powerful tool.

My first specific goal was attained (see The CHILDES Project, Tools for Analyzing Talk editions). Regarding point 1, from the three stages of collected data, the ones belonging to stages one and two have already been sent, with bullets for analyzing each utterance, and phonetic and morphological description of the child’s utterances. Data belonging to the third stage was already sent, without bullets and phonetic and morphological description of the child’s utterances. We are in the process of putting the bullets on the third stage and formalizing the grammar for automatic analysis of the Portuguese language. All the data can be heard through audio. Points 3 and 4 may be evaluated through the list of scientific production appended to this letter, referring to the period 2003-2007.

The new proposal to be presented for a renewal of NIH funding is of particular interest for further research on automatic morphological analysis of the acquisition of the Portuguese language.

Sincerely yours

Prof. Emeritus Leonor Scliar-Cabral


Dear Dr. MacWhinney,

I wish to let you know that, in my Ph.D. work on literal and non-literal motion in adult and child Greek (Selimis 2007), I have benefited from the CHILDES data base. Two chapters in my dissertation compare literal or non-literal uses of motion verbs in spontaneous conversations between young children and adults in Greek and English. The English data was extracted from the corpora of two girls, Eve and Sarah (Brown 1973), available from the CHILDES data base. In addition, my advisor, D. Katis, and I have given several papers which are partly based on these corpora (see below).

I would also like to take this opportunity to note that I have a longitudinal corpus of one Greek child in the age range of 1;8-2;10 years. I hope to contribute it, although it has not been transcribed yet.

With best regards,

Stathis Selimis

Selected references


February 6, 2008

Dear Brian,

This is a brief note to thank you for your work on the CHILDES project. As I said in my letters in the past, CHILDES has been an indispensable part of my research and teaching. Here’s the publication list of research papers of the last 6 years (2002-2007) that made use of the CHILDES system (database, CHAT or CLAN):


In addition, CHILDES has been an important tool in my teaching. I am now teaching a directed research course for two linguistics undergraduate student using CHILDES L2 data from European Science Foundation. I also regularly encourage graduate students to analyze CHILDES data for their term project, one of which was presented at BUCLD as a poster: Kwan, C. & Shirai, Y. (2005). Adjectives in early developmental Cantonese: Time-stability and the use of property concept terms. Poster presented at the 30th Boston University Conference on Language Development, Boston, MA, Nov. 3-5.

CHILDES has been invaluable to me and to my students. I sincerely hope for the further development of the project.

Sincerely,

Yasuhiro Shirai
Professor of Linguistics
Dear Brian

I am glad to hear that you are working on the application for the continuation and further development of CHILDES. It is an invaluable resource for research on language development and also for my teaching. In the past two years, I have used it mostly to pull out examples that I can use in preparing teacher trainees here in Singapore. Because of the diversity in the CHILDES data, I can find examples from children of different ages, different places, different languages. This is particularly helpful for showing future teachers how the linguistics development of Singaporean children is similar/different to of others internationally.

CHILDES has also been used by our Child Language Special Interest Group (part of the Singapore Association of Applied Linguistics) to post information about work on child language in Singapore. (Dr. Madalena Cruz-Ferreira has posted a bibliography of work done in Singapore on Child Language.) CHILDES is an essential part of our sharing with the larger international community of scholars.

In a few cases my BA students have done some independent searching on CHILDES but we have not been able to do as much as this as I would like. At this point, students are unsure of how to proceed without guidance and I am not always able to provide the support for individual projects. Web-based tutorials and more friendly, up-to-date user interface would be particularly helpful for facilitating individual research at the undergraduate level. Further support for conversational analysis (CA) would also be most welcome as our students are learning to do preliminary CA analyses in their Master’s coursework – this is an area in which our MA students could certainly make better use of CHILDES with additional complimentary tools. I see the possibility of collaborative commentary (with blogs) as a very useful future tool for sharing with scholars internationally but also, more immediately, for working with future teachers and with current teachers on action research projects. This would allow for on-line commentary/discussion in an accessible, convenient manner.
For my own research, I am particularly excited by the possibility of assisted transcription (audio and video). The main cost of many of our research project on classroom language use is in transcription. For example, my current research on how children in lower primary use language during peer work requires that individual and group transcript to be created. We spend 8-12 hrs to transcribe 1 hr of classroom talk – a heavy cost of doing this type of research.

I look forward to hearing about a positive outcome on the grant and to seeing the further development of CHILDES.

Best regards
Rita Silver
Dear Professor MacWhinney,

I have been using the Child Language Data Exchange System (CHILDES) as a member of the research group *Emergent linguistic productivity*, directed by Professor Emeritus Leonor Sciliar-Cabral, and sponsored by the Brazilian National Council of research (CNPq).

My work consisted of feeding the CHILDES database with the corpora of Sciliar-Cabral’s doctoral thesis, discussing at the regular seminars of the research group theoretical and methodological issues, which are crucial for the field of language acquisition and coding the morphological categories of Paulo’s utterances (stages one and two). Since 2005 I presented several papers on the subject, as you can follow at the list bellow. The new proposal to be presented for a renewal of NIH funding is of particular interest for my further research on automatic morphological analysis of the acquisition of the Portuguese language.

Jocieli Sinigaglia
Universidade Federal de Santa Catarina
Florianópolis, SC
Brazil
jocielis@hotmail.com

Papers


___. (2006b). “Dos países, três formas de hablar”. Trabalho apresentado na Jornada Acadêmica de Espanhol III (Jaleo III). UFSC. 08-09/05/06.
____. (2006c). “Freqüência lexical numa criança aos 20 meses e 21 dias adquirindo o PB”. Apresentação (em forma de painel) no XVI Seminário de Iniciação Científica (SIC), UFSC.


Dear professor MacWhinney,

We are a team of linguists, psychologist and speech therapist, who have been involved into the research of child language acquisition in Slovak language, having been worked on it approximately since the beginning of the 21st century. From the beginning, inspired by a member of our team S. Kapalková, we have used the CHILDES system as a tool for transcribing our longitudinal video and DVD records of 5 Slovak-speaking children from their birth till now 5 – 6 years (every month – one hour record). Moreover, there are several semi-longitudinally observed children, as well as student papers, made as a part of the courses of Child Language Development, led by member of our team at two universities in Slovakia (Comenius University in Bratislava and Prešov University in Prešov); teaching how to work with CHILDES is a part of our courses. Our corpora of transcribed materials archive now about 200 hours of texts.

Working with CHILDES has been a part of the following research projects:

*Vytvorenie kurzu Vývin detskej reči pre humanitne zamerané fakulty na Slovensku* (Creating of the course „Speech Development“ for the Humanities in Slovakia).
Vzdělávací nadace Jana Husa CI 140 (Cursus Innovati) 2002-2004

*Vývin reči slovensky hovoriacich detí* (Language Development of Slovak-speaking Children)
*Vedecká grantová agentúra Ministerstva školstva SR a Slovenskej akadémie vied 2005-2007*

Verbálne a neverbálne správanie slovensky hovoriacich detí v raných štádiách vývinu ako východisko vytvorenia skriningového testu vývinu komunikačného správania (*Verbal and Non-Verbal Behaviour of Slovak-speaking Children in Early Development as the Basis of Creation First Screening Communicative Development Inventory*)
*Agentúra na podporu výskumu a vývoja Slovenskej republiky 2007-2009*

Bellow are the most important publication units (with bibliographical data), in which the CHILDES system was used:


PhD. disertations:

BÉREŠ, R.: Morfologické kategórie v ranej ontogenéze detskej reči.
HARČARIKOVÁ, P.: Vývin textových naratívnych štruktúr v detskej reči.
CHABOVÁ, M.: Lexikálnosémantická analýza plnovýznamových slovných druhov v reči slovensky hovoriacich detí.
ONDRÁČKOVÁ, Z.: Komparatívny výskum detských slov v slovenčine a v angličtine.
ZAJACOVÁ, S.: Formovanie komunikačného registra v rolových hrách detí.

Diploma theses:

DANÍKOVÁ, K.: Fonologické procesy v detskej reči.
MIHALČOVÁ, M.: Opakovanie v komunikácii dospelého a dieťaťa.
Bachalor theses:

BALUCHOVÁ, J.: Dieťa a bilingvizmus.
CUPEROVÁ, D.: Dieťa a bilingvizmus.)
ČECHOVÁ, M.: Dialóg u detí mladšieho školského veku.
FĽAKOVÁ, L.: Relačnosť v ranej ontogenéze.
KRAMÁRIKOVÁ, T.: Slovnodruhová charakteristika reči orientovanej na dieťa.
OCILKOVÁ, A.: Vývinové paralely detského dialógu a hry.
ŠTEFANCOVÁ, A.: Priestor a čas v ranej ontogenéze.

The members of our research team:

prof. PhDr. Daniela Slančová, CSc. (Prešov University, Prešov, Slovak Republic)
Mgr. Svetlana Kapalková, PhD. (Comenius University, Bratislava, Slovak Republic)
doc. PaedDr. Jana Kesselová, CSc. (Prešov University, Prešov, Slovak Republic)
doc. PhDr. Marina Mikulajová, CSc. (Comenius University, Bratislava, Slovak Republic)
PhDr. Iveta Bónová, PhD. (Prešov University, Prešov, Slovak Republic)
PhDr. Zuzana Ondráčková (Žilina University, Žilina, Slovak Republic)
Mgr. Stanislava Zajacová (Prešov University, Prešov, Slovak Republic)

With our best wishes,

Daniela Slančová
Professor of Slovak language, Prešov University, Slovak Republic

Svetlana Kaplaková
Comenius University, Bratislava, Slovak Republic
Dear Dr. MacWhinney,

As a member of the Department of Psychology at Alma College, I have offered a Spring Term Workshop in Children’s Language three times since 2002; we are currently registering students who will be taking it this May. With no other class on their schedule, a dozen undergraduates meet three hours a day, five days a week, read about language acquisition and aspects of early language use, then formulate questions and conduct their own inquiries using the CHILDES data base and associated analytical tools. The first part of the workshop is substantially scaffolded with handouts that introduce the syntax of a few basic CLAN commands (mlu, freq, +s, kwal) and identify a few particularly inviting and accessible corpora (Brown, Nelson, Frog stories); after that, workshop members are asked to go further afield, to explore other corpora and extend their repertoire of commands (MLT, columns, lines, combo) in the service of answering questions of their own. Their projects have ranged from the utterly predictable (Do color words make up a greater proportion of the words used by little girls than by little boys?) to the poignant (a search for color words in a father’s conversation with a blind child) to the surprising (a native Russian speaker’s comparison of parents’ use of affectionate diminutives in English and Russian). A student project that began two years ago as a small scale inquiry into possible differences in mothers’ questions to Down Syndrome and typically developing children using the Rondal corpus blossomed into a senior thesis that is nearing completion.

The CHILDES data base and analytic commands and the generosity of the CHILDES community have made possible a unique educational opportunity for students. They not only learn about language acquisition; they learn firsthand how a research community functions. Eye deep in their own inquiries, they spontaneously come up with questions about the reliability of categorizations and the generalizability of findings that I find hard to motivate in a research methods class. They are willing to entertain questions about sampling variability. They are thrilled to receive an e-mail response from chibolts or a coding scheme in draft format from a researcher who caught their query on info-childes. Presenting her project, an African American student told the group that this was the first time she had ever heard BVE discussed seriously: her high school teacher had sent students to stand in the corner “if they talked like that.” Easy to guess what she looked at and wrote about.

Although CHILDES has been used by researchers in psychology and is cited in books that we use in our classes, queries on info-childes and PSYTEACH lead me to believe that it is not likely to be included in undergraduate psychology teaching. Perhaps workshops could be organized for people who teach undergraduate courses in psychology, along the lines of the Chatauqua series for other science teachers.

Carol Slater
Dana Professor of Psychology, Alma College
Alma, MI 48801
Dear Brian,

I am writing in enthusiastic support of your application to NIH for continuing funding of the Child Language Data Exchange System (CHILDES). Before I retired in 2003, my own teaching and research would simply have been impossible without CHILDES—and without your own personal energy and initiative in keeping the international research community in our field involved in a standardized procedure for the responsible contribution, storage, sharing, and accessing of data. The continuing growth of the archived materials, along with the expanding flexibility of research tools—now multimedia—is an inspiring scientific model.

For the moment, I will only briefly highlight my uses of CHILDES over the past decade or so:
I regularly used CHILDES in both undergraduate and graduate instruction. Even lower-division students can have the thrill of accessing actual data and carrying out original studies. Student projects were often publishable and sometimes published, often presented at professional meetings. CHAT gives students experience in careful transcription and preparation of data; CLAN gives them experience in searching for patterns and testing theories. And, from the instructor’s point of view, CHILDES provides just the right examples needed to make a point in a lecture or to provide a problem set in a homework assignment.

My own research is explicitly crosslinguistic, and I am continually engaged in research projects that make it necessary to analyze comparable data across languages. In the past few years I’ve been working with a team of researchers at the Max Planck Institute for Psycholinguistics in Nijmegen, using CHILDES data in about a dozen languages to carry out new research on children’s developing understanding of space and causality (Slobin et al., in press b).

CHILDES houses most of my own previous datasets, including elicited narratives (“the frog story”) in a special crosslinguistic directory. I continue to draw on these materials for new analyses (Slobin, in press a).

You have been exceptionally helpful in supporting the development of a new system of transcribing sign language data at the morphological level—the Berkeley Transcription System (BTS), now incorporated in the CHILDES system. Together with colleagues in the US, the Netherlands, Germany, the UK, and Denmark, we are following your lead in transcribing data of sign language acquisition in a format that is compatible with CHILDES norms, allowing for contribution of a large and rich database of sign language materials to CHILDES. The multimedia capacities of CHILDES are essential to the further progress of research on signed languages (as well as co-speech gesture in hearing children and their caregivers) (Hoiting & Slobin, 2007; Slobin, 2005, 2008; Slobin et al., 2001).

CHILDES is wonderfully democratic and open, while adhering to the highest scientific standards. The system is well known and used on all continents. In fact, I can’t imagine research in our field without CHILDES—which means not only continuing support for the existing system, but also support for your many efforts to expand, modernize, and add new facilities and strengths. NIH should realize what a good investment they have made in this project, with a tremendous multiplier effect and a leader who has truly revolutionized the way research is done in this field.

With warm regards, and respect,

Dan I. Slobin
Emeritus Professor of Psychology and Linguistics

Rhoda and Richard Goldman Distinguished Professor
in Undergraduate and Interdisciplinary Studies

Research Psychologist
  Institute of Human Development
  Institute of Cognitive and Brain Sciences

References


Some recent publications and presentations based on CHILDES facilities


Dear Brian,

I am writing to express my support for the continuation of funding for the CHILDES Project. This database and related programs and tools has been an invaluable resource over the last several years for my research in early grammatical development in infants. I made considerable use of the CHAT/CLAN programs in transcribing a corpus of infant-directed speech by mothers of two young infants, and I intend to submit the corpus (and accompanying audio recordings) very soon to the database in order that this resource may be available to other researchers who are interested in the acoustic and grammatical properties of maternal speech. I also made use of the same system in analyzing a smaller sample of speech to the a third infant, which was the basis and the inspiration for some recent work on disfluency in maternal speech, and infants' perception thereof. The references for these projects are listed below. Aside from playing these major roles in my research, I have also made use of the CHILDES database in many smaller ways, including for example quick checks of the relative frequency of various lexical items and types, and providing examples in my teaching.

Not to be underestimated is the importance of the community that has grown up around the CHILDES database, which is evidenced by the large membership and activity on its various mailing lists. These lists are an important source of help, information, references and discussion. Indeed, anyone who is doing research in language acquisition is likely to make use of CHILDES system and community in one way or another over the course of their research and teaching. Its influence is so widespread and endemic that we are not even always aware that a particular piece of data originated there, despite the best efforts to document and reference its use. Any attempts to calculate a citation score will likely be a gross underestimate of its significance in the research community.

Despite this great influence, there are definitely areas that could use improvement or attention, especially given the rapid technological advances that have taken place since its inception. Dr. MacWhinney has outlined a number of important areas for further work or change over the coming years, including but not limited to improvements on the user interface, automatization of morphosyntactic analysis and syntactic measures such as IPSyn and DSS, blogging capabilities, and of course continued expansion of the database. I urge...
you to support the continuation and improvement of this valuable and successful
database project with a renewal of its funding.

Sincerely,
Melanie Soderstrom, Ph.D.
Brown University
Dept. of Cognitive and Linguistic Sciences

Relevant References:


Presentations:


Dear Brian MacWhinney,

This letter is in support of the CHILDES project as you apply for continuing funding. I am a doctoral student currently completing my dissertation and the availability of the CHILDES database has been an integral part of my dissertation work. My study investigates the role of lexical factors such as word frequency and phonological neighborhood density on young children’s production of known words. Currently, most studies of this type use adult-based measures of word frequency and neighborhood density. I was interested in developing more age-appropriate measures for my analysis and so have been working for the past two years on developing age-specific lexicons for children between 1 and 5 years old, from which we have derived age-specific measures of word frequency and neighborhood density. Using the CHILDES database, we have been able to generate a lexicon of almost 800,000 words for young children. We have presented preliminary results of this project and my dissertation research at three conferences and hope to soon make results easily available for all researchers and clinicians who would like to make use of the new age-specific measures of word frequency and phonological neighborhood density.

Presentations:

Sincerely,

Anna V Sosa
Dear colleague,

time has come to report to you on my own and my students’ use of the CHILDES tools since 2003, when I last documented this use. Altogether I have continuously been using CHILDES for almost 20 years in my research on Greek first language acquisition and German and Greek second language acquisition and I cannot imagine what I would have done and would do without this wonderful tool.

As I already mentioned in my 2003 report, in the DaZ-AF project funded by the German Science Foundation at Bonn and directed by myself together with Dr. Christine Dimroth from the Max Planck Institute for Psycholinguistics at Nijmegen, the Netherlands, approximately 120 hrs. of longitudinal observational data on the acquisition of German as a Second Language were transcribed, coded and analyzed using the CHILDES tools. Two doctoral dissertations using these data and analyzing them with the help of CHILDES have been completed in the meantime (Bast 2003, Pagonis 2007) and several others are being prepared. Furthermore, Dr. Dimroth and myself have been and are teaching seminars on the acquisition of German L2 at the Department of Linguistics of the University of Cologne at regular intervals introducing students to the analysis of the DaZ-AF data with the help of the CHILDES tools (e.g. winter semesters of 2006/07 and 2007/08).

CHILDES has also been used and continues to be used by more than twenty international researchers united in the Crosslinguistic Project on Pre- and Protomorphology directed by Wolfgang U. Dressler from the University of Vienna and Austrian Academy of Sciences, who explore the development of early grammar in typologically and genetically diverse languages, such as Turkish, Finnish, Estonian, Croatian, Russian, Greek, (Austrian) German, Italian, Spanish, French, Palestinian Arabic, Yucatec Maya. The latest publications and papers in which I am involved are Christofidou & Stephany (2003), Dressler et al. (2007), Savickienė et al. (2007), Stephany (2002), Stephany (2006), Stephany and Christofidou (to appear), Stephany and Voiekova (2003), Stephany and
Voiekova (eds.) (to appear), Thomadaki and Stephany (2007), and Xanthos et al. (submitted).

Although I have not yet been able to keep my promise of completing the transcriptions of my Greek child data included in the CHILDES Data Base and to connect the transcripts to a sound track, this will be one of my research projects of the near future.

Besides doing research on the development of grammar of Greek L1, Greek L2 and German L2 with the help of the CHILDES tools, I found them to be also excellently suited to the analysis of Greek children’s narrative texts (Stephany, to appear).

For many years, I have been teaching postgraduate seminars on the acquisition of Greek as a second language at the Philosophical Faculty of the University of Athens. For the empirical work my students have to do for their term papers, I introduce them to the use of the CHILDES tools (data transcription in CHAT format, lexicon-based grammatical coding of the learners’ utterances, data analysis with the help of the CLAN programs). I prepared a brief introduction to CHILDES with special attention to Greek (Stephany 2003). CHILDES has also been used and is being used for the empirical work done in my seminars on aphasia (together with Juliane Klann, M. A.).

When I think of my Athenian students and their agony when I leave the country and they are left with the challenge of using the CHILDES tools in their research after a necessarily brief introduction to handling the CLAN programs they need (mainly MOR, FREQ, KWAL, and COMBO), the planned web-distributed tutorials would be a great help. But I think that even for those familiar with the programs they use daily, an introduction to the revised versions of the programs or entirely new ones would be most welcome.

Although I remember having been told by the CHILDES responsible at the Max Planck Institute for Psycholinguistics at Nijmegen, NL, many years ago, that less user-friendliness is the price that has to be paid for the enormous flexibility of CHILDES and the CLAN programs, the creation of a new, user-friendly interface for CLAN will certainly be welcome on the condition that the flexibility of the programs is maintained. Also, it would be most welcome if the new interface were such that ‘old’ users of CHILDES will not have to put too much effort in learning how to handle it. A big incentive for familiarizing oneself with the new interface will, of course, be the prospect of using more powerful search programs that work with it.

Any improvement of CHILDES for doing syntactic analysis is, of course, most welcome. I have not yet used DSS for child Greek, a highly inflected language (at least if compared to English) with ‘free’ word order, so that I cannot say how much preliminary work would be needed for preparing a Greek corpus for (a fully automatic) analysis.

Although extension of automatic morphosyntactic analysis to more languages is certainly an advantage, the possibility of doing lexicon-based automatic coding of speech should by all means be maintained in parallel to automatic rule-based coding. The reason is that
lexicon-based coding allows the researcher to work with personal ways of coding adapted to particular research questions. I found the possibility of establishing a coded lexicon for the analysis of reference in narrative texts very helpful.

The other plans you mention all sound very promising, especially so computer-assisted methods for very rapid first-pass transcription. Will this be limited to English or applicable to other languages as well?

As you might know, work on DGS (Deutsche Gebärdensprache) is being pursued in Germany at the Rheinisch-Westfälische Technische Hochschule Aachen (Prof. Dr. Walter Huber, Neurolinguistics) as well as the University of Cologne among others. I have forwarded your letter to Dr. Claudia Becker (claudia.becker@uni-koeln.de) at the Faculty of Human Sciences of the University of Cologne, who would be interested in the possibility of coding German Sign Language with links to ELAN.

I do hope that the NIH will grant you the possibility of continuing your work on the CHILDES Data Base and the CLAN programs so that your plans from which the scientific community of linguists and psycholinguists in the whole world can profit will materialize.

Thanking you for the work you have invested in establishing and continually improving CHILDES,

with best wishes for the future,

Ursula Stephany,
Professor em. of Linguistics

References


Xanthos, Aris, Laaha, Sabine, Dressler, Wolfgang U., Stephany, Ursula et al. (submitted). On the role of morphological richness in the early development of noun and verb inflection. First Language.
February 4, 2008

Dear Brian,

I am writing to you in support of the CHILDES project. I cannot thank you enough for this valuable resource. I am currently a post-doctoral associate at Rutgers University. From 2001 to 2007, I was a Ph.D. student at Northwestern University. My research focuses on semantics and issues at the syntax-semantics interface in child language acquisition.

I have used CHILDES as a research numerous times in my work, and turn to it routinely to investigate topics and questions that arise in the course of my own research and conversations I have both with other researchers and parents. Below, please find a list of conference presentations and publications that involve the presentation and analysis of data from transcripts in CHILDES.


Syrett, K., & Lidz, J. (2007). Not all adjectives are the same: Children’s (and adults’) knowledge of semantic differences among gradable adjectives. Paper presented as part of the Symposium ‘Constraining the range of adjective meanings and their discourse referents: How young children learn and process adjectives’ (chairs K. Syrett and J. Lidz) at the Biennial Meeting of the Society for Research on Child Development, March 30-April 1, Boston, MA.

Again, thank you for offering and maintaining this unique and valuable resource for the field of child language development. I look forward both to continuing to use it as resource and seeing its development in years to come.

Best regards,

Kristen Syrett
130. Szagun, Gisela

Dear Brian,

in support of your application for the continuation of your research grant from the NIH I would like to give an outline how I and my research group at Oldenburg University have been using CHILDES.

We have collected the largest data base of German child language. This was done in a longitudinal study with samples of normally hearing children and deaf children with cochlear implants. The study was funded for four years, 1996-2000, by Deutsche Forschungsgemeinschaft (German Science Foundation). All 420 speech samples were transcribed with CHILDES. They are in the CHILDES data base as the SZAGUN Corpora. A large number of publications, invited talks and conference publications is based on this data set (see attached list), and work with the data is ongoing.

The data have also been used by other research groups, i.e. at Max-Planck Institute, Leipzig.

In a new research project, again funded by Deutsche Forschungsgemeinschaft, from 2003-2005, a series of grammatical analyses were carried out with the data using the CHILDES system. The special focus was the acquisition of gender marking, the use of the optional infinitive and the acquisition of verb marking.

Since 2005 our research group has been working on the construction of a German parental questionnaire and the collection of normative data using this questionnaire. Again, this study is funded by Deutsche Forschungsgemeinschaft, 2005 ongoing. Item selection for the vocabulary section of the questionnaire is based on the spontaneous speech samples which we transcribed with CHILDES. The questionnaire has an extensive grammar section including a detailed section on inflectional morphology. Item selection is again based on the speech samples and the grammatical analyses with the SZAGUN corpora. CHILDES programmes were used for these analyses.

Our samples cover typical development and language development in children with cochlear implants and represent the largest set of spontaneous speech samples for this latter population. They are also available in the CHILDES database.

Using the CHILDES programmes has made these transcriptions and analyses possible.

A list of publications, research grants, diploma and doctoral theses is attached.

In my view, it is essential that research on the CHILDES programmes continues which enables the programmes to be continually improved and become more useful to researchers studying languages other than English.

Yours sincerely,

Gisela
Publications based on data transcribed and analyzed using CHILDES – in English


Invited Talks


Presentations at conferences


Publications based on data transcribed and analyzed using CHILDES – in German


**Invited Talks**


**Research Grants to Gisela Szagun**

by: Deutsche Forschungsgemeinschaft (German Science Foundation)

1) Language Acquisition in children with normal hearing and children with cochlear implants
Project Nos: Sz 41/5-1, 1996-1998, and Sz 41/5-2, 1998-2000:

2) The acquisition of morphosyntactic schemata
Project No: Sz 41/10, 2003-2005

3) Variability in early language acquisition: a parental questionnaire study
Project Nos: Sz 41/11-1, 2005-2007, and Sz 41/11-2, 2007 ongoing

**Diploma Theses**


**Doctoral Theses**


131. Tang, Gladys

28 Feb 2008

Brian MacWhinney
Director, Child Language Data Exchange System
Professor, Department of Psychology
Carnegie Mellon University

Dear Prof. MacWhinney,

On behalf of the Centre for Sign Linguistics and Deaf Studies (CSLDS), Department of Linguistics and Modern Languages, the Chinese University of Hong Kong, I would like to extend my sincerest gratitude to you for your support in our UGC-funded project “Grammatical Development of Hong Kong Sign Language by deaf Children”. As you know, documenting sign language has been a challenge to any sign language researcher due to features of signing that are modality specific, leading to linguistic consequences that require modifications in current coding and computer programming procedures for purpose of data documentation. As it is the first time at The Chinese University of Hong Kong to conduct a sign language acquisition project, we are constantly bogged down by a lack of technology that is geared towards documenting child sign language. Therefore, your prompt response to our call for assistance in 2007 is deeply appreciated, and we are indebted to your generosity of sharing the technology of CHILDES with us and guiding us through the process of adaptation.

Through the various meetings we had in the summer of 2007 and subsequent email correspondences, we have learned that there are many adaptations that need to be made since we have been using ELAN, a software developed by Max Planck Institute of Psycholinguistics at Nijmegen, to capture the raw data. In fact, this software is designed for capturing spoken language discourse and gesture, but sign language researchers in the field have identified ELAN to be a more readily accessible software for coding sign language discourse. Hence we hope that with your continued guidance and technical support of your team members, we can line-up ELAN and the CLAN program in CHILDES, and to make the raw data in ELAN format compatible to CLAN.

In addition to working towards format compatibility, we hope to devise a new transcription system for sign language data, both child and adult, which is so far few and far between. We believe that CHILDES provides the best platform available to date for its goal of developing technology to accommodate child and adult’s production. We are also designing an auto-tagging system for coding the grammatical categories of sign language with links to ELAN as well as that of CLAN. We understand that there is still a long way to go as it is the first time we try to link up signed and spoken language in data documentation. It is our goal that our data will contribute to the field of child language and facilitate especially cross-linguistic and cross-modal comparison. In this regard, I
would like to inform you that the Centre is preparing for another child sign language project – bilingual acquisition of Hong Kong Sign Language and Cantonese by deaf children. It is going to be another 7-8 years of hard work. However, with your guidance, we are prepared to accept the challenge ahead and we hope that ultimately our data, the software we design, and the sign language transcription system are made available to the child language researchers in the field.

Yours faithfully,

Gladys Tang
Director, Centre for Sign Linguistics and Deaf Studies
Chairperson, Department of Linguistics and Modern Languages
The Chinese University of Hong Kong
Dear Brian,

Thank you for the opportunity to write in support of your CHILDES project. As you know, I have been using CHILDES since my dissertation research in 1993 and have contributed 3 separate corpora (one still "in progress") on Mandarin-speaking children and their caregivers to the database. I have published several papers based on the two early corpora, as listed below, and am currently working on analyses of both lexical and syntactic diversity on the "in progress" corpus of 300 Mandarin-speaking children who were recorded when they were 3-4 and 4-5 years of age narrating a version of the "Frog Story" picture book. For this latter corpus, I have worked intensely with graduate students and yourself at further developing the Mandarin lexicon available on CHILDES and at developing automated MOR coding for these files. We are still in the process of calibrating the MOR coding with the human coder(s) and will be able to send you an update of these analyses soon. Overall, though, it looks like the automatically coded data is correctly coding over 90% of all word types and this is remarkable given that we are dealing with child speech. With updates and further corrections to the lexicon, I fully expect this percentage to be considerably higher. I also look forward to writing up this process together with you for publication on how automated coding can work for Mandarin Chinese.

My own publications that have used the CHILDES Mandarin data (note that underlined authors are students who have been trained in the use of CHILDES and these corpora as part of these projects):

Tardif, T. (in progress). *Different words for different worlds: Early language development in Mandarin and Cantonese.* Accepted for publication as Essay in Developmental Psychology, Psychology Press.


verbs. Oxford University Press.


CONFERENCE PRESENTATIONS:


Tardif, T., & Chan, C. (March, 2005). *Word learning in Mandarin: Same or different?* Invited presentation to Department of Psychology, University of British Columbia. Vancouver, BC.


In addition to the above projects, I also use CHILDES in my classes on Language Acquisition and generally require students to choose a corpus and do intensive analysis of a single phenomenon in English or any other language that they are competent. I have consistently received positive feedback on the learning experiences that students have taken away from this assignment.

Finally, as you know, I have also submitted a recent grant proposal to collect new longitudinal data on Mandarin and English-speaking caregivers and their infants in both controlled laboratory and naturalistic settings from 6 to 24 months of age. If awarded, this project would also make extensive use of both the "standard" CHILDES features as well as its new Mandarin lexicon, parsing rules, and video synching capabilities.

Thus, I very much hope the CHILDES project will continue to receive strong support from NIH and that it will continue to develop as a wonderful tool for language acquisition research.

Sincerely,

Twila Tardif
Dear Dr. MacWhinney,

I recently had a paper come out for which we used CHILDES transcripts. Here is the citation:


Thank you for your work on the database!

Medha Tare
Doctoral Student
University of Michigan
Dear Brian,

I write in support of your application for further funding to support the CHILDES system. Over the past 5 years I and my students have made extensive use of the transcription system and analysis programs. For example, from 2005-2007 I held a grant from the British Economic and Social Research Council (ESRC) along with Dr. Caroline Rowland at the University of Liverpool, UK to fund the collection of a longitudinal data set from British-English speaking children aged 2;10-3;6. These data were transcribed using the CHILDES system and will in the longer term be donated to the CHILDES database. Analysis is ongoing and relies extensively on use of the CLAN programs. In addition students working on a range of different projects in child language regularly use the CLAN software to conduct searches of naturalistic data sets to establish patterns of use, and input frequencies. For example, we are currently conducting a large study into children's use of past tense morphology that relies on CLAN-based analyses of CHILDES corpora for estimates of the distribution of forms in child-directed speech. Another project investigating the acquisition of questions relies on analyses of CDS to establish high and low frequency lexical combinations available to children, while yet another project investigating the acquisition of case marking uses distributional information from CDS to select experimental stimuli. Finally, one of our PhD students has just completed her thesis on the acquisition of complement clauses. This work was exclusively based on naturalistic data, some of which was from the CHILDES database, and all of which was analysed using the CLAN programs. This is merely a subset of the work that is previously ongoing which relies on CHILDES.

In my opinion continued support for the CHILDES system is absolutely essential to support child language research across the globe.

Best wishes,

Anna Theakston
University of Manchester, UK
135. Thomadaki, Evangelia

Dear Professor MacWhinney,

As this is the first time that I have had the opportunity to contact you, I feel obliged to express my sincere admiration as well as my gratitude as a linguist for the work you and your colleagues have accomplished all the previous years concerning this very ‘important component of the scientific infrastructure’ for the study of language. Anyone who has benefited by using these programs feels that they certainly deserve further financial support.

I got acquainted with CHILDES and CLAN programs a few years ago, thanks to Prof. Dr. U. Stephany (who was also my “Magistermutter”), but I had not really had any real working experience with them until 2004/5, when we (i.e. Prof. Stephany and myself) engaged in a common research project concerning the development of diminutives (s. publications below). I used some of the search commands then to search through Clan files, made accessible to us by Prof. Dimitra Katis. I also have my own taped material from my two daughters, which unfortunately I had not been able to transcribe and codify yet, but hope to be able to do so in the near future. My very limited expertise with these programs has nevertheless taught me to appreciate the substantial help that a researcher can get from these programs. This is apparent even to someone like me who has used just a fraction of their potential.

Therefore I am more than glad to offer my support for the continuation of these challenging, steadily evolving and updated, indispensable and enjoyable programs. Please, keep up the good work.

Best wishes

Dr. Evangelia Thomadaki, Assistant Professor of Linguistics
Democritus University of Thrace
Department of Language, Literature and Culture of the Black Sea Countries
ethomada@bscc.duth.gr
ninthom@otenet.gr

Publications making use of the CHILDES/CLAN
136. Tonelli, Livia

From: "tonelli.fabris" <tonelli.fabris@libero.it>
Date: February 15, 2008 10:58:09 PM GMT+01:00
To: "macw" <macw@cmu.edu>
Cc: "tonelli.fabris" <tonelli.fabris@libero.it>
Subject: CHILDES usage

Dear Brian,
In the last five years we have been cooperating in the completion of ITAL MOR. Furthermore, we have been using CHILDES in our university courses on second language acquisition: by the use of CHILDES we could teach students the bases to automatically analyze spoken language, code errors and realize statistics. We have also presented the philosophy, functioning and scope of CHILDES during talks given at University of Genova and Sassari.
At the moment we are using ITAL MOR for two papers, one on the acquisition of diminutive in different languages and the other on the acquisition of clitics in German and Italian.

Livia Tonelli and Maurizio Fabris (University of Genova)
Dear Prof. Brian MacWhinney,

I am a Ph.D. student at the University of Hawaii currently working a dissertation project on the Vietnamese child language. I started using CHILDES about four years ago. In the last one and a half years, I have been using CHILDES on a daily basis for transcribing, coding, and running analyses for Vietnamese, a tonal language. With the help of the Unicode system provided in CHILDES, I was able to transcribe Vietnamese including all of its tonal diacritics. I did not have to change the tones to numbers, which makes the transcription much easier to read and analyze. With help of the automatic MOR program that exist for Cantonese and Mandarin in the CHILDES database, I was able to extend the MOR program to Vietnamese and could run automatic morphosyntactic analyses on Vietnamese successfully. My dissertation, future conference presentations, and papers will be based on transcripts and analyses for Vietnamese that I would have not been able make without the existence of CHILDES.

Thanks so much for creating, maintaining and improving CHILDES! Since there is no Vietnamese child data in the CHILDES yet, I will contribute all of my longitudinal Vietnamese data collected over a period of one year in Vietnam to CHILDES in the very near future. The availability of the Vietnamese data in CHILDES data corpus will then be easily accessed, shared, and disseminated among language acquisition researchers who might wish to conduct research on Vietnamese child language or to use the data for cross-linguistic comparisons.

Jennie Tran
Ph.D. Candidate
Department of Linguistics
University of Hawaii
Dear Dr. MacWhinney,

In the past six years, I have used CHILDES for two research projects I carried out while being a graduate student at the University of Hawai‘i at Manoa. The first project was on the production of determiners and determiner fillers by a French Canadian child from the York Corpus. The second project, which I conducted with Dr. Katherine Demuth (Brown University), was on the production of French determiners by two French children from the Lyon Corpus. Below is a list of conference presentations and publications that came out of this work. I feel that the experience I gained from working on these projects, which I could not have done without the database, was extremely beneficial and likely contributed to my obtaining a tenure-track Assistant Professor position straight out of graduate school.

I sincerely hope that the CHILDES project will continue to receive financial support.

Best regards,

Annie

Conference Presentations:


Publications:


Annie Tremblay,
Assistant Professor
Department of French
University of Illinois at Urbana-Champaign
2090 FLB, 707 S. Mathews Ave.
Urbana, IL 61801
Phone: (217) 265-0764
Fax: (217) 244-2223
Dear Prof. MacWhinney,

As a user and a beneficiary of the Child Language Data Exchange System, I would like to take this opportunity to report to you about how we have been using the system and our appreciation and expectations for the continuous development of the system.

My research in the past ten years has been heavily related to and replying on CHILDES. Using the platform of CHILDES, I was able to construct a child language corpus of Taiwanese, under the support the National Science Council, Taiwan. The Taiwanese Child Corpus contains longitudinal data of children acquiring Taiwanese (Southern Min Chinese spoken in Taiwan) as their first language. There are about 300 hours of recording of 14 children when they were between two and four years old. The size of the corpus is about 2 million words. Since CHILDES has been the most useful and efficient system for child language data exchange in the world, we have been strictly following its format. This has allowed us to take advantage of the great power of the user-friendly CHAT editor and CLAN checking/searching software, an, more recently, the PHON project. This platform has made it so much easier for researchers to collect, store, and analyze their own data, as well as to share and exchange with other researchers in the world. We are currently in the final stage of checking the morpho-syntactic annotation of our corpus and will donate this corpus to CHILDES after its completion.

My colleagues, research assistants, and students have witnessed a lot of progress of CHILDES in the past years. We hope that this system will continue to develop. In addition to the general development of all aspects of this system, in particular, we would like: (1) to see the extension of automatic morphosyntactic analysis to more languages. We are interested in constructing a grammar of Taiwanese that will facilitate automatic morphological analysis. Since Taiwanese is an under-studied non-Indo-European language, this grammar will also contribute to our understanding of linguistic universals. (2) The representation of phonology is another aspect that we hope the system will improve in the future, especially in respect to lexical tone. (3) We also are interested in the development of a fuller support for coding of sign language, since sign language is another highlighted research area of our Institute. I have been in charge of the Taiwan Sign Language Browser (on-line dictionary). My colleagues and I also have been working on the writing of a Taiwan Sign Language reference grammar. We hope that we
can share with more people our collection of Taiwan Sign Language data with the help of an extended support for video-based transcription.

Obviously, CHILDES has been an indispensable part of child language research. In appendix is a list of research grants and publications/presentations that have used either data or programs from CHILDES (including data from Taiwanese Child Language Corpus which is based on CHILDES) by my students, collaborators, and myself. We would not have been able to do any of these without CHILDES.

Finally, I would also like to take this opportunity to thank you and your research team for your extremely efficient help over the years. It is unbelievable how quickly my questions were answered and the problems were taken care of. We really appreciate it very much.

With best regards,

Jane S. Tsay, Ph.D.
Professor
Graduate Institute of Linguistics
National Chung Cheng University
Min-Hsiung, Chia-Yi 62102
Taiwan, R.O.C.
Phone: 886-5-2428239
Fax: 886-5-2721654
E-mail: Lngtsay@ccu.edu.tw
Appendix

RESEARCH GRANTS

Tsay, Jane S. (PI) 2006-2009
"A Corpus-based Study of the Acquisition of Prosody in Taiwan Southern Min," National Science Council, Taiwan (NSC 95-2411-H-194-022-MY3).

Tsay, Jane S. (PI) 2003-2006

Tsay, Jane S. (PI) 2000-2003
"Taiwan Child Language Corpus," National Science Council, Taiwan (NSC 89-2411-H-194-067, NSC 90-2411-H-194-031, NSC 91-2411-H-194-001)

Tsay, Jane S. (PI) 1997-2000

PUBLICATIONS/PRESENTATIONS


**MA THESES:**


140. Tsou, Chi-Zong

Dear Dr. MacWhinney,

I am an assistant professor and a speech language pathologist in Taiwan, and my research area is mainly about language communication problems of children with special needs.

For the past two years, I used CHILDES a lot and have published a study in an academic journal in Taiwan.


The same research had been presented in the following conference:


As for the features I hoped to be enhanced for the CHILDS includes:

1. After coding, could we search for the coded texts and pulled the coded texts together for further analysis or examination? For example, if I coded several files for relative clauses or NP with modifiers, could I pulled all the coded texts together(including the original filename and lines) for other uses?

Thank you so much for putting a lot of efforts for developing, maintaining and enhancing CHILDES.

Sincerely,

Chi-Zong Tsou
Assistant professor, Teacher Education Center
Taipei Physical Education College
141. Ullman, Michael

From: Michael Ullman <michael@georgetown.edu>
Date: February 4, 2008 8:49:29 PM GMT+01:00
To: macw@cmu.edu
Subject: Re: CHILDES renewal support letter

Dear Brian,

Thank you for all your work on CHILDES. It has proven to be very valuable for our research. In particular, we published a paper in 2006 which crucially relied on CHILDES data:


I give my strongest support to you and CHILDES, and to its continued funding.

All my best,

Michael
Dear Brian,

I wish to start my letter by thanking you for your devotion and life long commitment to the enormous and so very important project called CHILDES. I wish to express my appreciation for your timeless efforts to maintain, to improve and to get financial support for CHILDES. Without CHILDES, research in child language acquisition would have surely lagged behind other scientific disciplines and would have not reached the high scientific standards, achievements and recognition which the field had gained to date.

I have been studying the acquisition of Hebrew child language for almost 12 years, using longitudinal naturalistic speech samples of Hebrew-speaking children that were transcribed, coded and analyzed using CHILDES (Berman and Weissenborn 1991, http://childes.psy.cmu.edu/manuals/12other.doc#_Ref503958473 ). I based my doctoral dissertation on these data as well as a number of papers and book chapters which I cite below.

For the past four years, Dr. Rachel Yifat from the Dept. of Communication Disorders at Haifa University and I have been collecting data for two new corpora: (1) a database of clinical interactions between linguistically-impaired Hebrew-speaking children and their speech therapists, and (2) a corpus of speech samples of Arabic-speaking children in naturalistic and clinical interactions. These databases, too, were transcribed in CHAT format, and will be coded and analyzed using CHILDES.

I consider the CHILDES system essential for the study of child language for the following reasons: (1) it is a computer-based system, which is a “must” at our technological era if we are to handle large quantities of data; (2) the system’s standardized transcription format facilitates data sharing, and makes cross-linguistic databases comparable; (3) the system offers a wide variety of accountable analytical tools designed specifically for the study of child language; (4) it provides free access to a large variety of databases allowing researchers who cannot afford to collect their own databases the opportunity to do research (5) its additional resources such as the CHILDES bibliography, and in particular, the INFO-CHILDES list are extremely useful for obtaining information, and providing a forum for discussion and exchange of ideas.

The advantages of the CHILDES system for the study of child language have led me to get actively involved in the adaptation of CHILDES to Semitic languages: (1) I was involved in developing CHILDES compatible transcription conventions for Arabic (http://semtalk.talkbank.org/arabictranscription.htm); (2) I have developed in collaboration with Bracha Nir-Sagiv (Tel Aviv University) the Hebrew MOR grammar, which is an automatic morphological analyzer for the Hebrew corpora currently available on the CHILDES website (http://childes.psy.cmu.edu/morgrams/). (3) I built and I am currently running a website that provides information and support for the application of CHILDES to Semitic languages (http://semtalk.talkbank.org ).

The importance of CHILDES as a methodological tool for training future child language researchers has also been recognized by Israeli universities and linguistic
associations. Evidence of that is a one-semester CHILDES course that I have been teaching for the past four years to first year students at the Dept. of Communication Disorders at Haifa University, and a growing number of requests for CHILDES workshops that I have been invited to conduct at local conferences (see detailed list below).

Despite the many advantages of the CHILDES system, some things still need to be improved or developed: (1) a more user friendly interface that make it easier to cope with the huge number of symbols and conventional notations currently in use; (2) a web-based interface for transcription, coding and data-analysis; (3) more analytical tools for the analysis of clinical interactions.

In conclusion, I find the CHILDES system extremely useful, and I see great importance in its continued development and improvement for the benefit of my own research, as for the benefit of the language researchers’ community at large. I believe this enterprise to be of great importance for the study of language, and particularly, for rendering the field of language research more data-oriented, with a sound empirical basis.

Sincerely,

Sigal Uziel-Karl

Sigal Uziel-Karl.

Dept. of Communication Disorders
Haifa University
Israel
sigal@alum.mit.edu

Talks and workshops

July 2004:
“The CHILDES System”
Workshop, 19th Script Conference, Zixron Ya’akov, Israel (with Bracha Nir-Sagiv)

July 2005:
“The CHILDES System – Innovations and Application for the study of Hebrew and Arabic”
Workshop, 20th Script Conference, Nir Etsyon, Israel (with Bracha Nir-Sagive and Dr. Rachel Yifat)

October 2005:
“The CHILDES System – Application for the study of child Hebrew”
32nd Conference of the Israel Association for Applied Linguistics, Jerusalem, Israel

March 2006:
The CHILDES system – an overview
Lecture at the School of Communication Disorders, Tel Aviv University

November 2007:
The CHILDES system – applications for the study of speech and language disorders
Lecture at the School of Communication Disorders, Tel Aviv University

**January 2008:**
The CHILDES system – applications for the study of speech and language disorders and for clinical intervention
Workshop, the Annual Conference of the Israeli Speech, Hearing and Language Association (ISHLA), Nazaret, Israel (with Dr. Rachel Yifat)

**Publications**

**Doctoral Dissertation**

**Refereed Journals**

**Non-refereed Journals**

**Book Chapters**

**Student CHILDES based seminar papers**
http://semtalk.talkbank.org/paperabstracts.htm
Dear Professor MacWhinney,

I have been using the Child Language Data Exchange System (CHILDES) as a member of the research group *Emergent linguistic productivity*, directed by Professor Emeritus Leonor Sciliar-Cabral, and sponsored by the Brazilian National Council of research (CNPq).

My work consisted of feeding the CHILDES database with the corpora of Sciliar-Cabral’s doctoral thesis, discussing at the regular seminars of the research group theoretical and methodological issues, which are crucial for the field of language acquisition and coding the morphological categories of Paulo’s utterances (stages one and two). Much of my work was dedicated to compare Spanish acquisition with Brazilian one.

Since 2005 I presented several papers on the subject, as you can follow at the list bellow.

The new proposal to be presented for a renewal of NIH funding is of particular interest for my further research on comparing Spanish acquisition with Brazilian one.

Santo Gabriel Vaccaro
Universidade Federal de Santa Catarina
Florianópolis, SC, Brazil
santogvaccaro@hotmail.com


Dear Brian,

Apologies-- Just noticed I hadn't done this:
However, I would like to add my support to this extremely valuable international resource that benefits research and it clinical and educational application throughout the world. It would be a complete travesty to the whole international community if support for CHILDES was not continued. More than ever now--- with fast access to data and the field moving so much so quickly-- we need this.

I do hope that this valuable resource has continued funding and is able to develop further

Kind regards
Heather

Heather K.J. van der Lely  
Professor & Director  
Centre for Developmental Language Disorders & Cognitive Neuroscience,  
University College London  
4th Floor  
123-128 Grey's Inn Road,  
London  
WC1X 8WD  

www.ucl.ac.uk/dldcn  
e-mail: h.vanderlely@ucl.ac.uk  
tel: (+44) (0)20 7905 1292  
Administrator: (+44) 020 7905 1212
Dear Brian,

This letter is meant to support very strongly the renewal of your NIH grant dedicated to the CHILDES project. As a member of the child language community since the late seventies, I have been able to watch the extraordinary improvements and great advances the CHILDES project has brought to our scientific community since its constitution and with exponentially growing efficacy. It has offered new and innovative research tools, encouraged homogeneity of methods and thus a greater possibility to confront results across different research projects. Thanks to the international respect and approval that the CHILDES project has acquired over the years, it has helped to set the discipline onto firmer and more widely recognized scientific grounds. The creation of the info-childes bulletin, that is alive every day of the week, has added opportunities to link people and ideas, and enhanced professional identity.

In my own work I have benefited from the tools developed by the CHILDES project and from the CHAT method of transcription. As many people in the field who have started working on child language in the early seventies, I have gone through several ways of transcribing the audio-video material, with sometimes a loss of research time. Thanks to the CHILDES tools the transcription work has become more efficient, transparent and sharable and has thus enhanced the opportunities for collaborative work. Moreover, the CLAN tools have allowed analyses that would have been extremely time-consuming otherwise. All the grants I have been able to obtain for the study of child language, both in Switzerland and in France (recently for the project EMERGRAM on the Emergence of Grammaticality funded by the French National Agency for Research), have been grounded on the methods of the CHILDES project. The research team of the EMERGRAM Project, in France, Spain, Mexico and the
United States, (http://www.modyco.fr/index.php?groupesrecherche:projet%20emergram ) makes extensive use of the CHILDES programs and use the data bank to check on more specific hypotheses. Needless to say, the CHAT format is the format that I am now using both to revise past transcriptions and to transcribe new data. This is done not only for early language acquisition data gathered at home, but also for other natural interaction data (for example, home interaction between autistic children, familiar adults, siblings and peers) and we are starting to use it also for a large corpus of narrative data in children between 4 and 11 years, presenting normal or atypical developments (SLI and autistic children).

Being flexible, the CHAT system can be adapted to the special needs of each research. For example, for early language, and in order to avoid freezing the data, the phonetic transcription is placed in the main line while the possible interpretation(s) figure in the dependent line. This doesn't seem to be a main problem to accomplish further analyses. For coding, in CHAT practically everything is possible. At present, for French-acquiring children, we are transforming past work on "fillers" and on early "verb morphology" into a CHAT coding system particularly adapted to transitional periods. If proven useful, and if the need will be felt, thanks to CHILDES these tools can be made readily available to other researchers.

I also use the CHILDES data for pedagogical purposes. For their research papers, students have a precious opportunity to extend the small data set they can collect themselves during the semester to test their hypotheses. The CHILDES project becomes irreplaceable when cross-linguistic hypotheses need to be tested since it allows an easy access to acquisition data in different languages, stored in standard ways, and ready for analysis. Moreover, the fact that CHAT can be neatly put into a small set of simple rules, that it has developed video linking techniques, and the fact that rules and techniques are internationally-accepted and can be generalized to other contexts of work, is definitely a motivating thrust for students to engage themselves in the hard work of transcription and analysis of early child language data.

Since I strongly believe in the great services that the CHILDES project offers to our scientific community, the EMERGRAM project intends to bring a number of the many longitudinal studies collected during the years, and the ongoing studies, to the standards set by CHILDES and to contribute data of children acquiring French in their natural environment to the CHILDES database (transcriptions and video linking for those studies we have permission).

A sample of publications, research grants and student work is attached below. As for some specific requests for the next period, the goals specified in the NIH grant are all very useful to pursue. I would like to add some further ones:

(i) I would encourage any effort towards improving CLAN, making it as close as possible to an usual text-editor. I would like to be able to apply CLAN programs directly to WORD files and to have the CLAN commands as easy as
possible to utilize.

(ii) Since many of the transcriptions have been gone through different transcribers (sometimes over the years and the needs of the specific research projects the data were used), it would ne helpful to have a standardized way to trace the different transcriptions (with identity of the transcriber, the data of the transcription, etc.) and signal the one that has gotten the highest consensus.

(iii) It would also help to have an output as friendly as possible (e.g., a clearly readable Excel worksheet on which other operations proper to Excel can be executed).

(iv) Also, make the video-linking as child-friendly as possible, both when placing the bullets and when going through the transcripts. For example, it would be already very helpful to have a program that places bullets at some defined places of the sound stream to have a first approximation of turn-taking boundaries.

(v) Although this seems more of a dream, it would be a great revolution to have at one's disposal tools that would facilitate transcription: for example, a program that could approximate a first transcription ready to be checked by a human transcriber.

(vi) As the demands above tend to indicate, even with grants available, there are transcription problems. Transcribers are not easy to find and to train and, even more crucial, they tend to last a short time. I would welcome a common approach to these common problems. Since a fundamental link between data collection and data contribution to CHILDES resides in the transcription efforts, it would be important to find in the CHILDES project the logistic support to find the most appropriate solutions. This could range from developing a short multimedia training course on miniCHAT to collectively discuss initiatives for fund raising in this badly needed sector.

For the reasons exposed, I cannot but thank you and your collaborators for initiating and developing the CHILDES project and hope that it will continue to be funded allowing further improvements and the continuation of the many services that you and the CHILDES project provide to the child language scientific community as a whole.

Université Paris Descartes-
CNRS (UMR 7114 and
UMR 7023)

Hereafter some work and grants that have greatly benefited from the existence of the CHILDES project.

2004-2009
1991-1995
Fonds national suisse de la recherche scientifique (FNRS) "Récits, projets et explications : leur genèse et leurs relations au début de l'acquisition du langage"
(subsides n.11-30972.91 et n. 11-37304.93).

Publications
Veneziano, E. (1999). Early lexical, morphological and syntactic development in French :
Some complex relations. International Journal of Bilingualism, 3 (2), 183-
Communications


Veneziano, E (10-12 septembre 2007). Processus inter et intra dans l'acquisition et le fonctionnement du langage. *3e colloque "Constructivisme et éducation"*, Geneve, Suisse


Veneziano, E. (août 2005). Comportements langagiers dans l'interaction mère-enfant avant 3 ans et l'émergence d'une théorie de l'esprit implicite. GPPL, Department of Linguistics, University of Sao Paulo (USP), Sao Paulo, Brésil.


Kwiotek, Tomasz *Les PAEs et l'émergence de la morphologie verbale*, Université de Nancy 2, 2001
Ris, Claire et Siegler, Rachelle *Relation entre l'émergence des PAEs et celle de la catégorie verbale* (5 volumes). Université de Nancy 2, 2001

Ruppel, Aline
*Emergence des éléments additionnels préfixés et morphèmes grammaticaux bien formés.*
Université de Nancy 2, 2002

Jeandel, Cédric
*L'apparition des marques du futur chez l'enfant de 25 à 35 mois,* Université de Nancy 2, 2002

Gauthier, Isabelle - *Persuader avant trois ans : Étude longitudinale des moyens employés par la dyade mère-enfant et de leurs effets en situation communicative de requête, Thèse de doctorat,* Université de Nancy 2, 2001

Catherine-Lebon Florence
*L'émergence de la référence au passé chez l'enfant en interaction avec l'adulte,* Université de Nancy 2, 2003

E. Herr-Israel, Rutgers University (co-direction, L. McCune Nicolich) : *“The development of dynamic event word use through analysis of expressive options in the transition from single to combinatory speech”*. 2006.

1 March 2008

Dr. Joshua Viau
Cognitive Science
Johns Hopkins University
237 Krieger Hall
3400 N. Charles St.
Baltimore, MD 21218

Dr. Brian MacWhinney
Psychology Department
Carnegie Mellon University

Dr. MacWhinney,

I am writing this letter in strong support of the CHILDES project and your upcoming grant renewal application. I find it hard to overstate how important this project has been to my own research and to the community of language development researchers both here in the U.S. and abroad.

Below (on p. 2) you will find a list of my own publications and presentations that have depended heavily on and benefited greatly from findings based on CHILDES data. I sincerely hope that this small contribution, when combined with those of other researchers, will have some positive impact on the future of the project.

Please let me know if there is anything more I can do in this regard. In the meantime, I would like to quote the last paragraph of my recent dissertation’s acknowledgments section, where I attempt to put into words how grateful I am to have a resource like CHILDES at my disposal:

“Finally, I acknowledge a great debt to all of the parents who allowed their children to be recorded for the CHILDES database or to humor me personally in one study or another. It’s humbling to reflect on the degree to which science—and so much else in life—depends on the kindness of strangers. All one can do is pay it forward.”

In retrospect, I should have gone further, thanking all of the countless students, staff members, and administrators associated with CHILDES—with you foremost among them, I imagine—who have labored behind the scenes for years in order to continually improve the project and thereby facilitate research. So, belatedly, thanks, and I wish you much continued success. As my wife and I have a baby boy on the way, you may well be
hearing much more from me (and him) in the near future. We intend to document his language development, of course!

Best wishes,

Dr. Joshua Viau
viau@cogsci.jhu.edu

Publications relying on CHILDES data:


Conference presentations and invited talks relying on CHILDES data:


Dear Professor MacWhinney, dear colleague!

We are happy to know that the CHILDES Project will be continued. For us this is an excellent opportunity to conduct our research in L1 and L2 acquisition of Russian and other languages of Russian Federation at the corresponding technical level. I may already have written to you that Russian data that we are using are collected mostly not in the Academy of Sciences where I work but at the Chair of Child Language in Herzen State Pedagogical University of Russia. Spontaneous data are coded in CHAT, whereas diary data may be in different formats. To make the morphological line, we use a specially designed (both rule- and lexicon-based) morphological program (MORCOMM) for Russian. This databank contains data from about 40 children.

For us, CHILDES became a principal format for the acquisition data. In the last three years I gave two courses on L1 Acquisition at the Philological Faculty of the State University in Saint Petersburg: “Early stages in the L1 acquisition of Russian morphology” for MA students, “Acquisition of Russian as a native language” for the BS students from other cities of Russia learning by correspondence. CHILDES as a system and its capacities were especially introduced in both courses. Lectures on “Language analysis with the help of CHILDES” were delivered by me for Russian university fellows coming to Saint Petersburg for their subbatical from province.

My following publications in the last three years would never be possible without CHILDES:


All the best for the future development of CHILDES

Doc. Dr. Maria Dmitrievna Voeikova
Institute of Linguistic Studies, Russian Academy of Sciences
Saint Petersburg Russia
Dear Dr. MacWhinney:

I am writing to strongly support your renewal application for a NIH grant to continue development of the CHILDES database. I am a professor in speech pathology, and am finding the comprehensiveness of the database able to permit needed research in our field that has otherwise eluded us because compilation of a data set of sufficient size has not been logistically reasonable for most researchers. I have several students who are conducting pilots as part of nonthesis research projects for a master's degree. They aim to determine the frequency with which various grammatical environments prompt a child to commit certain errors in language production at various ages. The purpose is to create a teaching hierarchy of grammatical error "triggers" that prompt commission of errors of copula verbs because the environment presumably overloads the child's cognitive system. One student project has been completed, titled "Predicting morpheme errors in young children: factors related to grammatical environment" authored by Sharon Nichols. A second study is underway titled "Effects of grammatical environment on errors of inflection in children with specific language impairment" by Jenna Umeno. It is our anticipation to do a third paper for publication in which the environments are compared between normal speakers and those with specific language impairment. Remedially, the research is important since it is customary in speech therapy to train a child along a hierarchy of elicitation contexts that progress from easier to more difficult, and the grammatical environment is one of the factors that contribute to that hierarchy. Because the amount of data needed for such an analysis is substantial (to generate sufficient frequency scores for a variety of contexts at a number of different age and MLU levels), the CHILDES database has been extraordinarily useful. I sincerely hope that you will be able to continue development of the database, especially with the addition of clinical corpora.

Randy Weirather
Assistant Professor
Speech Pathology and Audiology
University of Hawaii
1410 Lower Campus Road
Honolulu, Hawaii 97822
Dear Brian,

My research team has utilized the CHILDES archives and the CLAN programs extensively in recent years. The research team includes my Polish colleague, Aleksandra Pawlak at Adam Mickiewicz University in Poznan, and numerous undergraduate students at SUNY Fredonia. We have used the system to obtain new data and to analyze existing data.

1) Obtaining new data.

The following describes the new data from the Fredonia project. When the initial investigation of this data is completed in 2008, the data will be submitted to the CHILDES archives.

The goal of this longitudinal project was to obtain caregiver-child interaction data from children aged 2 to 5, in order to capture the children’s language in an early phase of the acquisition of their event time (ET) system and continue observations through the emergence of their reference time (RT) system. Whenever possible, the children were audio tape recorded in either a laboratory setting or in their homes twice a month for approximately 30 minutes. The audio tapes were transcribed into the CHAT format (MacWhinney, 2000), and then the transcriptions were completely checked for accuracy. The children are as follows together with their starting ages, ending ages, and number (n) of transcripts: Ben (2;4 – 3;3 / 11), Emily (2;6 – 4;5 n = 23), Emma (2;7 – 4;7 n = 28), Jillian (2;1 – 2;10 n = 22), Matty (2;3 – 5;0 n = 56), and Roman (2;2 – 4;7 n = 42).

2) Analyzing existing data.

The following lists the recent publications that utilized the CHILDES archives extensively. The first three studies are cross-linguistic studies involving Polish and English.


Sincerely Yours,

Richard M. Weist, Distinguished Professor, SUNY Fredonia
150. Wellman, Henry
From: "Henry M. Wellman" <hmw@umich.edu>
Date: February 19, 2008 4:10:16 PM GMT+01:00
To: Brian MacWhinney <macw@cmu.edu>
Subject: Re: CHILDES usage

Dear Brian--

We do continue to use the CHILDES. Most recently we conducted an analysis of young children's causal question-asking (via questions such as 'why?' and 'how come?'). We examined (a) children's questions, (b) parents' responses to those questions, and (c) children's reaction to and satisfaction/dissatisfaction with parents' responses. The analyses (including 1000s of exchanges) would have been impossible without the CHILDES resources, or would have been completely nondefinitive because undertaken on utterances for only a single parent-child dyad with only a tiny sample of exchanges. The results were very compelling and provide the first clear and comprehensive evidence of how children react to and learn from their causal questioning. The results also provided the inspiration for developing a laboratory task that has allowed us to look at these explanatory-learning transactions still further. An initial ms. including the CHILDES study and the follow-up lab study has been submitted to Child Development (Frazier, Gelman & Wellman) and the data have already been presented at SRCD and at CDS.

Over the years, I and my students have used the CHILDES consistently and to great avail--some of our studies using this database are very highly cited. The CHILDES is an important (irreplaceable) resource for researchers conducting research into language and/or cognitive development. It has transformed the field (from a typical study with one or two children, to a typical study with much more generality across children, languages, development). It has served as a model for database sharing in other domains. More selfishly, it has allowed me and my collaborators to conduct research that would be impossible otherwise.

I fully support its continuation and continued support.

Sincerely, Henry Wellman
Professor of Psychology
University of Michigan
Dear Brian,
I have used the Thai database available through CHILDES in the last 6 years. The collection and transcription of this database was organised by Jordan Zlatev and Peerapat Yangklang at the Center for Research in Speech and Language Processing (CRSLP) at Chulalongkorn University, in conjunction with the Director of the research centre, Dr Sudaporn Luksaneeyanawin. This has resulted and contributed to the following publications and conference presentations:

**Publications**

**Conference Presentations**

Thanks very much for organising this invaluable database and the user-friendly interface and tools to access and analyse the data,

Best wishes,
Heather
February 21, 2008

To whom this may concern,

I understand that Professor Brian MacWhinney is applying for renewal of his NICHD grant for CHILDES. I am writing in support of his application.

My area of expertise is language learning and development in young Chinese-speaking children, especially those with developmental language disorders. In the past four years, I have been referring to archive data from CHILDES, the HKU and LeeWongLeung data files in particular, for teaching purposes. For research, I have used programs from CHILDES for management and analysis of language samples collected for typically-developing Cantonese-speaking children and Cantonese-speaking children with specific language impairment. I currently have two research projects involving the analysis of two sets of language samples with tools from CHILDES. The set that was collected from 100 typically-developing Cantonese-speaking children is being processed for eventual upload to the CHILDES website. The other set was collected for a study to validate how language sample measures can differentiate Cantonese-speaking children with and without specific language impairment.

Here are some of my research outputs using archive data and CHAT programs in CHILDES

BSc Honours Dissertation Supervision


Peer-reviewed presentations


Peer-reviewed publications


There is still a lot to be done with the CHILDES. For research in Cantonese, we need more powerful programs for accurate coding of the word categories from character scripts. We also need more user-friendly programs that would allow us to conduct prepositional and conversational analyses, to name a few. I look forward to learning more about language learning and development with other researchers in the CHILDES community.

Anita Wong, Ph.D.
Assistant Professor
amywong@hkusua.hku.hk
To whom it may concern:

I am Dr Wong Kwok Shing, Richard, Deputy Head of the Centre for Early Childhood Research and Development at the Hong Kong Institute of Education (HKIEd). I am writing to point out how much the CHILDES database has revolutionized the way teaching and research conducted in the East Asian context, and I sincerely hope that there will be continual financial support for this fundamental database so that it can continue to expand and include more languages of the world.

Being in the field of early childhood education, I had often noticed that teaching practice might bear a weak link to developmental norms. Prior to the availability of the two Cantonese corpora – CANCORP and the Hong Kong Bilingual Corpus, early childhood practitioners often had to rely on their personal judgment when creating a language curriculum for kindergarteners. They often felt that once the whole language approach is adopted and games are used in the class, children’s language would be automatically enhanced. The problem with such an approach is that the curriculum itself may not be age-appropriate at all. For example, I had witnessed teachers setting unrealistic language goals for 4-year-old children – expecting these children to produce stories with connectives instead of expecting them to produce complete sentences.

Since 2004 the afore-mentioned corpora were introduced to almost all of the early childhood language modules delivered at the HKIEd (such as Emergent Literacy, Language and Literacy in Early Childhood, Children with Language Difficulties, Promoting Language and Literacy in Early Childhood). Teachers have since become more aware of the language characteristics of children aged between age 2 and 6, and what would happen to children if they have a diverse language background. As a result of this change introduced, early childhood practitioners are now better able to verbalize why a certain language curriculum is age appropriate.

Between 21st and 25th February this year I was invited by the Macau government to travel to this Chinese city with a blooming economy in order to evaluate the language curricula adopted in the kindergartens there. During my visit, I shared with the Government officials I met the insights I gained from the CHILDES database, and the
Macau government is now working towards a more age-appropriate language curriculum for kindergarteners there.

In terms of research, prior to the availability of CHILDES, there was very little data-sharing among researchers. A lot of resources were wasted because researchers often had to start their own data collection or re-tag the data employed in other studies (as different studies tended to employ different classification systems) each time they wanted to evaluate a theoretical claim. But with the availability of these electronic corpora and the use of a common classification system, it has become much simpler to evaluate theoretical claims – with the click of a few buttons, you can extract the data you need from different languages. Such a computerized system has greatly facilitated comparisons across age and across languages. In fact, in the past 15 years, academics in Hong Kong have been working very hard to contribute their data to the CHILDES database. These researchers include Prof Paul Fletcher (former Chair of the Division of Speech and Hearing Sciences at the University of Hong Kong), Prof Stephen Matthews (Head of the Linguistics Department at the University of Hong Kong), Prof Virginia Yip and Prof Gladys Tang (Department of Linguistics and Modern Languages at the Chinese University of Hong Kong). Their academic papers which were published in a wide array of international refereed journals, ranging from developmental psychology, cognitive science, medicine, linguistics, etc. often made reference to the electronic database. These researchers also encouraged their doctoral students and master’s students to make use of these electronic corpora. Following their footsteps, in the past two years when I was supervising my own master’s students, I often emphasized the need to make use of existing corpus data during the stage of formulating a theoretical framework (before testing the framework with an experiment). When browsing through the library catalogues of theses relating to language development across the universities in Hong Kong, it is rare to spot a thesis that has not employed the CHILDES database.

The CHILDES database has greatly enhanced the teaching practice of language teachers and facilitated data-sharing among researchers in East Asia. Being an administrator myself, I understand that it takes great art and a vast amount of financial resources in order to sustain and strengthen a system. I should be very grateful if the relevant funding body could kindly continue to support this unique and fundamental database so that more practitioners and academics can receive benefits from it.

Should you have any question, please do not hesitate to contact me at

kswong@ied.edu.hk

Yours faithfully,

Dr Wong Kwok Shing, Richard
Dear Brian, 18 February 2008

We are writing in connection with your proposal for renewed funding to further expand and develop the CHILDES Project. It is our pleasure to share our experience using and working with CHILDES over the last five years.

As researchers and teachers in the field of childhood bilingualism, we would like to make a strong statement that the CHILDES project has done the field a tremendous service in promoting data exchange and research in language acquisition. Under the dynamic leadership of Professor Brian MacWhinney, CHILDES has evolved into an invaluable, user-friendly archive and platform used by researchers from all over the world to address a wide range of theoretical and empirical issues.

Just as the original CHILDES database greatly enriched the repertoire of languages for which developmental data are available, the current database has dramatically broadened the coverage of data documenting bilingual development across diverse language pairs. CHILDES has played an invaluable role in facilitating the construction of bilingual and multilingual corpora, giving a huge impetus to research in bilingual acquisition.

Alongside the steady development in coverage of languages and language pairs, we have witnessed the equally impressive development of technical tools. We are proud to have contributed in modest ways: for example, our team member Uta Lam produced a user-friendly web-based tutorial to introduce the archive and how the commands work. The tutorial has been translated into Spanish. Her materials are also used in our undergraduate and graduate courses conducted at the Chinese University of Hong Kong, giving new users a valuable introduction to the system. We take pride in a recent product of our collaboration released in 2007: a video interview featuring a discussion between Brian MacWhinney and Virginia Yip on the evolution and milestones of CHILDES overtime and the construction of bilingual corpora. The video can be found on Youtube as well as on website for our book The Bilingual Child.

The tools provided by CHILDES, especially the CHAT format and CLAN programme, have been invaluable for our team in the development of the Hong Kong Bilingual Child Language Corpus which currently contains longitudinal data seven Cantonese-English bilingual children between age 1;03-4;06. This is the first bilingual corpus involving two major languages of the world, namely Cantonese and English. For the first time, a Chinese language is represented in a bilingual child language corpus that is now freely downloadable via the CHILDES archive. It is also the largest bilingual...
corpus available to world-wide access, thanks to the ingenuity and dedication of the CHILDES team. Our recent book The Bilingual Child: Early Development and Language Contact, published by Cambridge University Press in 2007, would not have been possible without the support of CHILDES and the monolingual Cantonese corpus “CANCORP”. Numerous references to CHILDES are made in the book in grateful acknowledgment of its importance to our work.

In terms of teaching, CHILDES has provided a rich resource for students and teachers. CHILDES is introduced early on in our courses on first language acquisition, bilingualism and research methodology at the Chinese University of Hong Kong (Department of Linguistics and Modern Languages) and University of Hong Kong (Department of Linguistics and Division of Speech and Hearing Sciences). Students at the undergraduate and graduate levels make use of the monolingual and bilingual databases in their term projects and theses. There is a wide range of languages represented in CHILDES and students in Hong Kong are especially happy to find Cantonese, Mandarin and other Asian languages represented along with other well-researched languages such as English, German, French and Spanish. One of the important milestones of CHILDES is the development of automatic morphosyntactic analysis for different languages including Chinese languages, Cantonese and Mandarin. CHILDES keeps breaking new ground in cracking the codes of different languages and successfully pioneering the tagging of Chinese characters making it easy for Chinese readers.

There is no platform in our field as all-embracing and widely recognized as CHILDES. Regardless of theoretical persuasion, researchers regularly make use of the invaluable databases to verify their theories, raise new questions and advance the development of the field of language acquisition.

In the long term, generations of students and researchers around the world stand to gain from the work of CHILDES. We strongly recommend that continued funding for CHILDES be prioritized so that this crucial component of the scientific infrastructure for the study of child language development can be taken to further heights.

Sincerely,

Virginia Yip
Professor
Department of Linguistics and Modern Languages
Chinese University of Hong Kong
Shatin, N.T.
Hong Kong
Email: vcymatthews@cuhk.edu.hk

Stephen Matthews
Acting Head
Associate Professor in Linguistics
School of Humanities
University of Hong Kong
Pokfulam, Hong Kong
Email: matthews@hkucc.hku.hk
155. Zhou, Jing

Dear Brian,

I have listed some books and articles including theses that we used CHILDES to do child language study in mainland China. Please see the attachment. There are 2 doctoral theses and three master theses related to CHILDES will come out in May of 2008.

I also have 40 term papers using CHILDES to do child language study that are not included in the list. Just want you to know how wonderful the CHILDES is and more and more Chinese scholars are using it.

Thank you so much for building up and keeping this great platform for child language study over the world.

Zhou Jing

Zhou Jing, Ph. D
Professor & Deputy Dean
Faculty of Preschool & Special Education
East China Normal University
P. R. China

Books


Articles

2005

2006
1. 周兢，从语言到语言转换阶段的语言运用能力发展 - 3岁前汉语儿童语用交流行为习得的研究：心理科学，2006，29(6),pp1370-1375

2. 计均，周兢，李晓燕，自闭倾向儿童语言运用特点的个案研究：中国特殊教育，2006年第7期，pp32-36

3. 周兢，朱从梅，母亲在亲子阅读中的语言运用特点分析：幼儿教育（教育科学版）2006年第9期，pp48-55


5. 陈臻辉，不同交流方式和语境中弱智儿童交流效果的个案分析：中国特殊教育，2006年第10期

2007


Dissertations
2001
Zhou, J. (2001). Pragmatic Development of Mandarin-speaking Children From 14 Months to 32 Months, Doctoral thesis of The University of Hong Kong

2002

孙秀荣：母亲的言语运用对儿童语用发展的影响:1-3岁儿童的母子互动研究,2002年南京师范大学硕士论文

2003

1. 朱从梅：图画书阅读中母亲和儿童的语用研究，2003年南京师范大学硕士论文
2. 欧阳新梅：儿童的语用发展对母亲言语运用的影响 - 3-6岁儿童的母子互动研究

3. 吕修锋：TalkbankData在中文儿童语言研究中的运用，2003年南京师范大学硕士论文

2004
1. 李晓燕，不同教育背景母亲的言语运用对儿童语用的影响 - 3-6岁儿童的母子互动研究，2004年南京师范大学硕士论文
Li X.Y. (2004). The Effect of Language using of Mothers with different Education on Children's Language Using-A Research on 3 to 6 Year-olds in Mother-child's interaction context, Nanjing Normal University Master thesis.

2007
1. 陈臻辉，4-5岁弱智儿童的言语交流行为：在母子互动中的语用研究

2. 徐勇，特定型语言障碍儿童叙事性语言的动词使用研究

3. 贺利中，4-6岁汉语重度听觉障碍儿童语用发展研究

4. 张海霞，学前儿童看图叙事顺序能力研究