The acquisition of numeral classifiers by Malay children

Khazriyati Salehuddin
B Ed. TESL (First Class Honours), Universiti Kebangsaan Malaysia
MA ELS (Linguistics), Universiti Kebangsaan Malaysia

A thesis submitted for the degree of Doctor of Philosophy
MARCS Auditory Laboratories
University of Western Sydney
September 2009
Dedication

To my parents and the Di(i)ns of my life
Acknowledgements

A lot of credit goes to my supervisor, Dr Heather Winskel for her dedication in guiding me through my PhD. Her extremely generous insights, time, and resources have made my transition to developmental psychology from linguistics a stimulating and meaningful journey. She is an outstanding and efficient supervisor who shows great passion for her work without compromising her compassion. Working with her has been truly motivating. Thanks also to my co-supervisors, Prof. Nigel Bond and Prof. Denis Burnham, and to Colin Schoknecht and Johnson Chen for their technical support and beyond.

My PhD studies have been made possible through a scholarship awarded by the Ministry of Higher Education of Malaysia and my employer, Universiti Kebangsaan Malaysia. My gratitude also goes to MARCS Auditory Laboratories who awarded me a scholarship to conduct the data collection in Malaysia. Credit also goes to the Prime Minister’s Department of Malaysia, Dewan Bahasa dan Pustaka, the caretakers and their children, and the principals, teachers, and students of the pre- and primary schools who participated in the study. My gratitude also goes to Syarfina Shahruddin, Muhammad Irfan Bakhtiar, Mazriza Mohamed Diah, and all my friends in Sydney and Malaysia.

The support and blessings from my family gave me the strength to pursue my PhD. I would also like to give a special appreciation to my family: To my parents, Mak and Ayah, for showering me with their love and prayer, and for making me who I am today; to my siblings, Kak Long, Bang Ngah, Kak Cik and Kak Dek, my in-laws, and my nephews and nieces for their continuous support, motivation, and thoughtfulness. I am immensely thankful to my husband, Shamsudin Lajis, for his love, patience, warmth, and encouragement. He provides me with comfort and strength during challenging moments and gives me support in every possible way. Finally, my special love goes to our children, Nur Ad-Diina, Ahmad Naiimuddin, and Ahmad Najjmuddin who have been the source of my inspiration. May we all be blessed, Amin.
Statement of Authentication

I hereby declare that this submission is my own work and to the best of my knowledge it contains no material previously published or written by another person. I hereby declare that I have not submitted this material, either in full or in part, for a degree or diploma at the University of Western Sydney, or at any other educational institution, except where due acknowledgement is made in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception is acknowledged.

Khazriyati Salehuddin
# Table of Contents

Dedication.............................................................................................................. ii  
Acknowledgements............................................................................................ iii  
Statement of Authentication................................................................................ iv  
Table of Contents.................................................................................................... i  
List of Tables.......................................................................................................... v  
List of Figures......................................................................................................... vi  
Abbreviations and Acronyms................................................................................ ix  
Abstract.................................................................................................................. x  

## CHAPTER 1  
Numeral Classifiers: An Overview ................................................................. 1  
1.1 Introduction.................................................................................................... 1  
1.2 Numeral Classifiers....................................................................................... 5  
1.3 The Syntax, Semantics, and Pragmatics of Numeral Classifiers............... 7  
1.4 Numeral Classifier Languages.................................................................... 13  
  1.4.1 Japanese.................................................................................................. 13  
  1.4.2 Chinese.................................................................................................. 16  
  1.4.3 Thai......................................................................................................... 19  
1.5 The Malay Numeral Classifier System..................................................... 21  
  1.5.1 The Syntax of Malay Numeral Classifiers........................................... 22  
  1.5.2 The Semantics of Malay Numeral Classifiers....................................... 24  
  1.5.3 The Pragmatics of Malay Numeral Classifiers....................................... 31  
1.6 Summary....................................................................................................... 37  

## CHAPTER 2  
The Categorisation and Acquisition of Numeral Classifiers....................... 38  
2.1 Categorisation, Cognition, and Numeral Classifiers.................................. 38  
  2.1.1 Categorisation and Cognition.............................................................. 38  
  2.1.2 Categorisation and Numeral Classifiers........................................... 42  
2.2 Typicality in Categorisation and Numeral Classifier Exemplars............. 44  
2.3 Categorisation in the Malay Numeral Classifier System......................... 49  
2.4 Numeral Classifier Acquisition................................................................. 53  
  2.4.1 Japanese Numeral Classifier Acquisition.......................................... 55  
  2.4.2 Chinese Numeral Classifier Acquisition.......................................... 59  
  2.4.3 Thai Numeral Classifier Acquisition................................................. 64  
2.5 Research Questions and Hypotheses......................................................... 67  
  Developmental Patterns of Malay Numeral Classifier Acquisition............. 67  
2.6 Summary....................................................................................................... 69  

## CHAPTER 3  
Pilot Study: The Production of Fourteen Numeral Classifiers ....................... 70  
3.1 Introduction.................................................................................................... 70  
3.2 Method.......................................................................................................... 72  
  3.2.1 Participants............................................................................................ 72  
  3.2.2 Stimuli.................................................................................................. 73  
  3.2.3 Procedure............................................................................................. 73
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.4</td>
<td>Practice Trial</td>
<td>74</td>
</tr>
<tr>
<td>3.2.5</td>
<td>The Experiment</td>
<td>75</td>
</tr>
<tr>
<td>3.3</td>
<td>Coding of Responses</td>
<td>76</td>
</tr>
<tr>
<td>3.4</td>
<td>Results</td>
<td>76</td>
</tr>
<tr>
<td>3.5</td>
<td>Discussion</td>
<td>83</td>
</tr>
<tr>
<td>3.6</td>
<td>Summary</td>
<td>85</td>
</tr>
<tr>
<td>CHAPTER 4</td>
<td>Preliminary and Supplementary Data Collection and Research Questions</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>87</td>
</tr>
<tr>
<td>4.2</td>
<td>Selection of Stimuli for the Experiments Based on Degree of Typicality</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Stimuli</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Procedure</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>91</td>
</tr>
<tr>
<td>4.3</td>
<td>Input Frequency and Usage</td>
<td>94</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Caretaker-Child Interactions</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Participants</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Stimuli</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Procedure</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Practice Trial</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>The Game</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>108</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The Children’s Corpus</td>
<td>108</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Dewan Bahasa dan Pustaka Corpus</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Method</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>111</td>
</tr>
<tr>
<td>4.3.4</td>
<td>The Malaysian Education System and the Formal Teaching of Numeral Classifiers</td>
<td>113</td>
</tr>
<tr>
<td>4.4</td>
<td>Research Questions and Hypotheses</td>
<td>114</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Developmental Patterns of Malay Numeral Classifier Acquisition.</td>
<td>114</td>
</tr>
<tr>
<td>4.4.2</td>
<td>The Categorisation of Objects into Numeral Classifier Categories..</td>
<td>117</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary</td>
<td>118</td>
</tr>
<tr>
<td>CHAPTER 5</td>
<td>Experiment 1: Elicited Production Task</td>
<td>120</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>120</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Across Numeral Classifier Category Development</td>
<td>121</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Within Numeral Classifier Category Development</td>
<td>123</td>
</tr>
<tr>
<td>5.2</td>
<td>Method</td>
<td>123</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Participants</td>
<td>123</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Stimuli</td>
<td>124</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Procedure</td>
<td>125</td>
</tr>
<tr>
<td>5.2.3.1</td>
<td>Picture-Familiarisation Session</td>
<td>125</td>
</tr>
<tr>
<td>5.2.3.2</td>
<td>Practice Trial</td>
<td>125</td>
</tr>
<tr>
<td>5.2.3.3</td>
<td>The Experiment</td>
<td>126</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Coding of Responses</td>
<td>128</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>130</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Across Numeral Classifier Category Development</td>
<td>130</td>
</tr>
</tbody>
</table>
9.1 Introduction........................................................................................................ 192
9.2 The Developmental Pattern of Malay Numeral Classifier Acquisition.......... 192
  9.2.1 The Order of Numeral Classifier Acquisition........................................... 192
  9.2.2 The Strategies Used for Unknown Numeral Classifiers......................... 197
  9.2.3 The Usage of Alternative Numeral Classifiers......................................... 198
  9.2.4 The Order of Acquisition Based on Semantic Complexity..................... 199
  9.2.5 The Order of Comprehension in Comparison to Production................. 200
  9.2.6 The Influence of the Children’s Linguistic Environment....................... 202
9.3 The Categorisation of Objects into Numeral Classifier Categories................. 206
  9.3.1 The Categorisation of Objects from Broader to Finer Distinctions into Numeral Classifier Categories............................................................... 206
  9.3.2 The Role of Typicality in the Categorisation of Objects into Numeral Classifier Categories............................................................... 209
  9.3.3 Summary of the Categorisation of Objects into Numeral Classifier Categories........................................................................................................ 211
9.4 Limitations of the Research............................................................................ 213
9.5 Future Directions.............................................................................................. 214
9.6 Conclusions...................................................................................................... 214

References............................................................................................................ 217

List of Appendices.................................................................................................. 229
Appendix 1 Malay Numeral Classifier List.............................................................. 230
Appendix 2 Research Approval from the Malaysian Government.................... 231
Appendix 3 Ethics Approval from University of Western Sydney...................... 233
Appendix 4 Test Materials for the Pilot Study..................................................... 234
Appendix 5 Typicality Survey Form........................................................................ 235
Appendix 6 Invitation to Participate in the Putar, Cari, & Kira Game.................. 239
Appendix 7 The Putar, Cari, & Kira Game Picture-book.................................... 240
Appendix 8 Rubric for Caretakers.......................................................................... 242
Appendix 9 Test Items for Matching Comprehension Task.............................. 245
Appendix 10 Published Journal Articles and Book Chapters.............................. 248
  Appendix 10.1 Salehuddin, K. (2007, pp. 82-93)............................................. 249
List of Tables

| Table 1.1. | Some Malay Numeral Classifiers and Examples ........................................ | 28 |
| Table 3.1. | Description of Participants in the Pilot Study ..................................... | 72 |
| Table 3.2. | Description of Practice Trial Items .................................................. | 75 |
| Table 3.3. | Description of Test Stimuli for the Pilot Study .................................... | 76 |
| Table 3.4. | Order of Malay Numeral Classifier Production Based on Age Group. ............ | 80 |
| Table 3.5. | The Frequency of the Three Most Frequently-used Default or Alternative Numeral Classifiers and the Numeral Classifiers They Substitute .................................................. | 83 |
| Table 4.1. | The Numerical Scoring System between 1 and 5 on the Scale from Very Atypical to Very Typical for the Numeral Classifier Exemplars | 90 |
| Table 4.2. | Summary of the Survey Responses ...................................................... | 93 |
| Table 4.3. | Significant Differences between the Typicality Types across All Numeral Classifiers .................................................. | 94 |
| Table 4.4. | Description of Participants in the “Putar, Cari, & Kira” Game .......... | 99 |
| Table 4.5. | Frequency of Numeral Classifier Usage among Caretakers ...................... | 104 |
| Table 4.6. | Frequency of Malay Shape-based Numeral Classifiers in the 178,000-word Children’s Corpus .................................................. | 110 |
| Table 4.7. | Frequency of Numeral Classifiers from the DBP Corpus ...... | 112 |
| Table 4.8. | Ranking of Numeral Classifier Frequency in the Children’s Corpus and the DBP Corpus .................................................. | 112 |
| Table 4.9. | Ranking of Numeral Classifier Frequency in the Children’s Corpus and their Sequence of Appearance in School Textbooks ........ | 114 |
| Table 5.1. | Description of Participants in the Study ............................................. | 124 |
| Table 5.2. | Description of Practice Trial Items .................................................... | 126 |
| Table 5.3. | Description of Test Stimuli for Experiment 1 ....................................... | 128 |
| Table 5.4. | Stage of Malay Shape-based Numeral Classifier Acquisition Based on the Mean Number of Correct Production Responses across All Age Groups .................................................. | 133 |
| Table 5.5. | The Frequency and Percentage of the Shape-based Numeral Classifiers Used as Alternative of Substitution Responses for Correct Responses .................................................. | 135 |
| Table 5.6. | Mean Number of Correct Responses Based on Typicality Type .............. | 136 |
| Table 6.1. | Description of Participants in Experiment 2 ....................................... | 144 |
| Table 6.2. | Description of Test Stimuli for Experiment 2 ..................................... | 146 |
| Table 6.3. | Description of Practice Trial Items for Experiment 2 .......................... | 149 |
| Table 6.4. | Stage of Malay Shape-based Numeral Classifier Acquisition Based on the Mean Number of Correct Matching Comprehension Responses across All Age Groups .................................................. | 152 |
| Table 6.5. | Percentage of Alternative Responses Based on the Total Number of Incorrect Matching Responses .................................................. | 154 |
| Table 6.5. | Mean Number of Correct Responses Based on Typicality Type .............. | 155 |
| Table 7.1. | Description of Participants ............................................................... | 166 |
| Table 7.2. | Description of Test Stimuli for the Strong Contrast Condition ........... | 168 |
| Table 7.3. | Description of Test Stimuli for the Weak Contrast Condition ............ | 170 |
| Table 8.1. | Description of Test Stimuli for the “Select the Odd-One-Out” Task .. | 184 |
| Table 1-A | Malay Numeral Classifier List Adapted from Osman (2002), and Othman (2004) .................................................. | 232 |
List of Figures

Figure 1.1. The distribution of numeral classifiers in the languages of the world (Aikhenvald, 2003, p. 122) .......................................................... 7
Figure 1.2. Japanese numeral classifier system as illustrated in Yamamoto and Keil (2000, p. 381) ................................................................. 15
Figure 1.3. Mandarin numeral classifier distribution (adapted from Hu, 1993, p. 8) ............................................................ 16
Figure 1.4. Hokkien numeral classifier distribution (adapted from Ng, 1989, p. 57) ................................................................. 16
Figure 1.5. Cantonese numeral classifier distribution (adapted from Tse, Li, & Leung, 2007, p. 498) ......................................................... 17
Figure 2.1. The Malay numeral classifier system from the perspective of the native speaker (Salehuddin & Winskel, 2008a). The dotted line under the “Specific” numeral classifier category indicates that the inventory list of specific numeral classifiers is more numerous than illustrated. ........................................ 50
Figure 3.1. Mean number of the three response types produced by all age groups ................................................................................. 77
Figure 3.2. Mean number of correct responses for the 14 numeral classifiers by age group. Adults gave 100% correct responses to all 14 numeral classifiers tested ........................................................................ 78
Figure 3.3. Mean number of correct responses for each numeral classifier. Circles indicate that the mean numbers were not significantly different from each other as indicated by pairwise comparisons (LSD = .05) .................................................................................. 79
Figure 3.4. Mean number of correct responses for all three numeral classifier types in each age group ............................................................. 82
Figure 4.1. Example of manner of responding to the survey. Bilah was used to illustrate how the survey should be responded to because bilah is not one of the numeral classifiers tested in the survey. .......... 91
Figure 4.2. The spinning wheels for ekor (wheel labelled “Y”) and orang (“Z”). The window on each spinning wheel revealed only one picture at a time. Each picture on the spinning wheel was labelled with the object name. ................................................................. 100
Figure 4.3. The A3-size picture book showing the setting of a park to contextualise the objects. “YZ” was labelled at the top right-hand corner of the book. Pictures of objects were glued on the setting to create a 3D effect. ................................................................. 101
Figure 4.4. A caretaker and a child interacting while playing the “Putar, Cari, & Kira”. The child is spinning the windowed spinning wheel. .......................................................................................... 102
Figure 5.1. A slide from the picture-familiarisation session. ......................... 125
Figure 5.2. Mean number of correct responses produced by the different age groups for each shape-based numeral classifier. .......................... 131
Figure 5.3. Mean number of the four response types used by each age group. “Others” includes ‘unitisers’, ‘collective nouns’, ‘measure words’, and ‘content words’. .................................................... 134
Figure 5.4. Mean number of correct responses by each age group based on typicality type. ................................................................. 137
Figure 5.5. Mean number of correct responses for each numeral classifier based on typicality type. ................................................................. 138
Figure 6.1. Matching comprehension task stimuli. Children were asked to match the numeral classifier names to the respective pictures of exemplars one column at a time. Pictures on the left column are very typical exemplars of the eight numeral classifiers tested while pictures in the right column are atypical exemplars of the numeral classifiers. While working on the pictures in the left column, those in the right column were covered. ................................. 147
Figure 6.2. Mean number of correct responses by each age group for each numeral classifier. ................................................................. 151
Figure 6.3. Mean number of correct responses by each age group based on typicality type. ................................................................. 156
Figure 6.4. Mean number of correct responses for each numeral classifier based on typicality type. ................................................................. 157
Figure 6.5. Mean number of correct responses for the comprehension task in comparison with the production task for each age group. .............. 158
Figure 6.6. Mean number of correct responses for the comprehension task in comparison with the production task for the numeral classifier categories. ......................................................................................... 159
Figure 7.1. Four exemplars used for each numeral classifier category: two typical exemplars, two atypical exemplars. In the strong contrast condition, typical exemplars of batang (pencil and tree) were paired with typical exemplars of buah (bus and box) respectively (top row). The atypical exemplars of batang (river and road) were each paired with atypical exemplars of buah (planet and robot) (bottom row). ................................................................. 167
Figure 7.2. To counterbalance the pencil-bus picture pair, pencils were positioned on the left column of one picture pair while in the other picture pair, pencils were positioned on the right. ..................... 168
Figure 7.3. In the weak contrast condition, typical exemplars of batang (pencil and tree) were each paired with typical exemplars of utas (necklace and rope) (top row) while atypical exemplars of batang (river and road) were each paired with atypical exemplars of utas (watch and chain link) (bottom row). .................................................. 169
Figure 7.4. Audio-visual displays for one of the 4 orang [animate: human] versus ekor [animate: animal] picture pairs used in the practice trial session. For these picture pairs, the audio prompt was either “orang” or “ekor”. The position of the red and green dots remained in this order (red on left, green on right) in all trials. .... 171
Figure 7.5. The overall protocol for the Strong-Weak Contrast Discrimination Task. ................................................................. 173
Figure 7.6. Mean number of correct responses for each age group. .......... 175
Figure 7.7. Mean number of correct responses for each numeral classifier based on contrast type. ................................................................. 176
Figure 7.8. Mean number of correct responses for each numeral classifier based on typicality type. ................................................................. 177
Figure 7.9. Mean number of correct responses for each numeral classifier based on strong-weak contrast and typical-atypical exemplars. ...

Figure 8.1. Test stimuli for the batang [1D: +rigid] category. Children had to select the odd-one-out from a set of three pictures. ............

Figure 8.2. A visual display of one exemplar of the bilah [specific: knife] numeral classifier. The picture was displayed simultaneously with an audio prompt “bilah” to indicate to the children that a knife is an exemplar of “bilah”. ........................................

Figure 8.3. A visual display of atypical and typical exemplars of the bilah [specific: knife] numeral classifier (i.e., axe, saw) and a perceptually similar non-member of the bilah numeral classifier category (comb). The visual display appeared simultaneously with the audio prompt “bukan bilah” (i.e., non-member of “bilah” numeral classifier). The position of the red, yellow, and green dots remained in this order (red on left, yellow in the middle, and green on right) for all trials. .........................

Figure 8.4. Overall protocol of the “Select the Odd-One-Out” Task. .........

Figure 8.5. Mean number of correct responses in the select the “Select the Odd-One-Out” Task for each age group. (The correct response is to select the perceptually similar non-member as the odd-one-out). ..............................................................
## Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation / Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+rigid</td>
<td>Rigid objects</td>
</tr>
<tr>
<td>-rigid</td>
<td>Flexible objects</td>
</tr>
<tr>
<td>1D</td>
<td>Saliently long objects</td>
</tr>
<tr>
<td>1D: +rigid</td>
<td>Long and rigid</td>
</tr>
<tr>
<td>1D: -rigid</td>
<td>Long and flexible</td>
</tr>
<tr>
<td>2D</td>
<td>Saliently flat objects</td>
</tr>
<tr>
<td>2D: +rigid</td>
<td>Flat and rigid</td>
</tr>
<tr>
<td>2D: -rigid</td>
<td>Flat and flexible</td>
</tr>
<tr>
<td>3D</td>
<td>Saliently rounded objects</td>
</tr>
<tr>
<td>3D: big</td>
<td>Rounded and big</td>
</tr>
<tr>
<td>3D: fine</td>
<td>Rounded and fine</td>
</tr>
<tr>
<td>3D: medium</td>
<td>Rounded and medium</td>
</tr>
<tr>
<td>3D: small</td>
<td>Rounded and small</td>
</tr>
<tr>
<td>A</td>
<td>Atypical exemplar of a numeral classifier category</td>
</tr>
<tr>
<td>AS</td>
<td>Associative marker</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>COM</td>
<td>Comitative marker</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
</tr>
<tr>
<td>DBP</td>
<td>Dewan Bahasa dan Pustaka</td>
</tr>
<tr>
<td>DI</td>
<td>Disposal marker</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive marker</td>
</tr>
<tr>
<td>M</td>
<td>Moderate exemplar of a numeral classifier category</td>
</tr>
<tr>
<td>N</td>
<td>Noun</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative marker</td>
</tr>
<tr>
<td>Num</td>
<td>Numeral</td>
</tr>
<tr>
<td>NumCl</td>
<td>Numeral classifier</td>
</tr>
<tr>
<td>PRT</td>
<td>Perfective marker</td>
</tr>
<tr>
<td>PST</td>
<td>Past tense marker</td>
</tr>
<tr>
<td>Q</td>
<td>Question marker</td>
</tr>
<tr>
<td>QUOT</td>
<td>Quotative marker</td>
</tr>
<tr>
<td>RTs</td>
<td>Reaction times</td>
</tr>
<tr>
<td>T</td>
<td>Typical exemplar of a numeral classifier category</td>
</tr>
<tr>
<td>TOP</td>
<td>Topic marker</td>
</tr>
<tr>
<td>Va</td>
<td>Very atypical exemplar of a numeral classifier category</td>
</tr>
<tr>
<td>Vt</td>
<td>Very typical exemplar of a numeral classifier category</td>
</tr>
</tbody>
</table>
Abstract

Numeral classifiers are typically used in counting or referring to objects. Numeral classifiers acquisition is of interest to researchers as researchers are able to examine how children learn to categorise and label objects in their environment using a constrained framework, and how this ability develops and becomes more refined with age. Cross-linguistic studies on numeral classifier acquisition provide researchers with greater insights into patterns that are both universal and language-specific. Malay numeral classifier acquisition has not been previously investigated. Thus, the current research aimed to investigate Malay numeral classifier acquisition to identify the acquisition patterns that are specific to Malay and also common with other numeral classifier languages previously studied.

Four experiments were conducted on 6- to 9-year-old Malay children and a comparison group of adults in this thesis: Experiment 1 (Elicited Production Task), Experiment 2 (Matching Comprehension Task), Experiment 3 (Strong-Weak Contrasts Discrimination Task) and Experiment 4 (Select The Odd-One-Out Task). Preliminary and supplementary data was collected. A survey on typicality of exemplars was initially conducted to select the test stimuli to be used with each shape-based numeral classifier. An analysis of numeral classifier frequency occurring in a general Malay database and a database constructed from materials relevant to children was conducted. In addition, numeral classifier usage occurring in caretaker-child interactions was investigated using a counting game.

Experiments 1 and 2 were conducted to investigate the acquisition order of Malay numeral classifier production and comprehension, and the factors that affect
their order. In Experiment 1, children were asked to count the items they saw in a range of picture stimuli depicting objects that require numeral classifiers in counting processes. In Experiment 2, children were asked to match a numeral classifier name to the items that require the use of the particular numeral classifier in counting activities. It was found that children used omission and substitutions as strategies in place of unknown numeral classifiers. These substitutions and omissions revealed that children depend, to a large extent, on perceptual features in categorizing objects. In addition, the relative complexity of semantic features of numeral classifiers and children’s linguistic environment, which includes both informal usage and formal instruction of numeral classifiers, affect Malay numeral classifier acquisition order.

Experiments 3 and 4 were conducted to investigate how children categorise items into particular numeral classifier categories. In Experiment 3, children were requested to select an item from a picture-pair stimulus that would match the numeral classifier name correctly when the stimulus was presented to them. In Experiment 4, children were requested to select an item from a series of picture stimulus that does not belong to the named numeral classifier category. It was found that children categorise objects with more correct responses when there are strong rather than weak contrasts between the exemplars of shape-based numeral classifiers. Children also sorted typical exemplars more correctly with faster responses in Experiment 3. In Experiment 4, the youngest children (the 6-year-olds) selected the perceptually different member as the odd-one-out rather than the perceptually similar non-member. In contrast, the older children predominantly selected the perceptually similar non-member as the odd-one-out. This indicates that young children are strongly influenced by the perceptual features in the categorisation of objects into numeral classifier categories, and that children’s knowledge of the semantics of
numeral classifier categories gradually develops and becomes more refined as they grow older.

The Malay numeral classifier system is shaped by both universal cognitive-based and language-specific categorisation processes. This thesis has shown that numeral classifier acquisition order and the way children classify objects into the numeral classifier categories are influenced by multiple factors, which includes semantic complexity, linguistic environment, exemplar typicality, and young children’s tendency to rely on perceptual features. This research further indicates that the acquisition of Malay numeral classifiers is a relatively prolonged process in comparison to other numeral classifier languages because numeral classifier usage is less obligatory in Malay in comparison to other numeral classifier languages.
1.1 Introduction

Human beings assign objects and items into particular categories, both consciously and unconsciously, during normal everyday life. The categorisation of objects into the appropriate category is a vital skill for humans to acquire (Ashby & Maddox, 2005). When a child encounters a novel or new object in the environment, s/he needs to decide which category to assign the object or item. The act of categorisation is an essential fundamental cognitive skill for humans to operate effectively in all cultures and societies. Apart from being a fundamental process in human cognition, the act of categorising and labelling objects is also one of the most basic processes in language development (Croft & Cruse, 2004; Lakoff, 1987). In order to categorise and label a particular object, stored information is retrieved and evaluated, and only when there is a satisfactory resemblance between new and stored information is an object accepted into the respective category (Barsalou, Huttenlocher, & Lamberts, 1998). Objects that are perceived as belonging to the same category are then assigned to the same label (Mervis & Rosch, 1981). Numeral classifier systems offer researchers an ideal opportunity to examine how children learn to categorise and label objects in their environment using a constrained framework or system.
Numeral classifiers are typically used in counting objects. Speakers of numeral classifier languages need to learn how to categorise objects in their environment and pair them with the appropriate numeral classifier, using the language-specific classification system. The aim of the current research was to examine the developmental patterns observed in both production and comprehension of Malay numeral classifiers. The acquisition of eight shape-based numeral classifiers in Malay was investigated through a range of different tasks; namely the Elicited Production Task (Experiment 1), the Matching Comprehension Task (Experiment 2), the Strong-Weak Contrast Discrimination Task (Experiment 3), and the Select the Odd-One-Out Task (Experiment 4). The content of the chapters in the thesis are outlined as follows.

Chapter 1, *Numeral Classifiers: An Overview* reviews numeral classifier systems from a cross-linguistic perspective, in particular the syntactic, semantic, and pragmatic functions of numeral classifiers. The Malay numeral classifier system is then discussed. Chapter 2, *The Categorisation of Numeral Classifiers* discusses numeral classifiers from a cognitive perspective. Categorisation theories and models that are relevant to numeral classifier systems are reviewed.

In Chapter 3, *Pilot Study: The Production of Fourteen Numeral Classifiers*, a preliminary study conducted on 3- to 12-year-old Malay children is reported. The patterns of development of fourteen Malay numeral classifiers were investigated through an elicited production task. An additional aim of the pilot study was to determine the appropriate age group and the particular numeral classifier categories to be focused on in the main experiments in this research.
In Chapter 4, the research questions, predictions and hypotheses of the current research are outlined. In order to select test stimuli for each of the eight numeral classifier categories, a survey on the typicality of Malay numeral classifier exemplars was conducted. Furthermore, it is believed that children tend to learn words earlier when they are produced more frequently in speech directed to them and when language occurs frequently in their environment (Gallaway & Richards, 1994; Snow & Ferguson, 1977). In order to investigate the speech directed to children, interactions between caretaker and child dyads were observed, recorded, and analysed while they participated in a counting game, the Putar, Cari & Kira game.

School-age children learn language both in informal settings (through indirect modelling and reinforcement) and formal instruction (Nippold, 1988). Results of word frequencies of the Malay numeral classifiers occurring in (1) a Children’s corpus made up of texts from children’s television programmes, children’s story books, folk stories and bedtime stories, and (2) a larger database from Dewan Bahasa dan Pustaka, are reported. Finally, the Malaysian education system and the sequence of teaching of numeral classifiers are discussed.

Chapter 5, Experiment 1: The Elicited Production Task investigates the production order of eight Malay shape-based numeral classifiers. Malay children between 6 and 9 years of age participated in this task, which aimed at examining if the acquisition order of numeral classifier production is affected by the relative complexity of semantic features of individual numeral classifiers and the typicality of exemplars. Picture stimuli were used to elicit responses from the children.

Chapter 6, Experiment 2: The Matching Comprehension Task investigates the comprehension order of the same eight numeral classifiers as used in Experiment 1. Children were required to match picture stimuli to the correct numeral classifier
name. This was also to examine if the relative complexity of semantic features of individual numeral classifiers and the typicality of exemplars affect the acquisition order of numeral classifier comprehension. The production patterns (Experiment 1) and the comprehension patterns of development were also compared.

Chapter 7, Experiment 3: The Strong-Weak Contrast Discrimination Task examines to what extent children’s category judgments are influenced by their perception. By manipulating picture pairing conditions based on dimensionality, rigidity, and size of the objects, this experiment aimed at investigating if categorising objects that differ in two features is easier and executed with faster reaction times than those that differ in only one feature. In addition, the effect of exemplar typicality on correct responses and reaction times in the different pairing conditions was also investigated.

Chapter 8, Experiment 4: The Select the Odd-One-Out Task examines if young children rely on perceptual features rather than semantic relations in determining category membership of the numeral classifier categories. Children had to select from a series of pictures, the object that did not fit in the numeral classifier category.

Finally, in Chapter 9, General Discussion and Conclusion, the results from the four experiments are reviewed and discussed in relation to cross-linguistic studies on human cognition and language development. The discussion focuses on different aspects of numeral classifier development, that is, its production and comprehension, both from an across (a comparison between numeral classifiers) and within (categorisation of objects into particular numeral classifier categories) category perspective.
1.2 Numeral Classifiers

Linguistic categorisation of nouns is commonplace in all languages and is often overtly manifested either in the form of noun classes or noun classifiers (Aikhenvald, 2003). However, noun classes are different from noun classifiers. Noun classes involve the grouping of all nouns into a fixed, limited number of classes via an obligatory morphological system that often requires concordances in their syntactic structure, for example, the use of an obligatory article in German and French (Dixon, 1986). Unlike noun classes, noun classifiers do not involve the classification of all nouns, despite the fact that there may be unlimited inventories of noun classifiers in a particular language (Dixon, 1986). There may be nouns in a classifier language that are left unclassified and, depending on which property of the noun is in focus, there may also be nouns that may have alternative choices of classifiers (Aikhenvald, 2003). Noun classifier usage may vary between different styles within a speech community (Dixon, 1986) and this often reflects the social status, competence, and cognitive categories of the speakers of a particular language (Adams, 1989; Lakoff, 1986). In addition, since sign languages classify objects like classifier languages do, sign languages are often regarded as classifier languages.

‘Numeral classifiers’ are one of the most commonly recognised types of noun classifier systems along with possessive classifiers and predicate classifiers. They are a syntactic-semantic category that is common in most Sino-Tibetan, Austronesian, some Indo-European, Uralic, Mayan, and Arawakian languages (Adams & Conklin, 1973; Aikhenvald, 2003; Allan, 1977; Craig, 1986; Croft, 1994; Goral, 1978; Kiyomi, 1992). They occur contiguously to numerals and expressions of quantity in a noun phrase. Being an instance of a linguistic device of categorisation (Craig, 1986) (cf. "language-inherent classification", Zhang &
Schmitt, 1998, p. 376), numeral classifiers explicitly classify nouns by denoting “some salient perceived or inputted characteristic” of the entity in question (Allan, 1977, p. 285). Syntactically, numeral classifiers – which, like other noun classifiers, predominantly occur as independent lexemes (Craig, 1986; Dixon, 1986) – form part of a noun phrase together with a noun and a numeral (Richards, Platt, & Weber, 1985). Semantically, numeral classifiers provide information about the physical properties (e.g., shape, size, thickness, and length), the conceptual properties (e.g., bendability, graspability, and moveability of objects), the functional properties, and the cognitive categories of objects in a particular culture. To a certain extent, numeral classifiers also provide information about the social status and the role of a particular noun, to the speakers within and across a particular speech community (Adams, 1989; Dixon, 1986; Lakoff, 1986; Zhang & Schmitt, 1998).

Regionally, the numeral classifier system is pervasive across languages of East, Southeast and South Asia, Oceania, and South America. The system is also found in scattered pockets across North and Central America (Aikhenvald, 2003). The presence of numeral classifiers all over the world (Figure 1.1) across various language families results in a vast difference in the inventories and the semantics of numeral classifiers (Adams & Conklin, 1973; Aikhenvald, 2003; Allan, 1977; Craig, 1986; Croft, 1994; Goral, 1978; Kiyomi, 1992).
Thirty-seven East Asian (e.g., Japanese, Chinese, and Korean) and Southeast Asian languages (e.g., Thai and Malay) are numeral classifier languages and most of these languages regard animacy, shape, and function as critical semantic features in the system (Adams & Conklin, 1973; Allan, 1977).

1.3 The Syntax, Semantics, and Pragmatics of Numeral Classifiers

Numeral classifiers most often occur as free morphemes in the same noun phrase as the nouns they qualify. Although in some cases they may sometimes be affixed to numerals, numeral classifiers may never form a morphological unit with a noun (Dixon, 1986). In some languages, the form of numeral classifiers may be homonymous to some nouns of the respective language (Hopper, 1986); however the similarity in the forms of both numeral classifiers and nouns is likely to be due to the metaphoric use of nouns as numeral classifiers (Goral, 1978). Unlike noun classes,
numeral classifiers do not affect the concordances of the sentence structure in which they occur since the references of numeral classifiers are limited within the boundaries of the noun phrase they qualify. Nevertheless, because numeral classifiers may sometimes be used anaphorically like a true pronoun (Downing, 1986), they may exist in a noun phrase without the need to mention the noun they classify (e.g., in 1.4.1).

Numeral classifiers are described as a closed class lexical category. Their membership is so exclusive that no new numeral classifiers can be invented, hence added to the existing inventory of numeral classifiers of a particular language (Uchida & Imai, 1999). Though structurally their grammatical role seems somewhat similar to the quantifiers in English such as a bucket of and a spoonful of; semantically, numeral classifiers are different from English quantifiers (or ‘measure words’ in Yamamoto & Keil, [2000]). Instead of quantifying mass nouns in measuring units, numeral classifiers qualify most and only count nouns based on their permanent qualities. Numeral classifiers are also different from adjectives in their semantic functions because the former does not restrict the properties of the nouns (Zhang & Schmitt, 1998). For example, the noun ‘bed’ will always be used with the Chinese numeral classifier zhang1 (i.e., numeral classifier for flat objects) “no matter whether they are genuinely flat or not” (Zhang & Schmitt, 1998, p. 377).

Numeral classifiers do not ‘unitise’ nouns as collective nouns do (e.g., a pride of lions, a school of fish). Instead, numeral classifiers categorise nouns into categories that carry semantic information. In addition, the assignment of numeral classifiers in counting activities is primarily dependent upon semantic properties of the noun they categorise, usually based on “the parameters of animateness, shape, or function which are attributed to the head noun” (Adams & Conklin, 1973, p. 1).
At a general level, most numeral classifiers do share some universal aspects in their classification of nouns (Allan, 1977). Most languages (but Chinese is an exception) begin their classification of nouns with the parameters of animateness, that is, via the distinction between animate and inanimate nouns (Adams & Conklin, 1973). Animate nouns in most languages are further categorised based on human versus non-human distinctions, while inanimate nouns are further classified based on the shapes and functions of the items. These human–non-human, and shape–function distinctions are known as ‘primary parameters’ of numeral classifiers (Adams & Conklin, 1973). Some of the properties of the primary parameters of inanimate nouns are inherent and time-stable properties, whereas some can be related to temporary states (Aikhenvald, 2003). For example, helai is a Malay numeral classifier used for leaves and other flat and flexible objects regardless of their condition, whereas tangkai [attached to a stalk], which is not a numeral classifier, on the other hand, is used with leaves only when the leaves are present together with their stalks. While the ‘flat and flexible’ property of leaves is time-stable in helai, it is not so permanent in tangkai because the ‘attached to a stalk’ property of leaves is merely a temporary state.

At a more specific level, substantial differences in numeral classifier languages exist and it is at this level that semantic features like rigidity and size are combined with the primary parameters to form further numeral classifier classes (Uchida & Imai, 1999). Rigidity and size are described as ‘secondary parameters’ (Adams & Conklin, 1973) because their semantic features are dependent on the semantic features of the primary parameters. The dependency of secondary parameters on primary parameters results in mixed semantic criteria of numeral classifier categorisation, making the numeral classifier category seemingly complex.
Allan’s (1977) description of shape-based numeral classifiers illustrates how numeral classifier categorisation is in general a complex and opaque categorisation process. Typically, the categorisation of shape-based numeral classifiers is based on saliently\(^1\) one-dimensional (henceforth 1D to indicate ‘long’), saliently two-dimensional (2D to indicate ‘flat’), and saliently three-dimensional (3D to indicate ‘round’ or ‘polyhedral’) parameters. In addition, secondary parameters such as the rigidity and the size of the objects have to be taken into consideration in determining which shape-based numeral classifier is to be assigned to the noun in question. Thus, in most languages, ‘stick-like’ objects usually get assigned ‘saliently 1D and rigid’ numeral classifiers (henceforth [1D: +rigid]), while ‘rope-like’ objects get [1D: -rigid] (Allan, 1977). ‘Plank-like’ objects, on the other hand, usually are assigned ‘saliently 2D and rigid’ numeral classifiers (henceforth [2D: +rigid]), while ‘fabric-like’ objects get [2D: -rigid]. Allan (1977) added that the saliently 3D subcategory is often associated with fruit and that this 3D subcategory is usually combined with the size parameter. Thus, ‘seed-like’ objects usually get ‘saliently 3D and small’ numeral classifiers (henceforth [3D: small]), and ‘bulky’ objects get [3D: big].

Functional parameters (e.g., tools, footwear, written materials) form even more language-specific numeral classifiers because their usage is dependent on how they are perceived and valued by the specific culture (Adams & Conklin, 1973). All these display language-specific partitioning differences. Unlike parameters of shape and animacy, functional parameters “are quite language-specific, reflecting the interests of members of the particular culture in which the language is spoken” (Hansen & Chen, 2001, p. 92).

---

\(^1\) ‘Saliently’ here means that an object would be naturally described as long (for saliently long) if it stood vertically from the ground.
Because of language-specific partitioning differences, the number of numeral classifier inventories and the manner in which numeral classifiers are subcategorised vary from one language to another. The numeral classifier inventories of North American languages are said to be typically small (Aikhenvald, 2003). The Blackfoot (Algonquian), for example, have only two sets of numeral classifiers, one for animate and one for inanimate nouns. East and Southeast Asian languages in contrast are known to have an extensive number of numeral classifiers. Vietnamese, for example, has 140 numeral classifiers and semantically, Vietnamese numeral classifier system takes animateness, shape, quantity, function, and degree of abstractness into account (Nguyen, 1966). Burmese, on the other hand, has 189 numeral classifiers in its numeral classifier inventory. Burmese numeral classifier system includes animateness and sacredness of objects, the dimensions of the objects in time and space, the manner the objects are used, the number the objects co-occur with, the length, weight, and height of the objects, and the presentation of the objects (Burling (1965) cited in Goral (1978)).

In most numeral classifier systems, there are a few unmarked general numeral classifiers that can be used broadly to refer to a large number of nouns. The rest of the numeral classifiers, on the other hand, are very specific in usage. These general classifiers are often used to substitute for specific numeral classifiers within the same ontological boundaries, without distorting the syntax and the semantics of the language. Examples of general classifiers are ขัน in Thai (Carpenter, 1991), -tsu/-ko in Japanese (Yamamoto, 2005), ge in Mandarin (Hu, 1993), and go3 in Cantonese (Loke & Harrison, 1986). The use of a general classifier does not make a person’s production erroneous. Its usage is not necessarily motivated by not knowing what numeral classifiers to use. Studies show that even highly educated people also
produce the general classifier; for example, the unmarked Mandarin Chinese general numeral classifier *ge* in the everyday conversation of educated adults in Erbaugh’s (2002) study.

From the semantic, cognitive, and cultural perspectives, numeral classifiers function as a means to communicate a few especially important classes that objects fall into in the manner a speech community perceives them (Craig, 1986). In addition, numeral classifiers may also function as “nominal substitutes, nominalisers of words in other form classes, markers of definiteness, relativisers, markers of possession, and as vocatives” (Craig, 1986, p. 7). Numeral classifiers may also be used in disambiguating sentences, establishing coherence in discourse, and marking registers and styles within a language (Goral, 1978).

Numeral classifiers may function as an anaphoric option for speakers to avoid excessive repetition of nouns in their discourse (Downing, 1986). Vietnamese numeral classifiers, for example, are used by speakers of this Southeast Asian language to refer to objects mentioned earlier in the discourse (Nguyen, 1957, p. 130). For example, in Example 1, *quyên* is used as a referring expression in the two consecutive noun phrases to avoid repeating the word *sách* (book) unnecessarily.

Example 1:

```
Tôi có hai quyên sách, mô подар quyên mong, mô подар quyên dây.
```

‘I have two books, one thin and one thick.’

Numeral classifiers are also used to mark the importance of a particular object in a text. This takes place when numeral classifiers are used with a full noun, “to foreground physical objects” (Hopper, 1986, p. 313). When a numeral classifier is used for this function, it usually indicates that the object that co-occurs with the
numeral classifier is part of the discourse, and consequently may become “a potential discourse topic” (Hopper, 1986, p. 313).

1.4 Numeral Classifier Languages

A large number of cross-linguistic studies on numeral classifier systems in different languages have been conducted. Studies by Allan (1977), Adams and Conklin (1973), and Goral (1978) were among the first few that rigorously analysed numeral classifier languages. Later in the 1980s, Craig (1986) collaborated with a number of researchers (e.g., Delancy, 1986; Downing, 1986; Hopper, 1986) to form a book on noun categorisation, which included cross-linguistic analyses of the various aspects of numeral classifiers. More recently, Aikhenvald (2003) conducted a comprehensive coverage on noun class systems of the world languages which included studies on numeral classifier systems. In the following sections, we review studies conducted on three numeral classifier languages, namely Japanese (1.4.1), Chinese (1.4.2) and Thai (1.4.3), before reviewing Malay numeral classifiers (1.5.). Japanese, Chinese and Thai are reviewed in depth because these languages are numeral classifier languages spoken in the same region as Malay. A review on these languages will enable us to make comparisons between Malay and other numeral classifier systems from the same region.

1.4.1 Japanese

Numeral classifier is an obligatory element in Japanese (Yamamoto, 2005), an East Asian language from the Japonic language family. They are in the form of bound morphemes in Japanese, and are suffixed to a numeral (Aikhenvald, 2003).
Japanese numeral classifiers occur either in pre-noun phrase position or a post-noun phrase position; for example,

Example 2:

\[
\text{ni-hon-no empitsu-o} \\
\text{two-NumCl-GEN pencil-ACC} \\
\text{‘two pencils’}
\]

Example 3:

\[
\text{empitsu-o ni-hon} \\
\text{pencil-ACC two-NumCl} \\
\text{‘two pencils’}
\]

(Matsumoto, 1993, p. 673)

Although average Japanese speakers may have a large inventory of numeral classifiers at their command (150 according to Yamamoto & Keil, 2000), only thirty of the numeral classifiers commonly play a role in their everyday language use (Downing, 1984).

Japanese numeral classifiers are divided into two major categories; animate and inanimate (Uchida & Imai, 1999; Yamamoto & Keil, 2000). Animate numeral classifiers are further divided into two subcategories; numeral classifiers for humans and those for animals. The inanimate numeral classifiers are also further divided into two subcategories; numeral classifiers for concrete objects and those for abstract objects. Concrete objects are further divided into two subcategories, which are shape-specific numeral classifiers and functional numeral classifiers. Plants and natural substances with solid shapes are also categorised in the same category as shape-specific numeral classifiers (Figure 1.2). With a varying degree of acceptability, the numeral classifier -tsu and the shape-specific classifier for 3D object–ko (Yamamoto & Keil, 2000) can be used in place of almost all inanimate objects (Matsumoto, 1993, p. 674).
Apart from performing syntactic-semantic functions, Japanese numeral classifiers are also used in spoken and written discourse to perform a pragmatic function. They are used anaphorically like a pronoun to refer to a particular noun as mentioned earlier (Downing, 1986, p. 349). For example,

Example 4:

```
Syuuui- TOP Singo- GEN son COP but Kikuko- NOM
kono yoo-ni site made Syuuui- to musubarete
to this extent Syuuui- COM be bound
inakereba naranai hodo, hutari-wa risoo-no
if not unacceptable extent 2-person-TOP ideal
huuuhu nano ka, Singo-wa utagai dasu to kagiri-ga
couple COP Q Singo-TOP doubt QUOT limit-
NOM
nakatta.
not exist-PST
```

“Even though Syuuui was his son, Singo couldn’t help wondering whether they were such an ideal couple that Kikuko should be linked to Syuuui to this extent”

(Downing, 1986, p. 349)

The numeral classifiers in Japanese may function as a “stylistic neutral anaphoric option for the speaker anxious to avoid the ponderous repetition of full
nouns…” (Downing, 1986, p. 345). Thus, with regard to the above excerpt, rather than repeating “Syuuiti” and “Singo” in the discourse, numeral classifiers -ri in hutari was used to avoid repeating the two proper nouns.

1.4.2. Chinese

Chinese languages, for example, Mandarin, Hokkien, and Cantonese, all have numeral classifier in their system; however, their characteristics appear to be quite different (see Figures 1.3, 1.4, and 1.5).

Figure 1.3. Mandarin numeral classifier distribution (adapted from Hu, 1993, p. 8).

Figure 1.4. Hokkien numeral classifier distribution (adapted from Ng, 1989, p. 57).
Structurally, numeral classifiers in Mandarin Chinese usually occur either after a numeral or after a determiner; and in both cases, they occur before the head noun in a noun phrase (Li, 2000). The same structure applies in Hokkien Chinese (Ng, 1989) and Cantonese Chinese (Wei & Lee, 2001).

Example 5:
\begin{verbatim}
yi-ben       shu
one-CL       book
\end{verbatim}
'a book'

Example 6:
\begin{verbatim}
zhe-zhi      bi
this-CL      pen
\end{verbatim}
'this pen'

(Li, 2000, p. 1115)

Example 7:
\begin{verbatim}
tsit       liap       pe
one-CL      pear
\end{verbatim}
'a pear'

Example 8:
\begin{verbatim}
yi  zh\textsuperscript{\ensuremath{a}}ng    piao
one  CL       ticket
\end{verbatim}
'the ticket'

(Ng, 1989, pp. 16, 19)
Despite the similarities in their structure, the number of numeral classifiers in different Chinese language varieties also differs. For example, there are over 200 numeral classifiers in Cantonese (Tse et al., 2007), out of which only 60 numeral classifiers are productively used (Wei & Lee, 2001). In Mandarin, out of 150 numeral classifiers known to adults (Erbaugh, 1984), only 22 are productively used (Wei & Lee, 2001, p. 346). Although there are differences in the number of productive numeral classifiers between Mandarin and Cantonese, the numeral classifiers used “make similar semantic distinctions in the two varieties of Chinese” (Tse et al., 2007, p. 498). All Chinese varieties have general classifiers. They are manifested as ge in Mandarin (Erbaugh, 1984; Hu, 1993) and Hokkien (Ng, 1989), and as go3 in Cantonese (Tse et al., 2007).

The semantics of Chinese numeral classifiers are apparently more complex and opaque in comparison to Japanese. This is because, while in Japanese, the classification between animate and inanimate entities is distinct (i.e., animate items are classified with different numeral classifiers from inanimate items) in Chinese, animateness does not appear to be feature that determines the classification of objects (Uchida & Imail, 1999). For example Mandarin Chinese classifies thin, long, and curvey objects, such as rivers and snakes as tiao, regardless of their animateness.

Among the three varieties of Chinese language, only the numeral classifiers of Mandarin Chinese have been analysed from a pragmatic perspective. In Mandarin
Chinese, numeral classifiers are used in discourse to indicate the thematic status and to foreground nouns (Li, 2000; Sun, 1988). For example, in the excerpt below, numeral classifiers *ge* and *zuo* are used in a Mandarin Chinese text to foreground objects namely, ‘monsters’ and ‘mountain’ (Li, 2000).

Example 10:

> Legend say be very old AS time, there-be one-CL called Youdu
> AS place all year not see sun, everywhere all pitch dark
> In there there-be one-CL big dark mountain mountain top live
> PRT many scary AS monster. Those monsters often descend
> mountain endanger people there-be one-CL giant named Kuafu, he
> use cane with monster fight PRT 9 day 9 night
> finally DI them beat dead

“Once upon a time, in a place called Youdu, people lived in darkness all year round. There was a big black mountain where many terrible beasts lived. The beasts often went out to harm people. There was a giant called Kuafu. He fought with the beasts with a stick for nine days and nine nights. Finally, he killed them all…”

(Li, 2000, pp. 1121-1122)

1.4.3 Thai

Quite a number of studies have investigated Thai numeral classifiers, for example, Haas (1942), Goral (1978), Allan (1977), Carpenter (1986), and Delancey (1986). Numeral classifier is an obligatory syntactic category in Thai, a Southeast Asian language from the Kra-Dai language family. Because Thai nouns cannot be quantified directly, the nouns must always be accompanied by a numeral classifier. They usually occur in the ‘N Num NumCl’ structure, and can also be used with quantifiers *lăj* (many), *kì* (how many), and *baŋ* (some). For example,
Example 11:

phû·jîŋ  să·m  khon
women  three  NumCl
‘three women’

(Caas, 1942, p. 202)

Example 12:

mi:  nòŋ  ki:  khon
have  younger-sibling  how many  NumCl
‘How many younger siblings do you have?’

(Carpenter, 1991, p. 94)

There are altogether 241 numeral classifiers in Thai (Goral, 1978); however, only 40 are commonly found in everyday contexts (Carpenter, 1986). The total number of Thai numeral classifier inventory reported by researchers varies because some numeral classifiers are very exclusive and are only used with one noun. These very specific types of numeral classifiers are homonymous with the nouns they classify, for example, kham (words), myan (city), and hŏŋʔ (room). As a result, these numeral classifiers are labelled as ‘repeaters’ by some researchers (e.g., Caas, 1942).

Example 13:

kham  sì  kham
word  four  word
‘four words’

(Caas, 1942, p. 203)

The semantics of Thai numeral classifiers is very complex. For example, there are five numeral classifiers that can be used with human nouns, and the distribution is dependent on the rank and status of the person. For example, ʔôŋ is used with nouns referring to royalty, rû·b for talapoins, thănʔ for nobles, khon for ordinary persons, and na·j for anyone between the nobles and ordinary persons (Caas, 1942). As for animals, while all animals are classified as tua, elephants, which play a special function in Thai ceremonies, are classified as chŷag.
Thai inanimate numeral classifiers are also complex. Shape-based numeral classifiers are distributed based on dimensionality, namely, 1D, 2D, and 3D. However, the classification of these shape-based numeral classifiers is more complex as for example, 1D objects can be used in five different forms, depending on their size, their volume, their flexibility, their functions, and whether or not the objects can be manipulated by hand (Carpenter, 1991; Haas, 1942). The Thai general numeral classifier ๒ can be used with any inanimate noun to correctly substitute for specific numeral classifiers (Carpenter, 1991; Gandour, Petty, Dardarananda, Dechongkit, & Mukngoen, 1984; Tuaycharoen, 1984).

Thai numeral classifiers can also be used as noun substitutes, that is, as referring expressions, which are semantically translatable as the English pronoun ‘one’, as illustrated in the examples below. In these examples, tua is used to refer to the noun ‘dog’ without the need to repeat the word ‘dog’ unnecessarily.

Example 14:

\begin{verbatim}
chăn mì- mā sŏŋ tua
\end{verbatim}

I have dog two NumCl

‘I have two dogs’

Example 15:

\begin{verbatim}
Tua-niʔ jurisdictions. Tua-nánʔ jurisdictions lég
\end{verbatim}

NumCl big NumCl little

‘This one is big. That one is little’.

(Haas, 1942, p. 204)

1.5 The Malay Numeral Classifier System

Malay, an Austronesian language which is the official language of Malaysia, is also known as Bahasa Melayu and Bahasa Malaysia. Spoken by 20 to 30 million native speakers, Malay is said to be one of the languages that has an extensive numeral classifier system (Richards et al., 1985). Known as penjodoh bilangan
(which literally means ‘matchmaker for number’) in Malay, the system of Malay numeral classifiers is very complex and is taught formally at school. The Malay numeral classifiers are also described as ‘coefficients’ and ‘numeral coefficients’ because of their rigid collocation between nouns and their respective numeral classifiers (Omar & Subbiah, 1995). In some cases, its absence in numeral phrases makes the structure of any formal Malay sentence ungrammatical (Omar, 1972).

1.5.1 The Syntax of Malay Numeral Classifiers

Malay numeral classifiers may sometimes be homonymous to some Malay nouns (Salehuddin & Winskel, 2009). However, despite being homonymous, they do not, unlike the Thai repeaters (Haas, 1942), collocate with each other. To illustrate, while there is a numeral classifier that is homonymous to the word ‘fruit’ in Malay (buah), the noun phrase *tiga buah buah never exists in Malay. Similarly, although the word biji in its nominal sense means ‘seed’, the noun phrase *tiga biji biji would clearly be an unacceptable noun phrase to native speakers of Malay. Instead, the correct noun phrases for both ‘three fruits’ and ‘three seeds’ are tiga biji buah and tiga butir biji respectively.

Structurally, Malay numeral classifiers, like Japanese and Chinese numeral classifiers, exist in a noun phrase, together with a numeral preceding it and a noun following it.

Example: 16:

<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>orang</td>
<td>anak</td>
</tr>
<tr>
<td>Three</td>
<td>NumCl</td>
<td>child</td>
</tr>
</tbody>
</table>
‘three children’
Example 17:
Num NumCl N
_Empat ekor_ arnab
Four NumCl rabbit
‘four rabbits’

When numeral classifiers are used in Malay, the noun that forms the head of the noun phrase appears in a singular form as illustrated above. In situations where numbers are not used to indicate plurality, numeral classifiers are obligatorily deleted (Omar & Subbiah, 1995, p. 23). Such instances happen usually when the nouns are reduplicated (Omar, 1974).

Example 18:
N -pl
_rumah-rumah_
house -pl
‘houses’

Example 19:
N -pl
_buku-buku_
book -pl
‘books’

In isolated cases, instead of appearing to the left of the head noun, numeral classifiers may also appear to the right of the noun, following a numeral. Such a structure is usually found when the different types of nouns are itemised, and when the nouns are emphasised (Goral, 1978).

Example 20:
N Num NumCl
_anak dua orang_
child two NumCl
‘Children: Two’

(DBP Corpus)
Example 21:
N kerbau balar tujuh ekor
umeral classifier
table

buffalo albino seven numeral classifier

‘Albino buffalos: Seven’

(DBP Corpus)

Example 22:
Ibu membeli ayam se ekor, djeruk 20 buah dan telur 10 butir

Mother buy chicken 1 numeral classifier pickle 20 numeral classifier and egg 10 numeral classifier

‘Mother bought a chicken, 20 oranges, and 10 eggs.’

(Goral, 1978, p. 28)

Despite the fact that they may occur contiguously to expressions of quantity
in a noun phrase (Craig, 1986) such as in beberapa orang doktor (a few doctors),
nouns preceded by indefinite quantifiers such as setiap (each, every), ramai, banyak
(many), segala, semua, seluruh, sekalian, para (all), sebahagian, sesetengah (some, several) are never classified in Malay (Hopper, 1986). Thus, although expressions
using definite quantifiers like sepuluh orang warganegara (10 numeral classifier citizens) and
seratus ekor binatang (100 numeral classifier animals) are correct, expressions like *setiap
orang warganegara (*each numeral classifier citizens) and *segala ekor binatang (*all numeral classifier animals) may never take place in any Malay noun phrase ("Tutorial Bahasa Melayu Online," 2000).

1.5.2 The Semantics of Malay Numeral Classifiers

Despite being syntactically different from the norm in some cases, what is
most significant in all the examples above is that the choice of numeral classifiers is
maintained and is consistent for each one of the nouns when they are used with
numerals. For example, the noun ‘children’ co-occurs only with the numeral
classifier orang [animate: human], and the noun ‘buffalos’ co-occurs only with the
numeral classifier ekor [animate: animal].
Although some researchers describe numeral classifiers as modifiers that unitise uncountable objects for quantifying purposes (i.e., measure words (Yamamoto & Keil, 2000) or unitising modifiers (Imail & Gentner, 1997)) such as ‘a spoonful of sugar’, ‘a bowl of rice’, and ‘a cup of coffee’ (Imai & Gentner, 1997), this does not apply to Malay numeral classifiers. While measure words are used to quantify mass nouns, numeral classifiers are used to qualify most\(^2\), and only count nouns. In fact, the kind of measure words mentioned by Imai and Gentner (1997) are also used in Malay expressions such as *sesudu gula*, *semangkuk nasi*, and *secawan kopi*, each bearing the meaning ‘a spoon of sugar’, ‘a bowl of rice’, and ‘a cup of coffee’ respectively. Malay numeral classifiers are also dissimilar to measure words as they have greater stability. The numeral classifier *orang*, for example, may only be used for human beings, and substitutions with any other numeral classifier will result in ungrammatical linguistic performance. On the other hand, the quantifier ‘a spoon of’ in ‘a spoon of sugar’ may always be substituted with other quantifying expressions like ‘a pinch of’, ‘a cup of’, ‘a bowl of’, and ‘a sack of’.

Osman (2002) and Othman (2004) give quite an extensive list of Malay numeral classifiers, together with their respective usage and exemplars (Appendix 1). There appears to be multiple forms of ‘numeral classifiers’ for certain nouns in the list, for example, *orang, angkatan, baris, gerombolan*, and *pasang* for human; *biji, tandan, pangsa*, and *sisir* for fruits; *buah, baris, deret*, and *pintu* for house; and *batang, ikat*, and *kotak* for pencil. A closer analysis of the words listed as ‘numeral classifiers’ by both Osman and Othman reveals that except for *orang, biji, buah*, and *batang* which are true classifiers for humans, fruits, houses, and pencils respectively, the rest of the words in the list of numeral classifiers listed in Osman (2002) and

\(^2\) There are nouns in Malay that are left unclassified, for example, the body parts ‘eye, nose, ears’.
Othman (2004) are not exactly true classifiers. While orang, biji, buah, and batang qualify the nouns by highlighting the characteristics of the respective nouns, words like angkatan, tandan, baris, and ikat do not. These words only describe the manner in which the objects in question may be presented when they appear in larger quantities. As a result, words such as angkatan, tandan, baris, and ikat cannot be used to indicate a particular noun as a single unit (e.g., sebaris pelajar means ‘a row of students’ and not ‘*a row of student’). Words like baris and angkatan are almost equivalent to the English ‘collective nouns’ and appear to be less restricted in their collocation in comparison to the numeral classifiers. Collective nouns can be used with any nouns, for example, sebaris pelajar, sebaris pokok, sebaris rumah (a row of students, a row of trees, a row of houses); but numeral classifiers can only be used with specific types of nouns, for example, seorang pelajar, sebatang pokok, sebuah rumah (a student, a tree, a house). These collective nouns (kata nama kelompok) syntactically never appear side-by-side with the Malay numeral classifiers. This is because, semantically, the presence of one form denies the presence of the other form (Hopper, 1986), that is, they are in complementary distribution.

Correct:

Example 23:

<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>dua</td>
<td>buah</td>
<td>meja</td>
</tr>
<tr>
<td>two</td>
<td>NumCl</td>
<td>table</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘two tables’</td>
</tr>
</tbody>
</table>

Example 24:

<table>
<thead>
<tr>
<th>Num</th>
<th>Collective N</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>dua</td>
<td>baris</td>
<td>meja</td>
</tr>
<tr>
<td>two</td>
<td>rows of</td>
<td>table</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘two rows of tables’</td>
</tr>
</tbody>
</table>
The classification of some items in Malay appears to be rather straightforward. In *dua orang kanak-kanak* (two NumCl children), *dua ekor kucing* (two NumCl cats), and *dua batang pokok* (two NumCl trees), the classification of object seems to be relatively semantically transparent. This is because the numeral classifiers for humans are classified with *orang*, which literally means ‘human’ in Malay, animals with *ekor*, which literally means ‘tail’, and trees with *batang*, which literally means ‘stem’.

However, its usage in *dua buah kereta* (two NumCl cars), *dua biji cawan* (two NumCl cups), and *dua kaki payung* (two NumCl umbrellas), is rather difficult to explain (Salehuddin, 2007). Because of this, the acquisition of other Malay numeral classifiers appears to be more opaque and cognitively demanding. At superficial level, there appears to be many exceptions to how objects are classified into the numeral classifiers. These are evident in, for example, the classification of cars, houses, and concepts with the numeral classifier *buah*; cups, and stones with *biji*; and umbrellas and some plants with *kaki*. The seemingly arbitrariness in the classification of objects is possibly due to the fact that these numeral classifiers are homonymous to Malay nouns. For example, while *buah* in Malay also means fruit; the numeral classifier *buah* does not classify fruits. Similarly, *biji* (Malay “seed”) is not used as a classifier for seeds and neither is *kaki* (Malay “leg”) used as a classifier for legs. This seemingly arbitrariness nature of the majority of Malay numeral
classifiers suggests that Malay numeral classifiers are often regarded as loose and opaque (or "semantically non-transparent" as described by Omar, 1972, p. 89), which presumably results in greater difficulty on the part of the learner in acquiring them.

Table 1.1 illustrates the collocation of the Malay numeral classifiers and the nouns.

Table 1.1 Some Malay Numeral Classifiers and Examples

<table>
<thead>
<tr>
<th>Malay numeral classifier</th>
<th>Semantic description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>batang</td>
<td>long things</td>
<td>river, candle, pen, tree, tooth, needle</td>
</tr>
<tr>
<td>beradak</td>
<td>small circular things</td>
<td>fishing hook, ear, ring, ring</td>
</tr>
<tr>
<td>bidang</td>
<td>items that are spread out</td>
<td>hand, cloth, carpet, sail, mat</td>
</tr>
<tr>
<td>bii</td>
<td>fruits, vegetables, small items</td>
<td>pumpkin, carrot, egg, cup, bull, balloon</td>
</tr>
<tr>
<td>bilih</td>
<td>sharp items</td>
<td>knife, scissors, scythe, machete</td>
</tr>
<tr>
<td>buah</td>
<td>big, solid items</td>
<td>car, country, house, books, box, radio</td>
</tr>
<tr>
<td>buku</td>
<td>squashish items</td>
<td>soap, bread,</td>
</tr>
<tr>
<td>butar</td>
<td>fine, spherical items</td>
<td>eggs, pebbles, beans, rice</td>
</tr>
<tr>
<td>dus</td>
<td>item accompanied by explosive sounds</td>
<td>shots,</td>
</tr>
<tr>
<td>ekor</td>
<td>all animals</td>
<td>mosquito, ant, horse, hare</td>
</tr>
<tr>
<td>helai</td>
<td>thin, spread out</td>
<td>leaf, shirt, paper,</td>
</tr>
<tr>
<td>kali</td>
<td>items with &quot;legs&quot;</td>
<td>umbrella, mushrooms</td>
</tr>
<tr>
<td>keping</td>
<td>thin items</td>
<td>wooden plank, card, picture</td>
</tr>
<tr>
<td>kepal</td>
<td>clunky</td>
<td>ice, meat, soap</td>
</tr>
<tr>
<td>hantung</td>
<td>flower</td>
<td>rose, all flowers</td>
</tr>
<tr>
<td>kawas</td>
<td>firearms</td>
<td>pistol, short gun</td>
</tr>
<tr>
<td>kebun</td>
<td>long, straight</td>
<td>paper, ware, thread</td>
</tr>
<tr>
<td>kakas</td>
<td>compiled</td>
<td>newspaper, brochures, magazine, books</td>
</tr>
<tr>
<td>orang</td>
<td>people</td>
<td>teacher, nurse, doctor</td>
</tr>
<tr>
<td>pepon</td>
<td>pieces</td>
<td>chocolate, petal, fire crackers</td>
</tr>
<tr>
<td>potong</td>
<td>words</td>
<td>words, bread</td>
</tr>
<tr>
<td>pucuk</td>
<td>sliced</td>
<td>meat, bread</td>
</tr>
<tr>
<td>puntung</td>
<td>thin &amp; fine; firearms</td>
<td>needles, bamboo shoots, letters, gun, rifle</td>
</tr>
<tr>
<td>rangkap</td>
<td>with flames</td>
<td>cigarette, firewood</td>
</tr>
<tr>
<td>rencon</td>
<td>verse</td>
<td>poem, proses</td>
</tr>
<tr>
<td>rongkol</td>
<td>fine, entangled, long</td>
<td>net</td>
</tr>
<tr>
<td>urot</td>
<td>corn</td>
<td>corn</td>
</tr>
<tr>
<td>urat</td>
<td>fine, long</td>
<td>hair, wire</td>
</tr>
<tr>
<td>utas</td>
<td>long</td>
<td>rope, chain</td>
</tr>
</tbody>
</table>

As shown in the semantic description column of Table 1.1, although the distribution of Malay numeral classifiers may seem arbitrarily motivated, the distribution is actually based on the nature, physical features, and types of objects. However, there may also be Malay nouns that may have alternative choices of numeral classifiers, depending on which property of the noun is in focus. For example, ‘bread’ is classified as bii if it is ‘bun-like’, as buku if it is ‘a loaf of bread’, and as keping if it is ‘a slice of bread’. 
Like Japanese, the Malay numeral classifiers are divided into two major categories, i.e. animate and inanimate categories. The semantic criteria for dividing the system of each Malay inanimate numeral classifier category however, appears to be complex. For example, while in Japanese shape-based numeral classifiers classify objects based on only the dimensionality of the objects, like the Chinese numeral classifiers, the classification of objects in Malay involves mixed semantic criteria in classifying members of a given category. However, the mixed semantic criteria of classification in Malay are not fully like Chinese. While the classification of items in Chinese disregards animateness (e.g., the same classifier *tiao* for thin, long, and curvey items like rope and snake), the mixed semantic criteria of the classification of items in Malay takes animateness into account. As a result, in Malay, rope is classified with *utas* whereas snake with *ekor*. This is further discussed in 2.3.

Most numeral classifier languages have general classifiers. The general classifiers are often used as default (one resorts to it when in doubt, see Pinker and Ullman, 2002) numeral classifiers in place of unknown classifiers. In most cases the use of these general classifiers is acceptable and regarded as grammatical. All Chinese varieties have general classifiers. They are manifested as *ge* in Mandarin (Erbaugh, 1984; Hu, 1993) and Hokkien (Ng, 1989), and as *go3* in Cantonese (Tse et al., 2007). In Thai, the general classifier *?an* may be used with any inanimate noun to correctly substitute for specific numeral classifiers (Carpenter, 1991; Gandour et al., 1984; Tuaycharoen, 1984). With a varying degree of acceptability, the Japanese numeral classifier *-tsu* and the shape-specific classifier for 3D object–*ko* (Yamamoto & Keil, 2000) can be used in place of almost all inanimate objects.

Unlike Chinese, Japanese, and Thai, Malay does not have an ‘unmarked general classifier’ that can be used to refer to a large number of nouns without
affecting the grammaticality of the sentence. Although Omar (1972) suggests that *buah can be considered as a general numeral classifier, this proposition does not hold because a noun phrase such as *tiga *buah *pensil (three pencils) and *tiga *buah *kertas (three pieces of paper) are regarded as grammatically incorrect. *Buah may appear to function like a general numeral classifier due to the fact that many objects are classified with *buah (a numeral classifier for rounded, big objects) and that *buah is a numeral classifier which is often used by speakers when they are in doubt as to which numeral classifier to use (Hopper, 1986; Omar, 1972). Hopper (1986) also claims that *suatu is also likely to be a Malay general classifier; however such a proposition is unacceptable as well. This is because *suatu is not used to categorise objects; in fact, *suatu is used more often with temporal words mostly in classical Malay (e.g., *hari [day], *masa *dahulu [long ago], and *petang [afternoon]), and with unspecified objects (e.g., *suatu *objek [an object], *suatu *lembaga [a figure]). Rarely are phrases such as *suatu *rumah (a house) and *suatu *telefon (a telephone) found in Malay discourse. Furthermore, if *suatu is said to be a general classifier which is derived from “*se-watu” (Hopper, 1986, p. 8), the usage of *dua *watu, *tiga *watu would be prevalent in Malay discourse. In addition, *suatu cannot be collocated with a plural noun, e.g. *tiga *watu *rumah (three *watu house) or *dua *watu *telefon (two *watu telephone) (Examples 26-29). Due to its inability to collocate with plural nouns, *watu (or *suatu) cannot be considered as a numeral classifier in Malay. As a result, there is no one particular Malay numeral classifier that may ‘fluidly’ be used to substitute other numeral classifiers without distorting the syntax and semantics of the language.
Example 26:
Correct:
<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>buah</td>
<td>rumah</td>
</tr>
<tr>
<td>Three</td>
<td>NumCl</td>
<td>house</td>
</tr>
</tbody>
</table>

(Three houses)

Example 27:
Correct:
<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dua</td>
<td>buah</td>
<td>telefon</td>
</tr>
<tr>
<td>Two</td>
<td>NumCl</td>
<td>telephone</td>
</tr>
</tbody>
</table>

(Three telephones)

Example 28:
Incorrect:
<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>watu</td>
<td>rumah</td>
</tr>
<tr>
<td>Three</td>
<td>NumCl</td>
<td>house</td>
</tr>
</tbody>
</table>

(Three houses)

Example 29:
Correct:
<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dua</td>
<td>watu</td>
<td>telefon</td>
</tr>
<tr>
<td>Two</td>
<td>NumCl</td>
<td>telephone</td>
</tr>
</tbody>
</table>

(Three telephones)

1.5.3 The Pragmatics of Malay Numeral Classifiers

Becker (1986) suggests that a numeral classifier is “an overlay of constraints - structural, generic, pragmatic, and referential ... this interplay between systems is quite particular and cannot be seen or described apart from particular contexts” (Becker, 1986, p. 331). Apart from Becker’s analysis of the pragmatic functions of Burmese numeral classifiers, other cross-linguistic studies have also shown that numeral classifiers of Thai, Japanese, and Chinese are also used in those languages to perform several pragmatic functions in various types of discourse (e.g., Carpenter, 1986; Downing, 1986; Li, 2000; Sun, 1988).
Malay numeral classifiers can also perform several pragmatic functions through the presence and absence of numeral classifiers in various contexts. In classical Malay, the presence and absence of numeral classifiers can indicate either the foregrounding-backgrounding functions, or the degree of definiteness-indefiniteness of objects in a particular discourse (Hopper, 1986). Using the Autobiography of Abdullah bin Abdul Qadir “Munshi” (i.e., Hikayat Abdullah) as his corpus, Hopper (1986) illustrates how Malay numeral classifiers have been used to foreground objects in a discourse. For example, in the following excerpt (Example 30), the numeral classifiers *buah* [3D: big] and *batang* [1D: +rigid] were used as a strategy to foreground elements that describe a place – ‘The Esplanade’. This was because the sentence was later followed by the sentence: “Mr. Farquhar walked all around the Esplanade and the Orang Laut came and looked at him…”.

Example 30:

_Maka se-belah arah ka-tepi sungai itu ada empat lima BUAH pondok kechil, serta ada tanaman-nya enam tu Joh BATANG pokok kelapa; dan lagi ada sa-BUAH rumah sedikit besar (138)_

‘And on the side leading toward the river there were four or five <CL> small huts, and in addition six or seven <CL> coconut trees in cultivation; furthermore there was a <CL> somewhat larger house…’

(Hopper, 1986, pp. 311-312)

On the other hand, Hopper suggests that the absence of numeral classifiers in the excerpt from _Hikayat Abdullah_ below are merely “incidental” and “props” to the setting. This is because the nouns (i.e., small huts, huts) were not developed further in the text (Hopper, 1986, p. 313).

Example 31:

_Maka ada ringgit hendak membeli, tiada dapat; ada-lah dua tiga <0> pondok kechil-kechil bersama-sama dekat rumah Temenggong, sakalian itu memakan tarok kayu, dan ikan kering, dan sagu rending, terkadang-kadang mendapat beras. Maka ada pula di-hujong Kampong Gelam dua tiga <0> pondok-pondok orang laut… (141)_
‘They had money to buy food, but there was nothing to buy. All there was were two or three \(<0\) small huts close together near the Temenggong’s house, and their occupants ate the sprouts of trees and dried fish and sago, and occasionally they got some rice. At the far end of Kampong Gelam there were two or three \(<0\) huts belonging to the Orang Laut (“Sea Gypsies”)…’

(Hopper, 1986, p. 312)

Hopper (1986) later illustrates that Malay numeral classifiers were also used in *Hikayat Abdullah* to create the sense of definiteness of the objects in question. Conversely, the absence of numeral classifiers creates the sense of indefiniteness of the objects mentioned in the discourse.

Example 32:
*Maka ku-dapati ada di-tengah rumah ada sa-ORANG orang Pelekat yang bernama Abdul Satar tengah makan* (44)

‘and I discovered that there was a <CL> man from Pelekat by the name of Abdul [sic] Satar in the middle of the room engaged in eating’

Example 33:
*Maka bapa-ku di-jadikan-nya \(<0\) nakhoda dalam sa-buah perahu terlalu besar* (5)

‘and they made my father <0> captain of a <CL> very big ship’

(Hopper, 1986, p. 314)

Using a 73,000-word modern day published corpus written by expert Malay language users, which includes a one-day publication of *Utusan Malaysia online*, an online version of mainstream Malay newspaper in Malaysia (December 4, 2007), 227 Malay folk stories from the web page of *Dewan Bahasa dan Pustaka*, and written data from thirteen in-class activity packages made available by *Berita Harian online*, Salehuddin, Winskel, and Maros (2008) show that the presence of numeral classifiers in modern Malay discourse also helps readers to make anaphoric and cataphoric reference to nouns i) within the same noun phrase boundary, ii) beyond the noun
Numeral classifiers are also used in Malay texts as a cataphoric reference, that is, a reference that refers to another word or phrase that will be used later in the same text (Richards et al., 1985). For example, in the following excerpt, “seorang” is used as a reference to the noun “lelaki” (men) within the same noun phrase “salah seorang daripada lelaki terbabit”. The use of the orang [animate: human] numeral classifier in the noun phrase enables the word “lelaki” to be omitted from what originally ought to be “salah seorang lelaki daripada lelaki terbabit”. The use of numeral classifiers as cataphoric reference within the same noun phrase is common in modern Malay when the noun phrase is translatable into the English structure as
’one of the N’. In such a case, ‘one’ is most of the time manifested in modern Malay as ‘salah se- NumCl’.

Example 35

Drug ketamin was found on bed in room house the while shabu was found in wallet one NumCl of men involved.

‘Catamin was found on the bed in the room of the house while shabu was found in the wallet of one of the men.’

("Adik VIP direman sehari,” 2007)

Salehuddin et al. (2008) also found that similar to the proposition made by Hopper (1986), the presence of numeral classifiers in modern Malay discourse also helps trigger the sense of definiteness and foregrounding of objects within a text. Conversely, as put forward by Hopper (1986), Malay numeral classifiers are also omitted in modern Malay discourse to create the sense of indefiniteness and backgrounding of objects within a text.

In addition to what Hopper (1986) has found in Hikayat Abdullah, Salehuddin et al. (2008) also found that the presence of Malay numeral classifiers in modern Malay texts may also be used to indicate the notion of indefiniteness of a noun in two different ways. To illustrate, in the following excerpt, “seorang” is used with the ‘generic’ sense similar to the meaning that is conveyed by the English indefinite article ‘a’. In the following sentence, although the person in the excerpt hoped to be a pilot, he made no specific indication as to the kind of pilot (i.e., a pilot who flies a jet in comparison to a pilot who flies a light aircraft) that he had in mind, and this was expressed via the word “seorang”.

- 35 -
Example 36
Menurutnya, beliau memang menyimpan cita-cita untuk menjadi seorang juruterbang dan tidak menolak apabila ditawarkan oleh Red Arrows kerana mahu merasai sendiri pengalaman itu.

‘According to him, he has always wanted to be a pilot and did not reject the offer from Red Arrows to experience flying their plane.’

("Sheikh Muszaphar pandu pesawat Red Arrows,” 2007)

Salehuddin et al. (2008) found that ‘menjadi seorang N’ (to become a N) structure is a widely used structure in Malay corpus that out of a 2.6 million-word Dewan Bahasa dan Pustaka corpus analysed (Salehuddin et al., 2008), a total of 84 instances of such a structure was identified, for example, “Ramlah bercita-cita hendak menjadi seorang pegawai penyelidik untuk mengkaji bukan sahaja tentang burung tetapi juga haiwan lain” (loosely translated as ‘Ramlah dreams of becoming a research officer to investigate not only about birds, but also about other animals as well.’) (Alaudin & Sulaiman, 1998).

In the following excerpt, the numeral classifier buah co-occurs with the prefix ‘sese-’ in “sesebuah”, bearing the ‘generic’ notion similar to the meaning that the English indefinite pronoun ‘any’ denotes. The kind of ‘indefiniteness’ that ‘any’ indicates will not be attained without the use of the ‘sese + NumCl’ structure.

Example 37:
Menurut Pengarah ISM, Prof. Madya Dr. Mohamed Fadzil Che Din, according to Director ISM, Prof. Assoc. Dr. Mohamed Fadzil Che Din, sains sosial merupakan bidang terpenting dalam proses pembangunan sesebuah tamadun, bangsa dan negara. Science social is field most important in process development any NumCl civilization, race and nation.
‘According to the Director of ISM, Associate Prof. Dr Mohamed Fadzil Che Din, social science is the most important field in the process of developing any civilization, race, and nation.’

("Sains sosial corakkan pembangunan insan," 2007)

One might argue that the ‘sese-NumCl’ form is not widely used in modern Malay corpus. However, Salehuddin et al. (2008) indicate that out of the 2.6million-word corpus they analysed, a total of 710 instances of ‘sesebuah’ and 1552 instances of ‘seseorang’ was identified. For example, “Dalam sesebuah karangan, penutup berfungsi sebagai kesimpulan kepada perbincangan yang dibuat.” (loosely translated as, ‘In any essay, the conclusion concludes the discussion made earlier in the essay.’) (Daud, 1998).

1.6 Summary

In this chapter we have provided an overview of numeral classifiers, including the syntax, semantics and pragmatic functions of numeral classifier languages, namely Japanese, Chinese and Thai, before comparing them with Malay. What follows next is a chapter on numeral classifiers and discussions on numeral classifiers as a manifestation of categorisation, and literature on previous numeral classifier acquisition studies.
CHAPTER 2
The Categorisation and Acquisition of Numeral Classifiers

2.1 Categorisation, Cognition, and Numeral Classifiers

The act of categorisation is one of the most basic processes of human cognition (Croft & Cruse, 2004; Lakoff, 1987) and the conceptualisation of human categorisation is manifested through language, overtly evident either in the form of noun classes or noun classifiers (Aikhenvald, 2003). Classifiers are, as Allan (1977) describes, “linguistic correlates to perception” (p. 308), since it is believed that “perceptual mechanisms underlie conceptual process” (Goldstone & Barsalou, 1998, p. 256). The general phenomenon of human categorisation may be grasped by studying classifier systems (Craig, 1986). Because categorisation in numeral classifier systems is usually determined by the inherent semantic properties of nouns, the acquisition pattern of numeral classifiers in any language can be linked to the underlying patterns of semantic and conceptual development of that particular language (Yamamoto & Keil, 2000).

2.1.1 Categorisation and Cognition

Categorisation takes place consistently in cognition. It happens whenever distinguishable objects or events are treated as one, usually through an object-naming process at the super-ordinate level of any semantic taxonomy (Mervis & Rosch, 1981). In the process of categorisation, ideas and objects are acknowledged, differentiated, and comprehended and as a result, the act of categorising further
illuminates the relationship between the subjects and objects of knowledge in a particular category (Harnad, 2003).

Although categorisation appears to take place quite effortlessly, it is not done arbitrarily. The correlations that exist within non-identical objects are usually discovered through the speakers’ own culture and through these correlations categories are formed (Mervis & Rosch, 1981; Rosch, 1987; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Besides the speakers’ own culture, categorisation is also influenced by a large array of different kinds of information (Keil, 2003) dependent on the amount of knowledge the speakers have about a particular object (Rosch et al., 1976). In short, categorisation refers to the organisation of lower-order categories into higher-order categories based on how the feature distributions overlap (Hills, Maouene, Maouene, Sheya, & Smith, 2009; Rogers & McClelland, 2004; Rosch et al., 1976).

Categorisation reduces the infinite differences among stimuli by simplifying the environment (E. M. Markman, 1989; Rosch et al., 1976). When the stimulus is simplified, the load on human memory is reduced; which results in more efficient information storage and retrieval (E. M. Markman, 1989). According to Hammer, Diesendruck, Weinshall and Hochstein (2009), category learning is achieved when children are able to identify the shared features among category members, and/or when they are able to identify the distinctive features among category non-members. In the process of categorisation, objects that are perceived as similar are sorted into the same category (E. V. Clark, 2003; Mervis, 1987) and concurrently, those that are considered dissimilar are categorised into different categories (Hampton, 1998). In the categorisation of subsequent or novel objects, the relevant information on any related object that has been stored earlier is retrieved and evaluated. Only when
there is a satisfactory resemblance between the new and the stored information will
these subsequent or novel objects be accepted into the respective categories
(Barsalou et al., 1998; Rogers & McClelland, 2004).

Preferential-looking, object-examination, and generalised imitation tests
conducted on infants indicate that categorisation begins to take place in human
cognition even before infants reach their first birthday (Madole & Oakes, 1999;
Needham, Cantlon, & Holley, 2006; Oakes & Madole, 2000; Saxe, Tzelnic, & Carey,
2006). During the course of their first year, infants are already able to demonstrate
their understanding of global conceptual categories by differentiating animals,
vehicles, furniture, and plants from each other (Mandler, 2004). Young children’s
and infants’ consistent discrimination of animate items from inanimate ones, and
their discrimination of inanimate items from human beings (Greif, Nelson, Keil, &
Gutierrez, 2006; Molina, Walle, Condry, & Spelke, 2004; Simons & Keil, 1995) are
described as an early type of categorisation. This kind of categorisation is referred to
as taxonomic categorisation, that is, “subdivision into kinds” (Tversky, 1989, p. 54),
which, according to Mandler (2004), forms a basis for language acquisition to take
place among infants. According to this viewpoint, only when children are able to
make between-category discrimination (e.g., vehicles from furniture) will they then
be able to make within-category discrimination (e.g., one vehicle from another
vehicle) (A. B. Markman & Wisniewski, 1997).

Research indicates that young children are strongly influenced by perceptual
features when they categorise objects (Gentner, 1978; Gentner & Imai, 1995; Imai,
Although young children’s categorisation of objects is strongly influenced by the
shape of the objects (Gentner, 1978; Gentner & Imai, 1995; Imai et al., 1994; Landau
The Categorisation and Acquisition of Numeral Classifiers

et al., 1988; L. B. Smith & Sera, 1992), shape is not the only aspect that influences categorisation. According to Keil (2008, p. 217), there are “rich and diverse array of processes in both perception and cognition” that have a varied degree of influence on categorisation and labelling, and influence how children acquire and use these labels (see also Booth & Waxman, 2008; Markson, Diesendruck, & Bloom, 2008). For example, for visually impaired individuals, the texture and hardness of objects play a more prominent role in categorising objects (Klatzky, Lederman, & Reed, 1987).

Maddox and colleagues (e.g. Maddox, 1992; Maddox & Ashby, 1993; A. B. Markman & Maddox, 2003) found that children categorise objects based on the number of perceptual feature differences; the more differences there are between the stimuli, the easier it is for children to categorise the objects. Thus, both similarities and differences in perceptual features help children to categorise objects into their respective categories.

The cumulative number of semantic features in a particular semantic category is also said to play a role in children’s categorisation patterns (E. V. Clark, 1973a; Daneman & Case, 1981; Matsumoto, 1987). According to this theory, words or categories with less semantic features are acquired earlier than those with more semantic features. For example, in a comprehension test on 1;6- to 6-year-olds, E.V. Clark (1973a) found that children comprehended the preposition in earlier than they comprehended on and under. Clark attributed her findings to the argument that in is cognitively simpler than either on or under because in requires minimal adjustments of children’s hypothesis about its meaning. In a labelling of novel actions experiment, Daneman and Case (1981) found that the novel word for a single-actor action was produced earlier than the novel word for actor-plus-spectator action,
which in turn was produced earlier than the novel word for actor-plus-double spectator action.

2.1.2 Categorisation and Numeral Classifiers

Numeral classifier systems can be viewed as a manifestation of human conceptualisation through linguistic categorisation (Schmitt & Zhang, 1998). Numeral classifiers acknowledge the membership of a referent in a particular category (Downing, 1984) via the use of the same label for individual objects (Mervis & Rosch, 1981). Speakers of numeral classifier languages need to learn how to categorise objects in their environment and pair them with appropriate labels that highlight the more prominent characteristics of the objects in question (Allan, 1977).

Numeral classifier acquisition patterns have often been viewed as “providing a potentially important source of information about underlying patterns of semantic and conceptual development” (Yamamoto & Keil, 2000, p. 381. See also Adams & Conklin, 1973; Carpenter, 1991; Craig, 1986; Matsumoto, 1987). This is because numeral classifiers represent a system that classifies nouns on the basis of their referents’ permanent attributes.

Children have been found to classify animate nouns earlier than inanimate nouns (e.g., Burling, 1959; Erbaugh, 1986; Gandour et al., 1984). This may be because, as previous studies have claimed, the concept of animacy is one of the first concepts to be acquired by children (Gelman & Coley, 1990; Mandler, Bauer, & McDonough, 1991).

Numeral classifier systems in languages like Japanese, Korean, and Thai classify nouns in terms of the parameter of animateness (i.e., they make a distinction
between animate and inanimate items), and because of this, this level of
categorisation in human cognition has been described as universal (Adams &
Conklin, 1973; Allan, 1977). However as is apparent from cross-linguistic research
on numeral classifiers, fairly large discrepancies in categorisation systems exist,
especially in relation to the manner numeral classifiers are further subcategorised.
The categorisation of numeral classifiers largely cuts across global taxonomic
categories based on the attributes of their referents from the native speakers’
perspective (T'sou, 1976) resulting in language partitioning differences (Yamamoto
& Keil, 2000).

The discrepancy between the conceptualisation of Chinese and Japanese
numeral classifiers (e.g., Figures 1.2 through 1.5), appears to disagree with the more
traditional view that the categorisation system is universally pre-structured. For
example, while Chinese does not distinguish objects based on their animateness,
Japanese does take animateness into account in classifying objects. In addition,
Chinese takes into account rigidity when classifying long and flat objects; Japanese
does not. The discrepancy in categorising the same objects in many different ways
suggests that learning is involved in the classification of objects by singling out some
features that are perceived via the senses, and ignoring others (Harnad, 2003). This
claim supports the view that linguistic expressions in different languages are actually
a manifestation of non-linguistic conceptual knowledge filtered by the specific
linguistic categories that exist in a particular language (Slobin, 1985); and that
language learning actually highlights what is semantically salient to native speakers
of the particular language (Bowerman, 1985; Hunt & Agnoli, 1991; Winskel, 2007).
A persistent interaction between non-linguistic concepts and semantic categories in
the input language exists in language acquisition (Choi & Bowerman, 1991). With
regard to numeral classifiers, according to E. V. Clark (1976, cited in Gentner & Bowerman, 2009, p. 468), “both language acquisition and numeral classifier semantics are influenced and constrained by the same cognitive biases”.

In linguistic category learning studies, researchers have also found that typical exemplars of the category are often acquired first (E. M. Markman, 1989; Mervis, 1980; Rosch, 1973; Shirai & Andersen, 1995). In addition, Rosch (1973) found that the reaction times (RTs) to typical exemplars were much faster than the RTs to atypical exemplars among 10-year-olds in comparison to adults, indicating that typical exemplars are learned earlier than the atypical ones. According to Shirai and Andersen (1995), only when typical exemplars are acquired can the less typical or marginal exemplars be learned. Since young children tend to categorise objects based on perceptual features (Gentner, 1978; Gentner & Imai, 1995; Imai et al., 1994; Landau et al., 1988; L. B. Smith & Sera, 1992), they tend to initially restrict the category label to the more typical members (Mervis, 1987; Mervis, Catlin, & Rosch, 1976; Mervis & Rosch, 1981; Rosch et al., 1976). As a result, in immature categorisation, atypical exemplars that do not share properties relative to adult category prototypes get excluded from the category, whereas out-of-category instances that do share these properties are inappropriately included (Rogers & McClelland, 2004).

2.2 Typicality in Categorisation and Numeral Classifier Exemplars

According to prototype and exemplar theories, the learning and representation of categorisation is claimed to be normally through exemplars or prototypes (Barsalou et al., 1998). The mental representation of a category is either based on a prototypical exemplar (e.g., ‘robin’ for the bird category), or based on the
composition of various exemplars of the category (e.g., robin, chicken, eagle, ostrich, penguin for the bird category) (Dopkins & Gleason, 1997). (See also Johansen and Palmeri (2002) for the various models of category learning which may be used depending on categorisation experience of the individual). As opposed to the classical theory of categorisation that takes into account only the criterial definitional features for membership qualification using category-defining rules, prototype theory, which was developed in cognitive psychology, presupposes a graded membership within a category (Shirai & Andersen, 1995).

There have been various approaches to describing prototypes. Jackendoff (1983), for example, proposes a two-type prototype condition, namely the centrality condition and the typicality condition. The centrality condition defines the central value for characteristics with continuous values on a scale based on the chosen characteristics; for example, the colour ‘red’ may be defined in terms of the values on a scale of hue, brightness, and saturation. In contrast, the typicality condition allows exceptions in defining categories. Thus, although a penguin cannot fly like the majority of birds can, a penguin is still considered ‘a bird’ because it shares other features that a typical bird has, for example, two legs, two wings, a beak, etc..

According to Rosch (1987), prototypes are determined by the speakers’ judgements of goodness of membership in a given category, with the membership of the category normally ranked from the prototype(s) to the peripheral(s). A member with more attributes in common with other members of the category, and with more dissimilarities with members of contrasting categories, is graded as a more prototypical member, or the most typical member (the best exemplar), of a particular category, and reflects the “redundancy structure of the category as a whole” (Rosch & Mervis, 1975, p. 602). Conversely, any members on the borderline, namely those
having less features in common with other members within the same category, especially with the most typical member, is graded as an atypical member of a category (Matsumoto, 1985). A penguin, for example, is likely to be an atypical exemplar of the ‘bird’ category in comparison to an eagle, because the former has fewer similarities with other members of the ‘bird’ category than the latter. Rather than being an atypical exemplar of the ‘fish’ category for its ‘ability to swim’, a penguin shares more features in common with members of the ‘bird’ category than with ‘fish’ category.

Typical exemplars of a category are usually distributed nearest to the prototype (Ashby, Boynton, & Lee, 1994). They are usually the first exemplars of a linguistic category learned by children (E. M. Markman, 1989; Mervis, 1980; Mervis, Catlin, & Rosch, 1975; Rosch, 1973; Shirai & Andersen, 1995). Studies have found that the degree of typicality of objects does play a role in categorising objects. According to Hampton and Mervis and colleagues (Hampton, 1998; Mervis & Pani, 1980; Rosch & Mervis, 1975), the redundancy of semantic features between typical exemplars and members of the same numeral classifier category facilitates categorisation. Objects are categorised easier and faster when there are more overlaps between the stored and the succeeding objects (Mervis & Pani, 1980; Rips, Shoben, & Smith, 1973; Rogers & McClelland, 2004; Rosch, 1973). This corresponds to the proposition that the learning and representation of categorisation is normally done through exemplars and prototypes (Barsalou et al., 1998). As a result, typical exemplars are likely to be named first when subjects are asked to recall all members of a category (Mervis et al., 1976). In addition, children have been observed to be more accurate in categorising tasks when the instances given are
The Categorisation and Acquisition of Numeral Classifiers

typical rather than atypical objects of the particular category (Gelman & Coley, 1990; Hampton, 1998).

Typical exemplars are also categorised relatively quickly in comparison to atypical exemplars due to their high similarity to members of the same category and their high dissimilarity with members of different categories (Rogers & McClelland, 2004). For example, typical exemplars usually receive the fastest yes responses when subjects are asked to verify if a stimulus is a member of a category or otherwise (Rips et al., 1973; Rosch, 1973). This is because, the amount of time taken to perform categorisation depends on to what extent the features of the stored and the succeeding objects overlap with each other; the more they overlap, the faster they are categorised. As a result, the more similar a stimulus is to the prototype of a category, the faster the reaction times for correct response is. The RTs to typical exemplars among the 10-year-olds in Rosch’s (1973) study were faster than the RTs to atypical exemplars, indicating that typical exemplars are learned earlier than the atypical ones.

Since both similarities (e.g. E. V. Clark, 2003; Mervis, 1987) and differences (e.g. Maddox, 1992; Maddox & Ashby, 1993; A. B. Markman & Maddox, 2003) in perceptual features help children to categorise objects into their respective categories, the disparity between typical exemplars and members of different categories also facilitates categorisation especially among children (Mervis & Pani, 1980; Rips et al., 1973; Rogers & McClelland, 2004; Rosch, 1973). This is because, according to Mervis and Pani (1980), the typical exemplars of a particular category have “maximal within-category similarity and minimal extra-category similarity” (p. 518)(p. 518)(p. 518)(p. 518)(p. 518). For example, categorising a car as a
type of vehicle is easier than categorising a *motor-home* as a type of vehicle since a
*motor-home* may also be categorised as a type of dwelling.

Young children tend to categorise objects based on perceptual features (e.g.,
Gentner, 1978; Gentner & Imai, 1995; Imai et al., 1994); and because of this, young
children tend to initially restrict the category label to the more typical members
(Mervis, 1987; Mervis et al., 1976; Mervis & Rosch, 1981; Rosch et al., 1976). This
is because, the “redundancy structure of the category as a whole” (Rosch & Mervis,
1975, p. 602) is not often reflected in atypical members of a category. As a result, in
immature categorisation, atypical exemplars that do not share properties relative to
adult category prototypes get excluded from the category, whereas out-of-category
instances that do share these properties are inappropriately included (Rogers &
McClelland, 2004).

Typical and atypical members of a category are also established through
familiarity (Croft & Cruse, 2004). According to Johnson and Mervis (1994),
children’s familiarity with objects increases through social everyday interactions.
Objects that are least frequently encountered by children are poorly recognised and
categorised (Lederman, Klatzky, Chataway, & Summers, 1990). Their ability to
categorise is usually expended to less typical exemplars only after the more typical
exemplars are acquired (E. M. Markman, 1989; Shirai & Andersen, 1995). Children
take an extended period of time before achieving adult-like performance in
categorisation, usually through actual exposure to usage in input (Matsumoto, 1985).

Quite a number of studies on numeral classifier acquisition show that the
acquisition of a numeral classifier begins with the prototype of the category and later
proceeds to the less prototypical exemplars (e.g., Carpenter, 1991; Matsumoto, 1985;
Uchida & Imai, 1999). An experimental elicitation task on Thai children between 1;8 and 11;3 showed that more typical exemplars of a numeral classifier category are learned earlier than the less typical exemplars (Carpenter, 1991). Likewise, a production experiment on the use of numeral classifiers for long objects by 5- to 7-year-old Japanese children found a significant difference between the correct responses for typical and atypical exemplars of the classifier (Matsumoto, 1985). Similarly, a series of error-detection experiments on 4- to 7-year-olds acquiring Japanese and Mandarin revealed that prototypical referents are learned earlier than less familiar items (Uchida & Imai, 1996). The test also found that subjects performed better at correcting mistakes when typical instances of a numeral classifier were tested than atypical ones (Uchida & Imai, 1999). All these studies indicate that younger children poorly recognise and categorise objects that they least frequently encounter (Matsumoto, 1985; Uchida & Imai, 1999). Children eventually achieve adult-like mental representation or schemata for numeral classifier categories, as they continue to modify and refine the semantic categories of numeral classifiers based on experience and input (Matsumoto, 1985).

2.3 Categorisation in the Malay Numeral Classifier System

Similar to Chinese, Japanese, and Thai, Malay numeral classifiers can be described using Allan’s (1977) framework for describing the numeral classifiers of other numeral classifier languages. As illustrated in Figure 2.1., like other numeral classifier languages, Malay nouns are subcategorised into ‘count nouns’ and ‘mass nouns’. The Malay mass nouns, however, do not co-occur with any numeral classifier; instead, they co-occur with measure words based on the quantity of the mass nouns that is needed to modify the nouns. To illustrate, mass nouns occurring
in a small quantity can be modified via measure words such as *cubit* (a pinch of) and *titik* (a drop of), while those occurring in a large quantity may be modified by other unitisers such as *guni* (a sack of) and *baldi* (a pail of).

The Malay count nouns in contrast, are categorised by the Malay numeral classifier system into two semantic categories on the basis of animateness, namely animate nouns versus inanimate nouns. The animate nouns are further subcategorised into two subcategories namely humans and animals, which are qualified by the numeral classifiers *orang* and *ekor* respectively. Inanimate nouns, on the other hand, are further subcategorised into two subcategories, namely shape and specific categories. The shape category is then further subcategorised into either...
Further analysis of the organisation of Malay numeral classifiers shows that to a large extent, Malay numeral classifiers can be described in a similar manner as the shape-based numeral classifiers described by Allan (1977). This is because, similar to most numeral classifier languages that Allan described, shape-based numeral classifier category of Malay is also subcategorised based on the dimensionality of the objects, namely 1D, 2D, and 3D.

The division of the shape-based numeral classifier category however, is not clear-cut as its partitioning is not mono-referential in nature. The poly-referential nature of shape-based numeral classifier subcategorisation involves a combination of different subcategories, that is, with mixed semantic criteria used to qualify a particular noun due to the dependency of secondary parameters on primary parameters.

In the assignment of nouns to their respective numeral classifiers, the 1D and 2D subcategories, for example, are not independent. This is because, like in the classification of Chinese classifiers, 1D and 2D numeral classifiers also have to be combined with another subcategory, that is, ‘rigidity’. The choice depends on whether the object is rigid or flexible. Thus, 1D objects that are rigid like a pen, a ruler, a tree, and a cane, co-occur with the numeral classifier batang [1D: +rigid], while 1D objects that are flexible like a chain, a strand of thread, and a bracelet, co-occur with the numeral classifier utas [1D: -rigid]. 2D objects, on the other hand, if they are rigid like a wooden plank, a slice of cake, and a photograph, co-occur with the numeral classifier keping [2D: +rigid], while those that are flexible like a piece of
cloth, a shirt, a piece of paper, and a pair of trousers, co-occur with the numeral classifier *helai* [2D: rigid].

Similar to the description of the shape-based numeral classifiers of Chinese, the 3D subcategory of Malay also does not occur independently. Its subcategorisation has to be combined with the ‘size’ subcategory. However, unlike the other numeral classifier languages described by Allan (1977), for example, Yurok (p.301), the size range for size subcategory in the Malay numeral classifier system is in four different forms, namely ‘fine’, ‘small’, ‘medium’, and ‘big’. Thus, for 3D objects that are fine like pebbles, precious stones, rice, and sand, the numeral classifier *butir* [3D: fine] is used, while those that are small (e.g., fruits, cups, and bulbs), the numeral classifier *biji* [3D: small] is used instead. 3D objects that are medium in size like chunks of chocolate, fist-sized stones, and lumps of clay co-occur with the numeral classifier *ketul* [3D: medium], while those that are big, (e.g., bags, computers, cars, and houses) co-occur with the numeral classifier *buah* [3D: big]. Because of the one-to-many references in the classification of shape-based numeral classifiers (e.g., *butir* for sand, rice, and pebbles, and *batang* for pens, rulers, and trees), the classification of Malay shape-based numeral classifiers is regarded as poly-referential in nature.

Similar to Micronesian languages, Malay does not have function-based numeral classifiers (Adams & Conklin, 1973). However, Malay shape-based numeral classifiers can be considered as having the closest resemblance to the function-based numeral classifiers of other numeral classifier languages. This is because, like function-based numeral classifiers, specific numeral classifiers are specific in nature. For example, in Japanese and Chinese, cars, boats and aeroplane are classified using numeral classifiers that are very specific in nature in the sense
that the respective numeral classifiers are not shared with other objects. Similarly, in Malay, knives and corn-in-cobs are classified using specific numeral classifiers that are exclusive to these objects. The specific subcategory however is not as complex as the shape-based numeral classifiers given that they involve a one-to-one mapping between the numeral classifier and the noun. For example, *kuntum* [specific: flower] is used to classify flowers, and *laras* [specific: firearm] can only co-occur with firearms. Because of the one-to-one references in its classification of objects (e.g., all flowers is classified with *kuntum*), Malay specific numeral classifiers are regarded as mono-referential in nature. However, specific numeral classifiers are also relatively difficult to be mastered since they involve the memorisation of a long list of items (Salehuddin, 2007). According to Omar (1972) when in doubt, children or adults tend to use *buah* or *biji* if they are not sure which classifier to use; however, the use of these numeral classifiers in their language use is still unacceptable.

### 2.4 Numeral Classifier Acquisition

In the East and Southeast Asian region, Japanese, Chinese, and Thai are three numeral classifier languages that have been most comprehensively investigated. The method of investigation ranges from cross-sectional studies (e.g., Carpenter, 1986; Chien, Lust, & Chiang, 2003; Gandour et al., 1984; Hu, 1993; Loke & Harrison, 1986; Mak, 1991; Matsumoto, 1985; Ng, 1989; Sanches, 1977; Uchida & Imai, 1999; Yamamoto & Keil, 2000) to longitudinal studies (e.g., Erbaugh, 1984; Murata, 1983; Naka, 1999; Okubo, 1967; Tuaycharoen, 1984). Some researchers have conducted comparative studies between numeral classifier acquisition of two different languages (e.g., Hansen & Chen, 2001; Uchida & Imai, 1999), whereas others have compared two different varieties of the same language (e.g., Loke &

Studies have investigated the production order of numeral classifier acquisition (e.g., Carpenter, 1991; Erbaugh, 1984; Gandour et al., 1984; Loke & Harrison, 1986; Mak, 1991; Matsumoto, 1985; Sanches, 1977; Tuaycharoens, 1984); the comprehension order (e.g., Chien et al., 2003; Yamamoto & Keil, 2000); and the acquisition order in both production and comprehension of numeral classifiers (Hu, 1993; Ng, 1989; Uchida & Imai, 1999; Yamamoto, 2005). The age groups of participants in these studies varied. Although some of the first language acquirers studied were as young as 2;0 (e.g., Carpenter, 1991; Erbaugh, 1984; Sanches, 1977), a majority of researchers have focused on children between 4 and 6 years old (e.g., Chien et al., 2003; Fang, 1985; Gandour et al., 1984; Hu, 1993; Loke & Harrison, 1986; Mak, 1991; Ng, 1989; Uchida & Imai, 1999; Yamamoto, 2005; Yamamoto & Keil, 2000). The oldest age group studied among first language acquirers were 11-year-olds for Thai (Carpenter, 1991) and 12-year-olds for Chinese (Ng, 1989) and Japanese (Sanches, 1977). While a great majority of numeral classifier acquisition studies were conducted on normal first language acquirers, there were some researchers who included in their studies participants who were (1) adult second language acquirers (Hansen & Chen, 2001; Liang, 2008), and (2) children diagnosed with learning disabilities (Mak, 1991). In general, researchers agree that numeral classifier acquisition takes place at a later stage in language development in comparison with, for example, noun and verb acquisition (Carpenter, 1991; Gandour et al., 1984; Mak, 1991; Ng, 1989; Sanches, 1977; Yamamoto, 2005).
The categorisation and acquisition of numeral classifiers reveal that general numeral classifiers are usually acquired relatively early by children and are used by children as default numeral classifiers before the mastery of more specific numeral classifiers (Carpenter, 1991; Chien et al., 2003; Wei & Lee, 2001; Yamamoto & Keil, 2000). For example, in Japanese, young preschoolers (1;6 – 4;0) use the more general numeral classifiers -tsu (inanimate: generic) and -ko (inanimate: 3D) in their production prior to the use of other numeral classifiers (Matsumoto, 1985).

2.4.1 Japanese Numeral Classifier Acquisition

The acquisition of the Japanese numeral classifier system has been widely studied (e.g., Hansen & Chen, 2001; Matsumoto, 1985; Sanches, 1977; Shirai, 2001; Uchida & Imai, 1999; Yamamoto, 2005). Among these studies, Hansen and Chen’s (2001) study was the only one associated with the Numeral Classifier Accessibility Hierarchy – NCAH. They hypothesised that numeral classifier acquisition follows the following order: Animate human > Animate nonhuman > Shape > Function. In their study on the acquisition and attrition order of numeral classifiers as a second language among adults, Hansen and Chen (2001) predicted that the animate human numeral classifier would be the first numeral classifier to appear and retained the longest. Conversely, function numeral classifiers would be the last numeral classifiers to appear. Hansen and Chen’s hypothesis was supported by their findings. Japanese human classifiers (-ri/-nin), along with numeral classifiers for small animals (-hiki), were the earliest numeral classifiers to appear among adult learners (Hansen & Chen, 2001). This was then followed by Japanese shape-based numeral classifiers (-hon, -mai, and -ko). Contradictory to their hypothesis, one functional numeral classifier, that is, the numeral classifier for books, was learned more quickly
than it was learned by participants in Downing’s study (1984). Hansen and Chen explained these results as due to their subjects being adult missionaries who were familiar with books.

Although Hansen and Chen’s study (2001) was conducted on adults, the same acquisition order was observed in the majority of Japanese numeral classifier acquisition studies on children (Matsumoto, 1985; Naka, 1999; Uchida & Imai, 1999; Yamamoto & Keil, 2000). For example, Matsumoto (1985), in an elicited production task, found that animate numeral classifiers are acquired prior to inanimate numeral classifiers. Previous research shows that Japanese general inanimate numeral classifier (-tsu/-ko) and the Japanese numeral classifier for humans (-ri/-nin) are the first two Japanese numeral classifiers to be acquired (Downing, 1996; Matsumoto, 1985; Sanches, 1977). This is followed by the general numeral classifier for animals (-hiki), large mechanical objects (-dai), long slender objects (-hon), and flat, thin objects (-mai). Uchida and Imai (1999) on the other hand, in an error-detection task on 4- to 6-year-olds, found that numeral classifiers with less features are counted earlier than those with more features.

Sanches (1977) and Matsumoto (1985) found that children’s acquisition “proceeds from an item which has a general application to an item which has a more restricted application” (Matsumoto, 1985, p. 81). For example, in a counting task on both familiar and unfamiliar objects, Japanese children between 5 and 7 years old tended to overgeneralise the use of -tsu and -ko general classifiers, and in their early stage of acquisition, children were also inclined to overextend the human numeral classifiers -ril-nin to animals as well (Matsumoto, 1985). These results were supported by findings from an error-detection task on 4- to 6-year-olds (Uchida & Imai, 1999). In addition, the acquisition of Japanese numeral classifiers begins with
the absence of use of a numeral classifier to an overgeneralisation in numeral classifier usage. Once children are able to differentiate the semantic rules of numeral classifiers, children stop overgeneralising and begin producing the correct numeral classifiers (Uchida & Imai, 1999).

According to Shirai (2001), although some 2-year-old Japanese children do show an early emergence of numeral classifiers, their production is limited to the general numeral classifiers -tsu, -ko, -nin/-ri. Children resort to using general classifiers in production because production involves a retrieval process from memory. According to Meyer and Bock (1992), the relatively more accessible lexicon such as the general classifiers would help children reduce semantic load (cf. Gandour et al., 1984).

However, studies have also shown that the order of Japanese numeral classifiers could also be influenced by the semantic complexity of individual numeral classifiers (Uchida & Imai, 1999). For example, through an error detection task, Japanese children have been found to acquire -ri/-nin earlier than -tou. According to Uchida and Imai (1999), this takes place because Japanese has only two numeral classifier forms for humans (-ri/-nin) – depending on the number of persons (i.e., -ri for 1 or 2 persons, and -nin for 3 or more). In contrast, there are three numeral classifier forms for animals in Japanese (-hiki, -wa, or -tou) – depending on the type of animals (i.e., -hiki for unmarked animals, -wa for birds, and -tou for large animals) (Uchida & Imai, 1999).

Children still have to refine their use of numeral classifiers in order to use them with atypical cases (Matsumoto, 1985). It has been argued that children only have a full understanding or knowledge of the underlying semantic criteria (which in
this case, the semantics of numeral classifiers) when they can apply the numeral classifier to new instances, which necessarily include atypical instances (Downing, 1984). Until they can do this, children tend to use general classifiers for atypical members of a numeral classifier (Downing, 1984). According to Matsumoto, refinement in the knowledge of the semantics of numeral classifiers is “presumably triggered by an actual exposure to such uses in the input (1985, p. 84). Thus, although children appear to have already mapped the semantics of numeral classifiers onto typical members, they have to keep modifying and refining the semantic categories of numeral classifiers based on the input data, so that they eventually conform to the adult-like convention of using the category (Matsumoto, 1985). He added that children will have to continue searching for the convention long after they appear to have acquired a “good productive control” over the numeral classifier (p. 85).

Although children may be able to acquire basic syntactic patterns of Japanese numeral classifiers very early, the semantics of the numeral classifiers seems to be acquired much later and is a slower process (Matsumoto, 1985, 1987; Sanches, 1977; Yamamoto & Keil, 2000). This is evident in children’s use of general classifiers prior to mastery of more specific numeral classifiers (e.g., Yamamoto & Keil, 2000). 6-year-olds in Yamamoto and Keil’s (2000) study, which involved a selection task in two different contrast conditions, comprehended only 60% of the numeral classifiers tested. In addition, 12-year-olds in Sanches’ (1977) study produced only 22 numeral classifiers in comparison to 36 common numeral classifiers by adults.

Input frequency has been cited to be an important factor in the acquisition of Japanese numeral classifiers (Matsumoto, 1985; Yamamoto, 2005). Yamamoto (2005), for example, found that higher frequency numeral classifiers in both speech
and written texts such as -tsu, -ko, -ri, -hiki, and -dai “emerged maturationally” earlier than lower frequency numeral classifiers (p. 119)

2.4.2 Chinese Numeral Classifier Acquisition

A number of studies have investigated the acquisition of Chinese numeral classifiers. For example, Hu (1993), Chien, Lust and Chiang (2003), and Uchida and Imai (1999) investigated the acquisition of Mandarin numeral classifier; Ng (1989) conducted a study on Hokkien numeral classifier acquisition; and Tse, Li and Leung (2007), and Wei and Lee (2001) conducted a study on Cantonese.

Previous studies on Chinese language have shown that the acquisition of Chinese numeral classifiers is a relatively slow process (Chien et al., 2003; Fang, 1985; Hu, 1993; Lee, 1996). For example, 4-year-old Mandarin Chinese acquirers produced only 2 numeral classifiers; 5-year-olds, 4 numeral classifiers; 6- and 7-year-olds, 12 numeral classifiers (Ying et al. (1990) cited in Lee, 1996)). This is slightly slower than the acquisition of Japanese numeral classifier. For example, children in Naka’s (1999) longitudinal study on a pair of twins, which was a longitudinal study on the natural speech data of a pair of twins, were able to produce 8 numeral classifiers by the age of two. The discrepancy however, could be due to the different method of data collection. Studies conducted on the acquisition of Chinese numeral classifiers were experimental studies, whereas the one conducted by Naka were natural speech data.

Uchida and Imai conducted an error detection task to investigate the acquisition of numeral classifiers by Japanese and Chinese children. They found that 5-year-old Mandarin Chinese children gave only 45% of correct responses in a
counting task, and in correcting errors, whereas Japanese children, by the end of 5
years old, could gave an almost 100% score in a similar. Their findings suggest that
Chinese children’s performance is significantly lower than Japanese children. They
concluded that the performance of 6-year-old Chinese was on the same level as the

All studies on the acquisition of Chinese numeral classifiers, like the
Japanese, have shown that children’s performance (which included number of correct
responses and numeral classifier inventories) increases with age (Chien et al., 2003;
Erbaugh, 1984; Hu, 1993; Loke, 1991; Tse et al., 2007; Uchida & Imai, 1999; Wei &
Lee, 2001). Results from these studies show that the acquisition of numeral
classifiers in Chinese is a gradual process, as even by the age of 12, children’s
numeral classifier repertoire is still not equivalent to an adults’ (Lee, 1996; Ng,
1989).

In a natural speech data collection, Erbaugh (Erbaugh, 1984, 1986) found that
animate numeral classifiers are acquired prior to inanimate numeral classifiers and
that numeral classifier for humans is acquired prior to numeral classifiers for
animals. Previous studies also show that Chinese numeral classifier acquisition
proceeds from general classifiers to specific classifiers (Erbaugh, 1984; Hu, 1993;
Ng, 1989; Tse et al., 2007; Wei & Lee, 2001). For example, findings from an
elicited production task on 3- to 5-year-olds in a toy-play context found that children
were already producing the Cantonese general classifier go3 as early as the age of 3
(Tse et al., 2007). A longitudinal study on Mandarin Chinese numeral classifier
usage from natural settings also found that the general classifier ge was used by
children as early as 2;6 (Erbaugh, 1984). This is similar to findings by Matusumoto
(1985) who found that children at their early stage of acquisition tend to
overgeneralise the use of the –tsu and –ko general classifiers (Matsumoto, 1985). Researchers suggest that children’s general classifier usage in their production exhibits their knowledge of the syntactic distribution of numeral classifiers, but their limited knowledge of the numeral classifier semantic distribution compels them to resort to using general classifiers as a place holder, until they learn the specific classifier (Hu, 1993, p. 54; Tse et al., 2007). According to Tse et al. (2007), usage of general classifiers are “syntactically motivated” while specific classifiers are more “semantically instigated” (p. 513).

Apart from using the general classifier in place of unknown numeral classifiers, children were also found to use some other strategies as place holders (Loke, 1991; Mak, 1991; Ng, 1989). For example, children tended to replace unknown numeral classifiers with known numeral classifier that share some similarities (Ng, 1989). In an elicited production task cross-sectional study on Hokkien Chinese numeral classifier acquisition, it was found that 5- to 12-year-old children replaced the Hokkien numeral classifier tiung [2D: -rigid] with tiao [1D: -rigid], and ki [1D: +rigid] with tiao [1D: -rigid] (Ng, 1989). Such a strategy is different from the use of general classifiers, because, as illustrated above, at least one of the features is shared by the two numeral classifiers (i.e., rigidity in the first, dimensionality in the second). The use of a general classifier in place of an unknown numeral classifier, on the other hand, does not need a common feature in the semantics of the numeral classifier.

If Japanese children in Matsumoto’s (1985) study were inclined to overextend the human numeral classifiers –ri/-nin to animals as well (both sharing the same feature i.e., both are animate), a similar strategy was also found in a longitudinal study conducted by Erbaugh (1984). Erbaugh’s subject, for example,
tended to overgeneralise the Mandarin *tiao* [1D] numeral classifier for a sword and a gun, two objects which require the *ba* numeral classifier (Erbaugh, 1984). According to Erbaugh (2002), the semantics of general classifiers in Mandarin is so flexible that even highly educated people produced *ge* in place of more specific numeral classifiers.

Like the Japanese acquirers in Uchida and Imai’s (1999) study, children have also been found to omit numeral classifiers as a strategy (Erbaugh, 1984; Wei & Lee, 2001). However, in older children, such a strategy was only found in Wei and Lee’s (2001) study, which was conducted on bilingual Cantonese-English speakers. Since it was generally found that Chinese children resorted to using a general classifier once they were aware of the syntactic structure of numeral phrases, children’s omission of numeral classifiers in Wei and Lee’s research is likely due to language transfer, as numeral classifiers are not found in English (Wei & Lee, 2001).

Previous research also reveals that the degree of typicality of exemplars plays a role in numeral classifier acquisition (Hu, 1993; Ng, 1989; Uchida & Imai, 1999). In an elicited production task, 5- to 12-year-old Hokkien Chinese children were found to give significantly more correct responses for typical objects than atypical ones (Ng, 1989). Similarly, 4- to 6-year-old children learning Mandarin Chinese were able to count and correct typical instances of numeral classifiers better than atypical ones (Uchida & Imai, 1999). Hu (1993) adds that in the production of specific classifiers (in comparison to a general classifier), children first use the specific classifiers with their respective prototypical members and memorise the limitation of the classifier-noun collocation.
In addition, Loke’s findings suggest that the more frequently a numeral classifier is used, the more children use the numeral classifier (Loke, 1991). For example, 64.78% of the production responses by Loke’s subjects, who were 5- to 6-year-olds, consisted of the Mandarin general classifier ge, and this numeral classifier was produced by the highest number of subjects (i.e., 21 subjects) (Loke, 1991, p. 108). In contrast, kuai was only produced 0.29% of the time, and this numeral classifier was only produced by 3 subjects. According to Hu (1993, p. 119), the acquisition of numeral classifiers “depends heavily on how much the child is exposed to the adult’s use of a particular classifier” and how relevant the nouns are to children with which a numeral classifier is used. Uchida and Imai (1999, p. 58) further add that

“in learning classifiers, extracting semantic rules is not enough for them to learn full use of numeral classifiers since most numeral classifiers, especially frequently used ones include members whose class membership may not always be transparent from the semantic rules … children then have to continue to pay close attention to the input in order to learn the proper classifiers for entities whose class membership cannot be so obviously detected from the rules they have extracted.” …

Hansen and Chen (2001) found that Chinese human classifiers gei, wei, and dui were the earliest to appear among adult learners, at about the same time the numeral classifier for small animals (zhi) appears. Hansen and Chen hypothesised that the numeral classifiers for books were early in acquisition because of their high frequency in missionary language.

Mak (1991) and Ng (1989) found that Chinese shape-based numeral classifiers are acquired prior to specific classifiers. This supports findings from
shape bias literature that children rely on the shape of the objects when classifying them (Gentner, 1978; Gentner & Imail, 1995; Imai et al., 1994; Landau et al., 1985; LB Smith & Sera, 1999). With regard to shape-based numeral classifiers, studies on Chinese numeral classifier acquisition also found that the dimensionality of numeral classifiers plays a role in acquisition order. However, there is a discrepancy in the order of acquisition. For example, while Hu (1993) found that 2D numeral classifiers are acquired prior to 1D and 3D numeral classifiers (2D > 1D, 3D); Erbaugh (1984) on the other hand found the order in the following sequence: 1D > 2D, 3D. Loke and Harrison (1986) however found the 3D > 1D >2D sequence. Rather than the dimensionality of numeral classifiers, Ng (1989), in contrast, found that the rigidity of the numeral classifiers instead determines the order of numeral classifier acquisition. She found that flexible numeral classifiers are acquired prior to rigid numeral classifiers. Hu (1993) suggests that the difference in results with regard to the acquisition order based on dimensionality is most likely to be due to the different research methods and criteria used for selecting shape-based numeral classifiers and their exemplars. For example, the representativeness of numeral classifier exemplars used in experiments is questionable, and thus affects the acquisition order (Hu, 1993).

2.4.3 Thai Numeral Classifier Acquisition

Thai is the only widely researched language with regard to numeral classifier acquisition studies among Southeast Asian languages (Carpenter, 1986, 1991; Gandour et al., 1984; Tuaycharoen, 1984). Whereas Carpenter’s and Gandour et al.’s studies were based on elicited production tasks, the study conducted by Tuaycharoen was a longitudinal study based on spontaneous speech. Despite the
dissimilarity in methodology, both Carpenter’s and Tuaycharoen’s participants were able to produce Thai numeral classifiers as early as the age of two. For example, in an elicited production task using puppets to stimulate numeral classifier usage from children between 1;8 and 11;3, two-year-olds were able to produce the numeral classifiers *khon* (numeral classifier for human) and *?an* (numeral classifier for things in general) with 10% correct responses (Carpenter, 1991). (However, 10% is not evidence of productivity. It is likely with such a small percentage that the children were remembering combinations as if they were single units.) However, in both Gandour et al. (1984) and Carpenter (1991), the general classifier *?an* was recorded as a correct response, not as an alternative response for unknown numeral classifiers.

Although children showed an early repertoire of Thai numeral classifier usage (e.g., two years old in Carpenter [1986] and Tuaycharoen [1984]), their performance suggests that their knowledge of numeral classifiers at that young age is limited to the system’s syntagmatic relations. For example, in a sentence completion task on children between 5 and 10 years old, children were found to overgeneralise the general classifier *?an* as a place holder for unknown numeral classifiers (Gandour et al., 1984). The same was also observed among Carpenter’s subjects (Carpenter, 1991). The use of the general classifier *?an* as a default numeral classifier, however, declines with age. For example, 5-year-olds in Gandour’s (1984) study used *?an* as a place holder 77.6% of the time; however, the 10-year-olds only used *?an* in place of unknown classifiers 24.2% of the time. This is similar to studies on Japanese (Matsumoto, 1985) and Chinese (Hu, 1993) numeral classifier acquisition. In addition, the production of Thai numeral classifier system appears to be a prolonged process. For example, 5-year-olds in Gandour et al.’s (1984) study gave only 15% correct responses in a sentence-completion task.
The use of the general classifier *tan* is not the only strategy Thai children use in place of unknown numeral classifiers. Apart from numeral classifier omission (Tuaycharoen, 1984), children in both longitudinal and cross-sectional studies have been found to use repeaters (which Gandour et al. termed as ‘overspecialisation’) instead of true numeral classifiers (Gandour et al., 1984; Tuaycharoen, 1984). The repeaters Thai children used were either in the form of a partial or full head noun in place of the missing numeral classifiers.

Example 38:

\[
\text{Daaw sìi } \text{daaw } \text{(duan)}
\]
stars four Repeater (NumCl)
‘four stars’

Example 39:

\[
\text{Khàjmúg } \text{hàa } \text{múg } \text{(méd)}
\]
pearls five Repeater (NumCl)
‘five pearls’

(Gandour et al., 1984, p. 462)

It was also found that Thai children tend to overextend the semantics of numeral classifiers as a strategy in numeral classifier usage. For example, older children in Carpenter’s study overextended the semantics of *chyag* (rope) to long, thin, flexible things (e.g., snake, gag, tube, spaghetti, belt) (Carpenter, 1991). This is similar to studies by Ng (1989) who found that children acquiring Hokkien Chinese tended to replace unknown numeral classifiers with known numeral classifier that share some similarities, for example, *tiung* [2D: -rigid] with *tiao* [1D: -rigid], and *ki* [1D: +rigid] with *tiao* [1D: -rigid] (Ng, 1989).

Thai children’s overgeneralisation, overspecialisation, and overextension strategies in place of more semantically specific numeral classifiers suggest that they have an awareness of the need to satisfy the syntactic requirements of Thai language (Carpenter, 1991). According to Carpenter (1991), this usually takes place before
children learn the limitations of the semantic relationship between a head noun and its specific numeral classifier. Relatively ‘semantically heavy’ or complex numeral classifiers, such as shape-based numeral classifiers, have been found to be relatively difficult for children to acquire (Gandour et al., 1984). By using the three strategies above as substitutes for more semantically specific or complex numeral classifiers, children can reduce the semantic load by making classification less dependent on inherent semantic characteristics of referents (Carpenter, 1991; Gandour et al., 1984).

In addition, it has been found that typical members of a numeral classifier category were acquired earlier than atypical members (Carpenter, 1991). Children’s performance across numeral classifier categories increased rather slowly, with 10-year-olds’ performance lagging behind adult’s performance (Gandour et al., 1984).

2.5 Research Questions and Hypotheses

The general aim of the study was to investigate the acquisition of Malay numeral classifiers. Based on a review of relevant literature, the acquisition of eight Malay shape-based numeral classifiers were investigated from an across and within numeral classifier category perspective.

**Developmental Patterns of Malay Numeral Classifier Acquisition**

*The first emergence of numeral classifiers in children.* Our review on crosslinguistic studies on numeral classifier acquisition indicates that the age when children begin to acquire numeral classifier varies. For example, in elicited production tasks, Japanese children were able to produce numeral classifiers at 2 years of age (Yamamoto, 2005), Chinese by the age of 2;6 (Hu, 1993) and Thai
The Categorisation and Acquisition of Numeral Classifiers

(Carpenter, 1991) by 1;8. It is predicted that by the age of 5, Malay children will be able to start using numeral classifier in their production.

The order of numeral classifier acquisition. Previous cross-linguistic numeral classifier acquisition studies have found that animate numeral classifiers are acquired prior to inanimate numeral classifiers (e.g., Erbaugh, 1984, 1986; Matsumoto, 1985; Ng, 1989). Furthermore, it has also been found that numeral classifiers for humans are acquired prior to those for animals (Erbaugh, 1984; Hansen & Chen, 2001). It has been found that shape-based numeral classifiers are acquired prior to specific numeral classifiers (Mak, 1991; Ng, 1989). On this basis, it was predicted that Malay numeral classifier development would follow the following order, namely animate numeral classifiers are acquired prior to inanimate numeral classifiers, and the numeral classifier for humans (orang) is acquired prior to the numeral classifier for animals (ekor). In addition, shape-based numeral classifiers will be acquired prior to specific numeral classifiers.

The strategies used for unknown numeral classifiers. In previous studies on the acquisition of numeral classifiers, young children tend to initially omit numeral classifiers altogether, or use general or alternative numeral classifiers in place of more specific ones (Carpenter, 1991; Erbaugh, 1984; Gandour et al., 1984; Hansen & Chen, 2001; Hu, 1993; Matsumoto, 1985; Ng, 1989; Tuaycharoen, 1984; Uchida & Imai, 1999; Wei & Lee, 2001). It was predicted that Malay children will use similar strategies; either omit numeral classifiers or use alternative forms of numeral classifiers in place of unknown numeral classifiers.
2.6 Summary

In this chapter, we have discussed how numeral classifiers can be viewed as a manifestation of human conceptualisation through linguistic categorisation. We have also reviewed how the degree of typicality of objects may play a role in human categorisation, particularly in assigning objects into numeral classifiers. Next, we reviewed the categorisation of objects into the Malay numeral classifier system. We then reviewed studies conducted on the acquisition of three numeral classifier languages, namely Japanese, Chinese, and Thai. Finally, the research questions and hypotheses of this research are outlined.
CHAPTER 3
Pilot Study: The Production of Fourteen Numeral Classifiers

3.1 Introduction

This chapter focuses on a preliminary pilot study, which was conducted to investigate the developmental patterns in the production of 14 common Malay numeral classifiers. An additional aim was to identify when Malay children begin to produce numeral classifiers so as to determine an appropriate age group for the experimental tasks and the particular numeral classifier categories to be focused on.

The pilot study consisted of an elicited production task. The objects used to elicit production responses were pictures of two animate numeral classifier exemplars, eight shape-based numeral classifier exemplars, and four specific numeral classifier exemplars. Each picture was randomly presented to the children and they were asked to count the objects in the pictures.

Based on the literature reviewed, the following predictions were formulated.

1. Previous cross-linguistic research involving elicited production tasks indicates that the age when children begin to produce numeral classifiers varies. For example, while Thai children began to produce numeral classifiers at the age of 2 (Carpenter, 1991), Mandarin Chinese (Hu, 1993) and Japanese (Yamamoto, 2005) 3 Pictures rather than real objects were used because a study by Fenson, Vella, and Kennedy (1989) has shown that children as young as 2 years old can match pictures with pictures as accurately as they can match objects with pictures. This suggests that pictures are very salient to young preschoolers (Gelman & Coley, 1990).
children began producing numeral classifiers at the age of 3 years old. Cross-linguistic research in natural settings also shows similar findings. While Mandarin Chinese children were observed to use numeral classifiers by 2;6 (Erbaugh, 1984), observations on Japanese children (Naka, 1999; Shirai, 2001) found that numeral classifiers were already produced by children by the age of 2. Based on this evidence, it was expected that by the age of 5 or younger, Malay children should be able to produce at least some of the Malay numeral classifiers.

2. Previous cross-linguistic studies on numeral classifier production indicate the usage of animate classifier(s) prior to inanimate ones (Burling, 1959; Erbaugh, 1986; Gandour et al., 1984). In relation to Malay numeral classifiers, it was predicted that Malay children would produce animate numeral classifiers, namely orang [animate: human] and ekor [animate: animal], earlier than inanimate numeral classifiers. Previous cross-linguistic studies on the acquisition of numeral classifier types indicate that shape-based numeral classifiers are produced prior to specific ones (Mak, 1991; Ng, 1989). Hence, it was predicted that the production of shape-based numeral classifiers among Malay children would take place prior to the production of specific numeral classifiers.

3. In previous studies on the acquisition of numeral classifiers, younger children tend to initially omit numeral classifiers altogether, or use general or alternative numeral classifiers in the place of more specific ones (Carpenter, 1991; Erbaugh, 1984; Gandour et al., 1984; Hansen & Chen, 2001; Hu, 1993; Matsumoto, 1985; Ng, 1989; Tuaycharoen, 1984; Uchida & Imai, 1999; Wei & Lee, 2001). It was predicted Malay children will use similar strategies; either omit numeral classifiers or use alternative forms of numeral classifiers in place of unknown numeral classifiers.
3.2 Method

3.2.1 Participants

One hundred and twelve children attending preschools and primary schools in Melaka, a state in the southern region of the Peninsular of Malaysia, participated in this experiment. The children speak the standard variety of Malay, the Johor-Riau dialect, which also derive from the southern region of the Peninsular Malaysia (Please see Appendix 2 for Research Approval from the Malaysian Government and Appendix 3 for Ethics Approval from the University of Western Sydney). The seven age groups each contained 16 children: (1) the <5-year-olds (the youngest age group), (2) the 5-year-olds, (3) the 6-year-olds, (4) the 7-year-olds, (5) the 8- & 9-year-olds, (6) the 10- & 11-year-olds, and (7) the >11-year-olds (the oldest age group). Participants were all native speakers of Malay and spoke the standard variety of Malay. 10 Malay adults also participated in the experiment as a comparison group. They were between 18 and 66 years of age (Mean age 34.7), 5 female and 5 male, and were from middle SES. They went through the same procedure as the children did. A description of the participants is given in Table 3.1.

Table 3.1. Description of Participants in the Pilot Study

<table>
<thead>
<tr>
<th>Age group</th>
<th>Age range</th>
<th>Mean age</th>
<th>No of participants</th>
<th>No. of males</th>
<th>No. of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5-year-olds</td>
<td>3;4 – 5;1</td>
<td>4.42</td>
<td>16</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>5;2 – 5;10</td>
<td>5.53</td>
<td>16</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>6-year-olds</td>
<td>5;11 – 7;1</td>
<td>6.43</td>
<td>16</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>7;2 – 7;11</td>
<td>7.46</td>
<td>16</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>8- &amp; 9-year-olds</td>
<td>8;2 – 9;5</td>
<td>8.92</td>
<td>16</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>10- &amp; 11-year-olds</td>
<td>9;6 – 10;7</td>
<td>10.06</td>
<td>16</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>&gt;11-year-olds</td>
<td>10;8 – 12;4</td>
<td>11.49</td>
<td>16</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Adult</td>
<td>18;3 – 66;5</td>
<td>43.67</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>122</td>
<td>63</td>
<td>59</td>
</tr>
</tbody>
</table>
3.2.2 Stimuli

Fourteen coloured pictures of objects familiar to children in their everyday environment (Appendix 4) were each printed on a 4cm x 5cm-card. Each one of the objects in the pictures requires the use of one numeral classifier from those listed in 3.2.3 in counting activities. All the items appeared in units of three so as to encourage the children to use a numeral, a numeral classifier, and a noun in counting the respective objects. If the items were presented only in one unit in the picture, naming the objects using only their respective names can be considered adequate. If the items were presented in units of twos in the picture, using the expression *sepasang* (“a pair of”) prior to the object names would be considered correct as well. Presenting the objects in units of threes in the picture should stimulate and motivate the children to count using a numeral classifier. (Varying the quantity of objects in each picture was not necessary because this task was not aimed at investigating the children’s ability to count).

3.2.3 Procedure

Children were asked to count pictures of objects that require the use of either animate, shape-based, or specific numeral classifiers in counting activities. Pictures of 14 objects were shown to them, one at a time, and the children were asked to count each one of the objects. Children first went through the practice trial session before proceeding to the experimental session. The 14 Malay numeral classifiers tested in this experiment were *orang* [animate: human], *ekor* [animate: animal], and 12 other numeral classifiers for inanimate objects, which were *batang* [1D: +rigid], *utas* [1D: -rigid], *keping* [2D: +rigid], *helai* [2D: -rigid], *buah* [3D: big], *ketul* [3D: medium], *biji* [3D: small], *butir* [3D: fine], *kuntum* [specific: flower], *tongkol*
The Production of Fourteen Numeral Classifiers

[specific: cob of corn], *kaki* [specific: umbrella], and *bilah* [specific: knife]. The first eight inanimate numeral classifiers are shape-based numeral classifiers while the last four are specific numeral classifiers. These numeral classifiers were chosen because they are some of the most commonly used numeral classifiers in Malay (Omar & Subbiah, 1995).

### 3.2.4 Practice Trial

Apart from the 14 pictures of numeral classifier exemplars listed in 3.2.3, four other pictures that require common numeral classifiers in the children’s linguistic environment were used for the practice trials. All the children were individually shown the practice trial cards described in Table 3.2. They were given examples of what to say, for example, “*Tiga bentuk cincin*” (“Three NumCl [specific: ring/hook] rings”) and “*Tiga pasang sayap*” (“Three NumCl [pair] wings”). A prompt in the form of the expression “*Berapa*?” (How many?) was given to the children if they merely mentioned the name of the object, for example, “*Cincin*” (“Ring”). The children were further prompted with “*Tiga apa cincin?*” (“Three what ring?”) if they merely gave “*Tiga cincin*” (“Three rings”) as their answers, with ‘*apa*’ deliberately stressed to signal to the children as well as to emphasise that they need to fill the gap between the numeral and the noun. Only when the researcher was convinced that the participants had understood the task did the experimental session commence.
3.2.5 The Experiment

The children were then asked to name the items they saw in the 14 experimental test stimuli (Table 3.3) before the counting task took place. This was to ensure that the children knew what the pictures represent. The counting task took place when all the 14 test items had been correctly named by the participants. The correct expected response in counting the objects in the test stimuli was to include the use of the correct numeral classifier between the numeral and the noun, for example, “Tiga orang kanak-kanak” (“Three NumCl children”). Dissimilar to the practice trial, the children were not prompted or doubly prompted with “apa” (“what”) in the experimental session. Pictures were presented in a fixed random order to all the children so that all participants get the same order and this would be fair to all participants. The experimental session lasted about 20 minutes for each child.

Table 3.2. Description of Practice Trial Items

<table>
<thead>
<tr>
<th>The stimulus</th>
<th>Sample response</th>
<th>Numeral classifier required</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 rings</td>
<td>3 bentuk cincin</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 hooks</td>
<td>3 bentuk mata kail</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 wings</td>
<td>3 pasang sayap</td>
<td>pasang</td>
</tr>
<tr>
<td>3 glasses</td>
<td>3 pasang cermin mata</td>
<td>pasang</td>
</tr>
</tbody>
</table>
Table 3.3. Description of Test Stimuli for the Pilot Study

<table>
<thead>
<tr>
<th>The stimulus</th>
<th>Correct response</th>
<th>Numeral classifier tested</th>
<th>Numeral classifier type</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 children</td>
<td>3 orang kanak-kanak</td>
<td>orang</td>
<td>[animate: human]</td>
</tr>
<tr>
<td>3 fish</td>
<td>3 ekor tekan</td>
<td>ekor</td>
<td>[animate: animal]</td>
</tr>
<tr>
<td>3 pencils</td>
<td>3 batang pensil</td>
<td>batang</td>
<td>[1D: +rigid]</td>
</tr>
<tr>
<td>3 boxes</td>
<td>3 buah kotak</td>
<td>buah</td>
<td>[2D: +rigid]</td>
</tr>
<tr>
<td>3 planks</td>
<td>3 keping papan</td>
<td>keping</td>
<td>[3D: big]</td>
</tr>
<tr>
<td>3 oranges</td>
<td>3 biji aren</td>
<td>biji</td>
<td>[3D: small]</td>
</tr>
<tr>
<td>3 necklaces</td>
<td>3 utas rantai</td>
<td>utas</td>
<td>[1D: -rigid]</td>
</tr>
<tr>
<td>3 leaves</td>
<td>3 helai daun</td>
<td>helai</td>
<td>[2D: -rigid]</td>
</tr>
<tr>
<td>3 chocolates</td>
<td>3 ketul coklat</td>
<td>ketul</td>
<td>[3D: medium]</td>
</tr>
<tr>
<td>3 rice (grain)</td>
<td>3 butir beras</td>
<td>butir</td>
<td>[3D: fine]</td>
</tr>
<tr>
<td>3 flowers</td>
<td>3 kuntum bunga</td>
<td>kuntum</td>
<td>[specific: flower]</td>
</tr>
<tr>
<td>3 cobs of corn</td>
<td>3 tongkol jagung</td>
<td>tongkol</td>
<td>[specific: cob of corn]</td>
</tr>
<tr>
<td>3 umbrellas</td>
<td>3 kaki payung</td>
<td>kaki</td>
<td>[specific: umbrella]</td>
</tr>
<tr>
<td>3 knives</td>
<td>3 bilah pisau</td>
<td>bilah</td>
<td>[specific: knife]</td>
</tr>
</tbody>
</table>

3.3 Coding of Responses

Each of the children’s production responses was classified into one of the following three categories:

(1) Correct response when children’s responses resemble those in Table 3.3 (Column 2);

(2) No classifier response when no numeral classifier was used in the response, (e.g., instead of “3 orang kanak-kanak”, “3 kanak-kanak” was used); and

(3) Alternative classifier response when an alternative numeral classifier was used in the response instead of the correct numeral classifier (e.g., for “3 keping papan”, “3 batang papan” was substituted).

3.4 Results

Figure 3.1 shows that the three youngest age groups (<5-year-olds, 5-year-olds, and 6-year-olds) were only able to produce a small number of correct numeral
classifier responses. Out of 224 responses from the <5-year-olds age group, the

group produced only one correct response (*orang*) and the 5-year-olds only produced

6 correct responses, namely *orang* (3), *biji* (2), and *ekor* (1). Even among the 6-year-

olds, only 32 of the 224 (0.14%) responses by the 6-year-olds were considered
correct. The two oldest age groups in this pilot study (i.e., the 10- & 11-year-olds

and the >11-year-olds) both produced 79.46% correct responses. The mean number

of correct responses gathered from all the children was 42.60%. All 10 adults who

participated in this experiment as a comparison group gave 100% correct responses
to all stimuli.

![Figure 3.1](image)

Figure 3.1. Mean number of the three response types produced by all age groups.

As can be seen in Figure 3.2, the *orang* [animate: human], *ekor* [animate:

animal], and *batang* [1D: +rigid] numeral classifiers were produced 100% correctly

by the 7-year-olds, and three other numeral classifiers were produced 100% correctly

by the 10- and 11-year-olds, namely *utas* [1D: -rigid], *helai* [2D: -rigid], and *kuntum*

[specific: flower].
In order to examine the main effect of numeral classifier, age group and gender on the correct responses, a 14 x 7 x 2 (Numeral Classifier X Age Group X Gender) mixed repeated measures ANOVA with numeral classifier (14 numeral classifiers) as a within-subjects factor and age group (<5-year-olds, 5-year-olds, 6-year-olds, 7-year-olds, 8- & 9-year-olds, 10- & 11-year-olds, >11-year-olds) and gender as between-subjects factors was conducted. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(90) = 243.88, p < .05$). As a result, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .72$). There was a significant main effect of numeral classifiers, $F(9.31, 912.43) = 59.51, p < .001$, partial $\eta^2 = .378$. Pairwise comparisons (LSD = .05) showed that orang [animate: human] numeral classifier was significantly the highest (mean = .71) whereas ketul [3D: medium] was significantly the lowest (mean = .06) (Figure 3.3).
There was a significant effect of age group, $F(6, 105) = 143.07, p < .001$, partial $\eta^2 = .378$. There was not a significant effect of gender, $p = .32$.

In order to determine the relative emergence order of the numeral classifiers, using Figure 3.2 as a reference, planned contrasts between the mean number of correct responses for the numeral classifiers in relation to the test value “0” were conducted for each age group. The planned contrast indicated that the production of *orang* among the <5-year-olds was not significant, $p = .33$. At the age of 5, Malay children were only able to produce three numeral classifiers with very low frequencies. The mean number of correct responses for the most frequently-produced numeral classifier (*orang*) by this age group was still not significant, $(p = \ldots)$

---

4 In an elicited production task, Yamamoto (2005) identified the stage of numeral classifier production by examining the first emergence of numeral classifiers for all age groups studied. Due to the cross-sectional nature of this pilot study, the mean number of correct responses for each numeral classifier between each age group fluctuates (e.g., the mean number of correct responses for *bijii* was lower among the 6-year-olds in comparison to the 5-year-olds). As a result, the Test Value “0” (i.e., no emergence of numeral classifier) was used to determine if the production of each numeral classifier was significant or not.
planned contrasts revealed that only at the age of 6 were *orang, ekor,* and *batang* significantly different from the test value “0”. Children’s production of *ketul* and *keping* was only significantly different from the test value “0” in the 10- & 11-year-old age group. Table 3.4 summarises the production order according to this analysis.

**Table 3.4. Order of Malay Numeral Classifier Production Based on Age Group**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Numeral classifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td><em>orang</em> [animate: human]</td>
</tr>
<tr>
<td></td>
<td><em>ekor</em> [animate: animal]</td>
</tr>
<tr>
<td></td>
<td><em>batang</em> [1D: +rigid]</td>
</tr>
<tr>
<td>7-year-olds</td>
<td><em>biji</em> [3D: small]</td>
</tr>
<tr>
<td></td>
<td><em>buah</em> [3D: big]</td>
</tr>
<tr>
<td></td>
<td><em>utas</em> [1D: -rigid]</td>
</tr>
<tr>
<td></td>
<td><em>helai</em> [2D: -rigid]</td>
</tr>
<tr>
<td></td>
<td><em>kimutum</em> [specific: flower]</td>
</tr>
<tr>
<td></td>
<td><em>bilah</em> [specific: knife]</td>
</tr>
<tr>
<td></td>
<td><em>kaki</em> [specific: umbrella]</td>
</tr>
<tr>
<td>8- &amp; 9-year-olds</td>
<td><em>butir</em> [3D: fine]</td>
</tr>
<tr>
<td></td>
<td><em>tongkol</em> [specific: cob of corn]</td>
</tr>
<tr>
<td>10- &amp; 11-year-olds</td>
<td><em>keping</em> [2D: +rigid]</td>
</tr>
<tr>
<td></td>
<td><em>ketul</em> [3D: medium]</td>
</tr>
</tbody>
</table>

Table 3.4 shows that animate numeral classifiers *orang* and *ekor* were the earliest to be produced (this is also evident in Figure 3.2), while all the other inanimate numeral classifiers were produced later, with a mixed order in production of shape-based and specific numeral classifiers.

In order to test the prediction that animate numeral classifiers are acquired prior to shape-based numeral classifiers, which are in turn acquired prior to specific numeral classifiers, a 3 (numeral classifier type: animate, shape-based, specific) X 7 (age group) mixed repeated measures ANOVA with numeral classifier type as a within-subjects factor and age group (<5-year-olds, 5-year-olds, 6-year-olds, 7-year-
olds, 8- & 9-year-olds, 10- & 11-year-olds, >11-year-olds) as a between-subjects factor was conducted. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 29.17, p < .05$), therefore Greenhouse-Geisser estimates of sphericity ($\varepsilon = .79$) was used to correct the degrees of freedom. There was a significant main effect of numeral classifier type, $F(1.59, 155.60) = 113.26, p < .001$, partial $\eta^2 = .536$. Pairwise comparisons (LSD = .05) showed that the production of animate numeral classifiers were significantly higher than shape-based and specific numeral classifiers, which were not significantly different.

There was a significant effect of age group, $F(6, 98) = 147.06, p < .001$, partial $\eta^2 = .900$ (Figure 3.4). The interaction between numeral classifier type and age group was also significant, $F(9.44, 165.21) = 10.50, p < .001$, partial $\eta^2 = .375$. Paired samples $t$-tests for each age group revealed that the three numeral classifier types were not significantly different from each other in the two youngest age groups. Among the 6-year-olds and the 7-year-olds, all the three response types were significantly different from each other whereas among the three oldest age groups, the animate numeral classifiers were significantly higher than shape-based and specific numeral classifiers; but the shape-based and specific numeral classifiers were not significantly different from each other. The overall results are similar to that of the three oldest age groups (i.e., Animate > Shape-based, Specific numeral classifiers).
In relation to incorrect responses, it can be seen from Figure 3.1 that the youngest children tended to give a “no numeral classifier response” whereas with age the children tended to give an “alternative classifier response”. A further analysis revealed that 163 out of 244 alternative classifier responses given by the children were *buah*, *batang*, and *biji* (see Table 3.5). In most cases, these numeral classifiers were used as the *default* numeral classifier for those that were acquired relatively late, namely *ketul*, *keping*, and *butir* (Table 3.4). As can be seen in Table 3.5, *buah* [3D: big] was used 14 times as a default numeral classifier for *ketul* [3D: medium], whereas *batang* [1D: +rigid] was used 28 times as a default numeral classifier in the place of *keping* [2D: +rigid]. In addition, *biji* [3D: small] was used as the default numeral classifier for *butir* [3D: fine] and *tongkol* [specific: cob of corn] each 19 times.
### 3.5 Discussion

The pilot study revealed that the <5-year-olds produced only one numeral classifier (*orang* [animate: human]), whereas the 5-year-olds produced three numeral classifiers, namely *orang*, *ekor*, and *batang* (Figure 3.2). Although previous cross-linguistic research indicates that in other languages, children as young as 2 and 3 years of age were already able to produce numeral classifiers in elicited production tasks (e.g., Carpenter, 1991; Hu, 1993; Yamamoto, 2005), the production of correct numeral classifiers appears to emerge relatively late in Malaysian children (Table 3.4). There is a relatively rapid development of numeral classifier acquisition between six and eight years of age among Malay children. Thus, except for three production responses, namely by the 5-year-olds, the hypothesis that Malay children would already be able to produce some Malay numeral classifiers by the age of 5 was not supported.

The findings of this pilot study support the second hypothesis, that is, Malay animate numeral classifiers were produced earlier than Malay inanimate numeral classifiers. The results found in the current study concur with previous numeral
classifier studies (Erbaugh, 1984, 1986; Matsumoto, 1985; Ng, 1989) and with findings from categorical studies that animate and inanimate items are consistently discriminated by young preschoolers (Simons & Keil, 1995). In addition, in this pilot study, the orang [animate: human] numeral classifier was produced prior to ekor [animate: animal] numeral classifier (Figures 3.2, 3.3 and Table 3.4). This production order also agrees with results found in Chinese (Erbaugh, 1984), and Japanese (Hansen & Chen, 2001) numeral classifier acquisition studies.

As can be seen from Table 3.4 some specific numeral classifiers were produced earlier than some shape-based numeral classifiers. This indicates that the production order of shape-based and specific numeral classifiers of Malay is not as clear cut as found in Cantonese Chinese (Mak, 1991) and Hokkien Chinese (Ng, 1989).

As previously mentioned in 2.3, conceptualisation of Malay specific numeral classifiers is relatively straight-forward as there is one-to-one mapping between the object and the numeral classifier, for example, all flowers take kuntum and all forms of knives take bilah. Therefore, the acquisition of specific numeral classifiers may be similar with how the name of an object is acquired. However, this is not reflected in the data as findings from the pilot indicate that the emergence of Malay specific numeral classifiers is not significantly different than the shape-based numeral classifiers. In contrast, Malay shape-based numeral classifiers are relatively more complex than the animate and specific numeral classifiers of Malay as the semantic criteria for categorisation comprise of either dimensionality and rigidity, or dimensionality and size (Salehuddin & Winskel, 2008a), i.e., mixed semantic criteria and poly-refential in nature.
Similar to findings from other numeral classifiers studies (e.g., Carpenter, 1991; Gandour et al., 1984; Hansen & Chen, 2001; Hu, 1993; Matsumoto, 1985; Ng, 1989; Uchida & Imai, 1999), Malay children tend to use alternative numeral classifiers in place of an unknown numeral classifier in their production responses. While alternative numeral classifiers in languages like Chinese, Thai, and Japanese tend to be those categorised as general numeral classifiers in the respective languages, the use of buah, biji, and batang as alternative numeral classifiers in Malay children’s production should be further investigated. This is mainly because Malay does not have a general numeral classifier whereas other numeral classifier languages do, for example, -tsu in Japanese, ge in Mandarin Chinese, and ๒ in Thai (see 1.4). However, as can be seen from the data, the use of the alternative numeral classifiers is not randomly overused. For example, in Table 3.5, buah [3D: big] was used in place of 3D numeral classifiers i.e. ketul [3D: medium] and biji [3D: small], whereas batang [1D: +rigid] was used in place of a [+rigid] numeral classifier i.e, keping [2D: +rigid]. This substitution appears to be systematic because children in the pilot study tended to substitute 3D numeral classifiers with a 3D numeral classifier and a [+rigid] numeral classifier with another [+rigid] numeral classifier.

3.6 Summary

Malay children initially produce numeral classifiers at about 4 to 5 years of age, and this begins with animate numeral classifiers before they proceed to inanimate numeral classifiers. They produce numeral classifiers for human (i.e., orang) first, before they produce the numeral classifier for animals (i.e., ekor). In place of unknown or new numeral classifiers, children appear to use an alternative numeral classifier, as the substitution responses appear to be systematic.
In conclusion, findings from the pilot study suggest that a suitable and appropriate age group for participants in the main experiments on numeral classifier acquisition is between 6 years and 9 years old. Shape-based numeral classifiers were selected as the numeral classifier type to be investigated because of the complexity of these numeral classifiers. In addition, several studies on other languages (e.g., Ng, 1989; Uchida & Imai, 1999) have investigated the production order of shape-based numeral classifier and hence form an interesting cross-linguistic comparison. In addition, it was decided to use more representative exemplars in the experimental tasks which are discussed in the following chapters.
CHAPTER 4
Preliminary and Supplementary Data Collection and Research Questions

4.1 Introduction

This chapter focuses on the preliminary and supplementary data collection in relation to the upcoming experiments. Data was collected on the degree of typicality of numeral classifier exemplars. In addition, analysis of the usage and frequency of numeral classifiers occurring in the children’s environment, including those occurring in caretaker-child interactions and the formal teaching of numeral classifiers, was conducted. Based on the literature reviewed (Chapter 2), results from the pilot study (Chapter 3), some preliminary and supplementary data collection (this chapter) was conducted to facilitate us in analysing data in the experiments.

4.2 Selection of Stimuli for the Experiments Based on Degree of Typicality

Typicality of exemplars to be categorised is of fundamental importance when children are learning to categorise and label objects. Language acquisition studies show that in general more typical members are learned earlier than less typical members of any given category (Carpenter, 1991; Erbaugh, 1984, 1986; Uchida & Imai, 1996, 1999). According to prototype theory, a prototypical member of a category is the best typical exemplar of a category (Rosch & Mervis, 1975). They are representative members of a semantic category (Lakoff, 1987). A typicality condition, that is, a categorical condition that allows exceptions, has been recognized
as a prototype (Jackendoff, 1983), which is usually established through familiarity (Rosch & Mervis, 1975).

More prototypical exemplars of a category have more attributes in common with other members of the category and at the same time have fewer attributes in common with members of contrasting categories (Rosch & Mervis, 1975). Because of this, prototypes are said to “reflect the redundancy structure of the category as a whole” (Rosch & Mervis, 1975, p. 602). Prototypes appear to be a major variable in category learning and development (Mervis & Pani, 1980) because children take advantage of the “semantic regularities” that result from structural redundancy in categorisation to organise category-assignment in their cognition (Carpenter, 1991, p.106). An individual’s familiarity with one entity in a category may represent an accurate and useful idea about a particular category and due to this, the notion of typicality (or representativeness) has a clear relation to frequency (Croft & Cruse, 2004).

The pilot study (discussed in Chapter 3) was conducted using only one exemplar for each numeral classifier category, which gives only limited information about the acquisition of numeral classifiers within a particular category. Exemplars of any given numeral classifier are not equal in the degree of their representativeness as a category member of the numeral classifier. In addition, the degree of acceptability as representatives of a numeral classifier varies according to the object and the speaker (Matsumoto, 1986, p.74). Instead of just using a single exemplar for each numeral classifier category, a range of exemplars from very typical to very atypical were used in the subsequent experiments. In order to determine the degree of representativeness of objects, a survey on the degree of typicality of eight shape-based numeral classifier exemplars was conducted, to provide a variety of items and
test stimuli associated with each numeral classifier category. The test stimuli used in the subsequent experiments are therefore more representative of the numeral classifier categories than just having a single test item for each category.

Participants

Thirty adult native speakers of Malay participated in the survey. The respondents were between 21 and 53 years of age (average 38.1 years old) and only 7 of the respondents were male. They were from various academic and SES backgrounds.

Stimuli

Participants responded to the survey electronically via email. Instructions were written in Malay on a word document and they were advised to seek further clarification regarding the instructions from the researcher (email address provided in the email) if necessary. The survey consisted of pictures of exemplars of eight Malay shape-based numeral classifiers which are common in the Malay children’s environment (Appendix 5). Participants were also required to fill in language background information about themselves before returning the survey form to the researcher.

Procedure

All respondents were asked to identify the degree of representativeness (or the typicality) of objects as members of 8 shape-based numeral classifier categories, namely batang [1D: +rigid], utas [1D: -rigid], keping [2D: +rigid], helai [2D: -rigid], buah [3D: big], ketul [3D: medium], biji [3D: small], and butir [3D: fine]. They
were given a set of pictures of objects that requires the use of the above-mentioned Malay shape-based numeral classifiers in counting activities. Participants were asked to mark number “5” if the respective object, from their point of view, was a very typical exemplar of the numeral classifier in question. Conversely, they were asked to mark “1” if the object above the number was a very atypical exemplar of the numeral classifier in question. They were asked to use the scale below (Table 4.1) as a guide. Since there were 8 shape-based numeral classifiers tested in this survey and for each numeral classifier, 5 objects were tested, there were altogether 40 pictures of objects that the respondents had to respond to in this survey. As shown in Appendix 5 (Typicality Survey), each of the pictures was labelled with the numerical score (1 through 5) for the respondents to mark.

Table 4.1. The Numerical Scoring System between 1 and 5 on the Scale from Very Atypical to Very Typical for the Numeral Classifier Exemplars

<table>
<thead>
<tr>
<th>Number</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>A very typical exemplar of the named numeral classifier.</td>
</tr>
<tr>
<td>4</td>
<td>A typical exemplar of the named numeral classifier.</td>
</tr>
<tr>
<td>3</td>
<td>A moderate (neither a typical nor an atypical) exemplar of the named numeral classifier.</td>
</tr>
<tr>
<td>2</td>
<td>An atypical exemplar of the named numeral classifier.</td>
</tr>
<tr>
<td>1</td>
<td>A very atypical exemplar of the named numeral classifier.</td>
</tr>
</tbody>
</table>

An example on how to respond to the survey was also given on the four-page survey form. Using the numeral classifier “bilah” as an example, respondents were shown pictures of a saw, a hammer, a knife, a chopper (a butcher’s knife), and a light sabre (the sword in Star Wars) as pictures of exemplars of bilah. Respondents were told that the knife was given as a very typical exemplar of bilah and because of this,
the number “5” underneath the picture of a knife was marked. In contrast, the *light sabre* was given as a very atypical exemplar of *bilah* and thus, the number “1” underneath the picture of the *light sabre* was marked (Figure 4.1). All pictures were labelled with their respective names.

![Penjodoh bilangan: Bilah](image)

<table>
<thead>
<tr>
<th>Gergaji</th>
<th>Kapak</th>
<th>Pisau</th>
<th>Pisau penolong dagub</th>
<th>Pedang Star Wars</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

Figure 4.1. Example of manner of responding to the survey. *Bilah* was used to illustrate how the survey should be responded to because *bilah* is not one of the numeral classifiers tested in the survey.

The survey form was sent to more than 50 native speakers of Malay; however only thirty respondents returned the form as a completed survey form. Out of the thirty respondents only four sought further information regarding the survey. The query was mainly regarding “what if their answers were wrong”; the respondents were politely told that there was neither right nor wrong answers to the survey as all that was needed from the survey was the point of view of individual native speakers of Malay regarding the degree of typicality of the numeral classifier exemplars in question.

Results

The responses from all 30 respondents were analysed. Because a “5” was given to a very typical exemplar of a particular numeral classifier and “1” to a very atypical exemplar, for each object, the possible summed score ranged between 150
(for a ‘very typical’ exemplar) and 30 (for a ‘very atypical’ exemplar). For all numeral classifiers, the object with the highest score was labelled as a “very typical” exemplar while the object with the lowest score was labelled as a “very atypical” exemplar. For example, ‘pencil’ scored the highest points (i.e., 138), and was then labelled as a ‘very typical exemplar of batang’ whereas ‘road’ scored the lowest (i.e., 77), and was labelled as ‘very atypical exemplar of batang’ (Table 4.2). Other objects with scores in between (e.g., ‘tree’, ‘broom’, and ‘river’ in batang) were ranked either as ‘typical’, ‘moderate’, or ‘atypical’ exemplars of the numeral classifier, depending on the score.
Paired samples $t$-tests were conducted on each numeral classifier to investigate whether or not the difference between typicality types was significant. Although overall scores showed that the scores from each typicality type were significantly different from each other (i.e., Very Typical > Typical > Moderate > Atypical > Very Atypical), the individual numeral classifier did not all have the same trend characteristics. Table 4.3 illustrates the differences.
Table 4.3. Significant Differences between the Typicality Types across All Numeral Classifiers

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>batang [1D: +rigid]</td>
<td>Vt, T &gt; M &gt; A, Va^a</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>Vt &gt; T, M, A &gt; Va</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>Vt, T &gt; M &gt; A, Va</td>
</tr>
<tr>
<td>helai [2D: -rigid]</td>
<td>Vt, T &gt; M, A, Va</td>
</tr>
<tr>
<td>buah [3D: big]</td>
<td>Vt, T &gt; T, M &gt; A, Va</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>Vt, T &gt; T, M &gt; A, Va</td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>Vt, T, M &gt; A, Va</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>Vt, T &gt; T, M &gt; A, Va</td>
</tr>
</tbody>
</table>

^a “Vt” refers to a very typical exemplar of the numeral classifier; “T” refers to a typical exemplar, “M”, a moderate exemplar; “A”, an atypical exemplar; and “Va”, a very atypical exemplar of the particular numeral classifier.

As illustrated in Table 4.3, Vt, T > M > A, Va in batang means that from the survey, the total score gathered from the respondents for very typical exemplar of batang (pencil) was not significantly higher than the typical exemplar (tree), but the scores for these two exemplars were significantly higher than the score for moderate exemplar (broom). The score for broom was significantly higher than the scores for river and road, but the scores for river and road were not significantly different from each other.

The information regarding typicality rating from this survey was used to select the test stimuli in the main experiments of this research, namely Experiments 1 through 4.

4.3 Input Frequency and Usage

Language development is influenced by the frequency and usage of lexical terms in the children’s linguistic environment (e.g., Goodman, Dale, & Li, 2008; Tare, Shatz, & Gilbertson, 2008). Varied conversational input through a variety of
linguistic and intentional contexts provides useful data for children to create early word-word mappings for non-object terms (Tare et al., 2008).

According to Mervis, Johnson and Mervis (1994), in the acquisition of subordinate categories, the kinds of linguistic information children are exposed to affect how readily children acquire those categories. The acquisition will be reinforced when children are exposed to a variety of linguistic constructions of those categories. For example, children will be able to learn the different types (i.e., subordinate categories) of shoes if the noun ‘shoes’ is used in children’s linguistic environment as compound nouns (running shoes), and nouns with relative clause or prepositional phrase adjuncts (shoes that are used for running) (Mervis et al., 1994). In addition, children’s interaction with adults also provides them with information about the form, the syntax, the semantics and the pragmatic functions of the particular words (E. V. Clark, 2003).

Research also indicates that the frequency of particular linguistic items occurring in the input to the child is a critical factor in language acquisition, particularly in early acquisition (e.g., Gallaway & Richards, 1994; Snow & Ferguson, 1977). This is because children acquire linguistic expressions which are constantly displayed in interactions within their social settings (E. V. Clark, 2003).

Some studies indicate that children tend to learn words earlier when they are produced more frequently in speech directed to them than less frequently occurring words (Gallaway & Richards, 1994; Snow & Ferguson, 1977). However, frequency alone does not necessarily influence the acquisition of items. Other factors interact with frequency. Brown (1973) for example, showed that frequency was not the
critical factor for grammatical morpheme (e.g., ‘the’ is frequent in English, but not acquired early).

Nevertheless, Tare, Shatz and Gilbertson (2008) examined the frequency and uses of subsets of colour and number words in mothers’ speech to toddlers, and of time words to preschoolers. Their findings support the argument that varied conversational input provides useful data for children to create early word-word mappings for non-object terms. In another study, Goodman, Dale, and Li’s (2008) analysis of MacArthur-Bates Communicative Development Inventory norming data and of parental input frequency from CHILDES database shows that within each of the six lexical categories studied, the more frequent the word is heard by the child, the earlier the word is produced. An important consideration which has been much debated is the best method to be used to quantify the input children are receiving. This is because the input being measured may not be a true representation of what children are actually exposed to in their linguistic environment (Hoff-Ginsberg, 1992; Pine, 1992).

Some studies suggest that preschool children’s language development can be influenced by various types of input, namely, children’s television programs (Ely & Gleason, 1995; Rice, 1990), storybook reading (Fisch, 2004; Snow, 1991; Snow, Tabor, Nicholson, & Kurkland, 1995), and non-interactive language such as overheard conversations (De Houwer, 1995). With regard to studies on numeral classifier acquisition, input frequency has been cited to be an important factor in the acquisition of Japanese numeral classifiers (Matsumoto, 1985; Yamamoto, 2005). Yamamoto (2005, p. 119), for example, found that higher frequency numeral classifiers in both speech and written texts such as -tsu, -ko, -ri, -hiki, and -dai “emerged maturationally” earlier than lower frequency numeral classifiers.
In order to examine the role of input in acquisition, research on early language development has mostly focused on data from “dyadic adult-child play settings” (Scott, 1988, p. 51). In the current study, an observation of caretaker-child interactions was also conducted to investigate numeral classifier usage in the interactions. According to Nippold (1988), for later language development, both spoken and written forms of communication are a significant source of stimulation that should be “scrutinised”. This is because, in later language development, school-going children learn language both in informal settings (through indirect modelling and reinforcement) and formal instructions. As numeral classifier acquisition takes place at a later stage in language development (especially in Malay as shown in the pilot study discussed in chapter 3) (Carpenter, 1991; Gandour et al., 1984; Mak, 1991; Ng, 1989; Sanches, 1977; Yamamoto, 2005), input from various forms of discourse and not limited to only those concerning caretaker-child interactions is likely to play a role in the acquisition of Malay numeral classifiers in children. In addition to the caretaker-child interaction data collection, an analysis of numeral classifier frequency occurring in a corpus generated from television programmes (25%), children’s story books (25%), and folk stories and bedtime stories (50%) (178,000 words) was also conducted. A similar analysis was also conducted using a larger corpus generated by Dewan Bahasa dan Pustaka (29 million words).

4.3.1 Caretaker-Child Interactions

Malay numeral classifiers are most prominently observed in counting activities. Since counting does not occur frequently in natural settings, a semi-structured elicitation procedure was used to examine numeral classifier usage in caretaker-child interactions. This task was conducted identify what numeral
classifiers caretakers would choose to talk about more in a situation where all classifiers had the same chance of being talked about. “Putar, Cari, & Kira” (“Spin, Seek & Count”) is a ‘made-up’ game designed to stimulate caretaker and child to interact with each other in a naturalistic setting, playing a counting game by spinning (putar) a spinning wheel, seeking (cari) the object that is revealed by the spinning wheel, and counting (kira) the objects in the picture book.

Participants

Thirty-five invitation letters were distributed to adults who had children aged between 3 and 5 years old living in the same house with them, to participate in the “Putar, Cari, & Kira” game (Appendix 6). However, only 11 caretakers responded and agreed to participate with their child / grandson / nephew in the game. They were all native speakers of Malay and spoke standard Malay as their first language. The caretakers were between 22 and 56 years old (mean = 37.55 years old) while the children were between 4;4 and 5;9 (mean = 5.14 years old). All participants were from middle SES. Only three of the caretakers were male and among the children, five were female. The caretaker-child relationships varied, as described in Table 4.4. However, all caretakers played a prominent or central role in bringing up the child, for example, the grandson in Pair 2 lived with the grandmother because his parents worked at an outstation, whereas the nephew in Pair 11 lived in the same house with his uncle and they both spent a lot of time together.
Table 4.4. Description of Participants in the “Putar, Cari, & Kira” Game

<table>
<thead>
<tr>
<th>Caretaker Year</th>
<th>Gender</th>
<th>Child (Year:Month)</th>
<th>Gender</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>38</td>
<td>F</td>
<td>5:7</td>
<td>M</td>
</tr>
<tr>
<td>Pair 2</td>
<td>56</td>
<td>F</td>
<td>4:7</td>
<td>M</td>
</tr>
<tr>
<td>Pair 3</td>
<td>42</td>
<td>F</td>
<td>5:6</td>
<td>M</td>
</tr>
<tr>
<td>Pair 4</td>
<td>38</td>
<td>F</td>
<td>4:7</td>
<td>M</td>
</tr>
<tr>
<td>Pair 5</td>
<td>42</td>
<td>F</td>
<td>5:9</td>
<td>F</td>
</tr>
<tr>
<td>Pair 6</td>
<td>47</td>
<td>F</td>
<td>5:6</td>
<td>M</td>
</tr>
<tr>
<td>Pair 7</td>
<td>26</td>
<td>F</td>
<td>5:8</td>
<td>F</td>
</tr>
<tr>
<td>Pair 8</td>
<td>28</td>
<td>M</td>
<td>5:0</td>
<td>F</td>
</tr>
<tr>
<td>Pair 9</td>
<td>29</td>
<td>M</td>
<td>4:4</td>
<td>F</td>
</tr>
<tr>
<td>Pair 10</td>
<td>45</td>
<td>F</td>
<td>5:0</td>
<td>F</td>
</tr>
<tr>
<td>Pair 11</td>
<td>22</td>
<td>M</td>
<td>5:0</td>
<td>M</td>
</tr>
</tbody>
</table>

Stimuli

“Putar, Cari, & Kira” consisted of a ten-page A3-size picture-book and a set of ten windowed spinning wheels. The ten-page picture-book comprised of pictures of five familiar settings to children, printed on glossy photo paper. The five different settings included a bedroom (labelled “AB”), a dining room (labelled “CD”), a kitchen (labelled “EF”), a highway (labelled “GH”), and a park (labelled “YZ”). Pictures of objects ranged from very typical to very atypical exemplars of Malay shape-based numeral classifiers (rated by adults in Table 4.2) were also printed on glossy photo paper and were later cut into picture cut-outs.

A picture of an exemplar of one numeral classifier was glued on a spinning wheel. Pictures of exemplars of the same numeral classifier (but from different typicality types) were also glued on the same spinning wheel. These pictures were positioned at an angle of 72° apart to make sure that the distance between one picture and the other was evenly distributed. There were altogether 10 spinning wheels (for 10 numeral classifiers), and each one of them was windowed so that at any one time, each spinning wheel would reveal only the picture of one object (Figure 4.2). Each

Figure 4.2. The spinning wheels for *ekor* (wheel labelled “Y”) and *orang* (“Z”). The window on each spinning wheel revealed only one picture at a time. Each picture on the spinning wheel was labelled with the object name.

The remaining picture cut-outs of the objects were glued in the picture-book on five different picture settings. The number of objects of each exemplar varied between two and four to encourage counting. Exemplars of *buah* [3D: big] and *utas* [1D: -rigid] were glued on the picture marked “AB” (bedroom setting), *keping* [2D: +rigid] and *ketul* [3D: medium] on “CD” (dining room setting), *helai* [2D: -rigid] and *butir* [3D: fine] on “EF” (kitchen setting), *biji* [3D: small] and *batang* [1D: +rigid] on “GH” (highway setting), and *ekor* [animate: animal] and *orang* [animate: human] on “YZ” (park setting) (Figure 4.3). The various picture settings were used to
contextualise the presence of the objects, and hence, stimulate conversation between caretakers and children (Appendix 7).

Figure 4.3. The A3-size picture book showing the setting of a park to contextualise the objects. “YZ” was labelled at the top right-hand corner of the book. Pictures of objects were glued on the setting to create a 3D effect.

Procedure

Each caretaker was first informed that the objective of the task was to investigate children’s development. All of them were told to interact with the child in as natural or normal a way as possible. In the process of instructing the caretakers what to do or what was expected from them, the word “penjodoh bilangan” (numeral classifier) was never mentioned by the researcher; although in demonstrating the task/game to the caretakers, the researcher used numeral classifiers when counting the objects, for example, satu ekor ayam, dua ekor ayam (one NumCl chicken, two NumCl chickens) and seorang bayi, dua orang bayi, tiga orang bayi (one NumCl baby, two NumCl babies, three NumCl babies). The interaction between each
caretaker and child were audio recorded and observed by the researcher. A clip-on microphone was attached to the child’s collar throughout the session. (See Appendix 8 for the Protocol for Caretakers).

Figure 4.4. A caretaker and a child interacting while playing the “Putar, Cari, & Kira”. The child is spinning the windowed spinning wheel.

Practice Trial

Using the “YZ” picture setting (a park) (Figure 4.3) and the “Y” and “Z” spinning wheels (exemplars of ekor [animate: animal] and orang [animate: human]) (Figure 4.2), caretakers were first shown how to play the “Putar, Cari, & Kira” game with the children. Caretakers were told that for each of the picture settings they played, the letters printed on the top right-hand corner of the picture-book must be matched with the letters that were printed at the centre of the spinning wheels; for example, the “YZ” picture setting must be played with only “Y” and “Z” spinning wheels. As a demonstration, the researcher first spinned (putar) the spinning wheel and when the wheel stopped at a particular picture (e.g., a chicken), the researcher
sought (cari) the picture cut-out of a chicken on page “YZ” and then counted (kira) the number of picture cut-outs of a chicken on page “YZ” (e.g., seekor ayam, dua ekor ayam). The researcher then handed the spinning wheel to the caretaker and asked the caretaker to repeat the procedure. Both researcher and caretaker took turns to do the spinning, seeking, and counting until all pictures on both “Y” and “Z” spinning wheels were accounted for. When the caretakers had fully understood the procedure they were told that they could now play the game with the child.

The Game

The same procedure administered in the practice trial took place in the experimental session, only this time, the game was played by the caretaker and child. Caretakers and children took turns to spin the “A” through “H” spinning wheels to seek and count the objects on pages “AB”, “CD”, “EF”, and “GH”. Each caretaker-child pair took between 12 and 22 minutes to complete the game (average 17.5 minutes).

Results

Out of the eleven caretakers, only five used at least one numeral classifier in the game, with two caretakers using all but two numeral classifiers either in counting the objects or prompting the children to count. None of the caretakers produced the butir [3D: fine] numeral classifier in their interactions. Out of the eight numeral classifiers tested, helai [2D: -rigid] was the most frequently-produced numeral classifier (12) followed by biji [3D: small] (8), and batang [1D: +rigid] (8) (Table 4.5).
Table 4.5. Frequency of Numeral Classifier Usage among Caretakers

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Pair 5</th>
<th>Pair 7</th>
<th>Pair 8</th>
<th>Pair 10</th>
<th>Pair 11</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>helai [2D: -rigid]</td>
<td>3</td>
<td></td>
<td>5</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>batang [1D: +rigid]</td>
<td>1 (2)*</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>2 (3)</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>atas [1D: -rigid]</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>buah [3D: big]</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>12 (2)</td>
<td>2</td>
<td>11 (1)</td>
<td>7</td>
<td>12 (2)</td>
<td>44 (5)</td>
</tr>
</tbody>
</table>

* Numbers in brackets represent the incorrect usage of numeral classifiers by caretakers, for example, 3 *keping piring instead of 3 biji piring.

A majority of caretakers used numeral classifiers as a prompt to get children to count. For example,

Caretaker 5 : *Ada berapa helai baju yang berwarna merah?*
there is how many NumCl shirt that colour red?
‘How many red shirts are there?’

Child 5 : *Satu, dua.*
one two
‘One, two.’

Caretaker 5 : *Ada berapa helai seluar?*
there is how many NumCl pants
‘How many pants are there?’

Child 5 : *Satu, dua*
one two
‘One, two.’

Some caretakers demonstrated counting using numeral classifiers to their children. For example,

Caretaker 7 : *Satu batang pensil, dua batang pensil, tiga batang pensil.*
one NumCl pencil two NumCl pencil three NumCl pencil
‘One pencil, two pencils, three pencils.’

Child 7 : *Satu pensil, dua pensil, tiga pensil.*
one pencil two pencil three pencil
‘One pencil, two pencils, three pencils.’
Some caretakers corrected their children when the latter wrongly used numeral classifiers, for example,

Child 11: *Pensil ada tiga *biji.*
  pencil there are three NumCl *(biji [3D: small])* 
  ‘There are three pencils.’

Caretaker 11: *Batang batang [1D: +rigid]*
  ‘Batang.’

Child 11: *Pensil ada tiga batang.*
  pencil there are three NumCl *(batang [1D: +rigid])* 
  ‘There are three pencils.’

Child 11: *Tali, tiga * biii tali*
  rope three NumCl *(biii [3D: small])* rope 
  ‘Rope, three ropes.’

Caretaker 11: *Tiga utas tali*
  three NumCl *(utas [1D: -rigid])* rope 
  ‘Three ropes.’

Child 11: *Tiga utas tali*
  three NumCl *(utas [1D: -rigid])* rope 
  ‘Three ropes.’

Only one caretaker (pair 11) insisted that the child should use numeral classifiers in their counting.

Caretaker 8: *Ada berapa keping gambar?*
  there are how many NumCl photograph? 
  ‘How many photographs are there?’

Child 8: *Satu, dua. Dua gambar.*
  one two two photograph 
  ‘One, two. Two photographs.’

Caretaker 8: *Ada berapa keping CD?*
  there are how many NumCl CD 
  ‘How many CDs are there?’

Child 8: *Satu, dua. Dua CD.*
  one two two CD 
  ‘One, two. Two CDs.’

Child 11: *Rantai, tiga rantai.*
  necklace three necklace 
  ‘Necklace, three necklaces.’
Preliminary and Supplementary Data Collection and Research Questions

Caretaker 11: *Tiga utas rantai.*
three NumCl necklace
‘Three necklaces.’

Child 11: *Tiga utas rantai.*
three NumCl necklace
‘Three necklaces.’

Child 11: *Seluar, dua seluar.*
pants two pants
‘Pants, two pants.’

Caretaker 11: *Dua helai seluar.*
two Num CL pants
‘Two pants.’

Child 11: *Dua helai seluar.*
two Num CL pants
‘Two pants.’

Two caretakers also demonstrated the use of numeral classifiers as a referring expression (anaphoric expression – refer to 1.5.3). Once the child mentioned the name of the object, the caretaker asked for the number of times the object appeared on the picture setting, using the numeral classifier, without naming the object itself.

Caretaker 5: *Ini apa?*
this what
‘What is this?’

Child 5: *Sungai.*
river
‘River.’

Caretaker 5: *Ada berapa batang?*
there are how many NumCl
‘How many are there?’

Child 5: *Satu, dua.*
one two
‘One, two.’

Caretaker 11: *Ini gambar apa?*
this picture what
‘What picture is this?’

Child 11: *Rambutan.*
rambutan
‘Rambutan.’

Caretaker 11: *Berapa biji?*
how many NumCl
‘How many?’
Only one child used numeral classifiers in his production voluntarily (i.e., not as a result of prompting or imitating the caretaker). However, this child only used *biji* [3D: small] in place of other numeral classifiers, suggesting that *biji* was functioning as a default numeral classifier for the child.

In caretakers, *biji* [3D: small] and *keping* [2D: +rigid] appear to be used as default numeral classifier. *Biji* was used as alternative numeral classifier in place of *butir* [3D: fine], whereas *keping* was used in place of *helai* [2D: -rigid] and *biji* [3D: small].
Discussion

As stated earlier, this task was aimed at identifying what numeral classifiers caretakers would choose to talk about more in a situation where all classifiers had the same chance of being talked about. The findings showed that although caretakers do use some numeral classifiers in the counting game, the accessibility of different numeral classifiers to different children varies, because of the huge individual variation (e.g., Pair 7 vs. Pair 11). This could be due to the fact that not all adults know the appropriate forms of numeral classifiers, or that they choose to use numeral classifiers randomly. Although this task is not able to show how it can influence numeral classifier acquisition order, it is able to show that caretakers do use numeral classifiers in colloquial languages when interacting with children. This task also shows that although the numeral classifier buah is the most frequent numeral classifiers found in texts, it appears to be the least frequently used numeral classifier in the counting game. The discrepancy is not explainable.

4.3.2 The Children’s Corpus

Since both spoken and written forms of communication have been argued to be a significant source of “stimulation” for older children (Nippold, 1988), written
texts were also scrutinised as another possible source of input to children in acquiring numeral classifiers. In order to gain an estimation of input frequency of the numeral classifiers in the children’s environment, a corpus from a range of texts relevant to preschool and school-aged children (henceforth ‘the Children’s corpus’) was analysed. The Children’s corpus consisted of children’s television programmes (25%), children’s story books (25%), and folk stories and bedtime stories available online in webpages by Dewan Bahasa dan Pustaka (50%), an official body in charge of promoting and preserving the Malay language. (The television programs include cartoons and educational programs targeted at children under 12 years old). From this 150,000-word corpus the word frequencies listed in Salehuddin and Winskel (2007) were obtained. Another analysis of a corpus consisting of 28,000-word children’s story books was conducted and together, the 178,000-word children’s corpus resulted in the word frequencies listed in Table 4.6. It can be seen that buah [3D: big] was the most frequently used shape-based numeral classifier followed by batang [1D: +rigid] and biji [3D: small], whereas butir [3D: fine] was the least frequently used shape-based numeral classifier in the children’s corpus.
4.3.3 Dewan Bahasa dan Pustaka Corpus

In addition to the children’s corpus, a larger corpus – not limited to children’s texts – was also used to give an estimation of numeral classifier word frequency input to the child. The corpus developed by Dewan Bahasa dan Pustaka (DBP), an authorised body that was established and managed by the government of Malaysia, was analysed for numeral classifier word frequencies. Since the size of the DBP corpus was extremely large (300 million at that time) and the corpus could not be accessed via the internet, a visit was made to the DBP Language Research Division (Bahagian Penyelidikan Bahasa DBP) to retrieve frequencies and concordances of the eight Malay numeral classifiers studied.

Method

After consultation with the officers in charge of managing the corpus, a decision was made to compile the frequency and the concordances of the eight
shape-based numeral classifiers from four different databases that would be most accessible to Malaysian readers. The databases were labelled as (a) DB3 (a compilation of books with a total of 11,137,717 words), (b) “Akhbar00 (a compilation of newspapers published in 2000 with a total of 6,800,502 words), (c) MBI (a compilation of non-academic magazines with a total of 8,128,816 words) and (d) MI (a compilation of academic magazines / journals with a total of 2,917,043 words).

Results

A search for the word frequencies of all eight Malay shape-based numeral was conducted from the 29 million-word corpus. From the search, a total of 38,713 instances were found to have one of the eight Malay shape-based numeral classifiers. However, analysis of the concordance of the words showed that only 28,884 of those instances were numeral classifiers while the rest were pseudo-numeral classifiers, which are words that are homonymous to the numeral classifiers (e.g., biji that literally means ‘seed’ and buah that literally means ‘fruit’. See 1.5.1). Since pseudo-numeral classifiers are homonymous to numeral classifiers, they are not easily tagged. As a result, the pseudo-numeral classifiers were hand searched in this analysis. As shown in Table 4.7, buah was the most frequently used numeral classifiers while utas was the least. The frequency of numeral classifier used in the 29 million-word corpus is as follows.
A Spearman’s rank correlation coefficient was conducted between the frequency of numeral classifiers in the children’s corpus (178,000) and those in DBP corpus. A strong positive relationship was found between the two databases, $r_s = .881$, $p$ (2-tailed) < .005. This indicates that to a large extent, the frequency of shape-based numeral classifiers existing in the children’s corpus can be associated in terms of ranking with the large DBP corpus (Table 4.8).

Table 4.8. Ranking of Numeral Classifier Frequency in the Children’s Corpus and the DBP Corpus

<table>
<thead>
<tr>
<th>Numeral classifier name</th>
<th>The Children’s corpus</th>
<th>DBP corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>buah [3D: big]</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>batang [1D: +rigid]</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>helai [2D: -rigid]</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>
4.3.4 The Malaysian Education System and the Formal Teaching of Numeral Classifiers

In the Malaysian education system, children begin primary school at the age of seven. Prior to this age, most children attend either private or public preschools as preparation for primary school. Children are only formally taught Malay numeral classifiers in Year 1 at 7 years of age, and at this age only 6 numeral classifiers are introduced in their textbooks, namely *orang* [animate: human], *ekor* [animate: animal], *batang* [1D: +rigid], *keping* [2D: +rigid], *buah* [3D: big], and *biji* [3D: small] (Yusof, Jeon, & Sukor, 2004). In Year 2 textbooks, three other numeral classifiers, namely *helai* [2D: -rigid], *ketul* [3D: medium], and *kuntum* [specific: flower] are introduced (A. J. Othman & Ismail, 2004). In Year 3 textbooks, the following additional numeral classifiers are introduced; *utas* [1D: -rigid], *butir* [3D: fine], *urat* [1D: +/-rigid], *pucuk* [specific: weapon, letters], *bentuk* [specific: ring/hook], and *kaki* [specific: umbrella] (Jeon, Sukor, & Yusof, 2005).

A Spearman’s rank correlation coefficient was conducted between the frequency of numeral classifiers in the children’s corpus (178,000 words) and the sequence the numeral classifiers appear in Malay textbooks. A strong positive relationship was found, \( r_s = .874, p \text{ (2-tailed)} = .005 \). Hence, the stages of teaching Malay numeral classifiers in Malaysian classrooms are closely associated with the usage frequency of numeral classifiers occurring in the Children’s corpus (Table 4.9).
Table 4.9. Ranking of Numeral Classifier Frequency in the Children’s Corpus and their Sequence of Appearance in School Textbooks

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>The Children’s corpus</th>
<th>Malay textbooks</th>
</tr>
</thead>
<tbody>
<tr>
<td>buah [3D: big]</td>
<td>1</td>
<td>Year 1</td>
</tr>
<tr>
<td>batang [1D: +rigid]</td>
<td>2</td>
<td>Year 1</td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>3</td>
<td>Year 1</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>4</td>
<td>Year 1</td>
</tr>
<tr>
<td>helai [2D: -rigid]</td>
<td>5</td>
<td>Year 2</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>6</td>
<td>Year 3</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>7</td>
<td>Year 2</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>8</td>
<td>Year 3</td>
</tr>
</tbody>
</table>

4.4 Research Questions and Hypotheses

The general aim of the study was to investigate the acquisition of Malay numeral classifiers. Based on a review of relevant literature (Chapter 2), results from the pilot study (Chapter 3), and the preliminary and supplementary data collection investigations (Chapter 4) conducted, the acquisition of eight Malay shape-based numeral classifiers were investigated from an across and within numeral classifier category perspective. The developmental patterns of Malay numeral classifier acquisition were examined, as well as how objects are categorised into particular numeral classifier categories by children. Several research questions and hypotheses were formulated and as a result, a series of experiments, which consisted of one production task and three comprehension tasks, were conducted. The participants were 140 6- to 9-year-old Malay children and a comparison adult group.

4.4.1 Developmental Patterns of Malay Numeral Classifier Acquisition

The numeral classifier acquisition order based on semantic complexity. In the pilot study, it was found that animate classifiers were produced prior to shape-
based and specific numeral classifiers and this confirms findings from previous cross-linguistic numeral classifier acquisition studies which found that animate numeral classifiers are acquired prior to inanimate numeral classifiers (e.g., Erbaugh, 1984, 1986; Matsumoto, 1985; Ng, 1989). Furthermore, it has also been found that numeral classifiers for humans are acquired prior to those for animals (Erbaugh, 1984; Hansen & Chen, 2001) and the pilot study confirms this order. It has also been found that shape-based numeral classifiers are acquired prior to specific numeral classifiers (Mak, 1991; Ng, 1989); however, this order is not supported by the pilot. Results from the pilot study indicate that Malay children produced some shape-based numeral classifiers earlier than other shape-based numeral classifiers. In order to investigate the acquisition order of the the eight Malay shape-based numeral classifiers, an Elicited Production Task (Experiment 1) and a Matching Comprehension Task (Experiment 2) were conducted. According to the semantic complexity theory, the cumulative number of semantic features in a particular semantic category plays a role in children’s categorisation patterns (E. V. Clark, 1973a; Daneman & Case, 1981; Matsumoto, 1987). Categories with less semantic features are acquired earlier than those with more semantic features. In relation to Malay shape-based numeral classifiers, it was predicted that shape-based numeral classifiers with complex features, which are the multi-form 3D numeral classifiers (four forms: fine, small, medium and big) will be acquired later than those with less complex features, which are the 1D and 2D classifiers (two forms: +rigid and –rigid).

The usage of alternative numeral classifiers. Substitutions and errors may also suggest which numeral classifier terms children have available and give clues of the acquisition order of terms within a semantic domain (Bernstein Ratner, 2000). Previous research indicates that more general numeral classifiers are usually used by
children as default numeral classifiers prior to mastery of more specific numeral classifiers, and are often overused for referents (Carpenter, 1991; Chien et al., 2003; Wei & Lee, 2001; Yamamoto & Keil, 2000). Even though Malay does not have a specific general numeral classifier per se, we can still predict that more general numeral classifiers (with a broader and more productive usage) will be used by Malay children as substitutes for other more specific numeral classifiers. The substitution of more general classifiers will give clues about the acquisition order. In addition, numeral classifiers that share the same dimensionality type and rigidity type are also used as substitutes (Loke, 1991; Mak, 1991; Ng, 1989). Thus, it was also predicted that children will substituted unknown shape-based numeral classifiers with numeral classifiers that share the same dimensionality type and rigidity type.

The role of typicality in the categorisation of objects into numeral classifier categories. According to prototype theory, children categorise typical exemplars more readily than atypical exemplars (e.g., Carpenter, 1991; Matsumoto, 1985; Uchida & Imai, 1996). Based on these studies it was predicted that more typical exemplars of a particular shape-based numeral classifier will be produced and comprehended earlier and will be classified with fewer errors than more marginal or peripheral (atypical) exemplars of the same shape-based Malay numeral classifier.

The influence of the children’s linguistic environment. Research indicates that the frequency of linguistic items occurring in the child’s environment is a critical factor in language acquisition (Gallaway & Richards, 1994; Hu, 1993; Matsumoto, 1985; Snow & Ferguson, 1977). Research conducted based on the Competition Model (MacWhinney & Bates, 1989) have illustrated that the exact recordings of words and patterns in different contexts and in many kinds of exposures influence the learning of language forms. Children will acquire a linguistic item quickly if the
particular linguistic item is reliably present in adult language input. Children, and even adults, will learn a linguistic item relatively late if the input is rare and unreliable. As a result, the linguistic item is also less strongly grasped in adults than the linguistic item that is reliably present in adult language input. It was expected that both linguistic input to the child through indirect modelling and reinforcement, and the formal teaching of numeral classifiers which begins in primary school (as discussed in 4.3), will affect the acquisition order of numeral classifiers.

4.4.2 The Categorisation of Objects into Numeral Classifier Categories

The categorisation of objects from broader to finer distinctions into numeral classifier categories. Previous studies suggest that children’s ability to sort or categorise objects proceeds through a differentiation of broader categories to much finer distinctions (e.g. Mandler et al., 1991; Mandler & McDonough, 1993). In an experiment on Japanese numeral classifier comprehension, Yamamoto and Keil (2000) found that children were able to point to the correct pictures at a younger age when the contrasts among the classifiers were maximised. When the contrasts among the classifiers were minimised, only the older children could do them correctly. In Experiment 3 (Strong-Weak Contrasts Discrimination Task) it was predicted that children would categorise shape-based numeral classifier exemplars with strong contrasts (with differences in dimensionality and rigidity or dimensionality and size) with more correct responses and faster reaction times than exemplars with weak contrasts (a difference only in either rigidity or size).

The role of typicality in the categorisation of objects into numeral classifier categories. According to prototype theory, children categorise typical exemplars more readily than atypical exemplars (e.g. Carpenter, 1991; Matsumoto, 1985;
Uchida & Imai, 1996). Thus, in Experiment 3, it was predicted that Malay children would sort typical exemplars with more correct responses and faster reaction times than atypical exemplars. This is because categorisation judgement can easily be made when objects are familiar and typical objects. In addition, research indicates that young children are strongly influenced by perceptual features and tend to categorise objects based on perceptual features rather than other semantic relations (e.g., biological relations) in determining group membership (Imai et al., 1994). According to Brooks (1978) and Allen and Brooks (1991), children start categorising items that require complex abstract rule by making use of the perceptual similarities between the new items and the known members of the category. In Experiment 4, it was predicted that young children would select the atypical exemplar of a numeral classifier category as a non-member rather than the perceptually similar non-member, whereas older children are more likely to have acquired a more refined or developed category membership representation, and consequently will be more likely to choose the perceptually similar non-member as the odd one out.

4.5 Summary

In this chapter we discussed findings from the preliminary and supplementary data collection for this research. We identified typical and atypical exemplars of the eight shape-based numeral classifiers for the main experiments in this research. This was to select a larger representative range of test stimuli for each numeral classifier. Input frequency from various databases and caretaker-child interaction was also analysed to further investigate if input plays a role in children’s acquisition of Malay shape-based numeral classifiers. Finally we presented the research questions and
hypotheses of this research, which are based on language acquisition and categorisation literature.
CHAPTER 5
Experiment 1: Elicited Production Task

5.1 Introduction

From the pilot study, it was found that Malay animate numeral classifiers are acquired prior to inanimate numeral classifiers and that numeral classifiers for humans are acquired prior to those for animals. These findings support findings from previous numeral classifier acquisition studies (Erbaugh, 1984; Hansen & Chen, 2001). Results from the pilot study also indicate that Malay children produced some shape-based numeral classifiers earlier than other shape-based numeral classifiers; and this does not support results found in studies by Mak (1991) and Ng (1989).

This chapter investigates the acquisition of Malay eight shape-based numeral classifiers that are common in Malay. Shape-based numeral classifiers were selected as the numeral classifier type to be investigated because of the complexity of these numeral classifiers. In order to investigate the acquisition order of the eight Malay shape-based numeral classifiers, an Elicited Production Task (Experiment 1) was conducted.

The aim of Experiment 1 was to investigate the production order of eight Malay shape-based numeral classifiers, namely batang [1D: +rigid], utas [1D: -rigid], keping [2D: +rigid], helai [2D: -rigid], buah [3D: big], ketul [3D: medium], biji [3D: small], and butir [3D: fine], through an enumeration task. Specifically, the
aim was to examine if the order of numeral classifier production is affected by the relative complexity, i.e. children’s categorisation patterns is influenced by the cumulative number of semantic features (E.V. Clark, 1973; Daneman & Case, 1981). Matsumoto (1987) and Uchida and Imai (1999), for example, found that individual numeral classifiers with more features are acquired later than those with less features. An additional aim was to examine how the degree of typicality of exemplars used in numeral classifier categories affects production order.

The stimuli used to elicit production responses in this experiment were pictures of the shape-based numeral classifier exemplars, each with five exemplars with different typicality ratings, for example, for *batang* [1D: +rigid], a very typical exemplar was ‘pencil’, a typical exemplar ‘tree’, a moderate exemplar ‘broom’, an atypical exemplar ‘river’, and a very atypical exemplar ‘road’. Each picture was presented to the children in random order, and they were asked to count the objects in the picture stimuli.

Based on previous literature, the following predictions were formulated:

5.1.1 Across Numeral Classifier Category Development

1. *The order of acquisition based on semantic complexity.* According to semantic complexity theory, the cumulative number of semantic features in a particular semantic category plays a role in children’s categorisation patterns (E.V. Clark, 1973a; Daneman & Case, 1981; Matsumoto, 1987). Categories with less semantic features are acquired earlier than those with more semantic features. Previous research suggests that the acquisition order of numeral classifiers within a particular language is affected by the complexity of the semantic features of the
individual numeral classifiers, that is, those with more features are acquired later than those with less features (Uchida & Imai, 1999). In relation to Malay shape-based numeral classifiers, it was predicted that shape-based numeral classifiers with complex features, which are the four-form 3D numeral classifiers (i.e., *buah* [3D: big], *ketul* [3D: medium], *biji* [3D: small] and *butir* [3D: fine]) would be produced later than those with less complex features, (i.e., the two-form 1D (i.e., *batang* [1D: +rigid] and *utas* [1D: -rigid] and 2D numeral classifiers (i.e., *keping* [2D: +rigid] and *helai* [2D: -rigid]).)

2. The usage of alternative numeral classifiers. Substitutions and errors may also suggest which numeral classifier terms children have available and give clues of the acquisition order of terms within a semantic domain (Bernstein Ratner, 2000). Previous research indicates that more general numeral classifiers are usually used by children as default numeral classifiers prior to mastery of more specific numeral classifiers, and are often overused for referents (Carpenter, 1991; Chien et al., 2003; Wei & Lee, 2001; Yamamoto & Keil, 2000). Numeral classifiers that share the same dimensionality type and rigidity type are also used as substitutes (Loke, 1991; Mak, 1991; Ng, 1989). Even though Malay does not have a specific general numeral classifier per se, we can still predict that more general numeral classifiers (with a broader and more productive usage) will be used by Malay children as substitutes for other more specific numeral classifiers. The substitution of more general classifiers will give clues about the acquisition order. As evident in the pilot study, children did fall back on particular numeral classifiers in place of unknown numeral classifiers. Eventhough Malay does not have a specific general classifier per se, we can still predict that more general numeral classifiers will be used by Malay children as substitutes for other more specific numeral classifiers, and this will give clues about
the Malay numeral classifier acquisition order. (ii) Furthermore, previous studies (e.g. Loke, 1991; Mak, 1991) have found that alternative or substitution responses for shape-based numeral classifiers tend to be ‘same dimension substitutions’ rather than ‘different dimension substitutions’. The pilot study also shows that children rarely substituted 3D numeral classifiers with 2D and 1D numeral classifiers, and vice versa. Eventhough this experiment has a larger number of exemplars, it was predicted that Malay numeral classifier substitution in this experiment would adhere to a similar trend.

5.1.2 Within Numeral Classifier Category Development

3. The role of typicality in the categorisation of objects into numeral classifier categories. According to prototype theory, children categorise typical exemplars more readily than atypical exemplars (e.g., Carpenter, 1991; Matsumoto, 1985; Uchida & Imai, 1996). Based on these studies it was predicted that more typical exemplars of a particular shape-based numeral classifier will be produced earlier and classified with fewer errors than more marginal or peripheral (atypical) exemplars of the same shape-based Malay numeral classifier. (i.e., Very Typical > Typical > Moderate > Atypical > Very Atypical).

5.2 Method

5.2.1 Participants

One hundred and forty children attending a preschool and a primary school in the same national school in Melaka, a southern state in the Peninsular of Malaysia, participated in the study. None of the children who participated in the pilot study participated in this experiment. The children were all native speakers of Malay and
spoke Malay as their first language. They spoke the standard variety of Malay, i.e. the Johor-Riau dialect of Malay due to the proximity between the state of Melaka to the state of Johor. All the 6-year-olds were preschoolers while the 7-, 8- and 9-year-olds were in their first, second and third year of the Malaysian primary education system respectively. There were thirty-one 6-year-olds, thirty-six 7-year-olds, forty-one 8-year-olds and thirty-two 9-year-olds. 20 adults also participated in the experiment as a comparison group. The adults lived in the vicinity of the school and were from a mixed educational background. Flyers were earlier sent to adults to invite them to participate in the experiment. The adults responded to the invitation either via face-to-face interaction or telephone conversation. All participants were from middle SES. A description of the participants is given in Table 5.1.

Table 5.1. Description of Participants in the Study

<table>
<thead>
<tr>
<th>Age group</th>
<th>Age range</th>
<th>Mean age</th>
<th>No of participants</th>
<th>No of males</th>
<th>No of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td>5.8 – 6.7</td>
<td>6.18</td>
<td>31</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>6.8 – 7.6</td>
<td>7.13</td>
<td>36</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>8-year-olds</td>
<td>7.9 – 8.8</td>
<td>8.25</td>
<td>41</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td>9-year-olds</td>
<td>8.11 – 9.8</td>
<td>9.28</td>
<td>32</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Adults</td>
<td>17.3 – 77.8</td>
<td>48.07</td>
<td>20</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>160</strong></td>
<td><strong>63</strong></td>
<td><strong>97</strong></td>
</tr>
</tbody>
</table>

5.2.2 Stimuli

The stimuli consisted of pictures of objects which were selected based on earlier ratings by 30 adult Malay native speakers (refer to 4.2). All exemplars rated as very typical through to very atypical exemplars were used as stimuli and appeared in units of three so as to encourage the children to use a number, a numeral classifier and a noun in naming the respective objects.
5.2.3 Procedure

5.2.3.1 Picture-Familiarisation Session

Prior to the experimental session, a slide display of all 48 pictures with an audio presentation of the names of the respective objects in the picture-display were shown to the children (Figure 5.1). In this picture-familiarisation session, children were asked to repeat the names of the objects after the audio presentation of the respective objects before proceeding to the next slide. This was to ensure that the children were familiar with the items presented to them. They could also ask for further clarification in this session.

![Figure 5.1. A slide from the picture-familiarisation session.](image)

5.2.3.2 Practice Trial

Four of the 48 pictures were used for practice trials (Table 5.2). Children were shown what to say when a prompt was given to them, for example, “Tiga bentuk cincin” (“Three NumCl [specific: ring/hook] rings”) and “Tiga pasang cermin mata” (“Three NumCl [pair] glasses”) respectively. The numeral classifiers used in the practice trial were not used in the experimental session. A prompt in the form of the expression “Berapa?” (“How many?”) was given to the children if they merely
mentioned the name of the object, for example, “Cincin” (“Ring”). The children were further prompted with “Tiga apa cincin?” (“Three what rings?”) if they merely gave “Tiga cincin” (“Three rings”) as their answers, with ‘apa’ (“what”) deliberately stressed to signal to the children as well as to emphasise that they were required to fill the gap between the numeral and the noun. Children were prompted with “Tiga apa cincin?” for a second time if they still gave “Tiga cincin” as their answers. If children’s answer remained without the use of numeral classifiers, e.g. “Tiga cincin” after two prompts of “apa”, the researcher will show them the correct way to say “three rings”, i.e. “tiga bentuk cincin”. Only when the researcher was convinced that the participants had understood the task did the experimental session commence.

Table 5.2. Description of Practice Trial Items

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Sample response</th>
<th>Numeral classifier required</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 rings</td>
<td>3 bentuk cincin</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 hooks</td>
<td>3 bentuk mata kail</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 wings</td>
<td>3 pasang sayap</td>
<td>pasang</td>
</tr>
<tr>
<td>3 glasses</td>
<td>3 pasang cermin mata</td>
<td>pasang</td>
</tr>
</tbody>
</table>

5.2.3.3 The Experiment

In the experimental session, children were asked to count the items that were presented on the Microsoft Powerpoint slide presentation one by one. The correct expected response was to count the items using the correct numeral classifier, for example, “Tiga batang pensil” (“Three NumCl pencil”). All the objects require Malay shape-based numeral classifiers in the counting activity. The pictures used as
test stimuli in this experiment are described in Table 5.3. Pictures were presented in a fixed random order to all the children. This means that the pictures were arranged in no particular order; however, all pictures were shown to the children in the same order.

In contrast to the practice trial, the children were not prompted or doubly prompted with “apa” in the experimental session as they were in the practice trial. The experimental session lasted about 20 minutes for each child.
Table 5.3. Description of Test Stimuli for Experiment 1

<table>
<thead>
<tr>
<th>Numeral classifier tested</th>
<th>Numeral classifier type</th>
<th>Stimulus</th>
<th>Typicality</th>
<th>Correct response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sepang</td>
<td>1D: rigid</td>
<td>pencils</td>
<td>Very Typical</td>
<td>3 balang pencil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>trees</td>
<td>Typical</td>
<td>3 balang pokok</td>
</tr>
<tr>
<td></td>
<td></td>
<td>brooms</td>
<td>Moderate</td>
<td>3 balang pongpu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rulers</td>
<td>Atypical</td>
<td>3 balang sungan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>roads</td>
<td>Very Atypical</td>
<td>3 baung jeban</td>
</tr>
<tr>
<td>Litar</td>
<td>1D: rigid</td>
<td>necklaces</td>
<td>Very Typical</td>
<td>3 was rameen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ropes</td>
<td>Typical</td>
<td>3 was tali</td>
</tr>
<tr>
<td></td>
<td></td>
<td>strands of thread</td>
<td>Moderate</td>
<td>3 was hong</td>
</tr>
<tr>
<td></td>
<td></td>
<td>watches</td>
<td>Atypical</td>
<td>3 was iam tungan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>chain link</td>
<td>Very Atypical</td>
<td>3 was rama kunz</td>
</tr>
<tr>
<td>Kepong</td>
<td>2D: rigid</td>
<td>photographs</td>
<td>Very Typical</td>
<td>3 kepong gambar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wooden plank</td>
<td>Typical</td>
<td>3 kepong papun</td>
</tr>
<tr>
<td></td>
<td></td>
<td>baskets</td>
<td>Moderate</td>
<td>3 kepong bidik</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CDs</td>
<td>Atypical</td>
<td>3 kepong CD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>slices of cakes</td>
<td>Very Atypical</td>
<td>3 kepong kah</td>
</tr>
<tr>
<td>Holau</td>
<td>2D: rigid</td>
<td>sheets of papers</td>
<td>Very Typical</td>
<td>3 holau kertas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hats</td>
<td>Typical</td>
<td>3 holau jila</td>
</tr>
<tr>
<td></td>
<td></td>
<td>leaves</td>
<td>Moderate</td>
<td>3 holau daun</td>
</tr>
<tr>
<td></td>
<td></td>
<td>handkerchief</td>
<td>Atypical</td>
<td>3 holau sayakang</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pairs of pants</td>
<td>Very Atypical</td>
<td>3 holau sekar</td>
</tr>
<tr>
<td>Biah</td>
<td>3D: big</td>
<td>boxes</td>
<td>Very Typical</td>
<td>3 biah bus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>boxes</td>
<td>Typical</td>
<td>3 biah kotak</td>
</tr>
<tr>
<td></td>
<td></td>
<td>families</td>
<td>Moderate</td>
<td>3 biah kebarja</td>
</tr>
<tr>
<td></td>
<td></td>
<td>planets</td>
<td>Atypical</td>
<td>3 biah planet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>robots</td>
<td>Very Atypical</td>
<td>3 biah robok</td>
</tr>
<tr>
<td>Ketul</td>
<td>3D: med</td>
<td>stones</td>
<td>Very Typical</td>
<td>3 ketul kantu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pieces of meat</td>
<td>Typical</td>
<td>3 ketul dagung</td>
</tr>
<tr>
<td></td>
<td></td>
<td>drumsticks</td>
<td>Moderate</td>
<td>3 ketul pajak ayam</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gold ingots</td>
<td>Atypical</td>
<td>3 ketul jongkong umas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>chocolates</td>
<td>Very Atypical</td>
<td>3 ketul cockat</td>
</tr>
<tr>
<td>Bitu</td>
<td>3D: small</td>
<td>rambutans</td>
<td>Very Typical</td>
<td>3 bitu rambutans</td>
</tr>
<tr>
<td></td>
<td></td>
<td>balls</td>
<td>Typical</td>
<td>3 bitu bola</td>
</tr>
<tr>
<td></td>
<td></td>
<td>oranges</td>
<td>Moderate</td>
<td>3 bitu oron</td>
</tr>
<tr>
<td></td>
<td></td>
<td>caps</td>
<td>Atypical</td>
<td>3 bitu arion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>plates</td>
<td>Very Atypical</td>
<td>3 bitu pinggan</td>
</tr>
<tr>
<td>Butir</td>
<td>3D: fine</td>
<td>grains of rice</td>
<td>Very Typical</td>
<td>3 butir beras</td>
</tr>
<tr>
<td></td>
<td></td>
<td>stars</td>
<td>Typical</td>
<td>3 butir basang</td>
</tr>
<tr>
<td></td>
<td></td>
<td>precious stones</td>
<td>Moderate</td>
<td>3 butir kati permata</td>
</tr>
<tr>
<td></td>
<td></td>
<td>seeds</td>
<td>Atypical</td>
<td>3 butir biji bendah</td>
</tr>
<tr>
<td></td>
<td></td>
<td>grains of sugar</td>
<td>Very Atypical</td>
<td>3 butir gula putir</td>
</tr>
</tbody>
</table>

5.2.4 Coding of Responses

Each of the children’s production responses was classified into either one of the two main categories of correct or incorrect response, and if incorrect, their responses were further classified into three subcategories.
(1) “Correct response” when children’s responses resembled those described in Table 5.3 (Column 5).

(2) “Incorrect response” when children’s responses were different from described in Table 5.3. This type of response was further divided into three subcategories:

a) “No classifier response” when no numeral classifier was used in the response, for example, instead of “3 *batang pensil*” (3 [stem] or [1D: +rigid] pencil), “*3 pensil*” (3 pencils) was used.

b) “Alternative classifier response” when an alternative numeral classifier was used instead of the correct numeral classifier, for example, for “3 *batang pensil*” (3 [stem] or [1D: +rigid] pencil), “*3 buah pensil*” (3 [3D: big] pencil) was substituted.

c) “Other response” when another response besides an alternative numeral classifier response was used,

  o “Unitisers” when a Malay expression to quantify mass nouns was used, for example, “3 *genggam gula pasir*” (3 handfuls of sugar) for “3 *butir gula pasir*” (3 grains of sugar). Although this is grammatically correct, it is not a numeral classifier.

  o “Collective Nouns” when a Malay expression to quantify multiple count nouns via a larger object was used, for example, “3 *tangkai daun*” (3 stalks of leaves) for “3 *helai daun*” (3 pieces of leaves). Although this is grammatically correct, it is not a numeral classifier.

  o “Measure words” when an international measuring unit was used, for example, “3 *kilogram gula pasir*” (3 kilograms of sugar)
*Elicited Production Task*

butir gula pasir” (3 grains of sugar). Although this is grammatically correct, it is not a numeral classifier.

- “Content words” when a noun, a verb or an adjective was used, for example, “*3 tajam batu*” (3 sharp stones) for “*3 ketul batu*” (3 chunks of stones).

### 5.3 Results

#### 5.3.1 Across Numeral Classifier Category Development

Only a small number of correct Malay shape-based numeral classifiers (5.25%) were produced by the youngest age group, (i.e., the 6-year-olds). As expected, the mean number of correct responses increased with age; the 7-year-olds gave 18.20% correct responses, the 8-year-olds 41.53%, and the 9-year-olds 51.25%. Despite showing the highest percentage of correct responses amongst the children, the responses given by the 9-year-olds were markedly different from the proportion of correct responses given by the adults (87.63%)

Figure 5.2 shows that the 6-year-olds only produced three correct numeral classifiers, namely *biji* [3D: small], *batang* [3D +rigid], and *helai* [2D: -rigid], whereas the 7-year-olds produced all but one of the shape-based numeral classifiers tested, that is, *butir* [3D: fine]. The 8- and 9-year-olds produced all the eight tested numeral classifiers.

---

5 Adults in Matsumoto’s (1985) experiment produced the expected numeral classifier response 85% of the time.
In order to compare the number of correct responses for the different numeral classifiers in the different age groups and across gender, a $8 \times 5 \times 4 \times 2$ (Numeral Classifier X Typicality Type X Age Group X Gender) mixed repeated measures ANOVA, with numeral classifier and typicality type as within-subjects factors, and age group (6-year-olds, 7-year-olds, 8-year-olds, 9-year-olds) and gender as between-subjects factors was conducted. (Adults’ responses were not analysed because the aim of this experiment was to investigate the developmental pattern in children). Mauchly’s test indicated that the assumption of sphericity for numeral classifier, typicality, and numeral classifier X typicality had been violated ($\chi^2(27) = 74.14, p < .001; \chi^2(9) = 36.85, p < .001; \chi^2(405) = 990.31, p < .001$ respectively). As a result, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .86; \varepsilon = .87; \varepsilon = .66$). There was a significant main effect of numeral classifier ($F(6.02, 794.99) = 90.62, p < .001$, partial $\eta^2 = .407$), a significant effect of typicality type ($F(3.48, 458.78) = 107.61, p < .001$, partial $\eta^2 = .449$), and a
significant main effect of age group ($F(3, 132) = 105.59, p < .001$, partial $\eta^2 = .706$). There was not a significant effect of gender ($p = .02$). There was also a significant interaction effect between numeral classifier and age group ($F(21, 924) = 12.07, p < .001$, partial $\eta^2 = .215$). Tukey’s post hoc analysis at $\alpha = .05$ revealed that correct responses for the different age groups were all significantly different from each other.

The mean number of correct responses for the eight shape-based numeral classifiers by the children is listed in Table 5.4. This also gives an indication of the acquisition order of Malay shape-based numeral classifiers based on the production task. Planned contrasts revealed that all numeral classifier pairs except *buah* and *biji* were significantly different from each other; *helai* and *batang* ($t(139) = 4.35, p < .001$), *batang* and *buah* ($t(139) = 3.46, p = .001$), *biji* and *utas* ($t(139) = 4.50, p < .001$), *utas* and *keping* ($t(139) = 2.08, p < .005$), *keping* and *ketul* ($t(139) = 3.65, p < .001$), and *ketul* and *butir* ($t(139) = 3.84, p = .005$). The order of shape-based numeral classifier production shows that while one of the 2D numeral classifiers (*helai*) was produced the earliest, the other 2D numeral classifier (*keping*) was only produced after the production of two 1D and two 3D numeral classifiers.
Table 5.4. Stage of Malay Shape-based Numeral Classifier Acquisition Based on the Mean Number of Correct Production Responses across All Age Groups

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Mean number of correct responses</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>helai [2D: -rigid]</td>
<td>.53</td>
<td>.40</td>
</tr>
<tr>
<td>batang [1D: +rigid]</td>
<td>.43</td>
<td>.27</td>
</tr>
<tr>
<td>buah [3D: +big], biji [3D: small]</td>
<td>.56, .55</td>
<td>.32, .27</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>.24</td>
<td>.32</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>.18</td>
<td>.24</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>.12</td>
<td>.22</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>.05</td>
<td>.17</td>
</tr>
</tbody>
</table>

If we examine individual classifiers, there was a significant main effect of age group for batang [1D: +rigid] \((F(3, 136) = 36.51, p < .001, \text{partial } \eta^2 = .446)\), for utas [1D: -rigid] \((F(3, 136) = 53.27, p < .001, \text{partial } \eta^2 = .540)\), for keping [2D: +rigid] \((F(3, 136) = 35.72, p < .001, \text{partial } \eta^2 = .441)\), for helai [2D: -rigid] \((F(3, 136) = 68.12, p < .001, \text{partial } \eta^2 = .600)\), for buah [3D: +big] \((F(3, 136) = 54.25, p < .001, \text{partial } \eta^2 = .545)\), for ketul [3D: medium] \((F(3, 136) = 14.69, p < .001, \text{partial } \eta^2 = .245)\), for biji [3D: small] \((F(3, 136) = 17.40, p < .001, \text{partial } \eta^2 = .277)\), and for butir [3D: fine] \((F(3, 136) = 9.04, p < .001, \text{partial } \eta^2 = .166)\). Tukey’s post hoc analysis at \(\alpha = .05\) revealed that the mean number of correct numeral classifier responses between the 6- and the 7-year-olds were significantly different for batang [1D: +rigid], helai [2D: -rigid] and buah [3D: +big], whereas the mean number of correct numeral classifier responses between the 8- and the 9-year-olds were significantly different for utas [1D: -rigid], buah [3D: +big], and butir [3D: fine]. Tukey’s post hoc analysis also shows that with the exception of butir [3D: fine], the
mean number of correct responses for the 7- and 8-year-olds were significantly different for all numeral classifiers.

In relation to the incorrect responses, it can be seen from Figure 5.3 that the youngest children tended to give a “No Classifier” response, whereas with age the children tended to give an “Alternative Classifier” response. In total, 21 numeral classifiers were used as alternative substitution responses. Sixteen ‘other responses’ given by the children included 10 unitisers (e.g., gelas (a glass of), cubit (a pinch of), and bungkus (a packet of)), 4 collective nouns (e.g., kumpulan (a group of) and baris (a row of)), 2 measure words (e.g., batu (a mile) and kilo (a kilogram/kilometer)), while the remaining 28 responses were nouns and adjectives (e.g., bingkai (frame) and lurus (straight)).

![Figure 5.3](image_url)

Figure 5.3. Mean number of the four response types used by each age group. “Others” includes ‘unitisers’, ‘collective nouns’, ‘measure words’, and ‘content words’.
As can be seen from Table 5.5, the numeral classifier *buah* [3D: big] (446 occurrences) and *biji* [3D: small] (422 occurrences) were the most frequently-used alternative numeral classifiers. *Keping, helai, batang, ketul*, and *utas* were used as alternative numeral classifiers to a lesser extent. Other numeral classifiers used as alternative numeral classifiers included *pasang* [arrangement: pair] (89 instances), *orang* [animate: human] (55), and *ekor* [animate: animal] (36). *Butir* (292), *ketul* (274), followed by *keping* (221) had the most substitutions made for them.

The numeral classifiers *buah* [3D: big] and *biji* [3D: small] were used predominantly in place of 3D numeral classifiers; *buah* was mainly used in place of other 3D numeral classifiers (51%) rather than replacing 1D (29%) or 2D (19%) numeral classifiers, and *biji* was used predominantly in place of other 3D numeral classifiers (77%) in contrast to 1D (8%) and 2D (15%) numeral classifiers. *Helai* [2D: -rigid] was used as an alternative numeral classifier predominantly in place of *utas* [1D: -rigid] (48%) and *keping* [2D: +rigid] (39%). *Batang* [1D: +rigid] was mainly used as a substitution response for *keping* [2D: +rigid] (61%).

Table 5.5. The Frequency and Percentage of the Shape-based Numeral Classifiers Used as Alternative or Substitution Responses for Correct Responses

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kepting, helai, batang, ketul, and utas were used as alternative numeral classifiers. Butir (292), ketul (274), followed by keping (221) had the most substitutions made for them.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>448</td>
</tr>
</tbody>
</table>
5.3.2 Within Numeral Classifier Category Development

Pairwise comparisons based on estimated marginal means (LSD = .01) indicated that in the overall results, correct responses for very typical exemplars were significantly higher than typical exemplars, which in turn were significantly higher than moderate exemplars. Correct responses for moderate exemplars were significantly higher than atypical exemplars, but correct responses for atypical exemplars were not significantly different from very atypical exemplars ($p = .24$). The overall pattern was Very Typical $>$ Typical $>$ Moderate $>$ Atypical, Very Atypical (Table 5.6).

Table 5.6. Mean Number of Correct Responses Based on Typicality Type

<table>
<thead>
<tr>
<th>Exemplar type</th>
<th>Mean number of correct responses</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Typical</td>
<td>.45</td>
<td>.29</td>
</tr>
<tr>
<td>Typical</td>
<td>.32</td>
<td>.23</td>
</tr>
<tr>
<td>Moderate</td>
<td>.27</td>
<td>.24</td>
</tr>
<tr>
<td>Atypical</td>
<td>.20</td>
<td>.20</td>
</tr>
<tr>
<td>Very atypical</td>
<td>.18</td>
<td>.19</td>
</tr>
</tbody>
</table>

There was also a significant interaction between typicality type and age group ($F(12, 528) = 9.41, p < .001$, partial $\eta^2 = .176$). Pairwise comparisons (LSD = .01) for typicality type for each age group showed that correct responses for very typical exemplars were significantly higher among all children except the 6-year-olds, whereas correct responses for atypical and very atypical exemplars were not significantly different from each other for all age groups. Correct responses for typical exemplars were significantly higher than moderate exemplars only in the 7-
year-olds, while correct responses for moderate exemplars were significantly higher than atypical exemplars only in the 6- and 8-year-olds (Figure 5.4).

Figure 5.4. Mean number of correct responses by each age group based on typicality type.

There was also a significant interaction between typicality type and numeral classifier \((F(18.54, 2447.01) = 24.53, p < .001, \text{partial } \eta^2 = .157)\) (Figure 5.5), and a significant interaction between typicality type, numeral classifier, and age group \((F(84, 3696) = 5.13, p < .001, \text{partial } \eta^2 = .104)\). Correct responses for atypical exemplars of utas (watch), typical exemplars of keping (wooden plank), moderate exemplars of buah (family), typical exemplars of biji (ball), typical exemplars of butir (star), and moderate exemplars of butir (precious stones) were significantly different from the overall numeral classifier X typicality type results.
5.4 Discussion

This experiment reveals that Malay numeral classifier acquisition pattern shows a particular order of development. The production order began with one 2D numeral classifier and followed by one 1D numeral classifier. This order was then followed by two 3D numeral classifiers. At this level, the production order appears to support the order predicted from the semantic complexity theory, i.e. the cumulative number of semantic features in a particular semantic category is said to play a role in the development of children’s categorisation ability (Clark, 1973a; Daneman & Case, 1981). Based on this theory, it was earlier predicted that in this experiment, children would produce the two-form 1D and 2D numeral classifiers earlier than the four-form 3D numeral classifiers.

However, the emergence of two 3D numeral classifiers (i.e. buah [3D: big] and biji [3D: small]) in the order of acquisition took place relatively earlier than the
predicted order. Although these two numeral classifiers emerged after one 2D numeral classifier \textit{helai} [2D: -rigid] and one 1D numeral classifier \textit{batang} [1D: +rigid], these 3D numeral classifiers appeared prior to the other 1D and 2D numeral classifiers, namely \textit{ketul} [3D: medium] and \textit{butir} [3D: fine]. This suggests that the production order of Malay shape-based numeral classifiers is not solely influenced by the complexity of the numeral classifier categories. There appears to be other factors that can also influence the order of shape-based numeral classifier acquisition, which results in the earlier emergence of \textit{buah} [3D: big] and \textit{biji} [3D: small] than the predicted order; and one of the factors could be the high frequency of these two numeral classifiers in the Malay children’s linguistic environment.

As indicated from the preliminary and supplementary data collection, \textit{buah} and \textit{biji} were the most frequently used numeral classifiers in the DBP corpus and the children’s corpus. \textit{Buah}, for example, appeared 21,468 times in the DBP corpus and 546 times in the Children’s Corpus. Although \textit{biji} appeared 4,095 times in the DBP corpus and only 74 times in the Children’s Corpus, its frequency is relatively high in comparison with the other numeral classifiers. The higher frequency of \textit{buah} and \textit{biji} could be the reason why these two numeral classifiers emerged earlier in children’s production despite the complexity of their semantics.

Apart from analysing the correct responses, children’s incorrect responses in the form of substitutions and errors may also suggest which numeral classifier terms children have available and give clues of the acquisition order of terms within a semantic domain (Bernstein Ratner, 2000). Our analysis of the substitutions and errors made by the participants indicated that children produced \textit{buah} [3D: big] most frequently in place of \textit{biji} [3D: small] (23%) and \textit{butir} [3D: fine]; and they produced \textit{biji} [3D: small] most frequently in place of \textit{butir} [3D: fine] (41%) and \textit{ketul} [3D:
medium] (26%). This suggests that buah was acquired prior to biji and butir, and biji was acquired prior to butir and ketul. This order also corresponds to the order of correct production responses showed in this experiment.

Previous research indicates that more general numeral classifiers are usually used by children as default numeral classifiers prior to the mastery of more specific numeral classifiers, and are often overused for referents (Carpenter, 1991; Chien et al., 2003; Wei & Lee, 2001; Yamamoto & Keil, 2000). As mentioned earlier, Malay does not have a general classifier that would semantically substitute numeral classifiers correctly. If we look at Table 5.5, we can see that about 60% of the time, children were using buah and biji. The substitutions they made suggests that children used these numeral classifiers as default classifiers; buah for big items, and biji for small items. The use of these two numeral classifiers as default classifiers by the children could also be due to the availability of these numeral classifiers in their linguistic environment. Since buah classifies big, rounded objects, children tended to classify big objects (e.g., tree, road, and river), as buah. Similarly, since biji classifies small, rounded objects, children tended to classify small objects (e.g., stone, seeds, and rice), as biji.

Apart from using buah and biji in place of unknown numeral classifiers, children also tended to use an alternative numeral classifier that shares the same dimensionality type (henceforth ‘within-the-same-dimension’ numeral classifier) in place of the correct numeral classifier. For example, biji [3D: small] was used as an alternative numeral classifier predominantly in place of other 3D numeral classifiers (77%). In addition, children also tended to use helai [2D: -rigid] as an alternative numeral classifier in place of utas [1D: -rigid] and batang [1D: +rigid] in place of keping [2D: +rigid]. This shows that the substitution was not only based on the
similarity of the dimensionality of objects; substitution was also conducted based on
the similarity of the rigidity of the objects.

Children produced correct responses on very typical and typical exemplars of
shape-based numeral classifiers more than they did on atypical and very atypical
exemplars. This pattern was observed in the overall shape-based numeral classifiers
order. However, there were exceptions to this trend in the acquisition of individual
classifiers. For example, correct responses for very atypical exemplar of *utas*
(watch), typical exemplar of *keping* (wooden plank), moderate exemplar of *buah*
(family), typical exemplar of *biji* (ball), and typical and moderate exemplar of *butir*
(star and precious stone) (Figure 5.5) were all significantly low in comparison to
their typical pattern of correct responses. The low number of correct responses for
these exemplars is possibly due to the different ways these objects may be perceived
by the children. For example, watch was most of the time classified as *buah* and this
could be due to the fact that a watch may be perceived by the children as closely
related to a clock, which is classified by the numeral classifier *buah*. Because of the
similarity between the face of the watch with that of a clock, children classified a
watch as *buah* rather than *utas*.

In this chapter, we have discussed the acquisition of shape-based numeral
classifiers by examining children’s production through an elicited production task.
The following chapter discusses the experiment conducted to investigate children’s
comprehension of Malay shape-based numeral classifiers.
CHAPTER 6
Experiment 2: Matching Comprehension Task

6.1 Introduction

This chapter focuses on comprehension of the same eight Malay shape-based numeral classifiers as used in the elicited production task, Experiment 1, through a matching comprehension task. The aim of this experiment was to explore the developmental patterns observed in the comprehension of Malay shape-based numeral classifiers. Specifically, it was aimed at examining if the complexity of the semantic features of individual numeral classifiers plays a role in numeral classifier comprehension order. This experiment also aims to investigate if the degree of typicality of numeral classifier exemplars has an effect on the comprehension order. An additional aim was to examine the relationship between production and comprehension of numeral classifiers.

The objects used in this experiment were pictures of very typical, typical, moderate, atypical and very atypical exemplars of the shape-based numeral classifiers. A picture of an exemplar of all shape-based numeral classifiers from the same typicality type (e.g., pictures of very typical exemplar of batang, very typical exemplar of utas, very typical exemplar of keping, helai, buah, ketul, biji, and butir) were simultaneously shown to the children on an A4 paper. Children were asked to match the numeral classifier’s name, which was read to them, to the correct picture. Based on the literature reviewed, the following predictions were formulated.
6.1.1 Across Numeral Classifier Category Development

1. *The order of acquisition based on semantic complexity.* Based on semantic complexity theory we can predict that correct comprehension responses to the 1D and 2D numeral classifiers will be higher than the 3D numeral classifiers (i.e. 1D, 2D > 3D).

5.1.2 Within Numeral Classifier Category Development

2. *The role of typicality in the categorisation of objects into numeral classifier categories.* The role of typicality was also examined in the comprehension of numeral classifiers. It was predicted that more typical exemplars of a particular shape-based numeral classifier will be comprehended earlier and matched to the correct numeral classifiers with fewer errors than more marginal or peripheral (atypical) exemplars of the same shape-based Malay numeral classifiers (i.e., Very Typical > Typical > Moderate > Atypical > Very Atypical).
6.2 Method

6.2.1 Participants

One hundred and forty-one children attending a preschool and a primary school in the same national school participated in the study. A majority of the children who participated in the elicited production task, which was conducted one week prior to the matching comprehension task, participated in this experiment.

The children were categorized into four age groups: thirty-one 6-year-olds, thirty-seven 7-year-olds, forty-two 8-year-olds and thirty-one 9-year-olds. They were all native speakers of Malay and spoke Malay as their first language and a majority of them participated in the elicited production task. The same 20 adult native speakers of Malay who participated in the elicited production task also participated in this experiment as a comparison group. A description of the participants is given in Table 6.1.

Table 6.1. Description of Participants in Experiment 2

<table>
<thead>
<tr>
<th>Age group</th>
<th>Age range</th>
<th>Mean age</th>
<th>No. of participants</th>
<th>No. of males</th>
<th>No. of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td>5;8 – 6;7</td>
<td>6.18</td>
<td>31</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>6;8 – 7;6</td>
<td>7.13</td>
<td>37</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>8-year-olds</td>
<td>7;9 – 8;8</td>
<td>8.25</td>
<td>42</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>9-year-olds</td>
<td>8;11 – 9;8</td>
<td>9.28</td>
<td>31</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>Adults</td>
<td>17;3 – 77;8</td>
<td>48.07</td>
<td>20</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>160</td>
<td>63</td>
<td>97</td>
</tr>
</tbody>
</table>
The matching comprehension task was chosen as a means to test children’s comprehension because matching exercises are commonly found in children’s fun activity books, workbooks and textbooks, and are familiar to Malay children.

6.2.2 Stimuli

This experiment involved the use of the same set of pictures of exemplars of all the eight shape-based numeral classifiers tested earlier in the elicited production task. All pictures of numeral classifier exemplars ranked as very typical, typical, moderate, atypical, and very atypical exemplars in the survey (4.2) were now grouped in five different sets of pictures so that each set had exemplars of all the eight numeral classifiers from the same typicality ranking (e.g., the first set had pictures of very typical exemplar of all eight numeral classifiers, the second set had pictures of typical exemplar of all eight numeral classifiers, etc.). In this matching comprehension task, all eight 25mm X 25mm pictures of shape-based numeral classifier exemplars from the same set were arranged randomly in vertical order and presented simultaneously in one column to the participants on a single A4-size paper (Figure 6.1). The names of the numeral classifiers tested in this experiment (e.g., batang, utas) were placed vertically in a fixed order in the middle column while pictures of shape-based numeral classifier exemplars from another set were randomly put in the third column. The test stimuli for this task are as described in Table 6.2.
Table 6.2 Description of Test Stimuli for Experiment 2

<table>
<thead>
<tr>
<th>Page 1</th>
<th>Page 2</th>
<th>Page 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left column (Very atypical)</td>
<td>Middle column</td>
<td>Right column (Typical)</td>
</tr>
<tr>
<td>Paper</td>
<td>Batting</td>
<td>Photograph</td>
</tr>
<tr>
<td>Rambutan</td>
<td>Keep</td>
<td>Ball</td>
</tr>
<tr>
<td>Chain</td>
<td>Usb</td>
<td>Box</td>
</tr>
<tr>
<td>Pencil</td>
<td>Helix</td>
<td>Star</td>
</tr>
<tr>
<td>Rice</td>
<td>Batch</td>
<td>Tree</td>
</tr>
<tr>
<td>Stone</td>
<td>Kettel</td>
<td>Skirt</td>
</tr>
<tr>
<td>Blank</td>
<td>Biji</td>
<td>Rope</td>
</tr>
<tr>
<td>Box</td>
<td>Batter</td>
<td>Mest</td>
</tr>
<tr>
<td>Left column (Typical)</td>
<td>Middle column (Clustering name)</td>
<td>Right column (Very atypical)</td>
</tr>
<tr>
<td>Road</td>
<td>Running</td>
<td>Watch</td>
</tr>
<tr>
<td>Robot</td>
<td>Keep</td>
<td>Cup</td>
</tr>
<tr>
<td>Cake</td>
<td>Usb</td>
<td>River</td>
</tr>
<tr>
<td>Chou Laks</td>
<td>Helix</td>
<td>Gold Ingat</td>
</tr>
<tr>
<td>Chocolate</td>
<td>Batch</td>
<td>CD</td>
</tr>
<tr>
<td>Pants</td>
<td>Kettel</td>
<td>Etmet</td>
</tr>
<tr>
<td>Sugar</td>
<td>Biji</td>
<td>Seed</td>
</tr>
<tr>
<td>Orange</td>
<td>Biji</td>
<td>Blank</td>
</tr>
<tr>
<td>Flute</td>
<td>Batter</td>
<td>Handkerchief</td>
</tr>
<tr>
<td>Precious Stone</td>
<td>Batter</td>
<td>Blank</td>
</tr>
</tbody>
</table>
Matching comprehension task stimuli. Children were asked to match the numeral classifier names to the respective pictures of exemplars one column at a time. Pictures on the left column are very typical exemplars of the eight numeral classifiers tested while pictures in the right column are typical exemplars of the numeral classifiers. While working on the pictures in the left column, those in the right column were covered.
6.2.3 Procedure

Children were asked to match each numeral classifier name in the middle column of each page to the appropriate exemplars in both the left and right columns of the first two pages of the experimental test stimuli and the left column of the last page of the experimental test stimuli. They were each given a pencil to do the task. While matching the numeral classifier names to the pictures in the left column, pictures in the right column were covered with a blank sheet of paper so that the children could focus on one typicality type at a time. Each of the numeral classifier names was first read out aloud to the children before they matched the numeral classifier name to the picture of the exemplar.

Since the majority of the children had already participated in the familiarisation session in Experiment 1, the same familiarisation session following the exact procedure earlier was only conducted on those children who did not participate in the elicited production task.

6.2.3.1 Practice Trial

Eight pictures were used in the practice trial prior to the experimental session to familiarise the children with the experiment. The pictures included 2 exemplars of orang [animate: human], 2 exemplars of ekor [animate: animal], 2 exemplars of bentuk [specific: ring/hook], and 2 exemplars of pasang [pair] numeral classifiers. They were placed in the position described in Table 6.3. Pictures on the right column were first covered with a blank piece of paper so that children would focus on those in the left column. One numeral classifier name was first read and shown to the children, and the children were then asked to point to one picture in the left column that uses that particular numeral classifier. Children were then asked to draw
a line using a pencil from the numeral classifier name to the picture they selected to match the numeral classifier name to its respective exemplar. Subsequently, another numeral classifier name was read out aloud to the children, and children were again asked to draw a line from the numeral classifier name to the picture they selected. The same steps were carried out with the pictures in the right column when all the numeral classifier names in the middle column were read out. Children’s understanding of the task was determined when they gave a 100% correct responses in the Practice Trial session. Only when the children had fully understood the task did the experimental session take place.

Table 6.3. Description of Practice Trial Items for Experiment 2

<table>
<thead>
<tr>
<th>Picture</th>
<th>Numeral classifier name</th>
<th>Picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>glasses</td>
<td>orang</td>
<td>fishing hooks</td>
</tr>
<tr>
<td>rings</td>
<td>ekor</td>
<td>women</td>
</tr>
<tr>
<td>men</td>
<td>pasang</td>
<td>fish</td>
</tr>
<tr>
<td>birds</td>
<td>bentuk</td>
<td>wings</td>
</tr>
</tbody>
</table>

6.2.3.2 The Experiment

The procedure was similar to the one described in the practice trial session, except that in the experimental session, each column had eight pictures – to accommodate all the eight shape-based numeral classifiers tested (Appendix 9). Each session lasted for approximately 20 minutes for the younger children and 10 minutes for the older children and adults. All the children received a merit sticker and a mechanical pencil at the end of the experiment.
6.2.4 Analysis of Results

The children’s correct responses were recorded. They were given 1 point for each correct match between the picture and the numeral classifier name. Thus, children would get a maximum of 8 points for every typicality type if they matched all the numeral classifier names to the correct exemplars from one typicality type (e.g., very typical exemplars column). For each of the numeral classifiers tested, children got a maximum of 5 points per numeral classifier if they matched the tested numeral classifier name to all the exemplars (very typical, typical, moderate, atypical, and very atypical) correctly. The maximum score a child could get in this experiment was 40 points.

6.3 Results

6.3.1 Across Numeral Classifier Category Development

As expected, the percentage of correct responses increased with age; 6-year-olds, 22.1%; 7-year-olds, 37.9%; 8-year-olds, 61.3%; and 9-year-olds, 76.4%. Despite showing the highest percentage of correct responses amongst the children, the responses given by the 9-year-olds were markedly different from the proportion of correct responses given by the adults (96.13%).
In order to compare the number of correct responses for the different numeral classifiers in the different age groups, an 8 (numeral classifier) X 5 (typicality type) X 4 (age group) X 2 (gender) mixed repeated measures ANOVA, with numeral classifier and typicality type as within-subjects factors, and age group (6-year-olds, 7-year-olds, 8-year-olds, 9-year-olds) and gender as between-subjects factors was conducted. Mauchly's test indicated that the assumption of sphericity for numeral classifier, typicality, and numeral classifier X typicality had been violated ($\chi^2(27) = 47.37, p < .01; \chi^2(9) = 26.01, p < .01; \chi^2(405) = 573.16, p < .001$ respectively). As a result, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .91; \varepsilon = .90; \varepsilon = .76$).

There was a significant main effect of numeral classifier, $F(6.36, 845.52) = 45.19, p < .001$, partial $\eta^2 = .254$; a significant effect of typicality type, $F(3.61, 480.07) = 47.09, p < .001$, partial $\eta^2 = .261$; and a significant main effect of age group, $F(3, 133)=86.13, p < .001$, partial $\eta^2 = .660$. Gender, however, was not
significant \((p = .75)\). Tukey’s post hoc analysis at \(\alpha = .05\) revealed that correct responses of the different age groups were all significantly different from each other (i.e., correct responses by the 6-year-olds were significantly lower than the 7-year-olds which in turn were significantly lower than the 8-year-olds. Correct responses by the 8-year-olds were significantly lower than the 9-year-olds). The mean numbers of correct scores for the eight shape-based numeral classifiers across age groups are listed in Table 6.4. Pairwise comparisons based on estimated marginal means (LSD = .05) revealed correct responses between helai and batang were not significantly different from each other, keping, utas, and ketul were not significantly different, ketul and buah were not significantly different, and biji and butir were not significantly different from each other. On the basis of these results, the stage of Malay shape-based numeral classifier comprehension order is as illustrated in Table 6.3. The order of shape-based numeral classifier comprehension shows that the 3D numeral classifiers were comprehended at an older age than the comprehension of the 2D and 1D numeral classifiers, although there is qualified support as ketul was not significantly different from keping and utas.

Table 6.4. Stage of Malay Shape-based Numeral Classifier Acquisition Based on the Mean Number of Correct Matching Comprehension Responses across All Age Groups

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Mean number of correct responses</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>helai [2D: rigid]</td>
<td>.66</td>
<td>.32</td>
</tr>
<tr>
<td>batang [1D: +rigid]</td>
<td>.65</td>
<td>.23</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>.51</td>
<td>.26</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>.51</td>
<td>.36</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>.47</td>
<td>.28</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>.47</td>
<td>.28</td>
</tr>
<tr>
<td>buah [3D: big]</td>
<td>.44</td>
<td>.35</td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>.37</td>
<td>.30</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>.36</td>
<td>.34</td>
</tr>
</tbody>
</table>
Matching Comprehension Task

In addition, there was a significant interaction effect between numeral classifier and age group \( F(21, 931) = 6.64, p < .001, \) partial \( \eta^2 = .130 \) (Figure 6.2). If we examine individual classifiers, there was a significant main effect of age group for batang \([1D: +rigid]\) \( F(3, 137) = 14.01, p < .001, \) partial \( \eta^2 = .235 \), for utas \([1D: -rigid]\) \( F(3, 137) = 48.65, p < .001, \) partial \( \eta^2 = .516 \), for keping \([2D: +rigid]\) \( F(3, 137) = 21.48, p < .001, \) partial \( \eta^2 = .320 \), for helai \([2D: -rigid]\) \( F(3, 137) = 77.57, p < .001, \) partial \( \eta^2 = .629 \), for buah \([3D: +big]\) \( F(3, 137) = 68.08, p < .001, \) partial \( \eta^2 = .599 \) for ketul \([3D: medium]\) \( F(3, 137) = 18.26, p < .001, \) partial \( \eta^2 = .286 \), for biji \([3D: small]\) \( F(3, 137) = 42.53, p < .001, \) partial \( \eta^2 = .482 \), and for butir \([3D: fine]\) \( F(3, 137) = 48.40, p < .001, \) partial \( \eta^2 = .515 \). Tukey’s post hoc analysis at \( \alpha = .05 \) revealed that the mean number of correct numeral classifier responses between the 6- and the 7-year-olds were significantly different for utas \([1D: -rigid]\), helai \([2D: -rigid]\), buah \([3D: +big]\), and ketul \([3D: medium]\), while the mean number of correct numeral classifier responses between the 8- and the 9-year-olds were significantly different for utas \([1D: -rigid]\), buah \([3D: big]\), and butir \([3D: fine]\). Tukey’s post hoc analysis also shows that except for batang \([1D: +rigid]\), utas \([1D –rigid]\), and ketul \([3D: medium]\), correct numeral classifier comprehension between the 7- and 8-year-olds was significantly different for all the other numeral classifiers.

Incorrect Responses

Analysis of children’s incorrect matching responses showed that children had the tendency to match a numeral classifier exemplar to a numeral classifier sharing one feature with the correct numeral classifier. As shown in Table 6.4, 33.04% of incorrect matching responses given for helai \([2D: -rigid]\) were the other 2D numeral
classifier (i.e., keping [2D: +rigid]), while 6.96% were utas [1D: -rigid]. Helai shares a similarity with keping in its dimensionality and with utas in its rigidity.

Table 6.5. Percentage of Alternative Responses Based on the Total Number of Incorrect Matching Responses

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Dimensionality</th>
<th>Rigidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>batang [1D: +rigid]</td>
<td>13.84%</td>
<td>8.48%</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>14.99%</td>
<td>13.83%</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>10.32%</td>
<td>6.49%</td>
</tr>
<tr>
<td>helai [2D: -rigid]</td>
<td>33.04%</td>
<td>6.96%</td>
</tr>
<tr>
<td>buah [3D: big]</td>
<td>42.03%</td>
<td></td>
</tr>
<tr>
<td>ketui [3D: medium]</td>
<td>42.47%</td>
<td></td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>53.48%</td>
<td></td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>61.27%</td>
<td></td>
</tr>
</tbody>
</table>

6.3.2 Within Numeral Classifier Category Development

Pairwise comparisons based on estimated marginal means (LSD = .01) indicated that in the overall results, correct responses for moderate, typical, and very typical exemplars were not significantly different from each other. Correct responses for moderate, typical, and very typical exemplars were significantly higher than atypical exemplars, which in turn were significantly higher than correct responses for very atypical exemplars. The overall pattern of exemplars was Moderate, Typical, Very Typical > Atypical > Very Atypical (Table 6.5).
Table 6.6. Mean Number of Correct Responses Based on Typicality Type

<table>
<thead>
<tr>
<th>Typicality type</th>
<th>Mean number of correct responses</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate</td>
<td>.59</td>
<td>.30</td>
</tr>
<tr>
<td>Typical</td>
<td>.56</td>
<td>.28</td>
</tr>
<tr>
<td>Very Typical</td>
<td>.56</td>
<td>.31</td>
</tr>
<tr>
<td>Atypical</td>
<td>.42</td>
<td>.25</td>
</tr>
<tr>
<td>Very Atypical</td>
<td>.34</td>
<td>.30</td>
</tr>
</tbody>
</table>

There was also a significant interaction between typicality type and age group ($F(12, 532) = 3.59, p < .001, \text{partial } \eta^2 = .075$). Pairwise comparisons (LSD = .01) for typicality type for each age group showed that correct responses for very typical exemplars were not significantly different from typical and moderate exemplars for the 7-, 8- and 9-year-olds; but for these age groups, correct responses for very typical, typical, and moderate exemplars were significantly higher than atypical and very atypical exemplars. Correct responses for atypical exemplars were significantly higher than very atypical exemplars in 6-, and 7-year-olds. Among the 8- and 9-year-olds, the difference between atypical and very atypical exemplars was not significant (Figure 6.3). In comparison, adults’ correct responses for very typical and moderate exemplars were significantly higher than very atypical exemplars ($p < .05$).
There was also a significant interaction between typicality type and numeral classifier \((F(21.29, 2831.66) = 12.59, p < .001, \text{partial } \eta^2 = .086)\). The interaction between typicality type, numeral classifier, and age group was also significant, \(F(84, 3724) = 2.85, p < .001, \text{partial } \eta^2 = .060\). The significantly higher correct responses for moderate exemplars in comparison to very atypical exemplars in the overall results were observed to be consistent in all numeral classifier (Figure 6.4). However, correct responses for very typical exemplars of *keping* (photograph), and very typical exemplars of *helai* (paper) were significantly different from the overall numeral classifier X typicality type results. This is because the mean number of correct responses for very typical exemplars of *keping* and *helai* were significantly lower than the correct responses for typical exemplars. The mean number of correct response for atypical exemplar of *helai* was significantly higher than very typical exemplar of *helai*.
To investigate if the mean number of correct responses in the Matching Comprehension Task was higher than that in the Elicited Production Task, a 2 (task type) X 8 (numeral classifier) X 4 (age group) repeated measures ANOVA with task type (comprehension, production) and numeral classifier as within-subjects factors, and age groups (6-, 7-, 8-, and 9-year-olds) as a between-subjects factor was conducted. (Gender was not analysed this time since gender has not shown any significant effect in both the production and comprehension tasks). Mauchly’s test indicated that the assumption of sphericity for numeral classifier was violated ($\chi^2(27) = 54.92, p = .001$). As a result, the degree of freedom was corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .89$). Assumptions of sphericity for task type, and task type X numeral classifier were not violated.
There was a significant effect of task type, $F(1, 130) = 464.79$, $p < .001$, partial $\eta^2 = .782$; a significant effect of numeral classifier, $F(6.23, 810.23) = 93.39$, $p < .001$, partial $\eta^2 = .418$; and a significant effect of age group $F(3, 130) = 114.82$, $p < .001$, partial $\eta^2 = .726$. Pairwise comparisons based on estimated marginal means (LSD = .05) revealed that the total mean number of correct responses for the matching comprehension task was significantly higher than the elicited production task ($p < .001$), whereas correct responses from all age groups were significantly different from each other (6-year-olds > 7-year-olds > 8-year-olds > 9-year-olds).

There was a significant interaction between task type and age group, $F(3, 130) = 4.46$, $p < .01$, partial $\eta^2 = .093$. Paired samples $t$-tests on the total number of correct responses in both comprehension and production tasks for each age group revealed that correct responses for the comprehension task was significantly higher than production task in all age groups at $p < .001$, namely 6-year-olds, $t(30) = 10.75$; 7-year-olds, $t(35) = 13.42$; 8-year-olds, $t(39) = 10.06$; and 9-year-olds, $t(30) = 10.74$ (Figure 6.5). Adults’ correct comprehension responses were significantly higher than production at $p = .001$.

Figure 6.5. Mean number of correct responses for the comprehension task in comparison with the production task for each age group.
The interaction between task type and numeral classifier was also significant, $F(7, 910) = 29.59, p < .001$, partial $\eta^2 = .185$. Paired samples $t$-tests on the correct responses for individual numeral classifiers in both comprehension and production tasks revealed that correct responses for the comprehension task was significantly higher than the production task in all numeral classifiers at $p \leq .001$, except for *biji* [3D: small], where the difference was not significant (Figure 6.6).

![Figure 6.6. Mean number of correct responses for the comprehension task in comparison with the production task for the numeral classifier categories.](image)

In addition, there was also a significant interaction between numeral classifiers and age group, $F(21, 910) = 11.77, p < .001$, partial $\eta^2 = .205$, and a significant interaction between task types, numeral classifiers and age groups, $F(21, 910) = 6.54, p < .001$, partial $\eta^2 = .131$. 
6.4 Discussion

In the current experiment, similar to Experiment 1, the Production task, it was found that the comprehension of Malay numeral classifier development is a relatively delayed and prolonged process. A significant difference ($p<.001$) between the performance of the 9-year-olds (76.4%) and the adults (96.13%) was found, indicating the acquisition process continues to develop throughout the school years. Similar to previous cross-linguistic studies, Malay numeral classifier acquisition was also found from this comprehension task to be relatively delayed (e.g., Gandour et al., 1984; Matsumoto, 1985; Uchida & Imai, 1999).

However, there are also some other cross-linguistic studies that indicate that children’s acquisition of numeral classifiers resembles adults’ at a much younger age than in the current study. For example, Chien et al. (2003), in a guessing game, found that 7-year-olds’ comprehension of Cantonese Chinese numeral classifiers reflected adults’ comprehension. The early adult-like comprehension of numeral classifier as shown in Chien et al.’s study could also be due to the selection of the numeral classifiers tested. The test items in Chien et al’s study included the use of measure words, i.e., words that are used to quantify mass nouns. Because measure words are more observable (e.g., a bowl, a cup), the use of these classifiers would provide clues to the younger children as to which nouns would collocate with which measure words in the guessing game.

In the current study, children comprehended the 1D and 2D numeral classifiers before comprehending the 3D numeral classifiers (Table 6.3). This supports our prediction which was based on semantic complexity theory. Since there are only two forms of 1D numeral classifiers (batang and utas) and two forms of 2D...
numeral classifiers (keping and helai), it was easier for children to comprehend the 1D and 2D numeral classifiers in comparison to the 3D numeral classifiers. The 3D numeral classifiers are more complex than the 1D and the 2D numeral classifiers because the 3D numeral classifier type is manifested in four forms (buah, ketul, biji and butir). For 3D numeral classifiers, children have to comprehend the differences between numeral classifiers for big, medium, small and fine items, whereas for 1D and 2D numeral classifiers, they only have to comprehend the differences between numeral classifiers for rigid or flexible items. Because of this, children comprehended 3D shape-based numeral classifiers at a later age than they comprehended the 1D and 2D numeral classifiers. There was no significant difference in their comprehension of 1D and 2D numeral classifiers because the complexity of the two numeral classifiers is similar, i.e., based on either rigid or flexible.

In relation to children’s errors, children tended to match 3D numeral classifiers exemplars with other 3D numeral classifiers (Table 6.4). There were fewer within-the-same-dimension mismatches for 1D and 2D numeral classifiers in comparison to 3D numeral classifiers. This shows that children used the dimensionality of the object as the basis for categorisation. When children are asked to categorise objects, they are able to “break down the dimensions that are already fused” (Goldstone & Barsalou, 1998 p. 22) and make judgements based on the similarities in theose dimensions. In the case of the Malay shape-based numeral classifiers, children were able to make judgements in deciding the dimensionality of the objects by matching them to the numeral classifiers that respresents the same dimensionality as the stimuli.
Finally, in general, very typical, typical and moderate exemplars of numeral classifiers were comprehended earlier than atypical and very atypical exemplars. This shows that to a large extent, what is typical to adults is also typical to children. However, there were two exceptions from the very typical exemplars used, i.e. keping (photograph) and helai (paper) (Figure 6.4), which are both 2D numeral classifiers differing in terms of rigidity. This discrepancy between the overall results and children’s performance on these two numeral classifiers is probably due to the way children perceive paper and photographs. Some paper can be perceived as hard (+rigid), thus, children classified papers with the numeral classifier keping [2D: +rigid]. Similarly, photographs may be perceived as flexible (-rigid) due to their bendability; as a result, children classified photographs with the numeral classifier helai [2D: -rigid].
CHAPTER 7
Experiment 3: Strong-Weak Contrast Discrimination Task

7.1 Introduction

This experiment aimed to investigate the strategies that Malay children use to sort or categorise objects into numeral classifier categories. Specifically, it aimed at examining if categorising pairs of objects that differ in terms of two features (e.g., in dimensionality and rigidity, or in dimensionality and size) is an easier task than categorising objects that differ in just one feature (e.g., in dimensionality only). It has been claimed that exemplars from broader categories (e.g., one human exemplar [man] versus one furniture exemplar [chair]) are easier to categorise than exemplars from finer or narrower categories (e.g., one human exemplar [man] versus another human exemplar [boy]) (e.g., Mandler et al., 1991; Mandler & McDonough, 1993; Mandler & McDonough, 1996). In an experiment on Japanese numeral classifier comprehension, Yamamoto and Keil (2000) found that children were able to point to the correct pictures at a younger age when the contrasts among the numeral classifiers were maximised (human vs. animals vs. machines) rather than when the contrasts among the numeral classifiers were minimised (small animals vs. big animals vs. birds). When the contrasts among the classifiers were minimised, children could only do them correctly at an older age.

In addition, this experiment aimed at investigating if the degree of typicality of numeral classifier exemplars has an effect on children’s performance in sorting the
objects. Rogers and McClelland (2004) claim that typical exemplars are categorised more quickly in comparison with atypical exemplars, because of the former’s high similarity to the members of the same category and high dissimilarity to the members of contrasting categories.

The objects for categorisation in this experiment were pictures of eight Malay shape-based numeral classifier exemplars, which were paired based on the semantic features of the numeral classifier categories they are members of. The broader categories condition (henceforth the strong contrast) was achieved by pairing numeral classifier exemplars with two differences in semantic features (e.g., pairing exemplars of *batang* [1D: +rigid] with *buah* [3D: big]), since dissimilarity between the two stimuli is enhanced when more dissimilar features are added to the stimuli (Maddox, 1992). The narrow categories condition (henceforth the weak contrast) was achieved by pairing numeral classifier exemplars with only one difference in semantic features (e.g., pairing exemplars of *batang* [1D: +rigid] with *utas* [1D: -rigid]). All exemplar pairs were also paired based on their typicality type, so that a typical exemplar of, for example, *batang*, was paired with a typical exemplar of *utas*; while an atypical exemplar of *batang* was paired with an atypical exemplar of *utas*.

Based on previous literature, the following predictions were formulated:

1. (a) It was predicted that Malay children would be more likely to sort exemplars of numeral classifiers correctly if the contrast between the two exemplars is strong in comparison to weak. For example, children would give more correct responses when the contrast is between exemplars from *batang* [1D: +rigid] and *buah* [3D: big], than between *batang* [1D: +rigid] and *utas* [1D: -rigid]. (b) It was
also predicted that children’s reaction times (RTs) for correct responses would be faster in the strong contrast than the weak contrast condition.

2. (a) It was predicted that children would give more correct responses when the items to be sorted were typical than atypical exemplars of each numeral classifier category. For example, children would give more correct responses when typical exemplars of *batang* were used rather than atypical exemplars. (b) Studies (e.g. Rips et al., 1973; Rogers & McClelland, 2004; Rosch, 1973; E. E. Smith, Shoben, & Rips, 1974) showed that atypical members of a category take longer to verify than typical members. It was also predicted that children’s reaction times for categorising typical exemplars would be faster than for atypical exemplars.

3. (a) Based on (1) and (2) above, it was also predicted that children would give the most correct responses for strong contrasts of typical exemplars (henceforth strong-typical exemplars) while weak contrasts of atypical exemplars (henceforth weak-atypical exemplars) would get the least number of correct responses. The number of correct responses for strong-atypical and weak-typical exemplars would fall in between. (b) It was also predicted that children’s reaction times to the correct responses would be the fastest for strong-typical exemplars and slowest for weak-atypical exemplars. The reaction time for strong-atypical and weak-typical exemplars would fall in between the fastest and slowest reaction times.

7.2 Method

7.2.1 Participants

One hundred and forty-seven children attending a preschool and a primary school in the same national school who were native speakers of Malay and spoke
Malay as their first language participated in the study. A majority of them participated in Experiments 1 and 2. The children were categorized into four age groups: thirty-one 6-year-olds, thirty-eight 7-year-olds, forty-one 8-year-olds and thirty-four 9-year-olds. 23 adult native speakers of Malay who lived in the vicinity of the school also participated in this experiment as a comparison group. A description of the participants is given in Table 7.1. This task was administered one week after the matching comprehension task (Experiment 2) and in the preceding experiments (Experiments 1 and 2), none of the participants were informed if the responses they gave were correct or incorrect.

Table 7.1 Description of Participants

<table>
<thead>
<tr>
<th>Age group</th>
<th>Age range</th>
<th>Mean age</th>
<th>No. of participants</th>
<th>No. of males</th>
<th>No. of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td>5;8 – 6;7</td>
<td>6.18</td>
<td>31</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>6;8 – 7;6</td>
<td>7.13</td>
<td>38</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>8-year-olds</td>
<td>7;9 – 8;8</td>
<td>8.25</td>
<td>41</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>9-year-olds</td>
<td>8;11 – 9;8</td>
<td>9.28</td>
<td>34</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>Adults</td>
<td>17;3 – 77;8</td>
<td>48.07</td>
<td>23</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>167</td>
<td>64</td>
<td>103</td>
</tr>
</tbody>
</table>

The majority of the children who participated in the previous two experiments participated in this strong-weak contrast discrimination task. The same group of adults also participated in this experiment as a comparison group. In the preceding experiments none of the participants were informed if the responses they gave were correct or not. This task was administered one week after the matching comprehension task (Experiment 2).
7.2.2 Stimuli

The stimuli used in this experiment were the same pictures of exemplars as used in the eight shape-based numeral classifiers tested in the previous experiments. However, in the current task, very typical and typical exemplars were reclassified as typical exemplars while very atypical and atypical exemplars were reclassified as atypical exemplars. The pictures of moderate exemplars of the eight shape-based numeral classifiers were excluded. Each numeral classifier was tested four times; twice with typical exemplars and twice with atypical exemplars (Figure 7.1).

![Figure 7.1. Four exemplars used for each numeral classifier category: two typical exemplars, two atypical exemplars. In the strong contrast condition, typical exemplars of batang (pencil and tree) were paired with typical exemplars of buah (bus and box) respectively (top row). The atypical exemplars of batang (river and road) were each paired with atypical exemplars of buah (planet and robot) (bottom row).](image)

Pictures were used as stimuli in the experiment. The first set had pairs of pictures of numeral classifier exemplars with no similarity in dimensionality of the
objects – the strong contrast pairs (Figure 7.1, and see Table 7.2, column 1). The second set had pairs of pictures of numeral classifier exemplars with a similarity in dimensionality of the objects – the weak contrast pair. The task of the children was to select the object that would match the numeral classifier name correctly when each picture pair was presented simultaneously to them on a laptop.

Table 7.2. Description of Test Stimuli for the Strong Contrast Condition

<table>
<thead>
<tr>
<th>Numeral classifier pair</th>
<th>Typical exemplars</th>
<th>Atypical exemplars</th>
<th>Very typical exemplars</th>
<th>Atypical exemplars</th>
</tr>
</thead>
</table>

Each picture pair appeared twice (but not one after another). The position of each exemplar was counterbalanced (Figure 7.2). As a result, each numeral classifier pair contained 8 audio-visual displays of 4-picture pairs.

Figure 7.2. To counterbalance the pencil-bus picture pair, pencils were positioned on the left column of one picture pair while in the other picture pair, pencils were positioned on the right.
To create a weak contrast condition, exemplars of numeral classifiers sharing the same dimensionality type were paired with each other, for example, exemplars of *batang* [1D: +rigid] were paired with *utas* [1D: -rigid] (Figure 7.3, and see Table 7.2, column 1). Similarly, each numeral classifier pair contained 8 audio-visual displays of 4-picture pairs to counterbalance the presentation of the stimuli.

![Figure 7.3. In the weak contrast condition, typical exemplars of *batang* (pencil and tree) were each paired with typical exemplars of *utas* (necklace and rope) (top row) while atypical exemplars of *batang* (river and road) were each paired with atypical exemplars of *utas* (watch and chain link) (bottom row).](image)
Table 7.3. Description of Test Stimuli for the Weak Contrast Condition

<table>
<thead>
<tr>
<th>Numeral classifier pair</th>
<th>Typical exemplars</th>
<th>Atypical exemplars</th>
<th>Very atypical exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Necklaces –</td>
<td>Ropes –</td>
<td>Watches –</td>
</tr>
<tr>
<td></td>
<td>Papers –</td>
<td>Shirts –</td>
<td>Pants –</td>
</tr>
<tr>
<td></td>
<td>Stones –</td>
<td>Meat –</td>
<td>Gold Ingots –</td>
</tr>
<tr>
<td></td>
<td>Rice –</td>
<td>Stars –</td>
<td>Seeds –</td>
</tr>
</tbody>
</table>

7.2.3 Procedure

This experiment was carried out using e-Prime, a research application suite, which allows experiment generation and millisecond precision data collection. Responses (both correct and incorrect) by the participants and their reaction times were recorded by ePrime software.

Since the majority of the children and adults had already gone through the familiarisation session in Experiments 1 and 2, the same familiarisation session following the exact procedure was only conducted on those who did not participate in any one of the preceding experiments.

7.2.3.1 Practice Trial

Six pairs of audio-visual displays were presented prior to the experimental session to familiarise the children with the experiment. The audio-visual displays included 4 pairs of orang [animate: human] versus ekor [animate: animal] exemplars and 2 pairs of bentuk [specific: ring/hook] versus pasang [pairs] exemplars, with an audio prompt in the form of the numeral classifier names, namely orang, ekor,
bentuk, or pasang, at the onset of each display. Children were required to select the item that would match the numeral classifier name correctly.

Figure 7.4. Audio-visual displays for one of the 4 orang [animate: human] versus ekor [animate: animal] picture pairs used in the practice trial session. For these picture pairs, the audio prompt was either “orang” or “ekor”. The position of the red and green dots remained in this order (red on left, green on right) in all trials.

A display screen welcoming the children to the practice trial session appeared on the laptop monitor. On this screen children were instructed to press the red dot (placed on the “z”-key of the laptop keyboard) if they thought the audio prompt they heard on the onset of the picture display was the numeral classifier name that was used to count the items with a red dot underneath it. For Figure 7.4 for example, if the child thought the numeral classifier audio prompt “orang” is used to count the children in the first picture pair, the red dot on the laptop keyboard would be pressed since the picture of children on the laptop monitor had a red dot underneath it. Children were asked to press the green dot (placed on the “m”-key) if they thought the audio prompt they heard is used to count the items with a green dot beneath it\(^6\). To illustrate, for Figure 7.4, if the child thought the audio prompt “orang” is used to count chickens in the first picture pair, the green dot on the laptop keyboard would be pressed since the picture of chicken on the laptop monitor had a green dot beneath

\(^6\) All the children were asked to use their left index finger to press the red dot and their right index finger to press the green dot.
it. The instruction on the introduction screen was read out to the children and explained further to ensure they fully understood it. A feedback display page appeared on the laptop monitor after they had keyed in their response to indicate whether or not they had responded correctly (in blue font) or incorrectly (in red font). If no keys were pressed, or if the children did not respond to the prompts given, the children were reminded by the feedback display page to press either the red or the green dot on the keyboard at the onset of each audio-visual display. Children were asked to press the white dot (the “spacebar” key on the laptop) to proceed to the next item. Only when the children had achieved 100% correct responses in the practice session could they proceed to the experimental session.

7.2.3.2 The Experiment

The procedure was similar to the ones in the practice trial session, except that in the experimental session, the feedback display page was not shown to the children. Once the children pressed either the red or the green dot, the press the ‘spacebar’ key page was displayed to indicate to the children that they could now proceed to the next item\(^7\). The 32 audio-visual displays in each contrast condition appeared on the monitor in a random order unique for each child. The entire experiment lasted approximately 15 minutes for each child. Figure 7.5 illustrates the overall protocol for both the practice trial and experimental sessions. Children received a merit sticker at the end of each experiment.

---

\(^7\) Since the experiment would not proceed to the next item unless the children pressed the spacebar, the pace of the experiment was controlled by the children.
7.3 Results

7.3.1 Correct Responses

As expected, the number of correct responses for both contrast types increased with age. An 8 (numeral classifier) X 2 (contrast type) X 2 (typicality type) X 4 (age group) X 2 (gender) repeated measures ANOVA with numeral classifier, contrast type (strong, weak), and typicality type (typical, atypical) as within-subjects factors, and age group (6-year-olds, 7-year-olds, 8-year-olds, 9-year-olds) and gender as between-subjects factors was conducted.

Mauchly’s Test of Sphericity indicated that the assumption of sphericity for the following within-subjects effects was violated: numeral classifier ($\chi^2(27) = 49.17, p < .01$), Numeral Classifier X Contrast Type ($\chi^2(27) = 135.37, p < .001$), Numeral
Classifiers X Typicality Type ($\chi^2(27) = 48.30, p < .01$), and Numeral Classifier X Contrast Type X Typicality Type ($\chi^2(27) = 57.33, p < .01$). As a result, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. The assumption of sphericity for the other within-subjects effects, namely contrast type, typicality type, and Contrast Type X Typicality Type, were not violated.

There was a significant main effect of numeral classifier, $F(6.37, 866.06) = 41.14, p < .001$, partial $\eta^2 = .232$, and a significant effect of contrast type, $F(1, 136) = 24.22, p < .001$, partial $\eta^2 = .151$. Correct responses for strong contrast exemplars were significantly higher than weak contrast exemplars. A significant effect of typicality type, $F(1, 136) = 144.95, p < .001$, partial $\eta^2 = .516$ was also found. Correct responses for typical exemplars were higher than atypical exemplars. There was a significant effect of age group, $F(3, 136) = 88.06, p < .001$, partial $\eta^2 = .660$. Tukey’s post hoc analysis at $\alpha = .05$ showed that correct responses by the 6-year-olds were significantly lower than the 7-year-olds, which in turn were significantly lower than the 8- and 9-year-olds, which were not significantly different.
Gender, however, was not significant, \( p = .87 \). Furthermore there was no interaction between contrast type and typicality type \( p = .08 \). There was a significant interaction between numeral classifier and contrast type, \( F(5.59, 759.58) = 21.74, p < .001 \), partial \( \eta^2 = .138 \). Paired samples \( t \)-tests revealed that the mean number of correct responses for strong contrast stimuli were significantly higher than weak contrast stimuli for \( biji \) [3D: small] \( t(143) = 10.75, p < .001 \), and \( butir \) [3D: fine] \( t(143) = 3.32, p = .001 \). There was no significant difference for the other numeral classifiers.
Figure 7.7. Mean number of correct responses for each numeral classifier based on contrast type.

There was a significant interaction between numeral classifier and typicality type, $F(6.42, 898.12) = 21.90, p < .001$, partial $\eta^2 = .135$. Paired samples $t$-tests revealed that the mean number of correct responses for typical exemplars were significantly higher than atypical exemplars for batang [1D: +rigid] $t(143) = 7.71, p < .001$, utas [1D: -rigid] $t(143) = 7.10, p < .001$, keping [2D: +rigid] $t(143) = 2.62, p = .001$, ketul [3D: medium] $t(143) = 6.22, p < .001$, and biji [3D: small] $t(143) = 12.48, p < .001$. The mean number of correct responses for the different typicality types was not significantly different for helai, buah and butir.
There was also a significant interaction between numeral classifier, contrast type and typicality type, $F(6.31, 857.80) = 7.90$, $p < .001$, partial $\eta^2 = .055$. Paired samples $t$-tests revealed that the total mean number of correct responses follows the following sequence: strong-typical $>$ weak-typical $>$ strong-atypical, weak-atypical. Further analysis of individual numeral classifiers revealed that only batang showed a similar resemblance to the overall results. No significant difference was found for helai. As can be seen in Figure 7.9, correct responses for strong-typical exemplars was not significantly different from weak-typical for utas, helai, buah, and ketul; and was not significantly different from strong-atypical exemplars for helai and butir. Correct responses for strong-typical exemplars were also not significantly different than weak-atypical for helai and buah. Weak-typical was not significantly different from strong-atypical for keping, helai, biji, and butir and neither was it significantly different from weak-atypical for keping, helai, buah, and butir. Unlike the overall
results however, correct responses for strong-atypical exemplars was significantly different from weak-atypical for buah, ketul, biji, and butir.

![Bar Chart](image.png)

Figure 7.9. Mean number of correct responses for each numeral classifier based on strong-weak contrast and typical-atypical exemplars.

### 7.3.2 Reaction Times

To investigate if children’s reaction times (RTs) for correct responses were faster in the strong contrast condition than the weak contrast condition, a univariate ANOVA was conducted on the mean RT for each correct numeral classifier response with contrast type (strong, weak) and typicality type (typical, atypical) as the fixed factors. There was a significant main effect of contrast type, $F(8, 319) = 4.80, p < .001$, partial $\eta^2 = .107$. The mean RTs for correct responses were significantly longer in the strong contrast condition in comparison to the weak contrast condition. There was also a main effect of typicality type, $F(8, 319) = 13.94, p < .001$, partial $\eta^2 = .259$. The mean RTs for typical exemplars were significantly shorter than for atypical exemplars.
7.4 Discussion

Research indicates that children’s ability to categorise objects proceeds through a differentiation of broader categories to much finer distinctions (Mandler et al., 1991, p. 264; Mandler & McDonough, 1993). In the current experiment, it was found that children categorised strong contrast pairs of numeral classifier exemplars with two differences in features (e.g., exemplars from *biji* [3D: small] and *utas* [1D: rigid]) more correctly than weak contrast pairs, those with only one difference in features (e.g., exemplars from *biji* [3D: small] and *butir* [3D: fine]). The results support the prediction that a greater number of contrasts in the dimensionality and rigidity, or the dimensionality and size of the stimuli facilitates categorisation by children. Hence, exemplars from broader categories are easier to categorise than exemplar from finer or narrower categories.

Interestingly, the significant response between the strong contrast and the weak contrast was only evident for *biji* [3D: small] and *butir* [3D: fine]. For the other classifiers *helai*, *batang*, *keping*, *utas*, *ketul*, and *buah*, children’s correct responses were not significantly different in both strong and weak contrast conditions possibly because children had better comprehension of the numeral classifiers in comparison to *biji* and *butir*. As shown in Experiment 2, *biji* and *butir* were the last two shape-based numeral classifiers to be comprehended by the children (Table 6.4). Children’s correct responses to *biji* and *butir* in the strong contrast condition were significantly higher than in the weak contrast condition probably due to the fact that they had not fully comprehended the underlying semantics of the numeral classifiers, and this is manifested in their performance on the atypical exemplars of the two numeral classifiers. Due to the fact that the number of perceptual feature differences have been found to influence the categorisation of objects in children (Maddox,
1992; Maddox & Ashby, 1993; A.B. Markman & Maddox, 2003), for biji and butir, the correct responses for the strong contrast condition (two differences) was significantly higher than the number of correct responses for the weak contrast condition (one differences).

Intriguingly, children’s RTs for exemplars with strong contrasts were significantly longer than those with weak contrasts. The longer RTs for exemplars in the strong contrast condition in comparison to those in the weak contrast condition could be due to the relative semantic complexity of exemplar pairs. For example, when presented with the audio-prompt batang [1D: +rigid] in the strong contrast condition, both dimensionality and rigidity of the two comparison exemplars had to be evaluated. However, when presented with the audio-prompt batang [1D: +rigid] in the weak contrast condition, only rigidity of each exemplar needs to be evaluated. This is because in the weak contrast condition, the exemplars in the picture pair share similarity in one of the dimensions. An additional considerations is that according to Folstein, Van Petten, and Rose (2008, p. 477), participants in multifeatured-stimuli experiments are not likely to give their response before completing their stimulus evaluation because accuracy (more correct responses) is considered more important than speed, which can result in a longer RT.

However, typical exemplars in the strong contrast condition did not receive the highest mean number of correct response from the children and neither did atypical exemplars in weak contrast condition receive the lowest, hence, rejecting the final prediction made for this experiment (i.e. strong-typical > weak-typical > strong-atypical, weak-atypical). This suggests that contrast condition and typicality type operate independently in the categorisation process.
CHAPTER 8
Experiment 4: Select the Odd-One-Out Task

8.1 Introduction

The aim of the experiment was to investigate children’s categorisation strategies in identifying a non-member of a numeral classifier category. Previous categorisation studies (e.g., Bowerman, 1978; E. V. Clark, 1973b; Gentner, 1978; Keil, 2008; L. B. Smith, 1989) indicate that shape is particularly salient to young children and primarily helps them in the process of learning lexical categories. Landau, Smith, and Jones (1988) for example, in their experiment on abstract objects, reported that young children focused more on shape than other perceptual attributes (e.g., size, texture, and colour) when they were asked to select an object that would share the same name with the stimulus.

According to Mervis and colleagues, children tend to initially restrict category labels to only typical members (Mervis, 1987; Mervis et al., 1976; Mervis & Rosch, 1981; Rosch et al., 1976). This results in “immature” categorisation, where atypical exemplars that do not share properties relative to adult category prototypes get excluded from the category, whereas out-of-category instances that do share these properties are inappropriately included (Rogers & McClelland, 2004).

In relation to numeral classifiers, Imai et al. (1994) found that young children tend to categorise objects based on perceptual features in determining group membership. In sorting numeral classifier exemplars, younger children have been
found to depend on perceptual similarities; whereas older children are able to synthesise pieces of knowledge of numeral classifiers and form them into a cohesive whole, which is a necessary skill in order to categorise objects using a numeral classifier system (Uchida & Imai, 1999). In addition, in order to effectively sort objects into their respective numeral classifier categories, children also need to be able to exclude perceptually similar non-members of the category. Thus, this experiment focuses on the strategies that Malay children use to sort objects into numeral classifier categories.

The objects for categorisation in this experiment were pictures of typical and atypical exemplars of eight Malay shape-based numeral classifiers and pictures of objects perceived to be similar in their shape or dimensionality with the numeral classifiers tested. Children were asked to select from a set of three pictures the object that did not fit into the numeral classifier category (Figure 8.1). Children had to choose the odd one out from the three pictures: a typical member of a numeral classifier category, an atypical member of a numeral classifier category, and a perceptually similar non-member of a numeral classifier category.

Figure 8.1. Test stimuli for the batang [1D: +rigid] category. Children had to select the odd-one-out from a set of three pictures.
Based on previous studies on categorisation, it was predicted that young children would select more atypical exemplars of numeral classifiers as the “odd-one-out”. As children mature their ability to synthesise pieces of knowledge of numeral classifier and form them into a cohesive whole increases and they are expected to select perceptually similar non-members instead.

8.2 Method

8.2.1 Participants

The same group of participants in Experiment 3 participated in this experiment. This “Select the Odd-One-Out” experiment was conducted one week after the Experiment 3. In the previous three experiments none of the participants were informed if the responses they gave were correct or otherwise. Since no responses (correct or incorrect) were given by the researcher to any of the participants in all experiments, there question of practice effect did not play a role.

8.2.2 Stimuli

The stimuli used in the current study were pictures of very typical, typical and atypical exemplars of all eight shape-based numeral classifiers tested in the previous experiments. Apart from these pictures, 8 other pictures of objects which are perceptually similar in shape or dimensionality with the typical exemplars of the tested numeral classifiers were also used as stimuli. Pictures of a very typical exemplar of each numeral classifier along with audio forms of the numeral classifier names were used as stimuli. Another set of pictures, which included pictures of typical and atypical exemplars of each one of the numeral classifiers, together with perceptually similar non-members of the respective numeral classifier categories and
another audio prompt, were also used as stimuli to elicit responses (Figure 8.1, and see also Table 8.1). For example, in relation to Figure 8.1, children were first shown a picture of 3 pencils (i.e. typical exemplar of batang [1D: +rigid]). After five seconds, pictures of three roads, three watches, and three trees were simultaneously shown to the children. Children had to select the odd-one-out from this second set of pictures.

Table 8.1. Description of Test Stimuli for the “Select the Odd-One-Out” Task

<table>
<thead>
<tr>
<th>Audio prompt</th>
<th>Visual prompt (Typical)</th>
<th>Typical</th>
<th>Atypical</th>
<th>Non-member*</th>
</tr>
</thead>
<tbody>
<tr>
<td>batang</td>
<td>Pencil</td>
<td>Tree</td>
<td>Road</td>
<td>Watch</td>
</tr>
<tr>
<td>notas</td>
<td>Necklace</td>
<td>Rope</td>
<td>Chain link</td>
<td>Worm</td>
</tr>
<tr>
<td>keping</td>
<td>Photograph</td>
<td>Wooden plank</td>
<td>Cake</td>
<td>Leaf</td>
</tr>
<tr>
<td>helai</td>
<td>Paper</td>
<td>Shirt</td>
<td>Handkerchief</td>
<td>Biscuit</td>
</tr>
<tr>
<td>buah</td>
<td>Bus</td>
<td>Box</td>
<td>Robot</td>
<td>Durian</td>
</tr>
<tr>
<td>ketil</td>
<td>Stone</td>
<td>Meat</td>
<td>Chocolate</td>
<td>Seed</td>
</tr>
<tr>
<td>biji</td>
<td>Rambutan</td>
<td>Ball</td>
<td>Plate</td>
<td>Planet</td>
</tr>
<tr>
<td>butir</td>
<td>Rice</td>
<td>Star</td>
<td>Sugar</td>
<td>Orange</td>
</tr>
</tbody>
</table>

* The non-member of the numeral classifier category chosen shares a similar shape and/or dimensionality with a typical member of the respective numeral classifier category. A watch, for example, is 1D in nature, sharing similar dimensionality with batang. A worm, to illustrate further, resembles the shape and dimensionality of the typical members of utas for its long and flexible structure. Children who were able to synthesise pieces of numeral classifier knowledge and form a cohesive whole would choose the non-member as the odd-one-out.

8.2.3 Procedure

This experiment was carried out using the e-Prime research application suite which recorded both correct and incorrect responses by the participants as well as their reaction times with millisecond precision.

Since the majority of the children and adults had already experienced the familiarisation session in Experiments 1, 2, and 3, the same familiarisation session following the exact procedure was only conducted on those who did not participate in any one of the preceding experiments.
8.2.3.1 Practice Trial

A display screen welcoming the children to the practice trial session appeared on the laptop monitor. On this screen children were told that they would be presented with two types of picture displays. In the first display, a picture of a numeral classifier exemplar and an audio form of the numeral classifier name were presented to the children on the laptop monitor (Figure 8.2).

![Figure 8.2. A visual display of one exemplar of the bilah [specific: knife] numeral classifier. The picture was displayed simultaneously with an audio prompt “bilah” to indicate to the children that a knife is an exemplar of “bilah”.

In the second display, a series of three pictures with colored dots underneath each of them and another audio display were presented to the children (Figure 8.3).
Select The Odd-One-Out Task

Figure 8.3. A visual display of atypical and typical exemplars of the bilah [specific: knife] numeral classifier (i.e., axe, saw) and a perceptually similar non-member of the bilah numeral classifier category (comb). The visual display appeared simultaneously with the audio prompt “bukan bilah” (i.e., non-member of “bilah” numeral classifier). The position of the red, yellow, and green dots remained in this order (red on left, yellow in the middle, and green on right) for all trials.

Children were instructed to press the red dot (placed on the “z”-key of the laptop keyboard) if they thought the audio prompt they heard at the onset of the second picture display was referring to the object with a red dot underneath it. Similarly, they were asked to press the yellow (placed on the “b”-key of the laptop keyboard) or the green dot (placed on the “m”-key) if they thought the audio prompt was referring to the object with a yellow or a green dot beneath it. The instruction on the introduction screen was read out to the children and explained further to ensure they fully understood the task. Audio-visual displays of two other numeral classifiers were presented to the children prior to the experimental session to familiarise the children with the task. The displays included pictures of a very typical exemplar of ekor [animate: animal] and bilah [specific: knife] numeral classifier, with an audio form of the respective numeral classifier names at the onset of the very typical exemplar display (Figure 8.2).

---

8 Children were asked to use their left index finger to press the red dot and their right index finger to press the green dot. They could use either index finger to press the yellow dot.
After a five-second audio-visual display of the very typical exemplar of one numeral classifier, a “press the spacebar” page appeared on the screen. Once the spacebar was pressed, pictures of typical and atypical exemplars of the same numeral classifier along with a picture of a perceptually similar non-member of the numeral classifier category (all in one screen) were displayed. This picture display appeared simultaneously with another audio prompt with the word “bukan” (i.e., not) preceding the numeral classifier name (Figure 8.3).

A feedback display page appeared after they had keyed in their responses. This was to indicate to the children whether or not they had responded to the prompts correctly (in a blue font) or incorrectly (in a red font). If no keys were pressed within 20 seconds, or if the children did not respond to the prompts given, the children were reminded by the feedback display page to press either the red, yellow, or green dot on the keyboard at the onset of each audio-visual display. Children were asked to press the white dot (the “spacebar” key on the laptop) to proceed to the next item. Only when the children had achieved a 100% correct score in their responses did they proceed to the experimental session.

8.2.3.2 The Experiment

The procedure was similar to the practice trial session, except that in the experimental session, the feedback display page was not shown to the children. Once the children pressed either the red, the yellow, or the green dot, the press the “spacebar” key page was displayed to indicate to the children that they could now proceed to the next item. The numeral classifier stimuli appeared in a fixed random order for each child. Figure 8.4 illustrates the overall protocol for both the practice
trial and experimental sessions. All the children received a merit sticker at the end of the experiment.

Figure 8.4. Overall protocol of the “Select the Odd-One-Out” Task.

8.3 Results

An 8 (numeral classifier) X 3 (response type) X 4 (age group) X 2 (gender) mixed repeated measures ANOVA with numeral classifier and response type (perceptually similar non-member, atypical numeral classifier member, typical numeral classifier member) as within-subjects factors, and age group (6-, 7-, 8-, and 9-year-olds) and gender as between-subjects factors was conducted. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 12.99, p < .01$ for response type) and ($\chi^2(104) = 372.16, p < .001$ for numeral classifier), therefore
degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .92$ and $\varepsilon = .74$).

There was a significant main effect of response type, $F(1.83, 249.14) = 309.35, p < .001$, partial $\eta^2 = .695$. There was no main effect of numeral classifier ($p = .67$), age group ($p = .81$) or gender ($p = .11$). Pairwise comparisons based on estimated marginal means (LSD = .01) revealed that selection of the perceptually similar non-member of numeral classifier categories were significantly higher than selection of atypical exemplars, which in turn were significantly higher than selection of typical exemplars.

Interaction between response type and age group was significant, $F(5.50, 249.14) = 30.76, p < .001$, partial $\eta^2 = .404$. Paired samples $t$-tests on the response types given by each age group revealed that among the 6-year-olds, responses for the perceptually similar non-member was not significantly higher than responses for the atypical member, but was significantly higher than responses for the typical member. For all other age groups, responses for the perceptually similar non-member (the correct response) were significantly higher than for the atypical member which was in turn significantly higher than the typical member.
The mean number of correct responses between each of the children’s age groups was significantly different from each other. The mean number of correct responses between the 9-year-olds and the adults were not significantly different, but the mean number of correct responses between the 6-, 7-, 8-, and 9-year-olds were all significantly different from each other.

### 8.4 Discussion

In the current experiment, we predicted that the youngest children would be more likely to be influenced by perceptual features of the objects and hence more likely to choose the atypical member of a numeral classifier rather than the perceptually similar non-member as the odd-one-out. Results from this experiment supports this prediction, as the the youngest age group, the 6-year-olds, had a tendency to choose the atypical exemplar of the numeral classifier category as the
odd-one-out in preference to the perceptually similar non-member (the correct response). With age the children tended to correctly choose the perceptually similar non-member. This indicates that the younger children are more influenced by perceptual features, whereas with age children become progressively more aware of the rules associated with the numeral classifier categories, and presumably through experience, refine their category representations to include atypical members. The oldest children and adults still occasionally incorrectly selected the atypical member as the odd-one-out.

This experiment shows that younger children rely on perceptual features of items in categorising them. Since the “redundancy structure of the category as a whole” (Rosch & Mervis, 1975, p. 602) is often not reflected in atypical members, atypical exemplars of a category are more likely to be excluded by younger children. For example, younger children tended to select atypical exemplar of biji [3D: small] ‘plate’ as a non-member of the biji numeral classifier category and yet included ‘planet’ in the category. This could be due to children basing their decision predominantly on perceptual features. A plate, although round, can be perceived as more like a flat object than a rounded object. Since biji is a numeral classifier label for rounded (3D) objects, a flat (2D) numeral classifier, from the children’s perspective, could co-occur with the plate. Furthermore, since learning of categorisation membership is normally done through exemplars and prototypes (Barsalou et al., 1998), the use of rambutans (typical exemplar) as the visual prompt might influence children to select another 3D object (planet) as a member of biji. Children might have not fully extracted the full semantic range of rules for biji, that is, usage of biji includes dinner sets and kitchenware.
CHAPTER 9
General Discussion and Conclusion

9.1 Introduction

In the current research, the developmental patterns of both the production and comprehension of eight Malay shape-based numeral classifiers were investigated through the four experiments discussed in previous chapters. This chapter discusses the findings from these experiments. We first discuss Malay numeral classifier developmental patterns in terms of production and comprehension, and compare our findings with those from previous cross-linguistic research. Our discussion focuses on the factors that affect the developmental patterns of shape-based numeral classifiers in both production and comprehension. Subsequently, we discuss how children categorise objects into the numeral classifier categories and the factors that affect this process.

9.2 The Developmental Pattern of Malay Numeral Classifier Acquisition

9.2.1 The Order of Numeral Classifier Acquisition

The current study reveals that children’s acquisition of Malay numeral classifiers begins with the production of animate numeral classifiers followed by inanimate numeral classifiers, and is followed by the production of shape-based and specific numeral classifiers. Results from the pilot study revealed that Malay children first use animate numeral classifiers appropriately from about 4 or 5 years
General Discussion and Conclusion

Although the age they began to use the numeral classifiers is not as early as children from other numeral classifier studies (Erbaugh, 1984, 1986; Matsumoto, 1985; Ng, 1989), the order of acquisition is similar as children classify animate nouns earlier than inanimate nouns and categorise animate items earlier than inanimate items (Simons & Keil, 1995). Children produced the numeral classifier for humans (orang) prior to the numeral classifier for animals (ekor). One child from the youngest age group (<5-year-olds) produced orang, and among the 5-year-olds, three children produced orang and only one child produced ekor. This order also concurs with results found in Chinese (Erbaugh, 1984) and Japanese (Hansen & Chen, 2001), that is, the numeral classifier for humans is acquired prior to the numeral classifier for animals. In addition, the findings of this research partially supports Hansen and Chen’s (2001) Numeral Classifier Accessibility Hierarchy (NCAH) i.e., Animate human > Animate nonhuman.

However, the second half of the NCAH, i.e. Shape > Function, is not supported by this research. Results were conflicting as some specific numeral classifiers were produced earlier than some shape-based numeral classifiers. This indicates that the production order between Malay shape-based and specific numeral classifiers is not as clear cut as found in previous cross-linguistic studies (Mak, 1991; Ng, 1989). In studies by Mak (1991) and Ng (1989), shape-based numeral classifiers were acquired prior to functional numeral classifiers. The discrepancy between this study and studies conducted by Mak (1991) and Ng (1989) could be due to the differences in the age group of the subjects and the number of numeral classifiers studied. Mak (1991) studied the developmental patterns of 6 shape-based and 4 functional numeral classifiers in 4- to 8-year-olds (122 children), whereas Ng (1989) studied the developmental patterns of 7 shape-based and 3 functional numeral
classifiers in 5-to 12-year-olds (137 children). The particular exemplars selected for
the experiments can also influence the results found. In the pilot study the typicality
of items to the children was unknown. Some items may be familiar items to the
children while some may not. In addition, there may also be some numeral
classifiers that are used more frequently than others in discourse. For example,
children are more likely to encounter knives in their everyday activities than they
would encounter necklaces; therefore bilah [specific – knife] is likely to be more
accessible to them than utas [1D: -rigid]. This could explain why we were not able
to determine the acquisition order of shape-based and specific numeral classifiers in
Malay.

One may argue that Malay has no functional classifiers; therefore, it is
inappropriate for us to treat specific numeral classifiers similar as functional
classifiers. On the contrary, although there is no numeral classifiers in Malay that
can be labelled as “functional classifiers”, the Malay specific classifiers outlined in
Chapter 2 have very close resemblances to the functional classifiers in other numeral
classifier languages. To illustrate, we will first have to compare the distribution of
Malay specific numeral classifiers and the functional classifiers of one numeral
classifier language, for example, Mandarin Chinese. The Malay specific classifier
for knives, bilah, is used for all types of knives, i.e. instruments that consisting of a
sharp blade attached to a handle that are used for cutting – regardless of their sizes.
On the other hand, in Mandarin Chinese (e.g., Hu, 1993, p. 8), all land vehicles are
classified as the functional classifiers liang. Because of the one-to-one mapping (i.e.,
mono-referential nature) between the numeral classifiers and the objects (e.g., bilah
for all knives in Malay, and liang for all land vehicles in Mandarin Chinese), it is
logical to say that specific numeral classifiers of Malay is similar to the functional
classifiers of other numeral classifier languages. The production of Malay numeral classifiers appears to be a relatively delayed and prolonged process. It is an intriguing question as to why numeral classifiers acquisition is not as early as in other numeral classifier studies (Erbaugh, 1984, 1986; Matsumoto, 1985; Ng, 1989). For example, Chinese children in Erbaugh’s (1986) study were able to produce a numeral classifier at 2;6, and Thai children in Tuaycharoen’s (1984) study started to make attempts to use numeral classifiers at 2 years of age. However, the methods of data collection used in different studies vary. The studies conducted by Erbaugh (1986) and Tuaycharoen’s (1984) were naturalistic longitudinal studies; in contrast, the current study and the studies by Gandour et al. (1984), Ying et al. (1983), and Matsumoto (1985) were experimental, cross-sectional studies (Shirai, 2001). As the type of task administered affects results, it is not possible to directly compare the age of acquisition in the current and previous studies.

For example, Uchida and Imai (1999) found that Japanese children showed a rapid improvement in their performance of detecting errors in numeral classifiers at the age of 4 and 5; whereas Chinese children showed an increase in their performance at the age of 6. However, if we compare the development between Malay and Japanese, and Malay and Chinese, Malay numeral classifier development appears to be more delayed. This relatively delayed development could be due to the choice of numeral classifiers used in this experiment. Although the study on the development of Chinese and Japanese by Uchida and Imai (1999) did include atypical exemplars similar to the current experiment, the atypical exemplars used in this experiment are possibly “more common” than in the study conducted by Uchida and Imai (1999). For example, in Uchida and Imai, atypical exemplars were snakes, whales, and little bears, whereas in this experiment, items such as gold ingots,
planets, and chain links were also included. Children in Carpenter’s (1991) study produced adult-like performance (80% correct response) in an elicited production task by the age of 9. However, in Carpenter’s study, the use of the Thai general classifier ?an was also considered as a correct response, which could be the reason why children have such an elevated performance in comparison to other cross-linguistic studies.

A plausible contributing factor for the relatively delayed numeral classifier acquisition process in Malay is the usage and obligatoriness of usage of Malay numeral classifiers. Malay numeral classifiers predominantly occur when counting or in referring to an already mentioned item. Although the absence of numeral classifiers in counting process results in ungrammatical use of language, speakers can still be understood and, as evident in the caretaker-child interaction in this research, are not typically corrected. In other languages, especially in languages like Thai, Japanese, and Chinese, numeral classifier are more obligatory in usage. The lesser degree of obligatoriness of numeral classifier in Malay in comparison to other numeral classifier languages is probably a contributing factor to why Malay children produce numeral classifier later than children from other numeral classifier languages.

The mixed semantic criteria of Malay numeral classifier could also be a factor to why Malay numeral classifier acquisition is a delayed and prolonged process. The mixed semantic criteria in numeral classifiers also occur in Thai. However, the degree of obligatoriness of numeral classifiers in Thai, in comparison to Malay appears to affect the acquisition of numeral classifiers. As a result, Malay numeral classifier acquisition is relatively more delayed and prolonged in comparison to Thai numeral classifier acquisition.
9.2.2 The Strategies Used for Unknown Numeral Classifiers

If we examine the error responses children made, it was found that in general the youngest children tended to omit numeral classifiers, whereas the older children tended to substitute an alternative numeral classifier in the correct syntactic order, that is, ‘Num NumCl N’. Similar results have been found in Chinese and Thai (e.g., Erbaugh, 1984; Tuaycharoen, 1984; Wei & Lee, 2001). In these studies, it has been found that young children tend to omit numeral classifiers when counting objects. In other numeral classifier studies (e.g., Carpenter, 1991; Gandour et al., 1984; Hansen & Chen, 2001; Hu, 1993; Matsumoto, 1985; Ng, 1989; Uchida & Imai, 1999), it has also been found that children around the same age (e.g., 5 years old) tend to substitute alternative numeral classifiers in their production responses. The differences could be due to the degree of obligatoriness of numeral classifiers in the different language. For example, in Malay, numeral classifiers are mainly used in counting processes and as references in discourse. However, in languages like Chinese, Japanese, and Thai, the use of numeral classifier is even more extensive. Malay children may regard omitting numeral classifiers in counting the object as acceptable whereas speakers of other languages may regard omitting numeral classifiers as unacceptable. Because numeral classifiers are less obligatory in Malay in comparison to other numeral classifier languages, children at 5 years of age still tend to omit numeral classifiers rather than substituting them in place of unknown numeral classifiers.
9.2.3 The Usage of Alternative Numeral Classifiers

Children also tended to use an alternative numeral classifier that shares the same dimensionality type (henceforth ‘within-the-same-dimension’ numeral classifier) in place of the correct numeral classifier, both in the production and comprehension tasks. This indicates that children were having difficulty in pairing the object with the correct numeral classifier when one of the dimensions of the object is shared by different numeral classifiers, e.g., *buah* [3D: big] in place of *ketul* [3D: medium]. The use of an alternative numeral classifier that shares a semantic feature with an unknown numeral classifier is also used as a strategy in Mandarin Chinese. In two different elicited production tasks (Loke, 1991; Mak, 1991), Chinese children used a shape-based numeral classifier that shares the same dimension with the unknown numeral classifier. Children also frequently used an alternative numeral classifier that shares the same rigidity type (henceforth ‘within-the-same-rigidity’ numeral classifier). For example, in the production task, children produced *helai* [2D: -rigid] most frequently in place of *utas* [1D: -rigid], and *batang* [1D: +rigid] in place of *keping* [2D: +rigid]. In the comprehension task, children selected *utas* [1D: -rigid] for *helai* [2D: -rigid], and *batang* [1D: +rigid] for *keping* [2D: +rigid]. It appears that children were also having difficulty in distinguishing between 1D and 2D, as dimensionality is a relative concept which is dependant on the manner objects are presented. For example, a wooden plank is 2D (flat) when it is laid flat; however it can also be 1D (long), if it is viewed from a vertical perspective. These substitution errors also indicate that children were having difficulty in making finer categorisation distinctions in terms of dimensionality and rigidity. The substitutions that young children made indicates that they depend to a large extent on perceptual features in categorising objects into linguistic categories.
9.2.4 The Order of Acquisition Based on Semantic Complexity

According to semantic complexity theory, the cumulative number of semantic features in a particular semantic category is said to play a role in the development of children’s categorisation ability (E. V. Clark, 1973a; Matsumoto, 1987). Lexical categories which involve using less semantic features to categorise objects are acquired earlier than those which involve more semantic features. Based on semantic complexity theory, the current study predicted that Malay numeral classifier acquisition order would follow the order that two-form 1D and 2D numeral classifiers would be acquired earlier than those with more complex features, the four-form 3D numeral classifiers (E. V. Clark, 1973a; Matsumoto, 1987). When matching 1D or 2D objects, categorisation is based on whether objects are ‘rigid’ or ‘flexible’, whereas matching 3D objects is more complex as it involves deciding whether objects are ‘big’, ‘medium’, ‘small’, or ‘fine’.

Results from the comprehension task in general support the order predicted from semantic complexity theory. It was found that children correctly matched 2D and 1D numeral classifiers more than the 3D numeral classifiers (although ketul [3D: medium] was not significantly different from utas [1D: -rigid] and keping [2D: +rigid]). However, the production order of shape-based numeral classifiers did not follow the predictions made based on semantic complexity theory. Although the production order began with a 1D and a 2D numeral classifier, the production of the other 1D and 2D numeral classifiers only took place after the production of two 3D numeral classifiers (i.e., buah [3D: big] and biji [3D: small]). In the production task these two 3D numeral classifiers, buah and biji, were commonly substituted in place of other numeral classifiers, which indicates that children are using them as default classifiers. This disparity in acquisition order for the production and comprehension
of numeral classifiers reveals that there are different processes or mechanisms involved in the comprehension and production of numeral classifiers.

9.2.5 The Order of Comprehension in Comparison to Production

In the current study, correct responses for the comprehension task were significantly higher than that of the production task in all age groups for all shape-based numeral classifiers, except *biji*. This supports the general claim that children comprehend a particular linguistic construction before they appropriately produce it (E. V. Clark, 1973b, 2003; Conroy & Lidz, 2007; Gennari & MacDonald, 2009; Ingram, 1974). Previous numeral classifier acquisition research also shows that children in general perform better on comprehension than production tasks (Hu, 1993; Uchida & Imai, 1999).

However, the comprehension order found in the matching task does not directly correspond to the production order. Although the order the first two and the last two numeral classifiers are acquired is the same for both production and comprehension, children produced two of the four-form 3D numeral classifiers (i.e., *buah*, *biji*) earlier than they could comprehend them. In the comprehension task, children had difficulty in matching the two 3D numeral classifiers (*buah*, *biji*) with the appropriate exemplars. This illustrates how children sometimes produce linguistic items before they comprehend them. As pointed out by R. Clark (1974), the production of a linguistic item does not necessarily mean that the item has been fully acquired. The production of certain linguistic items may actually be instances of unanalysed, but well-practiced routines that children pick up from adults (cf. E. V. Clark and Hecht’s (1983, p. 326) “production without comprehension”). In the production task, *biji* and *buah* were frequently used as default numeral classifiers,
probably as a result of their availability in the children’s linguistic environment and their broader application or usage in comparison with other Malay shape-based numeral classifiers.

Previous numeral classifier acquisition studies suggest that default classifiers may be used as placeholders to satisfy syntactic, rather than pragmatic or semantic requirements (Carpenter, 1991; Yamamoto, 2005). For example, Japanese children overused and overextended Japanese general numeral classifiers -tsu and -ko without fully comprehending their semantics (Yamamoto, 2005). According to Yamamoto, the precedence of the two Japanese general numeral classifiers in production over comprehension was due to the relatively high frequency of the two general numeral classifiers in the children’s linguistic environment.

In the case of the current research, usage of buah and biji as default numeral classifiers helps reduce the semantic load involved in learning numeral classifiers (cf. Gandour et al., 1984). According to E. V. Clark (2003), the usage of a default numeral classifier is similar to children’s early overextensions of meaning of lexical terms (e.g., dog is initially used to refer to all four-legged animals). As evident in Experiments 1 and 2, children usage of buah and biji for all 3D objects helps them to reduce the semantic load. As evident in Table 5.5, children also used buah also for big objects and biji for small objects. Children’s act of overextending the usage of buah and biji will decrease when the appropriate numeral classifiers are acquired, and when children begin to restrict and refine the semantic rules for the numeral classifiers buah and biji (c.f. E. V. Clark, 2003). Once children know that they must classify for example a grain of rice with butir and not with biji, and that butir is used for all [3D: fine] objects, children will cease using biji in place of butir and show more adult-like performance in their use of numeral classifiers.
The categorisation of some objects into a particular numeral classifier category is not always clear-cut as there is a degree of flexibility or choice in how some objects are classified. Items can be paired with different classifiers depending on their shape, appearance, or presentation, as categorisation is dependent on how the item is perceived by the speaker. For example, *bintang* (a star) is classified with *butir* if it is referred to from a distance (e.g., if the speaker is looking at the stars and counting them from the earth, since the size of stars is very fine from the earth). However, if a star is being discussed in relation to other objects in the outer space, the numeral classifier *buah* is used instead. This is due to the massive size of stars in relation to other objects in the world. Children may not be aware of these exceptions. They may learn the numeral classifier name for one object and overextend it for the same objects regardless of their conditions. Thus, children who first encounter the *bintang* as a huge object in the outer space would more likely classify *bintang* with *buah*, whereas children who first encounter of the word *bintang* as an object from the perspective of person who gazes at the stars in the sky, would more likely classify *bintang* with *butir*.

In sum, the order of acquisition found in the comprehension task in general conformed to the predictions made based on semantic complexity theory. However, in relation to the order found in the production task, *buah* and *biji* in particular did not conform to the predicted pattern. In order to explain this acquisition order we also need to consider additional contributing environmental factors.

### 9.2.6 The Influence of the Children’s Linguistic Environment

Language development is influenced by the frequency and usage of lexical terms in the children’s linguistic environment (e.g., Goodman et al., 2008; Tare et al.,
The frequency of particular linguistic items occurring in the input to the child is a critical factor in language acquisition, particularly in early acquisition (e.g., Gallaway & Richards, 1994; Snow & Ferguson, 1977).

Research conducted based on the Competition Model (MacWhinney & Bates, 1989) has shown that the exact words and patterns of usage in different contexts and in different kinds of exposures influence the learning of language forms. Children are more likely to acquire a linguistic item quickly if it is reliably present in adult language input. Children, and even adults, will learn a linguistic item relatively late, if the input is rare and unreliable.

Through interactions with caretakers, children are able acquire information about the form, syntax, semantics, and pragmatic functions of particular words (E. V. Clark, 2003). Varied conversational input through a variety of linguistic and intentional contexts provides useful data for children to create early word-word mappings for non-object terms (Tare et al., 2008). However, as shown in caretaker-child interactions (Chapter 4), adults do not frequently use numeral classifiers when they have to count objects. Even if they do, adults also tended to use alternative classifiers in classifying objects.

Since Malay numeral classifier do not occur reliably in adults’ language (as indicated in the low frequency in input), and because of the fact that adults also show alternative numeral classifier usage in their language use, the kind of language input that children get from adults cannot be considered to be reliable. Thus, numeral classifiers are acquired relatively late, which is evident in both the comprehension and production of Malay shape-based numeral classifiers. In Thai, Japanese, and Chinese, because of the obligatoriness of numeral classifiers in these languages,
children often receive reliable input from adults in the form of correct numeral classifier usage, which presumably explains why numeral classifiers are produced earlier in other languages in comparison to Malay.

Since Malay numeral classifiers are acquired relatively late, development continues in later child development, that is, through primary school and adolescence. In later language development, both spoken and written forms of communication are significant sources of language stimulation as school children learn language both in informal settings (through indirect modelling and reinforcement) and formal instruction (Nippold, 1988). Children’s television programmes (Ely & Gleason, 1995; Rice, 1990), storybook reading (Fisch, 2004; Snow, Tabors, Nicholson, & Kurkland, 1995), and noninteractive language such as overheard conversations (De Houwer, 1995) can all play a role in later language development.

In the current study, evidence that frequency plays a role in the acquisition of numeral classifiers was found as both buah and biji, the most frequently occurring numeral classifiers in the Malay Children’s Corpus and the Dewan Bahasa dan Pustaka’s Corpus, were produced relatively early. Furthermore, the relatively late acquisition of utas [1D: -rigid], ketul [3D: medium], and butir [3D: fine] in both the production and comprehension tasks corresponds to the low word frequency found for these numeral classifiers in the Malay Children’s Corpus and the Dewan Bahasa dan Pustaka’s Corpus. The current study also reveals that numeral classifiers are not used frequently in everyday language and speech directed to Malay children. In the “Putar, Cari, & Kira” game, it was found that not all adults produce numeral classifiers in their production. Furthermore, caretakers rarely insisted that children use numeral classifiers in the counting game. This indicates that the usage of Malay
numeral classifiers in everyday colloquial language is fairly optional in usage (see also Salehuddin & Winskel, 2008b). Similarly, Japanese numeral classifiers are regarded as “communicatively marginal items” since non usage of the numeral classifiers does not “entail a breakdown in communication” (Yamamoto, 2005, p. 179). The semantic information contained in numeral classifiers is rather redundant and is not semantically essential for communication in striking contrast to nouns and verbs.

The more frequently used numeral classifiers are also taught earlier in primary school than the less frequently used numeral classifiers. The significant increase in usage of the numeral classifiers batang and buah between the preschool children and the Year 1 children, as well as the significant increase in usage of utas and butir between Year 2 and Year 3 children, correspond with the introduction of these numeral classifiers in Year 1 workbooks (batang and buah) and in Year 3 workbooks (utas and butir) respectively. The formal teaching of numeral classifiers is probably particularly important in relation to the acquisition of the less commonly used numeral classifiers such as ketul and butir.

Children’s linguistic exposure, which includes both informal usage and formal instruction of numeral classifiers, plays an important contributing role in the acquisition of Malay shape-based numeral classifiers. Together with the semantic complexity of individual numeral classifiers, these multiple factors affect the acquisition pattern of Malay shape-based numeral classifiers.
9.3 The Categorisation of Objects into Numeral Classifier Categories

9.3.1 The Categorisation of Objects from Broader to Finer Distinctions into Numeral Classifier Categories

Research indicates that children’s ability to categorise objects proceeds through a differentiation of broader categories to much finer distinctions (Mandler et al., 1991, p. 264; Mandler & McDonough, 1993). Mandler and colleagues propose that objects from different taxonomic (“subdivision into kinds” (Tversky, 1989, p. 54)) categories are categorised more easily than objects from the same taxonomic category; for example, chairs versus cars are more readily categorised than objects from the same taxonomic category, for example, chairs versus tables. This is because objects from the lower level of a taxonomic category have “a high degree of within-category similarity and a high degree of between-category dissimilarity” (A. B. Markman & Wisniewski, 1997, p. 55). In the process of categorisation, objects that are similar are sorted into the same category (E. V. Clark, 2003; Mervis, 1987) and concurrently, those that are dissimilar are categorised into different categories (Hampton, 1998).

In young children, perceptual features play a prominent role in assigning category membership (Gentner & Imai, 1995; Goldstone & Barsalou, 1998; Hampton, 1998; Landau, Smith, & Jones, 1998; Mandler et al., 1991; Namy & Gentner, 2002). Research indicates that young children determine the membership of a particular category mainly based on the shape of the object (Gentner, 1978; Gentner & Imai, 1995; Imai et al., 1994; Landau et al., 1988; L. B. Smith & Sera, 1992). In addition, Maddox and colleagues (e.g. Maddox, 1992; Maddox & Ashby, 1993; A. B. Markman & Maddox, 2003) found that the categorisation of objects by children is influenced by the number of perceptual feature differences; the more
differences there are between the stimuli the easier it is for children to categorise the objects. According to Mervis et al. (1994, p. 229), “[d]omains whose subordinate categories differ along several dimensions might be easier for young children to master than domains whose subordinate categories differ along very few dimensions”. A. B. Markman and Wisniewski (1997) suggest that objects at a subordinate level are more difficult to acquire than objects at the basic level because of the homogeneous features they share with objects at the same level of hierarchy. For example, children acquire “chairs” earlier than they acquire the various types of chairs, namely “sofas”, “armchairs”, “benches”, and “recliners” because “chairs” are perceptually different from “tables”, “cupboards”, and other forms of furniture. In contrast, “sofas”, “armchairs”, “benches”, and “recliners” share many similarities in their features that discriminating them is a more challenging task among younger children (Hammer et al., 2009).

Based on previous research the current study predicted that categorisation of numeral classifier exemplars would be an easier task for children if the contrasts between exemplars of the two shape-based numeral classifiers were strong (with two differences in dimensionality and rigidity, or dimensionality and size) rather than weak (one difference in either rigidity or size). Results from the Strong-Weak Contrast Discrimination Task (Experiment 3, Chapter 7) support the prediction that children categorise objects more readily when there are strong rather than weak contrasts between the exemplars of shape-based numeral classifiers. Children categorised pairs of numeral classifier exemplars with two differences in features (e.g., exemplars of biji [3D: small] and utas [1D: -rigid]) with more correct responses than those with only one difference in features (e.g., exemplars of biji [3D: small] and butir [3D: fine]).
A similar trend was found for Japanese children as young children in general differentiated category members at a higher taxonomic level whereas older children were able to proceed downwards in the hierarchical categorical system to lower taxonomic levels (Yamamoto & Keil, 2000, p. 404). For example, younger Japanese children categorised butterflies, cups and TVs, (i.e., objects from different taxonomic categories) much more readily than butterflies, horses and hens (i.e., objects from the same taxonomic categories). In comparison, older children did not show a difference in their performance.

Intriguingly, children’s RTs for exemplars with strong contrasts were significantly longer than those with weak contrasts. The longer RTs for strong contrast exemplars in comparison to weak contrast exemplars could be due to the relative semantic complexity of exemplar pairs. For example, when presented with the audio-prompt *batang* [1D: +rigid] in the strong contrast condition, both dimensionality and rigidity of the two comparison exemplars have to be evaluated. However, when presented with the audio-prompt *batang* [1D: +rigid] in the weak contrast condition, only rigidity of each exemplar needs to be evaluated. This is because in the weak contrast condition, the exemplars in the picture pair share similarity in one of the dimensions. An additional consideration is that according to Folstein, Van Petten and Rose (2008, p. 477), participants in multifeatured-stimuli experiments are not likely to give their response before completing their stimulus evaluation because accuracy (more correct responses) is considered more important than speed, which can result in a longer RT.
9.3.2 The Role of Typicality in the Categorisation of Objects into Numeral Classifier Categories

According to prototype theory, members with more attributes in common with other members of the category, and with more dissimilarities with members of contrasting categories, are graded as more prototypical or typical members (the best exemplars) of a particular category. Conversely, members on the borderline (i.e., those having less features in common with other members within the same category, especially with the most typical member) are graded as atypical members of a category (Matsumoto, 1985).

According to Mervis and colleagues (Mervis & Pani, 1980; Rosch & Mervis, 1975), the redundancy of semantic features between typical exemplars and members of the same numeral classifier category facilitates categorisation. This is because the more the stored and the succeeding objects overlap with each other, the easier and faster the objects are categorised (Mervis & Pani, 1980; Rips et al., 1973; Rogers & McClelland, 2004; Rosch, 1973). This corresponds to the proposition that the learning and representation of categorisation is normally done through exemplars and prototypes (Barsalou et al., 1998). Since differences in semantic features also helps categorisation, the disparity between typical exemplars and members of different categories also facilitates categorisation especially among children (Mervis & Pani, 1980; Rips et al., 1973; Rogers & McClelland, 2004; Rosch, 1973).

In the current study, children categorised objects more accurately when instances given were typical exemplars of the particular category. Children sorted typical exemplars more correctly and with faster responses in the Strong-Weak Contrast Discrimination Task. This concurs with previous research on numeral classifier acquisition (e.g. Carpenter, 1991; Matsumoto, 1985; Uchida & Imai, 1996,
1999) and previous categorisation studies (Gelman & Coley, 1990; Hampton, 1998). Objects that are least frequently encountered by children are poorly recognised (Lederman et al., 1990, p. 58), and are categorised less accurately by younger children (Matsumoto, 1985; Uchida & Imai, 1999).

In the “Select the Odd-One-Out” task, we predicted that the youngest children would be more likely to be influenced by perceptual features of the objects and hence more likely to choose the atypical member of a numeral classifier rather than the perceptually similar non-member as the odd-one-out. Results from the “Select the Odd-One-Out” task support this prediction as the youngest children (the 6-year-olds) selected the perceptually different member as the odd-one-out rather than the perceptually similar non-member. In contrast, the older children (the 7-, 8-, and 9-year-olds) predominantly selected the perceptually similar non-member as the odd-one-out. This reveals that the older children have developed a greater awareness of the rules associated with category membership.

Since the “redundancy structure of the category as a whole” (Rosch & Mervis, 1975, p. 602) is often not reflected in atypical members, atypical exemplars are more likely to be excluded by younger children. For example, younger children tended to select atypical exemplar of biji [3D: small] ‘plate’ as a non-member of the biji numeral classifier category and yet included ‘planet’ in the category. In categorising numeral classifier exemplars, it appears that younger children are less able to extract the semantic rules of numeral classifiers and hence rely on perceptual similarities of exemplars per se to determine numeral classifier membership.

In relation to categorising typical exemplars more readily than atypical exemplars, there were some exceptions. For helai and butir there was no significant
difference between children’s correct responses and RTs for typical and atypical exemplars in the Strong-Weak Contrast Discrimination Task. As the numeral classifier helai was the first shape-based numeral classifier to be acquired by children both in the comprehension and production tasks, children may have fully extracted the semantic rules of helai and correctly applied the rules regardless of the degree of typicality of the exemplars. In the strong contrast condition, exemplars of butir were paired with exemplars of helai, a numeral classifier that children acquire relatively early. Thus, a correct response for helai would also produce a correct response for butir.

The atypical exemplar of biji (plate) was frequently selected as the non-member of the numeral classifier category in the Select the Odd-One-Out Task. This could be due to children basing their decision predominantly on perceptual features. A plate, although round, can be perceived as more like a flat object than a rounded object. Since biji is a numeral classifier label for rounded (3D) objects, a flat (2D) numeral classifier, from the children’s perspective, could co-occur with the plate. Furthermore, since learning of categorisation membership is normally done through exemplars and prototypes (Barsalou et al., 1998), the use of rambutans (typical exemplar) as the visual prompt might influence children to select another 3D object (planet) as a member of biji. Children might have not fully extracted the full semantic range of rules for biji, that is, usage of biji includes dinner sets and kitchenware.

9.3.3 Summary of the Categorisation of Objects into Numeral Classifier Categories

Children’s knowledge of the semantics of numeral classifier categories becomes more refined as they grow older. “Triggered by an actual exposure to such
uses in the input” (Matsumoto, 1985, p. 84), older children are able to synthesise pieces of knowledge of numeral classifiers and form them into a cohesive whole. It appears that the acquisition of numeral classifiers depends not only on how much the child is exposed to the adult’s use of the numeral classifiers, but also, how relevant the nouns used with the numeral classifiers are to the children (Hu, 1993). For example, in Malay, although *buah* is a frequently available numeral classifier, children had difficulty in matching *buah* to a ‘robot’ because a robot is not a common object in the Malay children’s environment. Children continue to modify and refine the semantic categories of numeral classifiers based on the input data. They eventually achieve an adult-like mental representation or schemata for the numeral classifier categories of their particular language (Matsumoto, 1985, p. 85).

In sum, young children are strongly influenced by the perceptual features in the categorisation of objects into numeral classifier categories. Young children tended to select the perceptually similar non-member as an exemplar from numeral classifier categories rather than the perceptually different member (Select the Odd-One-Out Task), indicating that perceptual features plays an important contributing factor in categorisation among young children. Furthermore, they exhibit difficulty in making finer categorisation distinctions when the objects/stimuli share semantic features (e.g., Strong-Weak Contrast Discrimination Task, Elicited Production Task & Matching Comprehension Task). Results also reveal that typicality of exemplars plays a significant role in categorisation as typical exemplars of numeral classifier categories were more readily categorised in comparison to atypical exemplars in the Strong-Weak Contrast Discrimination Task. In addition, in the Strong-Weak Contrast Discrimination task, it was also found that there was not an interaction
effect between contrast type and typicality, which suggests that these mechanisms or processes operate independently.

9.4 Limitations of the Research

This study has several limitations. Firstly, the exemplars chosen in these experiments were selected by the researcher. It would be more informative if the exemplars chosen were those identified by the respondents. This however would not be easy since there would be a very wide range of exemplars identified by respondents. Since the number of exemplars of shape-based numeral classifiers is unlimited (due to the poly-referential nature of shape-based numeral classifiers), leaving the selection of exemplars to the respondents would make it difficult for the researcher to select exemplars as the stimuli for this research.

Secondly, the degree of typicality of exemplars in this research is identified by adult respondents. Results could have been different if children rated the typicality of exemplars. This is because items that are typical to adults may not be typical to children, and vice versa. However, adults were recruited as their knowledge of language the system of language forms the eventual benchmark or target for the children’s developing language.

Finally, this research was conducted on speakers of standard Malay. It would be interesting to investigate if there are differences in response if speakers of Malay who speak dialects other than the standard Malay participated in this research.
9.5 Future Directions

Malay numeral classifiers are used as referring expressions as well for counting in spoken and written discourse. It would be interesting to examine when children start to use numeral classifiers as referring expressions and how usage develops. Furthermore, it would be interesting to investigate how native speakers of other languages who speak Malay as their second language acquire Malay numeral classifiers. It would also be interesting to investigate the acquisition pattern of bilingual speakers of two numeral classifier languages and how learning two different systems affects the learning process.

9.6 Conclusions

The Malay numeral classifier system is shaped by both universal cognitive-based and specific culture-based categorisation processes. This research indicates that acquisition order is affected by both animateness and semantic complexity. In addition, the Malay linguistic environment, which includes both informal usage and formal instruction of numeral classifiers, as well as typicality of the object to be classified, influences the categorisation process.

Perceptual features play a prominent role in categorisation by children proceeding from broader to finer distinctions. Substitution errors in the production and comprehension tasks reveal that children are having difficulty making finer distinctions when selecting or deciding which numeral classifier to use with a particular item. Evidence also indicates that typicality plays a prominent role in young children’s categorisation, which gives support to prototype theory, that is, the idea that categorisation is normally done through exemplars or typical members (Barsalou et al., 1998). Mental representation of a category develops based on both
prototypical exemplars and on experience with different exemplars (typical and atypical members) of that particular category. This experience assists the child in developing a fuller representation of the numeral classifier category (Dopkins & Gleason, 1997). Johansen and Palmeri (2002) suggest that mixed or hybrid representations involving rules, prototypes and/or exemplars are used to mentally represent categories of objects, which is dependent on categorisation experience of the individual. The child gradually refines their semantic representation of numeral classifier categories based on experience and input (Matsumoto, 1985). They eventually achieve an adult-like mental representation for the numeral classifier categorisation system.

The categorisation of an object involves the retrieval and evaluation of stored information on related objects and only when there is a satisfactory resemblance between new and stored information is an object accepted into the respective category (Barsalou et al., 1998). Some objects are readily categorised into a particular numeral classifier category, whereas other objects are more difficult to categorise as they can be placed into several different categories based on sometimes quite fine perceptual distinctions. As categorisation within the Malay numeral classifier system involves building up a comprehensive knowledge of complex inherent semantic characteristics of the different numeral classifier categories, it takes an extended period of time for children to acquire. Even in adults it can be seen that there is some ambiguity in which classifiers to use with a particular object. It depends on how the object is viewed or interpreted. In conclusion, this study shows that the acquisition pattern of Malay numeral classifiers is a complex and relatively prolonged process (which continues to develop into the school years) with
interaction of multiple factors, which includes semantic complexity and linguistic environment, affecting the acquisition order.
References


References


Malaysia pengeluar daging boer terbesar. (2007, 4 December). *Utusan Malaysia*.


Sains sosial corakkan pembangunan insan. (2007, 4 December). *Utusan Malaysia*.


References


List of Appendices

<table>
<thead>
<tr>
<th>Appendix 1</th>
<th>Malay Numeral Classifier List</th>
<th>230</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 2</td>
<td>Research Approval from the Malaysian Government</td>
<td>231</td>
</tr>
<tr>
<td>Appendix 3</td>
<td>Ethics Approval from University of Western Sydney</td>
<td>233</td>
</tr>
<tr>
<td>Appendix 4</td>
<td>Test Materials for the Pilot Study</td>
<td>234</td>
</tr>
<tr>
<td>Appendix 5</td>
<td>Typicality Survey Form</td>
<td>235</td>
</tr>
<tr>
<td>Appendix 6</td>
<td>Invitation to Participate in the <em>Putar, Cari, &amp; Kira</em> Game</td>
<td>239</td>
</tr>
<tr>
<td>Appendix 7</td>
<td>The <em>Putar, Cari, &amp; Kira</em> Game Picture-book</td>
<td>240</td>
</tr>
<tr>
<td>Appendix 8</td>
<td>Rubric for Caretakers</td>
<td>242</td>
</tr>
<tr>
<td>Appendix 9</td>
<td>Test Items for Matching Comprehension Task</td>
<td>245</td>
</tr>
<tr>
<td>Appendix 10</td>
<td>Published Journal Articles and Book Chapters</td>
<td>248</td>
</tr>
</tbody>
</table>
## Appendix 1 Malay Numeral Classifier List

Table 1 - A Malay Numeral Classifier List Adapted from Osman (2002) and Othman (2004)

<table>
<thead>
<tr>
<th>Malay numeral classifier</th>
<th>Usage</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>A ngkatan</td>
<td>groups of people or things</td>
<td>ship, soldiers, artists, writers</td>
</tr>
<tr>
<td>Baris</td>
<td>rows of things / people</td>
<td>people, houses, trees, sentences, tables, chairs</td>
</tr>
<tr>
<td>batang</td>
<td>long things</td>
<td>river, candle, column, pen, tree, tooth, chalk, tapioca</td>
</tr>
<tr>
<td>batas</td>
<td>Rows</td>
<td></td>
</tr>
<tr>
<td>bentuk</td>
<td>small circular things</td>
<td>fishing hook, ear ring, ring</td>
</tr>
<tr>
<td>berkhas</td>
<td>items tied together</td>
<td>keys, sticks</td>
</tr>
<tr>
<td>bidang</td>
<td>items that are spread out</td>
<td>land, cloth, carpet, sail, mat</td>
</tr>
<tr>
<td>biji</td>
<td>fruits, vegetables, small items</td>
<td>pumpkin, carrot, egg, cup, ball, balloon</td>
</tr>
<tr>
<td>bilah</td>
<td>sharp items</td>
<td>knife, scissors, scythe, machete</td>
</tr>
<tr>
<td>buah</td>
<td>big, solid items</td>
<td>car, country, house, books, box, radio</td>
</tr>
<tr>
<td>baku</td>
<td>square-shaped items</td>
<td>soap, bread</td>
</tr>
<tr>
<td>batir</td>
<td>fine, spherical items</td>
<td>eggs, pebbles, beads, rice</td>
</tr>
<tr>
<td>carik / cebis</td>
<td>torn items</td>
<td>paper, cloth</td>
</tr>
<tr>
<td>cubi</td>
<td>Pinch</td>
<td>salt, sugar, turmeric powder</td>
</tr>
<tr>
<td>cucuk</td>
<td>skewed items</td>
<td>satay</td>
</tr>
<tr>
<td>das</td>
<td>item accompanied by explosive sounds</td>
<td>shots</td>
</tr>
<tr>
<td>deret</td>
<td>items that are arranged</td>
<td>buildings, houses, shops, cars</td>
</tr>
<tr>
<td>ekor</td>
<td>all animals</td>
<td>mosquito, ant, horse, hare</td>
</tr>
<tr>
<td>gerombolan</td>
<td>a group of bad people</td>
<td>bad people, bandits, rioters</td>
</tr>
<tr>
<td>geneggam</td>
<td>items held firmly by squeezing the hand</td>
<td>rice, sand</td>
</tr>
<tr>
<td>gugus</td>
<td>items tied together</td>
<td>grapes, coconuts, keys</td>
</tr>
<tr>
<td>galung</td>
<td>rolled items</td>
<td>film, carpet, mat, wire</td>
</tr>
<tr>
<td>belai</td>
<td>thin, spread out</td>
<td>leaf, shirt, paper</td>
</tr>
<tr>
<td>ilhat</td>
<td>items tied together</td>
<td>sticks, sugar cane,</td>
</tr>
<tr>
<td>iris</td>
<td>items cut into small pieces</td>
<td>onion, cucumber, meat</td>
</tr>
<tr>
<td>jambak</td>
<td>items in bunches</td>
<td>flower, keys</td>
</tr>
<tr>
<td>kaki</td>
<td>items with &quot;legs&quot;</td>
<td>umbrella, mushrooms</td>
</tr>
<tr>
<td>kawan</td>
<td>a group of</td>
<td>animals</td>
</tr>
<tr>
<td>kelompok</td>
<td>a group of</td>
<td>people, animals, clouds</td>
</tr>
<tr>
<td>kepala</td>
<td>squeeze in the hand to form a solid entity</td>
<td>rice, sticky rice, clay</td>
</tr>
<tr>
<td>keping</td>
<td>thin items</td>
<td>wooden plank, card, picture</td>
</tr>
<tr>
<td>kerat</td>
<td>slice</td>
<td>bread</td>
</tr>
<tr>
<td>ketel</td>
<td>chunky</td>
<td>ice, meat, soap</td>
</tr>
<tr>
<td>kotak</td>
<td>box</td>
<td>matches, pencils, shoes</td>
</tr>
<tr>
<td>kumpulan</td>
<td>group</td>
<td>students, workers</td>
</tr>
<tr>
<td>kuntum</td>
<td>flower</td>
<td>rose, all flowers</td>
</tr>
<tr>
<td>laras</td>
<td>firearms</td>
<td>pistol, short gun</td>
</tr>
<tr>
<td>lembar</td>
<td>long, straight</td>
<td>paper, wire, thread</td>
</tr>
<tr>
<td>longkok</td>
<td>piled up</td>
<td>durian, clothes, soil</td>
</tr>
<tr>
<td>musnihub</td>
<td>compiled</td>
<td>newspaper, brochures, magazine, books</td>
</tr>
<tr>
<td>orang</td>
<td>people</td>
<td>teacher, nurse, doctor</td>
</tr>
<tr>
<td>papan</td>
<td>pieces</td>
<td>chocolate, sator, fire crackers</td>
</tr>
<tr>
<td>pasang</td>
<td>pairs</td>
<td>shoes, husband-wife, suit</td>
</tr>
<tr>
<td>pangga</td>
<td>parts in fruit</td>
<td>durian</td>
</tr>
<tr>
<td>pasu</td>
<td>pots</td>
<td>flower</td>
</tr>
<tr>
<td>pasukan</td>
<td>team</td>
<td>police, soldiers, scouts</td>
</tr>
<tr>
<td>patah</td>
<td>words</td>
<td>words</td>
</tr>
<tr>
<td>potong</td>
<td>sliced</td>
<td>meat, bread</td>
</tr>
<tr>
<td>pacuk</td>
<td>thin &amp; fine; firearms</td>
<td>needles, bamboo shoots, letters; gun, rifle</td>
</tr>
<tr>
<td>puntung</td>
<td>with flames</td>
<td>cigarette, firewood</td>
</tr>
<tr>
<td>ranggal</td>
<td>verse</td>
<td>poem, prose</td>
</tr>
<tr>
<td>rawan</td>
<td>fine, entangled, long</td>
<td>net</td>
</tr>
<tr>
<td>ruas</td>
<td>long with smaller parts</td>
<td>bamboo</td>
</tr>
<tr>
<td>rumpon</td>
<td>grown in groups</td>
<td>bamboo, sugar cane,</td>
</tr>
<tr>
<td>sikat / sisir</td>
<td>for bananas (in combs)</td>
<td>bananas</td>
</tr>
<tr>
<td>tandan</td>
<td>fruits in big bunches</td>
<td>bananas, coconuts</td>
</tr>
<tr>
<td>tangkai</td>
<td>in branches</td>
<td>fruits, flowers,</td>
</tr>
<tr>
<td>titik</td>
<td>in drops</td>
<td>water, ink, dew</td>
</tr>
<tr>
<td>tongkol</td>
<td>corn in cobs</td>
<td>corn</td>
</tr>
<tr>
<td>alai</td>
<td>fruits with smaller fruit flesh with seeds</td>
<td>durian, jackfruit</td>
</tr>
<tr>
<td>urat</td>
<td>fine, long</td>
<td>hair, wire</td>
</tr>
<tr>
<td>atas</td>
<td>long</td>
<td>rope, chain</td>
</tr>
</tbody>
</table>
Appendix 2  Research Approval from the Malaysian Government

UNIT PERANCANG EKONOMI
Economic Planning Unit
IABATAN PERDANA MENTERI
Prime Minister's Department
BLOK B5 & B6,
PUSAT PENTADBIRAN KERAJAAN PERSEKUTUAN
62502 PUTRAJAYA
MALAYSIA

Ruj. Tuan:
Your Ref:

Ruj. Kami:
Our Ref:

Tarikh:
Date:

10 August 2006

KHAZRIYATI SALEHUDDIN
22 Jalan 3/10 Bangi Perdana
43650 Bandar Baru Bangi
Selangor

APPLICATION TO CONDUCT RESEARCH IN MALAYSIA

With reference to your application dated 13 June 2006, I am pleased to inform you that your application to conduct research in Malaysia has been approved by the Research Promotion and Co-Ordination Committee, Economic Planning Unit, Prime Minister's Department. The details of the approval are as follows:

Researcher's name : KHAZRIYATI SALEHUDDIN
Passport No. / I. C No: 720801-04-5424
Nationality : MALAYSIA
Title of Research : "THE ACQUISITION OF MALAY CLASSIFIERS BY MALAY-SPEAKING CHILDREN"

Period of Research Approved: TWO YEARS

2. Please collect your Research Pass in person from the Economic Planning Unit, Prime Minister's Department, Parcel B, Level 4 Block B5, Federal Government Administrative Centre, 62502 Putrajaya and bring along two (2) passport size photographs. You also required to comply with the rules and regulations stipulated from time to time by the agencies with which you have dealings in the conduct of your research.

3. Kindly note that you must not involve Year Six Students who will be sitting for UPSR exam very soon.

4. I would like to draw your attention to the undertaking signed by you that you will submit without cost to the Economic Planning Unit the following documents:
Appendices

a) A brief summary of your research findings on completion of your research and before you leave Malaysia; and

b) Three (3) copies of your final dissertation/publication.

5. Lastly, please submit a copy of your preliminary and final report directly to the State Government where you carried out your research.

ATTENTION

This letter is only to inform you the status of your application and cannot be used as a research pass.

Thank you.

Yours sincerely,

Munirah ABD. MANAN
b.7. Ketua Pengarah,
Unit Perancang Ekonomi,
(Seksi Ekonomi Makro)
Email: munirah@eppu.jupiter.my
Tel: 88892809/2818/2827

C.c:

Pengarah
Bahagian Perancangan Penyelidikan & Dasar Pendidikan
Kementerian Pelajaran Malaysia
Ara 1-4, Blok E8
Kompleks Kerajaan Parcel E
Pusat Pentadbiran Kerajaan Persekutuan
62604 Putrajaya
(u.p: Dr. Amir bin Salleh @ Mohd. Saleh) (Ruj. Tuan: KP(BPPDP)603/008 (20)

Pengarah,
Pusat Pengurusan Penyelidikan,
Universiti Kebangsaan Malaysia,
43600 UKM, Bangi, Selangor.
Appendix 3  Ethics Approval from University of Western Sydney

8 September 2006

Khzariyati Salehuddin
No 2A Alison Road
Randwick NSW 2031

Dear Khazriyati

HREC 06/083 The acquisition of Malay classifiers by Malay-speaking children

The Committee has reviewed the responses to the issues raised and has now agreed to approve the above mentioned project.

You are advised that the Committee should be notified of any further change/s to the research methodology should there be any in the future. You will be required to provide a report on the ethical aspects of your project at the completion of this project. The form is located on the Research Services Ethics Web Page.

The Protocol Number HREC 06/083 should be quoted in all future correspondence about this project. Your approval will expire 28 February 2010. Please contact the Human Ethics Officer, Kay Buckley on tel: (02) 47 360 883 if you require any further information.

The Committee wishes you well with your research.

Yours sincerely

[Signature]

Associate Professor Louise O’Brien
Acting Chairperson & Deputy Chairperson
UWS Human Research Ethics Committee
Cc Dr Heather Winskel
Appendix 4  Test Materials for the Pilot Study

Trial 1: bentuk
Trial 2: pasang
Trial 3: bentuk
Trial 4: pasang

orang  ekor  batang  buah
keping  biji  utas  helai
ketul  butir  kuntum  tongkol
kaki  bilah
Appendix 5 Typicality Survey Form

Saudara/Saudari yang saya hormati

Dibimpinkan dalam soal-soal ini gambar objek-objek yang menerangkan penggunaan penjodoh bilangan Melayu dalam aktiviti-aktiviti menghitung. Sila tandai 5 jika objek yang dipaparkan, pada pendangan saudara/saudari, diukirkan sebagai contoh yang paling tipikal (biasa) bagi penjodoh bilangan yang dibarkan. Sebaliknya, tandai 1 jika objek yang dipaparkan, pada pendangan saudara/saudari, diukirkan sebagai contoh yang paling atipikal (tidak biasa). Skala yang digunakan adalah seperti berikut:

1. Contoh yang paling tipikal (biasa) bagi penjodoh bilangan yang dibarkan.
2. Contoh yang agak tipikal bagi penjodoh bilangan yang dibarkan.
3. Contoh yang menjadi tipikal dan menjadi juga tidak tipikal bagi penjodoh bilangan yang dibarkan.
4. Contoh yang atipikal (tidak biasa) bagi penjodoh bilangan yang dibarkan.
5. Contoh yang amat atipikal bagi penjodoh bilangan yang dibarkan.

Contoh:

<table>
<thead>
<tr>
<th>Penjodoh bilangan: Bilaik</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geologi</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Peralatan dinding</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Peralatan daging</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Bilangan Star Wars</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Bagi contoh yang dibarkan ini, objek "pisau" merupakan contoh yang paling tipikal bagi saya untuk penjodoh bilangan "bilaik". Ia bermakna, apabila penjodoh bilangan "bilaik" disebutkan, secara spontan imej "pisau" yang hampir sama dengan yang dipaparkan di atas akan tergambar dalam pemikiran saya. Sebaliknya, "perang Star Wars" adalah objek yang tidak mengkaitkan tergambar dalam pemikiran saya jika penjodoh bilangan "bilaik" disebutkan, walaupun saya turut merasakan penjodoh bilangan "bilaik" untuk aktiviti menghitung. (Mohon disenarkan bahwa saudara/saudari tidak memilih satu-satu nomor itu lebih dari satu kali. Bagi "bilaik" di atas sahaja, "pisau" mendapat 5, "pisau penolong daging" 4 dan bulatnya 5; walaupun saya fikirkan kedua-duanya adalah agak serupa.)

Kerjasama saudara/saudari untuk memahami maklumat percubaan yang tertera di balaman 2 turut dipinda di sampah maklum balas ke atas lapan penjodoh bilangan Melayu di balaman 3 dan 4 bagi tujukan Kajian PhD saya.

Sekian, terima kasih.

Khāzīyyah Salehuddin
Cajob PhD
MARCS Auditory Laboratory
University of Western Sydney
Jantina:

Tariik lahir:

Keturuan:

Bahasa pertuturan di rumah:

Bahasa Melayu sebagi:  (1) Bahasa Perutama sahaja  (2) Bahasa Ikut sahaja
(3) Bahasa Perutama dan juga Bahasa Ibuanda
### Penjodoh bilangan: Batang

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image of a tree]</td>
<td>[Image of a pencil]</td>
<td>[Image of a road]</td>
<td>[Image of a pen]</td>
<td>[Image of a river]</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Otak

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image of a brain]</td>
<td>[Image of a rope]</td>
<td>[Image of a necklace]</td>
<td>[Image of a watch]</td>
<td>[Image of a wire]</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Keping

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image of a pen]</td>
<td>[Image of a cake]</td>
<td>[Image of a cake]</td>
<td>[Image of a CD]</td>
<td>[Image of children]</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Helat

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image of a leaf]</td>
<td>[Image of a shirt]</td>
<td>[Image of a paper]</td>
<td>[Image of pants]</td>
<td>[Image of a drawing]</td>
</tr>
</tbody>
</table>
### Penjodoh bilangan: Buah

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Box</td>
<td>Robot</td>
<td>Bus</td>
<td>People</td>
<td>Planet</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Ketal

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Ket</td>
<td>Tongkat &amp; Brush</td>
<td>Pinecone</td>
<td>Bone</td>
<td>Daging</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Biji

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>Rose</td>
<td>Coffee</td>
<td>Ball</td>
<td>Orange</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

### Penjodoh bilangan: Batu

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>Coral</td>
<td>Bedug</td>
<td>Bone &amp; Texture</td>
<td>Bij &amp; Stone</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>
Appendix 6  Invitation to Participate in the *Putar, Cari, & Kira* Game

Ibu / Bapa / Penjaga yang dihormati

**Pelawaan melibatkan diri dalam penyelidikan**

Dengan hormat saya dirahsiakan bahwa saya, Khazriyati Salehuddin, penyelidik dari University of Western Sydney dan Universiti Kebangsaan Malaysia, ingin mengajak para ibu / bapa / penjaga anak-anak berasal 3, 4 atau 5 tahun serta anak-anak untuk melibatkan diri dalam satu penyelidikan yang melibatkan penyertaan ibu / bapa / penjaga dan anak-anak.

Penyelidikan ini bertujuan untuk membina carta ke atas salah satu aspek perkembangan kanak-kanak Melayu melalui permainan antara ibu / bapa / penjaga dan anak. Penyelidikan ini telah diperoleh kebenaran daripada Unit Perancangan Ekonomi, Jabatan Perdana Menteri (Rujukan UPE: 40200/19/1/306).


Sehubungan ini, saya mohon kepada para ibu / bapa / penjaga yang berminat untuk melibatkan diri dalam penyelidikan ini agar menghubungi saya (Khazriyati Salehuddin) di talian: 019 6209653 untuk membantu menyiapkan Pelawaan ini teknik kepada 10 ibu / bapa / penjaga yang pertama yang melaharkan minat untuk melibatkan diri dalam penyelidikan ini.

Kerjasama daripada para ibu / bapa / penjaga selanjut adalah amat saya harap dan saya dahu dengan ucapan tertima karib.

Sekian, vasalakan.

Khazriyati Salehuddin, 019 6209653
Appendix 7  The Putar, Cari, & Kira Picture-book

“AB” – Bedroom setting

“CD” – Dining room setting
“EF” – Kitchen setting

“GH” – Highway setting
Appendix 8 Protocol for Caretakers

Ibu / Bapa / Perjaga yang dinormati sekalian

Terima kasih kerana bersetuju melibatkan diri dalam penyelidikan ini.

Tuan / Puan bersama anak / anak jagaan Tuan / Puan akan bersama-sama bermain sejenis permainan interaktif yang dinamakan **Putar, Cari dan Kirai**.

Dalam permainan ini, Tuan / Puan bersama anak / anak jagaan Tuan / Puan akan bergilir-gilir

1. Memutar “roda beringkap” sambil memejamkan mata;
2. Mencari barang yang tertera di “tingkap roda”;
3. Mengira bilangan barang yang tertera di “tingkap roda” tersebut.

Saya akan memberi contoh apa yang mungkin akan berlaku dalam permainan ini.

1. Bagi halaman yang bertanda “YZ” ini, saya akan mulakan permainan dengan **memutar “roda beringkap” bertanda “y” ini sambil memejamkan mata saya. Apebila saya berhenti memutar roda beringkap ini, saya dapat gembir “burung” tertera di tingkap roda.**
2. Saya pun mencari gambar burung dalam halaman yang bertanda "YZ".

3. Saya kemudian akan mengira bilangan “burung” dalam halaman yang bertanda “YZ”. Contohnya:
   a. “Satu ekor burung, dua ekor burung, tiga ekor burung, empat ekor burung.”

4. Kemudian, saya akan memberitahu anak saya supaya membuat perkara yang sama bagi roda beringkap yang bertanda “y”. Saya akan menyarankan
   a. Menutup roda beringkap bertanda “y” sambil memejamkan matanya,
   b. (jika gambar yang tertera di roda beringkap bertanda “y” adalah gambar ayam), saya akan meminta anak saya mencari gambar “ayam” di halaman “YZ”
   c. Saya kemudian akan meminta anak saya mengira bilangan “ayam” dalam halaman “YZ” (Anak saya akan mengira “satu ekor ayam, dua ekor ayam”)

5. Saya kemudian akan memutar “roda beringkap” bertanda “z” pula dan akan ulangi proses 1 hingga 4.
Tuan / Puan akan mulakan permainan dengan halaman “AB” menggunakan “roda beringkap” “a” dan “b”, seterusnya halaman “CD” menggunakan “roda beringkap” “c” dan “d”, halaman “EF” menggunakan “roda beringkap” “e” dan “f”, dan halaman “GH” menggunakan “roda beringkap” “g” dan “h”.

Pertualan antara Tuan / Puan bersama-sama anak / anak jagaan Tuan / Puan sepanjang sesi permainan ini akan dirakam, dan mikrofon akan dipasangkan pada kolar baju anak / anak jagaan Tuan / Puan.

Di akhir sesi ini, Tuan / Puan akan diberi sahut sihat berupa wang tunai berjumlah $20.00 bersama-sama cenderamata dari Australia. Anak / anak jagaan Tuan / Puan pula akan diberikan sijil penyertaan setelah butiran lengkap mengenai dirinya diperolehi.

Sekian, terima kasih.

Yang benar

Khazriyati Salehuddin
(Penyelidik dari University of Western Sydney & Penyarah Universiti Kebangsaan Malaysia)
Appendix 9  Test Items for Matching Comprehension Task

Salam sejahtera Adik-Adik

Suaikan penjodoh bilangan di tengah-tengah kotak di bawah dengan gambar-gambar yang menggunakan penjodoh bilangan tersebut.

Contohnya,

1. Tiga orang lelaki.
2. Tiga orang wanita.
3. Tiga ekor burung.
4. Tiga ekor ikan.
5. Tiga bentuk cincin.
6. Tiga bentuk mata kail.
7. Tiga pasang cermin mata.
8. Tiga pasang sayap
<table>
<thead>
<tr>
<th>Object</th>
<th>Image</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kertas</td>
<td><img src="image1" alt="Image" /></td>
<td>BATANG</td>
</tr>
<tr>
<td>Rambutan</td>
<td><img src="image2" alt="Image" /></td>
<td>BERAPA</td>
</tr>
<tr>
<td>Rantai</td>
<td><img src="image3" alt="Image" /></td>
<td>KANAK</td>
</tr>
<tr>
<td>Pensil</td>
<td><img src="image4" alt="Image" /></td>
<td>BINTANG</td>
</tr>
<tr>
<td>Beras</td>
<td><img src="image5" alt="Image" /></td>
<td>POKOK</td>
</tr>
<tr>
<td>Batu</td>
<td><img src="image6" alt="Image" /></td>
<td>BAJU</td>
</tr>
<tr>
<td>Papan</td>
<td><img src="image7" alt="Image" /></td>
<td>TALI</td>
</tr>
<tr>
<td>Biris</td>
<td><img src="image8" alt="Image" /></td>
<td>Daging</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Biskut</td>
<td>BATANG</td>
<td></td>
</tr>
<tr>
<td>Benang</td>
<td>KEPING</td>
<td></td>
</tr>
<tr>
<td>Pala Ayam</td>
<td>UTAS</td>
<td></td>
</tr>
<tr>
<td>Penyapu</td>
<td>HELAI</td>
<td></td>
</tr>
<tr>
<td>Doman</td>
<td>BUAH</td>
<td></td>
</tr>
<tr>
<td>Keluarga</td>
<td>KETUL</td>
<td></td>
</tr>
<tr>
<td>Oren</td>
<td>BUI</td>
<td></td>
</tr>
<tr>
<td>Buta Permata</td>
<td>BUTIR</td>
<td></td>
</tr>
</tbody>
</table>

Terima kasih, Adik

Nama: __________
Kelas: ______
Tarikh lahir: ______
Jantina: ______


Penjodoh Bilangan Melayu: Lakaran Representasi Sistem Konseptualisasi Penutur Asli Melayu

Khaziriyati Salehuddin

Pendahuluan

(kerana "gula", misalnya, boleh digunakan dengan kata "cubik", "sudo", "cawan", "mangkuk", "guni", dan sebagainya).

Penjodoh bilangan juga turut digambarkan sebagai satu kategori leksikal yang tertutup yang dilampirkan kepada satu-satu kata nama dengan bilangan yang tertentu (Uchida & Imai 1999:51). Penjodoh bilangan dianggap sebagai satu kategori tertutup memandangkan tidak ada "ahli baharu" yang dimasukkan dalam senarai kata penjodoh bilangan Melayu yang seda ada.


Asmah Haji Omar & Rama Subbiah (1995: 23) menatakan bahawa penjodoh bilangan Melayu adalah tidak sama, bergantung pada bentuk kata nama yang hadir bersama-sama dengan penjodoh bilangan tersebut. Tambah mereka lagi, apabila penjodoh bilangan digunakan, kata nama yang bermakna sebagai kepal frasa dan yang hadir bersama-sama dengan penjodoh bilangan tersebut haruslah berada dalam bentuk mufrad, contohnya seperti "dua orang guru" dan bukan "dua orang guru-guru". (Perbezaan ini penting ditatakan di sini kerana tanda kata bilangan dan penjodoh bilangan, "dua orang guru" tersebut haruslah dinamakan sebagai "guru-guru").

Pada zahirnya, pemilihan penjodoh bilangan dalam mencantum pengelasa satu-satu kata nama Melayu ini kelihatan agak rambang. Walaupun penggunaannya mungkin dapat dijelaskan lagi "dua orang kanak-kanak", "dua ekor kucing", dan "dua batang pokok", kehadirannya dalam struktur "dua buah kereka", "dua petak perkebunan", dan "dua bentuk cincin" agak sukar dijelaskan dan memerlukan penjelasan yang lebih terperinci. Selain daripada penjodoh bilangan untuk marusia
(orang) dan hewan (ekor) yang dapat dikatakan agak mudah, pemerolehan penjodoh bilangan Melayu yang lain tampaknya melibatkan proses-proses penghapalan yang agak banyak untuk mengakui satu-satu penjodoh bilangan dengan kata nama-kata nama yang digunakan bersama-sama. Pada zahirnya, terlalu banyak pengecualian yang terlibat dalam pemilihan penjodoh bilangan yang perlu dipelajari, contohnya seperti biji (penjodoh bilangan bagi semua jenis biji, telur dan beberapa unit barangan rumah), dan kaki (penjodoh bilangan yang digunakan untuk payung dan beberapa jenis bumbuan).

Penjodoh Bilangan dan Kognisi
Dari perspektif psikolinguistik, penjodoh bilangan adalah penting memandangkan penggunaannya menggambarkan cara penutur satu-satu bahasa itu mengkategorikan objek-objek yang dapat dihingga (Yamamoto & Keil 2000:379). Hal ini demikian kerana penjodoh bilangan dikatakan sebagai satu unit yang dapat mengklasifikasikan kata nama berdasarkan sifatsifat kata nama yang dimiliki. Oleh yang demikian, dalam usaha mendapatkan maklumat mengenai corak-corak yang mendasari perkembangan konseptual dan semantik, pemerolehan penjodoh bilangan kini sering dijadikan bahkan pembincangan dalam bidang psikolinguistik.


Cara sistem penjodoh bilangan membahagi-bahagikan ruang semantik sebenarnya memberikan peluang yang lebih besar untuk para pengkaji psikolinguistik menjalankan kajian terhadap hubungan
Appendices


Walaupun penjodoh bilangan merupakan elemen sintaksis dan semantik yang penting, sebahagian besar daripada penjodoh bilangan Melayu tidak diperoleh secara semulajadi sehingga terpaksa dijarkan di sekolah secara formal. Analisis terhadap bahasa Melayu puturapan 12 orang kanak-kanak prasekolah yang djalankan oleh Juriah Long dalam kajian ini telah menunjukkan wujudnya hanya satu penggunaan penjodoh bilangan daripada kurus yang diperolehnya, iaitu "lelaki worang aje menang". Sehingga hari ini, proses pengkategorian penjodoh bilangan bahasa Melayu yang berlaku dalam sistem kognisi penutur aslinya masih belum pernah dikaji.

Pemerolesan penjodoh bilangan mengambil masa yang agak lama memandangkan sifat semantik penjodoh bilangan yang agak kompleks. Rembang semantik dibahagikan-bahagikan oleh sistem penjodoh bilangan secara rambang dan juga sebaliknya. Selain itu, terdapat juga beberapa keadaan yang mempersulit kriteria semantik yang bercampur-campur digunakan untuk mengkategorikan objek-objek. Percampuran kriteria semantik ini melibatkan ciricir biologi, saiz, bentuk, dan fungsi satu-satu objek (Imai & Uchida 1993). Hal ini sesuai dengan yang disarankan oleh Zainal Abidin Ahmad (1993:133) yang menjelaskan:

Berhubung dengan sifat-nama membilang ini ada terpakai beberapa perintah khas di-namakan "Pemadaan Bilangan" atau semula-nya nama di-kawarkan dengan nama masing-masing benda yang dibilang itu kut isyarat-nya.

Metodologi

(1) 1436 din, 47, dan dikurniakan lima anak – Ami, 20, Amira, 18, Azrin ... (BHCR49)

Satu lagi masalah yang mungkin timbul akibat daripada menggunakan penjodoh bilangan sebagai permulaan carian ialah penjodoh bilangan, secara individunya, mempunyai maknanya yang tersendiri jika tidak digunakan sebagai penjodoh bilangan. Hal ini demikian kerana, seperti yang dinyatakan oleh Asmah Haji Omar (1992), dari segi sintaksis, penjodoh bilangan dapat berfungsi sebagai (a) kata nama penutur, dan (b) penjodoh bilangan sahaja. Penjodoh bilangan orang yang digunakan sebagai penjodoh bilangan bagi manusia, misalnya, dapat disamakan dengan maklud manusia yang dibawa oleh kata orang itu sendiri. Sebagai contohnya, kata orang dalam (2) yang berikut bukanlah satu bentuk penjodoh bilangan tetapi adalah separah kata yang membawa maklud "manusia".

(2) 626 menindas atau merampas hak orang. Tidak mencaulkan badat ... (BHAR56)

Setelah mengambil kira kekangan-kekangan serta masalah-masalah yang mungkin dihadapi dalam kerja-kerja menganalisis data korpus, pengkaji memutuskan untuk mencari kata nama sahaja bagi menyelesaikan masalah yang mungkin menyebabkan pengkaji tidak akan berjaya melihat penggunaan kata nama tanpa sebarang penjodoh bilangan. Namun demikian, mencari kata nama yang sesuai tidak mudah memandangkan bukan semata kata nama relevan bagi kedua-dua zaman yang berbeza. Sebagai contohnya, kata "komputer" tidak
mungkin wujud dalam data korpus koleksi klasik sementara kata “huitbatang” tidak mungkin hanya digunakan dalam *Berita Harian* 2004. Oleh itu, penyelidik telah mengambil keputusan untuk mendapatkan carian bagi kata-kata nama yang relevan bagi kedua-dua zaman. Dua daripada kata nama yang dipilih adalah anak dan kebun, memandangkan kedua-duanya masih wujud bagi kedua-dua zaman.

**Dapatkan Kajian**

Analisis yang diajukan menunjukkan bahawa walaupun penjodoh bilangan merupakan satu elemen penting dalam struktur bahasa Melayu, seperti yang dijelaskan oleh Zainal Abidin Ahmad (Za’ba) dalam *Pelita Bahasa*, penjodoh bilangan telah jelas mula diingatkan oleh penulis bahasa Melayu sejak dari zaman klasik lagi.

Salah-lebih pada ka’edah Melayu jika di-kata ‘setu kerbau, dua teior, tiga chinchin, tujoh bunga’ dan, sobagainya, kerana tiada memakai Pemadan Bilangan. (1952)

Namun demikian, kekerapan peninggalan penjodoh bilangan dalam penulisah klasik adalah jauh lebih rendah jika dibandingkan dengan peninggalan penjodoh bilangan bagi penulisah akhbar dalam tahun 2004. Walau bagaimanapun, kekerapan peninggalan penjodoh bilangan yang agak tinggi dalam penulisah moden ini sebenarnya tidak bererti bahawa penjodoh bilangan sudah dilupakan atau diingatkan sama sekali dalam sinaksis bahasa Melayu moden. Sebagai contohnya, walaupun terdapat peninggalan penjodoh bilangan dalam struktur (3) di bawah ini, masih terdapat himpunan perkataan yang menggunakan penjodoh bilangan *meng* bagi dipadankan dengan kata “anak” seperti struktur (4) di bawah ini.

(3) 2572 r bellau. Pun bagiti, empat anak Ahadiat teltalu mentah dan
ma (BHR82)

(4) 8092 san cukai sehingga lima orang anak. Bagaimanapun,
lanklah itu (SHL867)

Namun demikian, bagi penjodoh bilangan untuk kata “anak” ini, daripada 578 kekerapan yang dicemarkan daripada akhbar *Berita Harian* 2004, misalnya, hanya terdapat 155 jenis nama yang mempunyai kata bilangan dan kata nama “anak” sahaja yang diselangi oleh penjodoh bilangan *meng*. Sebanyak 423 kata nama “anak” yang lain, yang didahului kata bilangan, tidak diselangi oleh apa-apa penjodoh
bilangan.

Perkara yang sama juga berlaku dengan kata “kerbau”. Walaupun ada beberapa keadaan yang memperlihatkan penjodoh bilangan ekor ditinggalkan dalam akhbar 2004 yang dianalisis seperti dalam struktur (5), ada situasi penjodoh bilangan ekor didekalkan, seperti yang ditunjukkan dalam (6) di tawau ini. Namun demikian, daripada 8 penggunaan “kerbau” yang muncul dalam Berita Harian 2004, hanya terdapat dua situasi sahaja yang memperlihatkan penjodoh bilangan ekor ditinggalkan dalam frasa nama yang dikepalai oleh kata “kerbau”.

(5) 13 juga turut menetap 25,412 kerbau, kambing (19,853)
(6) 15 baban “seekor atau dua ekor kerbau” yang membawa lumpur itu suda BHRD07 bai-bai BHC880

Seperti yang dinyatakan sebelum ini, fenomena peninggalan penjodoh bilangan bukan sahaja berlaku pada zaman ini. Analisis data korpus menunjukkan bahawa hal ini telah mula terjadi sejak dari zaman penulisan klasik lagi seperti dalam (7) yang berikut.

(7) 1046 boleh cuba Kepada itu seriatus anak dara yang baik-baik nipanya. B40001

Namun, kebanyakannya kata nama “anak” yang muncul dengan kata bilangan dalam data klasik digunakan bersama-sama dengan penjodoh bilangan orang seperti di bawah ini.

(8) 455 anak dan empat puluh orang anak celeria-celeria duduk memegan B06006

Walau bagaimanapun, gaya bahasa yang berlainan yang digunakan pada zaman penulisan klasik menunjukkan bahawa kata nama yang menjadi kepala frasa tidak semestinya berada pada akhir frasa nama. Dalam beberapa situasi, kata nama yang ingin dinyatakan bilangannya dituruti oleh kata bilangan dan penjodoh bilangannya seperti:

(9) 9034 Yentuan Raja Suain, depet anak dua orang, maka sacang serta B0403

Analisis yang dijalankan ke atas kata “anak” daripada data korpus klasik menunjukkan daripada sejumlah 571 kata “anak” yang digunakan bersama-sama dengan kata bilangan, 387 daripadanya digunakan
bersama-sama dengan penjodoh bilangan orang – sama ada digunakan sebelum kata nama tersebut seperti dalam (8) di atas, mahupun yang digunakan selepas kata nama "anak", seperti yang ditunjukkan dalam (9) di atas. 184 penggunaan "anak" yang lain menunjukkan bahawa penjodoh bilangan orang telah pun mula ditinggalkan.

Perkara yang sama juga dikesan bagi kata nama "kerbau". Terdapat peninggalan penjodoh bilangan ekor bagi "kerbau" dalam struktur (10) yang berikut:

(10) 107 in Lanang empat puluh empat Kerbau Jaya, yang di-bela ci-jamu B00094

Namun, kekerapan peninggalan penjodoh bilangan ekor daripada data klasik ini tidaklah unggul dan kebanyakannya didahului oleh penjodoh ekor jika bilangan yang hendak dinyatakan oleh penulisnya. Hal ini dapat dilihat dalam struktur (11) yang berikut:

(11) 18 kur ikan raya dan sa-ekur kerbau bernama Si-banuang, maka sam B06003

Kedaiian yang memperlihatkan kata nama disusuli dengan kata bilangan berserta dengan penjodoh bilangan seperti yang berlaku dengan kata "anak" dalam "anak dua orang" di atas urut juga berlaku dengan kata nama "kerbau" seperti dalam struktur (12) yang berikut:

(12) 416 ar tujuh ekor, Sembelehan kerbau balar tujuh ekor itu, Tana B00455

Analisis yang dijalankan terhadap kata "kerbau" daripada data korpus klasik menunjukkan daripada sejumlah 60 kata "kerbau" yang digunakan bersama-sama dengan kata bilangan, 39 daripadanya digunakan bersama-sama dengan penjodoh bilangan ekor (atau ekor) – sama ada digunakan sebelum kata nama tersebut seperti dalam (11) di atas, atau pun yang digunakan selepas kata nama "kerbau", seperti yang ditunjukkan dalam (12) di atas. 21 penggunaan "kerbau" yang lain menunjukkan bahawa penjodoh bilangan ekor telah pun mula ditinggalkan.

Analisis terhadap kedua-dua kata nama yang dipilih menunjukkan bahawa walaupun terdapat kecenderungan untuk meninggalkan penjodoh bilangan dalam penulisan bahasa Melayu, penjodoh bilangan tetap digunakan dalam kedua-dua era. Yang penting ialah
penggunaannya adalah stabili, dalam erti kata, penjodoh bilangan bagi manusia adalah orang sementara penjodoh bilangan bagi haiwan adalah ekor. Kestabilan ini tidak dapat ditukar ganti. “Kerbau” tidak sekali-kali akan dipadankan dengan orang dan begitu juga “anak”, yang sekali-kali tidak akan dipadankan dengan ekor, melainkan anak tersebut merupakan anak binatang seperti

(13) 4895 an daripadanya keluar sa-ekor anak ayam jantan, sedang
ditempuh B00035.

Aplikasi Dapatan

Seperti yang dinyatakan sebelum ini, kajian terhadap penjodoh bilangan sebenarnya dapat membayangkan cara satu-satu masyarakat itu berfikir dan sistem konseptualisasi masyarakat Melayu sebenarnya dapat dikesajikan dengan menganalisis cara penjodoh bilangan dipadankan dengan cara-cara nama tertentu. Oleh yang demikian, berpandukan dapatan korpus yang menunjukkan bahawa walaupun wujud usaha-usaha untuk meninggalkan penjodoh bilangan dalam penulisan bahasa Melayu, keadaan itu tetap tidak menjadi satu kekurangan cara pemikiran orang Melayu. Sehingga itu, lakaletans representasi sistem konseptualisasi penutur asli Melayu yang dicadangkan ini diharapkan menjadi satu titik tolak bagi mengenal pasti cara orang-orang Melayu mengkonseptualisasikan objek-objek, yang dari sudut lingoistiknya, diterjemahkan sebagai kata nama-kata nama yang stabil dan diterima pakai sejak zaman-berzaman.

Bab ini bercadang untuk menyampaikan beberapa perkara. Pertama sekali, setiap kata nama dapat dilihat oleh masyarakat Melayu sebagai sesuatu yang dapat dihitung dan tidak dapat dihiung. Bagi setiap kata nama yang tidak dapat dihiung, satu unitizing modifier digunakan bagi membolehkannya dihiung seperti “seribu garam” dan “setitik nila” (bagi menandakan sedikit) sembari “sebukul baju” dan “sekahah rasi” digunakan untuk menandakan banyak.

Bagi kata nama terhitung pula, masyarakat Melayu mengkategorikannya sebagai bernyawa dan tak bernyawa. Untuk kata nama yang bernyawa, kata ini dapat dibahagikan kepada manusia dan haiwan, yang masing-masing dipadankan dengan penjodoh bilangan orang dan ekor.
Sebaliknya, bagi kata nama yang tak bernyawa pula, secara amnya, kata nama itu dikategorikan oleh masyarakat Melayu mengikut tiga kategori, iaitu (1) saiz, (2) jenis, dan (3) bentuk. Kategori saiz dilihat oleh masyarakat Melayu dalam Lima saiz yang berbeza. Objek-objek biasanya dikategorikan sebagai kecil, kecil, nipis, panjang, dan besar, bergantung pada pengalaman mereka terhadap saiz satu-satu barang. Masyarakat Melayu turut melihat objek-objek dalam pelbagai bentuk dengan adanya objek-objek yang diperspesikan mereka sama ada sebagai lurus, bertepi, rata, bercelurat, bulat, beloid, berketal, berbendera, dan tajam. Selain itu, masyarakat Melayu juga cenderung mengkategorikan objek-objek mengikut jenis-jenis mereka, bergantung pada jenis kata nama tersebut dan kegunaannya.

Penutup
Lakaran representasi sistem konseptualisasi penjodoh bilangan yang dipaparkan di atas hanyalah merupakan satu permulaan bagi mengenal pasti cara masyarakat Melayu berfikir dan memanifestasikan persepsi mereka. Lakaran representasi sistem konseptualisasi ini sebenarnya penting bagi memulakan kajian lanjutan terhadap pemikiran masyarakat Melayu. Antara kajian yang dapat bertuik tolak daripada lakaran representasi sistem konseptualisasi penjodoh bilangan Melayu termasuklah kajian terhadap proses pencerahan penjodoh bilangan
Melayu lagi menjavab persoalan yang diutarakan oleh hipotesis Whorf. Kajian sempanamatinipenting, lebih-lebih lagi memandangkan penilaian penjodoh bilangan sebenaranya mampu menjelaskan cara masyarakat satu-satu bahasa itu berfikir dan menterjemahkan pandangan mereka terhadap apa-apa yang berlaku di sekelilingnya.

**Rujukan**


MALAY NUMERAL CLASSIFIERS:
SKETCHING CONCEPTUAL REPRESENTATION
FROM A NATIVE SPEAKER’S PERSPECTIVE

Khazriyati Salehuddin
MARC5, Auditory Laboratories
<kh.salehuddin@uws.edu.au>

Heather Winskel
University of Western Sydney
<h.winskel@uws.edu.au>

Abstract
Numerical classifiers are defined as a syntactic-semantic category denoting “some salient perceived or imputed characteristic” of a particular count noun (Allan 1977: 298). Similar to 36 other East and Southeast Asian languages, Malay has a rich numeral classifier system (Adams & Conklin 1973). Numerical classifiers in Malay syntactically exist in a noun phrase together with a numeral preceding it and an obligatory noun following it. Malay numeral classifiers are semantically significant as they provide information about the physical and functional properties of count nouns to the speakers within the Malay speech community. The semantic criteria for dividing the system of each classifier category appear to be complex and opaque, with mixed semantic criteria involved in classifying members of a given category. Though the motivation may be obvious in cases like “dua orang kanak-kanak”, “dua ekor kelinci” and “dua batu pelitik”, (literally in English ‘two human children’, ‘two tail cats’, and ‘two stem trees’ respectively), the motivation seems somewhat arbitrary in cases like “dua buah kereta”, “dua buah perangai” and “dua bentuk cincin” (literally ‘two fruit cars’, ‘two broken words’ and ‘two shape rings’ respectively). Other than the classifiers for human, animals and trees—which are rather straightforward in their usage—the acquisition of the other Malay classifiers seems to involve a lot of memorization in determining which classifier should co-occur with a particular noun; as there seem to be many exceptions or non-prototypical examples which need to be learnt, for example, “baby” (the classifier that literally means ‘seed’ for all types of fruits, eggs, some household items) and “kaki” (the classifier that literally means ‘leg’ for umbrellas and some kinds of plants) (Asmah Haji Omar & Rama Subbiah, 1995: 22). Despite sharing some universal characteristics with other classifier languages such as Thai, Japanese and Korean (Carpenter, 1991; Uehara & Inui, 1999; and Yamamoto & Keil, 2000), the subcategorization of the Malay classifiers has characteristics that are specific to Malay. This paper aims to illustrate the conceptual system of Malay classifiers from a native speaker’s perspective and discuss its classification vis-a-vis Allan’s (1977) description of other classifier languages of the world.

Numerical Classifiers
Numerical classifiers form one syntactic-semantic category that is common in Sino-Tibetan and Austronesian languages – so common that 37 of the Southeast Asian and East Asian

© Khazriyati Salehuddin & Heather Winskel

languages have them as an obligatory category in their linguistic structure. They are also described as a closed class lexical category as their membership is so exclusive that no members are not added to the existing list of numeral classifiers (Uchida & Inai 1999: 51). Syntactically, numeral classifiers form part of a noun phrase within which, this syntactic-semantic category must be present together with a numeral and a head noun. Numeral classifier examples in Thai, Japanese and Chinese are illustrated as follows.

Thai

N   Adj   Num   CL
Dok  nan  rak  soong  khon
Child lovable  two  CL
‘Two cute children’ (Simpsons, n.d.)

Japanese

Num   CL   N
Ni mai kami
Two   CL   paper
‘two sheets of paper’ (Yamamoto & Keil 2000)

Chinese

Num   CL   N
Saan1 go3  gasi1 zel
Three   CL   older sister
‘three older sisters’ (Wei & Lee 2001)

Semiantically, numeral classifiers provide information about the physical and functional properties, and to a certain extent, the social status of a noun, to the speakers within and across a particular speech community. Because of this, classifiers are actually a grammatical system that reflects the manner countable objects are counted, quantified and categorized by the particular speech community (Yamamoto & Keil 2000).

Though structurally their grammatical role seems somewhat similar to quantifiers in English such as 'a piece of' and 'a portion of', classifiers are different from the English quantifiers due to the fact that rather than 'quantizing' them, classifiers categorize nouns into categories that carry semantic information. In other words, instead of quantifying mass nouns in measuring units, classifiers qualify all count nouns based on their permanent qualities. In addition, classifiers are stable because unlike English quantifiers, substitutions of classifiers results in the ungrammatical use of language.

Malay, an Austronesian language, is said to be one of the languages that has an extensive system of numeral classifiers (Richards et al. 1985). The system of Malay numeral classifiers is very complex and has to be formally taught in schools. This is indicated by numerous publications of reference books solely focusing on Malay classifiers. The Malay numeral classifiers are also described as coefficients and numeral coefficients (Asmah Haji Omar & Rama Subbiah 1995: 23) and its absence makes the structure of any formal Malay sentence ungrammatical. Structurally, Malay numeral classifiers are somewhat similar to Japanese and Chinese, as Malay numeral classifiers are preceded by a numeral and followed by a head noun.
Appendices

Malay Classifiers

Nun CL N
Tiga orang kanok-kanok
Three CL child
‘three children’

Nun CL N
Empat ekor arwak
Four CL rabbit-pl
‘four rabbits’

As illustrated above, when numeral classifiers are used in Malay, the noun that forms the head of the noun phrase appears in a singular form. In situations where numbers are not used to indicate plurality, numeral classifiers are obligatorily dropped (Asmah Haji Omar & Rani Subiah 1993):

N -pl N -pl
Rumah-rumah Buku -buku
House -pl Book -pl
‘houses’ ‘books’

In isolated cases, instead of appearing to the left of the head noun, classifiers can be used to the right of it, following a numeral.

N Nun CL (DBF Corpus: B00403)
Anak dua orang
Child two CL
‘Two children’

N Nun CL (DBF Corpus: B00455)
Kambu belas nyuk ekor
Buffalo albino seven CL
‘Seven albino buffalos’

Despite being syntactically different from the norm, what is most significant in the two examples above is that the choice of classifiers is maintained and consistent for each type of noun, i.e. the noun ‘children’ co-occurs with the (animate, human) “orang” classifier while the noun ‘buffalos’ co-occurs with the (animate, animal) “ekor” classifier.

Imai and Gentner described numeral classifiers as similar to modifiers that utilize uncountable objects for quantifying purposes (i.e. unizing modifiers) such as “a spoonful of sugar”, “a bowl of rice” and “a cup of coffee” (1997: 174). This however does not truly describe Malay classifiers as while unizing modifiers are used to quantify mass nouns, numeral classifiers are used to quantify all count nouns. Imai and Gentner’s unizing modifiers are indeed used in Malay expression such as “sesuatu gula”, “semangkuk nasi” and “secup kopi”, each means “a spoon of sugar”, “a bowl of rice” and “a cup of coffee” respectively. Unizing modifiers, however, should not be equated with numeral classifiers also because as illustrated above, unlike unizing modifiers, numeral classifiers are stable. To illustrate, the numeral classifier “orang”, may only be used for human beings and
substitutions with any other classifier will result in ungrammatical use of language. On the other hand, utilizing modifiers are inflexible in nature. To illustrate, ‘sugar’ can collocate with a variety of expressions like ‘a packet of’, ‘a spoon of’, ‘a cup of’, ‘a bowl of’, ‘a sack of’ etc.

At a superficial level, the selection of numeral classifiers in classifying Malay nouns seems to be arbitrary. Although its usage in “dua orang kanak-kanak”, “dua ekor lucung” and “dua batang pokok” for ‘two CL children’, ‘two CL cats’ and ‘two CL trees’ seems rather straightforward, its usage in “dua buah kereta”, “dua biji cawan” and “dua kaki payung” for ‘two CL cars’, ‘two CL cups’ and ‘two CL umbrella’ is rather difficult to explain.

Apart from the numeral classifiers for human (“orang” – literally ‘human’), animal (“ekor” – literally ‘tail’) and trees (“batang” – literally ‘stem’) that seem to be relatively semantically transparent, the acquisition of other Malay classifiers appears to involve a more cognitively demanding task. There appear to be many exceptions to the rule in the selection of some classifiers, e.g. the classifiers “buah” that literally means ‘fruit’ for cars, houses, concepts but not fruit, the classifier “biji” that literally means ‘seed’ for cups, stones but not seeds, and the classifier “kaki” that literally means ‘leg’ for umbrellas, and some plants, but not legs. This kind of arbitrariness in the majority of Malay numeral classifiers means that Malay classifiers are often regarded as loose and opaque which presumably results in greater difficulty on the part of the learner in acquiring them.

**Numeral Classifiers and their Cognition**

The study on numeral classifiers is significant in psycholinguistics as its usage reflects the manner a particular speech community categorizes countable objects (Yamamoto & Keil 2000).

Numeral classifiers are said to be a unit that classifies nouns based on the characteristics of the nouns in question. Generally, numeral classifier system shares some universal features among different classifier languages such as the distinction between animate and inanimate objects.

However, the classifier system also exhibits the manner an individual language categorizes objects which is specific to that particular language. The way classifiers divide the semantic space gives psycholinguistic researchers the opportunity to carry out research on the relationship between language and cognition, such as the Whorfian hypothesis and whether linguistic categories affect the way humans think or perceive of the world, or in turn to what extent does the way that humans think or perceive influence the linguistic categories formed in languages of the world. A less deterministic version of the Whorfian hypothesis argues that language may affect classification during the process of thinking-for-speaking (Slobin 1996, 2003). Children focus their attention on the language-specific characteristics of their language from a very early age. These aspects may affect online thinking-for-speaking or may have longer term consequences on non-linguistic cognition.

Though classifiers are an important syntactic-semantic element of Malay linguistics, the process involved in the categorization of Malay classifiers that takes place in the cognitive system of its native speakers has never been studied.

The collocation between Malay numeral classifiers with particular nouns appears to occur in both a systematic and fairly arbitrarily manner. Furthermore, there are cases where
mixed semantic criteria are used in object categorization. These mixed semantic criteria involve biological features, size, and form of objects.

This paper aims to illustrate the conceptual system of Malay numeral classifiers from a native speaker's perspective and discuss its classification vis-à-vis Allan's (1977) description of other classifier languages of the world.

Similar to other languages of the world, Malay nouns are subcategorized into "count nouns" and "mass nouns".

The Malay mass nouns do not co-occur with any classifier; instead, they co-occur with utilizing modifiers based on the quantity of the mass nouns that is needed to modify the noun. To illustrate, mass nouns occurring in a small quantity can be modified by quantifiers "cube" (a pinch of) and "trik" (a drop of), while those occurring in a large quantity may be modified by other quantifiers such as "gumi" (a sack of) and "baldi" (a pair of).

Malay count nouns can further be categorized by the Malay numeral classifier system into two semantic categories on the basis of animacy, i.e., animate versus inanimate. The animate nouns are then subcategorized into two subcategories namely human and animal which are qualified by the classifiers "orang" and "clor" respectively.

Inanimate objects on the other hand are subsequently subcategorized into two subcategories i.e., shape and the specific categories. The shape category is then further subcategorized into rigidity, dimensionality and size.
Further analysis on the organization of Malay classifiers shows that to a large extent Malay classifiers can be described in a similar manner as the shape-based classifiers described by Allan (1977). This is because, like most classifier languages that Allan described, shape-based classifiers of Malay can also be subcategorized based on the dimensionality of the objects, i.e. one-dimensional, two-dimensional and three-dimensional classifiers.

The division of the shape-based classifier category however, is not clear-cut as its partitioning is not mono-referential but poly-referential in nature. This subcategory of classifier involves a combination of different subcategories, i.e. with mixed semantic criteria used to qualify a particular noun due to the dependency of secondary parameters on primary parameters.

Each of the one- and two-dimensional categories, for example, is not independent as each of them has to be combined with a different subcategory, i.e. 'rigidity' - depending on whether the object is rigid or flexible.

Thus, one-dimensional nouns that are rigid like a pen, a ruler, a tree and a case, co-occur with the classifier ‘batang’ while one-dimensional nouns that are flexible like a chain, a strand of thread, and a bracelet co-occur with the classifier ‘unas’. Two-dimensional nouns, on the other hand, if they are rigid like a plank of wood, a slice of cake, and a photograph, co-occur with the classifier ‘keping’ while those that are flexible like a piece of cloth, a shirt, a piece of paper and a pair of trousers, co-occur with the classifier ‘helai’.

Similar to the description of the shape-based classifiers for the majority of classifier languages, Malay’s three-dimensional subcategory also does not occur independently. Its subcategorization has to be combined with the ‘size’ subcategory. However, unlike the other classifier languages described by Allan (1977), the size range for size classification is from ‘fine’, ‘small’, ‘medium’ and ‘big’.

Thus, for three-dimensional objects that are fine like pebbles, precious stones, rice and sand, the classifier ‘buiir’ is used while those that are small e.g. fruits, cups, and bulbs,
the classifier “biji” is used instead. Three-dimensional objects that are medium in size like chunks of chocolate, fist-sized stones and lumps of clay co-occur with “bentuk”, while those that are big e.g. bags, computers, cars and houses co-occur with the classifier “bush”. The shape-based classifiers mentioned above are summarized below.

Table 1: Malay shape-based numeral classifiers

<table>
<thead>
<tr>
<th>Dimensionality</th>
<th>One-Dimensional</th>
<th>Two-Dimensional</th>
<th>Three-Dimensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other References</td>
<td>Rigid</td>
<td>Not applicable</td>
<td>Not applicable</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>Long</td>
<td>(not applicable)</td>
</tr>
<tr>
<td></td>
<td>Big</td>
<td>(not applicable)</td>
<td>(not applicable)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>(not applicable)</td>
<td>(not applicable)</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>(not applicable)</td>
<td>(not applicable)</td>
</tr>
<tr>
<td></td>
<td>Fine</td>
<td>(not applicable)</td>
<td>(not applicable)</td>
</tr>
</tbody>
</table>

Similar to Micronesian languages, Malay does not have function based classifiers (Adams & Coultham 1973). Instead, specific classifiers exist in Malay language. The specific sub-category however is even more complex than the other classifier sub-categories since unlike those of shape, rigidity and size, the mapping of specific classifiers is not based on perceptual saliency. Specific categories are unique in their use as their usage is only limited to the nouns they collocate with. Flowers, for example, co-occur only with “bentuk” while furniture only with “bentuk”. The following illustrates the possible Malay noun phrases with their respective classifier.

<table>
<thead>
<tr>
<th>Num</th>
<th>CL</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>orang</td>
<td>guru</td>
</tr>
<tr>
<td>Three CL (animate; human)</td>
<td>(three teachers)</td>
<td>teacher</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Num</th>
<th>CL</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>ekor</td>
<td>kucing</td>
</tr>
<tr>
<td>Three CL (animate; animal)</td>
<td>(three cats)</td>
<td>cat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Num</th>
<th>CL</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>batang</td>
<td>kaleng</td>
</tr>
<tr>
<td>Three CL (1D; rigid)</td>
<td>(three pencils)</td>
<td>pencil</td>
</tr>
<tr>
<td>Num</td>
<td>CL</td>
<td>N</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tiga</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendices

Malay Classifiers

<table>
<thead>
<tr>
<th>Num</th>
<th>CL</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>hihah</td>
<td>pisau</td>
</tr>
<tr>
<td>Three</td>
<td>CL (specific knife)</td>
<td>knife</td>
</tr>
<tr>
<td></td>
<td>'three knives'</td>
<td></td>
</tr>
</tbody>
</table>

Overall, the Malay Numeral Classifier system may be sketched as illustrated in Figure 1 by the majority of its adult native speakers.

Conclusion

This sketch on the conceptual representation of the numeral classifier system is only the starting point in identifying the manner Malays think and manifest what they perceive. It is hoped that this paper will stimulate further investigations into the issues raised in the language and cognition field.

Figure 1: The Malay numeral classifier system
References


Dewan Bahasa dan Pustaka Corpus.


Bab 14

PEMEROLEHAN BAHASA KANAK-KANAK DI MALAYSIA: ANTARA TEORI DAN REALITI

Khazriyati Salehuddin & Heather Winskel
MARCS Auditory Laboratories, Australia


Pengenalan

Proses penerolehan bahasa merupakan satu proses semulajadi yang dilalui oleh setiap orang kanak-kanak yang membolehkan setiap seorang daripada mereka menguasai satu-satu bahasa. Proses ini berlaku tanpa disedari oleh kanak-kanak itu sendiri kerana
ianya tidak melibatkan proses pembelajaran secara formal yang disertai dengan arahana-
rahin spesifik. Melalui proses inilah, perkembangan bahasa yang meliputi penggunaan
rumus-rumus linguistik sama ada yang mudah mahupun yang kompleks dicapai kanak-
kanak kanan kurang perlu adanya usaha yang bersungguh-sungguh dari pihak kanak-kanak
itu sendiri. Proses pemerolehan ini adalah berlainan sama sekali dengan proses
pembelajaran bahasa memandangkan proses pembelajaran satu-satu bahasa itu
biasanya berlaku dalam konteks yang formal yang dilengkapi dengan pengajaran secara
terperinci rumus-rumus linguistik kepada kanak-kanak bagi membekalkan mereka
menguasai bahasa tersebut.

Kajian proses pemerolehan bahasa sebenarnya telah mula mendapat tempat
dalam kajian perkembangan kanak-kanak seawal tahun 1787 lagi, bermula dengan
kajian yang dijalankan oleh ahli biologi Jerman, Tiedemann (Malmkjær, 2000). Sementak
itu, pelbagai bentuk kajian ke atas proses pemerolehan bahasa kanak-kanak — dari
kajian diari kanak-kanak individu yang berbentuk longitudinal kepada kajian bersampel
besar yang berbentuk cross-sectional — telah dijalankan ke atas kanak-kanak di serata
pelosok dunia. Kajian ke atas proses pemerolehan bahasa kanak-kanak ini semakin
dianggap penting sehingga di percaya antara antara antara arah atau jurnal serta buku-buku
ilmiah yang membicarakan proses ini diterbitkan dalam bilangan yang tinggi setiap
tahun.

Bilangan penerbitan hasil kajian yang menerokai proses perkembangan bahasa
kanak-kanak ini menggambarkan bahawa bidang ini bukan sahaja tidak seharusnya
dipinggirkan, tetapi sewajarnya diberikan perhatian yang serius. Ini kerana para
pengkaji proses pemerolehan bahasa kanak-kanak sebulat suara berceteju bahawa
pemahaman ke atas proses tersebut memandu kanak-kanak dalam menguasai satu-
satu bahasa. Masyarakat tidak boleh menolak, sebaliknya menerima hakikat bahawa
lebih banyak bahasa yang mampu dikunakan seorang, lebih banyak peluang baginya
untuk meluaskan serta mendalami ilmu yang ada padanya. Persoalannya, menguasai
satu-satu bahasa itu bukanlah satu yang mudah — jadi cara yang lembut untuk membantu
kanak-kanak menguasai satu-satu bahasa adalah dengan memahami bagaimana
caranya proses pemerolehan bahasa berlaku serta faktor-faktor yang dapat membantu
melicinkan proses tersebut.

Secara amnya, kajian ke atas proses pemerolehan bahasa kanak-kanak
merupakan satu kajian yang menarik memandangkan kajian-kajian terdahulu
menunjukkan bahawa peringkat-peringkat yang dilalui oleh kanak-kanak dalam usaha
mereka menguasai satu-satu bahasa sesuai dengan proses semula jadi perkembangan
konseptual kanak-kanak (Uchida & Imai, 1999; Yamamoto & Keil, 2000). Pengkaji-
pengkaji perkembangan bahasa kanak-kanak seterusnya mendapati bahawa
perkembangan bahasa kanak-kanak sebenarnya adalah hasil interaksi antara
perkembangan konseptual yang bukan berbentuk linguistik dengan kategori-kategori
linguistik yang unik bagi satu-satu bahasa (Choi & Bowerman, 1991).
Kajian Pemerolehan Bahasa di Malaysia


Walaupun proses pemerolehan bahasa hampir dianggap sebagai satu proses yang fiktif disebabkan oleh persamaan di antara peringkat-peringkat yang dilalui oleh kanak-kanak dalam penguasaan satu-satu bahasa itu, kepercayaan persekitaran yang Slobin kenali sebagai pengetahuan konseptual inilah yang menjadikan proses pemerolehan bahasa kanak-kanak di Malaysia iringi jika dibandingkan dengan proses pemerolehan bahasa kanak-kanak di lain-lain tempat di dunia ini.

Pelbagai aspek boleh dikaji oleh penyelidik untuk melihat proses pemerolehan bahasa kanak-kanak. Antara yang sering kali dikaji ialah proses kanak-kanak menguasai kategori leksikal. Kajian rentas bahasa yang telah dijalankan ke atas perkembangan kata-kata awal yang dihasilkan oleh kanak-kanak menunjukkan kata nama menguasai sebahagian besar pemerolehan kata awal (Gentner, 1982). Ini kerana, jika dibandingkan dengan kategori-kategori bahasa yang lain seperti kata kerja dan kata sifat, kata nama bersifat lebih stabil dari segi makna yang dibawa olehnya. Selain daripada itu, pemetaan semantik kata nama adalah lebih terus dengan dunia persepsi-konseptual kanak-kanak berbanding kata leksikal yang lain (Gentner, 1982; Gentner & Boroditsky, 2001; Goldfield, 2000).

Kajian ke atas proses pemerolehan kata nama seterusnya menunjukkan bahawa pada peringkat-peringkat awal proses pemerolehan kata nama, antara yang dilakukan oleh kanak-kanak adalah dengan meluaskan penggunaan kata nama yang sama untuk
pelbagai objek yang berlain. Proses yang dikenal sebagai overextension ini memberi gambaran kepada para pengkaji perkembangan bahasa kanak-kanak bahawa kanak-kanak kecil memberi tumpuan yang lebih kepada kualiti yang dapat diterima oleh orang lain. Clark dan Bowerman menyatakan bahawa kualiti ini berdasarkan bentuk objek-objek tersebut (Bowerman, 1978; Clark, 1973). Pemerhatian Clark (1973) misalnya, menunjukkan bahawa seorang daripada subjek kajianya menamakan semua objek yang berbentuk bulat (contohnya, kelip, bulatan, dan huruf ‘O’) sebagai ‘bulan’. Gentner, dalam eksperimen ‘bentuk fungsi’ yang diajalkannya juga menunjukkan bahawa 80% daripada subjek kajianya yang berusia antara 2 dan 5 tahun memberi tumpuan yang lebih kepada bentuk (Gentner, 1978). Begitu juga dengan kajian Landau dan rakan-rakannya yang turut melaporkan bahawa kanak-kanak yang berusia 2 hingga 3 tahun memberi tumpuan lebih kepada bentuk berbanding dengan saiz, tekstur dan warna apabila diminta mengkategorikan objek-objek yang diberikan kepada mereka (Landau et al., 1998). Imai dan rakan-rakan membuat rumusan bahawa kesamaan dari segi bentuk adalah penting dalam proses menganas kata nama (Imai et al., 1994). Dapatkan ini kesemua menggambarkan bahawa ‘bentuk’ diperspesi oleh kanak-kanak sebagai satu aspek yang begitu menonjol sehingga memainkan peranan yang penting dalam membantu mereka mempelajari kategori leksikal pada peringkat awal proses pemerolehan bahasa.

Apa yang berlaku dalam proses pemerolehan kata nama yang melibatkan pengkategorian secara tidak langsung melalui proses overextension ini sebenarnya memberi gambaran kepada pengkaji bahawa kanak-kanak telah mula mengkategorikan objek-objek semajak awal lagi. Fenomena ini seterusnya mencah yang bahawa proses pengkategorian merupakan satu proses fritrah yang wujud seiring dengan perkembangan kognitif manusia. Pengkategorian ini dimanifestasikan dalam pelbagai bentuk oleh pelbagai bahasa di dunia ini, antaranya pengkategorian dalam bentuk jantina, kasus dan sebagainya. Namun, bagi setengah-setengah masyarakat di dunia ini, pengkategorian ini dimanifestasikan dalam bentuk penggunaan penjodoh bilangan dalam sintaksis bahasa yang mereka gunakan.

**Pengkategorian dalam Penjodoh Bilangan**


Pada dasanya, sistem penjodoh bilangan berkongsi sifat-sifat universal seperti proses membezakan antara benda-benda hidup dan yang bukan hidup, yang masing-masingnya dirintahkan sebagai kata nama bermakna dan kata nama tak bermakna dalam bidang linguistik. Namun demikian, sistem penjodoh bilangan turut memperkukuhkan cara satu-satu bahasa membahagi-bahagikan objek-objek berdasarkan persamaan yang mereka miliki.

Cara sistem penjodoh bilangan membahagi-bahagikan ruang semantik sebenarnya memberi peluang yang lebih besar untuk para pengkajian psikolinguistik melakukan kajian ke atas hubungan antara bahasa dan kognisi seperti hipotesis Whorf yang mengetengahkan itu yang sering diperdebatkan sama ada kategori linguistik mempengaruhi cara pemikiran manusia, atau pun cara pemikiran manusia mempengaruhi kategori-kategori linguistik yang wujud dalam bahasa mereka. Hubungan antara bahasa dan kognisi inilah yang menjurut kanak-kanak semenjak dari awal usia mereka lagi memberikan perhatian kepada perkara-perkara yang wujud di persekitaran mereka yang lazimnya diterjemahkan dalam bahasa yang mereka gunakan.

Pemerolehan Penjodoh Bilangan Jepun, Cina & Thai


Walaupun terdapat perbezaan akan bila penjodoh bilangan mula digunakan bagi ketiga-tiga bahasa di atas, usia kanak-kanak mula menghasilkan penjodoh-

Kajian-kajian di atas turut menunjukkan bahawa walaupun proses pemerolehan penjodoh bilangan subjek-subjek mereka berlaku agak lambat di awal proses pemerolehannya, proses tersebut berlangsung dengan agak pantas sekitar usia 4 dan 5 tahun. Ini selari dengan peningkatan dalam perkembangan ilmu yang berkaitan rapat dengan rumus-rumus semantik bagi kategori-kategori yang berlainan (Uchida & Inai, 1999). Pada usia inilah kanak-kanak, menurut kajian yang melihat kebolehan kanak-kanak mencerminkan hal berlaku dengan masa, mampu beranjak dari sistem ‘eka-rujukan’ iaitu satu kebolehan yang membolehkan kanak-kanak memberi fokus hanya kepada satu peristiwa pada satu masa, kepada sistem ‘dvi-rujukan’, iaitu kebolehan yang membolehkan mereka memberi fokus kepada dua peristiwa serentak dalam satu-satu masa (Weist et al., 1997). Kebanyakan pengkaji proses pemerolehan penjodoh bilangan terdakui berpendapat bahawa arus keupayaan kognitif ini mungkin memainkan pengaruh yang besar ke atas kebolehan kanak-kanak mengkategorikan objek, khasnya melalui pergerakan klasifikasikan objek-objek berdasarkan ciri-ciri yang berbeza. Ini seterusnya mampu menjelaskan mengapa dalam proses pemerolehan penjodoh bilangan, sebalik sahaja mencapai usia 6 tahun, kanak-kanak mampu menggunakan penjodoh bilangan berfrekuensi rendah dan menggunakan penjodoh bilangan-penjodoh bilangan tersebut bersama-sama objek-objek tidak tipikal.

Walau bagaimanapun, adakah dapaan yang diperoleh dari pada kajian yang dilakukan ke atas kanak-kanak yang menguasai bahasa Jepun, Cina dan Thai ini sudah mencukupi untuk kita memahami proses perkembangan bahasa kanak-kanak di Malaysia seperti kanak-kanak Melayu misalnya? Jika dilihat dari susur galur bahasa Melayu, ketiga-tiga bahasa Jepun, Cina dan Thai itu tidak mempunyai pertalian kekeluargaan dengan bahasa Melayu yang merupakan bahasa Austronesian. Hal ini sudah pasti menjadi titik tolak kepada perlunya dijalankan kajian pemerolehan penjodoh bilangan Melayu pula. Selain dari itu, memandangkan sistem bahasa banyak dipengaruhi oleh sistem konseptual bukan linguistik masyarakat penuturnya, ketidaksemaan dari segi budaya antara masyarakat Jepun, Cina, dan Thai dengan masyarakat Melayu menguatkkan lagi keperluan kajian pemerolehan penjodoh bilangan Melayu dijalankan.

**Pemerolehan Penjodoh Bilangan Melayu**

Sebuah kajian rintis ke atas 112 orang penutur asli bahasa Melayu yang berusia antara 3 tahun 4 bulan dan 12 tahun 4 bulan telah dijalankan untuk melihat kebolehan mereka

Untuk menjalankan analisis ke atas maklumbalas yang diberikan oleh keseluruhan 112 orang kanak-kanak terbabit, maklumbalas yang mereka berikan telah dikategorikan sebagai "Maklumbalas tepat" dan "Maklumbalas Tidak Tepat". "Maklumbalas Tidak Tepat" ini termasuklah kegagalan kanak-kanak terbabit menggunakan sebarang penjodoh bilangan (contohnya 3 *kanak-kanak*) dan penggunaan penjodoh bilangan alternatif (contohnya 3 *batang papan*) dalam proses pengiraan separjyang eksperimen ini dijalankan.

Secara keseluruhannya, purata maklumbalas yang tepat yang diperoleh dari keseluruhan 112 orang kanak-kanak hanyalah 42.60%. Keputusan dari analisis yang dijalankan menunjukkan bahawa kanak-kanak antara usia 3 dan 5 tahun tidak mampu menghasilkan banyak penjodoh bilangan dengan tepat. Daripada 224 maklumbalas yang diterima daripada kanak-kanak antara usia 3 tahun 4 bulan dan 5 tahun 1 bulan itu, hanya satu maklumbalas sejauh yang tepat manakala dalam kelangan kanak-kanak antara usia 5 tahun 11 bulan dan 7 tahun 1 bulan, hanya 32 daripada 224 maklumbalas yang diterima adalah tepat. 79.46% daripada maklumbalas yang diterima daripada kanak-kanak antara usia 9 tahun 6 bulan dan 12 tahun 4 bulan pula merupakan maklumbalas yang tepat.

Data yang diperoleh turut menunjukkan bahawa kanak-kanak pada usia antara 3 tahun 4 bulan dan 5 tahun 1 bulan hanya mampu menghasilkan penjodoh bilangan *orang*, dan antara usia 5 tahun 1 bulan dan 5 tahun 11 bulan, mereka mula mampu menghasilkan penjodoh bilangan *ekor* dan *biji*. Antara 5 tahun 11 bulan dan 7 tahun 1 bulan, kanak-kanak mula menghasilkan 4 lagi penjodoh bilangan, iaitu *batang*, *helai*, *buah* dan *kuntum*. Namun demikian, analisis terperinci yang dijalankan menunjukkan bilangan penjodoh bilangan yang tepat yang berjaya dihasilkan oleh keseluruhan kanak-kanak di atas adalah da bawah paras 75% ketepatan.

Analisis varians yang dijalankan menunjukkan terdapatnya kesan pengaruh usia ke atas maklumbalas yang dijalankan sementara analisis *post hoc* Tukey pada $\alpha=0.01$ menunjukkan bahawa maklumbalas yang diterima daripada kanak-kanak yang berusia 7 tahun adalah secara signifikannya berlainan daripada kanak-kanak yang lebih muda mahupun yang lebih tua daripada mereka.
Penelitian ke atas setiap satu daripada penjodoh bilangan yang dihasilkan oleh kanak-kanak terbit memunjukakan kanak-kanak hanya mampu menghasilkan penjodoh bilangan dengan 100% ketepatan bagi setiap orang kanak-kanak pada usia 7 tahun. Ini membuktikan penjodoh bilangan orang, ekor dan batang. Tiga lagi penjodoh bilangan iaitu atas, helai dan kuning hanya dihasilkan dengan 100% ketepatan pada usia 10 tahun. Secara keseluruhannya, daripada 14 jenis penjodoh bilangan yang didinjan, 11 daripadanya boleh dihasilkan oleh kanak-kanak menjelang usia 11 tahun sementara tiga lagi masih tidak lagi dapat dihasilkan oleh mereka - walaupun pada usia 12 tahun.


Namun demikian, analisis ke atas setiap satu penjodoh bilangan menunjukkan proses pemerolehan penjodoh bilangan bersasakan bentuk dan penjodoh bilangan bersasakan bentuk yang dihasilkan oleh penyelidik-penyelidik terdahulu (Erbaugh, 1986; Matsumoto, 1985; Ng Bee Chin, 1989). Ini kerana tidak semua penjodoh bilangan bersasakan bentuk yang diperoleh lebih awal daripada penjodoh bilangan bersasakan bentuk. Variasi dari sebagian pemerolehan penjodoh bilangan ini mungkin disebabkan oleh faktor yang dicadangkan oleh Hansen dan Chen iaitu pendedahan kepada penjodoh bilangan tertentu menyebabkan kanak-kanak mempunyai kecenderungan untuk memperoleh penjodoh bilangan tersebut berbanding dengan yang lain (Hansen & Chen, 2001).
Perbincangan


Kajian-kajian terdahulu menyatakan bahawa proses pemerolehan penjodoh bilangan sebenarnya bervariasi dipengaruhi oleh perkembangan sistem kognitif kanak-kanak. Pengkajian pengkajian proses pemerolehan penjodoh bilangan bahasa Jepun, Cina dan Thai menunjukkan perkembangan sistem kognitif yang mendahuk pada usia 5 tahun sebenarnya membolehkan penutur-penutur bahasa Jepun, Cina dan Thai menghasilkan lebih dari 4 penjodoh bilangan pada usia 5 tahun. Jika dilihat pada kanak-kanak Melayu, bilangan 1 penjodoh bilangan yang dihasilkan dengan tepat pada usia 5 tahun adalah amat membimbingkan. Namun adaah statistik seumpamanya ini sudah mencukupi untuk kita membuat rumusan bahawa jika dibandingkan dengan sistem kognitif kanak-kanak Jepun, Cina dan Thai, perkembangan kognitif kanak-kanak Melayu adalah jauh lebih perlahan?

Dalam membimbingkan itu proses pemerolehan bahasa, satu perkara yang tidak seharusnya diingatkan dalam perbincangan adalah pengaruh input dalam perkembangan bahasa kanak-kanak. Hasil datapan kajian rintis yang disalakan penulis di atas mendorong penulis menunja kembali jenis-jenis input bahasa yang sedia ada untuk diteladani oleh kanak-kanak Melayu hari ini. Analisis korpus yang disalakan ke atas bahasa-bahasa yang wujud di persekitaran kanak-kanak Melayu hari ini, termasuklah bahasa-bahasa bacaan awal, bahasa-bahasa yang diterbitkan oleh media elektronik seperti rancangan berita, siri kartun dan rancangan kanak-kanak dalam bahasa Melayu mesti, memenjukkan penggunaan penjodoh bilangan yang amat kurang. Jika dibandingkan dengan bahasa-bahasa lain yang menggunakan penjodoh bilangan dalam sintaksis mereka, frekuensi penjodoh bilangan dalam bahasa-bahasa
yang diterbitkan dalam bahasa Melayu adalah rendah walau pun penjodoh bilangan dianggap sebagai unsur sintaksis-semantik yang wajib dalam bahasa Melayu.

Namun demikian, kajian susulan yang dilakukan menunjukkan bahawa sistem wacana bahasa Melayu sebenaranya membenarkan penjodoh bilangan ditinggalkan, walau pun pada hakikatnya sintaksis bahasa Melayu mewajibkan penggunaannya. Peninggalan penjodoh bilangan dibenarkan jika pengguna bahasa tersebut berharap untuk tidak membuat sebarang penekanan serta memberi ulasan lanjut ke atas objek-objek yang dihitung mereka. Pemerhatian penulis ini adalah selari dengan hasil kajian korpus yang dilakukan oleh Hopper ke atas Hikayat Abdullah (Hopper, 1986).

Berdasarkan analisis korpus dan analisis wacana yang dijalankan oleh penulis, kajian ke atas proses pemerolehan penjodoh bilangan bahasa Melayu ini menunjukkan proses perkembangan bahasa kanak-kanak bukan sahaja dipengaruhi oleh perkembangan sistem kognitif kanak-kanak, malah, pengaruh persekitaran juga memainkan peranan yang penting. Proses pemerolehan penjodoh bilangan yang lantak dalam kalangan kanak-kanak Melayu memang jelas disebabkan oleh faktor kurangnya persekitaran yang kaya dengan penggunaan penjodoh bilangan Melayu dan ini mencadangkan bahawa kelewatian ini bukanlah disebabkan oleh kelewat dalam proses perkembangan kognitif kanak-kanak Melayu. Ini kerana, di sebelah kelewat dalam menghasilkan penjodoh bilangan Melayu pada peringkat awal, terdapat peningkatan yang terdapat dari segi bilangan penjodoh bilangan yang dihasilkan sebaik sahaja kanak-kanak Melayu mencapai usia 6 hingga 7 tahun.

Kajian ke atas kanak-kanak Melayu ini menunjukkan kanak-kanak Melayu mampu menghasilkan 6 penjodoh bilangan pada usia 6 tahun - sama seperti bilangan penjodoh bilangan yang dihasilkan oleh kanak-kanak lepung yang seusia dengan mereka. Sementara itu, pada usia 7 tahun, bilangan penjodoh bilangan yang dihasilkan oleh kanak-kanak Melayu adalah 12, sama seperti bilangan yang dihasilkan oleh kanak-kanak Cina.

Satu perkara yang boleh kita pelajari dari kajian proses pemerolehan penjodoh bilangan Melayu di atas ialah pengaruh input ke atas perkembangan bahasa kanak-kanak. Sehubungan itu, ibu bapa serta pengguna-pengguna bahasa yang lain haruslah membantu kanak-kanak dalam proses perkembangan bahasa mereka dengan menyediakan persekitaran linguistik yang kaya dengan unsur-unsur penting bagi bahasa dan budaya mereka. Tanpa persekitaran linguistik yang kaya untuk kanak-kanak lalu, perkembangan kognitif sahaja tidak dapat membantu melicinkan proses perkembangan bahasa kanak-kanak.

**Penutup**

Perbincangan yang diadakan dalam bab ini menunjukkan perlu adanya kumpulan penyelidik-penyelidik tempatan menjalankan kajian ke atas perkembangan bahasa
kanak-kanak di Malaysia ke atas pelbagai aspek linguistik yang dimiliki oleh bahasa-bahasa yang wujud di Malaysia. Ini kerana penyelidik-penyelidik tempatan sahaja yang memahami persekitaran bahasa yang wujud di persekitaran kanak-kanak dan mampu menafsirkan keputusan-keputusan kajian yang berkemungkinan tidak sama dengan dapatan kajian luar. Oleh itu, adalah menjadi tanggung jawab penyelidik-penyelidik tempatan untuk bersama-sama berganding bahu menjalankan kajian-kajian lanjut ke atas proses pemerolehan bahasa kanak-kanak di Malaysia, bukan sahaja untuk membantu proses pengurusan satu-satu bahasa, malah, membenarkan daripada disalah tafsirkan oleh penyelidik-penyelidik bukan tempatan yang kurang faham akan aspek-aspek linguistik mahupun aspek-aspek bukan linguistik yang mampu mempengaruhi proses pemerolehan bahasa kanak-kanak.

Bab ini menunjukkan bahawa pengetahuan dari segi teori sahaja tentang perkembangan bahasa kanak-kanak adalah amat tidak memadai untuk memahami proses tersebut kerana realiti menunjukkan bahawa proses ini, walaupun mempunyai persamaan jika dilihat secara global, mempunyai perbezaan yang banyak jika dibandingkan satu persatu dengan proses perkembangan bahasa kanak-kanak yang lain. Selubung ini, adalah wajar jika di Malaysia ini, satu institusi yang khusus menjalankan kajian secara aktif ke atas perkembangan bahasa kanak-kanak yang lengkap dengan alat terkini seperti MRT dan PET/SCAN (alat canggih yang membolehkan penyelidik mengimbas aktifiti-aktifiti yang berlaku dalam otak kanak-kanak serta bahagian-bahagian otak yang aktif semasa mengendalikan aktiviti aktifiti linguistik) diwujudkan bagi membolehkan penyelidik-penyelidik tempatan menjalankan kajian-kanak-kanak yang pesat ke atas proses pemerolehan bahasa-bahasa tempatan di Malaysia ini.

Rujukan

Appendices

Pemerolehan Bahasa Kanak-Kanak / 213

(pp. 05-110). New York: Academic Press.


An investigation into Malay numeral classifier acquisition through an elicited production task

Khazriyati Salehuddin, University of Western Sydney
Heather Winskel, University of Western Sydney

ABSTRACT
The act of categorization and labeling is fundamental in human cognition and language development. By studying numeral classifier acquisition, researchers are able to examine how children learn to categorize and label objects in their environment using a constrained framework. The current study investigated the acquisition of eight shape-based numeral classifiers in Malay through an elicited production task in 140 6-to 9-year-old children. The aim was to examine the developmental patterns observed in the production of Malay shape-based numeral classifiers. Results indicated that the ability to produce the correct numeral classifiers is a relatively prolonged process that involves an interaction of a variety factors, including semantic complexity, input frequency, and the formal teaching of numeral classifiers in school.

KEYWORDS
Categorization; language acquisition; Malay; numeral classifiers; production

The categorization and labeling of objects is one of the most basic processes in human cognition and language development (Croft & Cruse, 2004; Lakoff, 1987). Categorization occurs early in development; young children and infants are able to consistently discriminate between animate and inanimate objects, and between
Inanimate objects and human beings (Greif, Nelson, Kell, & Gutterrez, 2005; Jipson & Geiman, 2007; Molina, Van de Walle, Condry, & Spelke, 2004; Simons & Kell, 1995). In order to categorize and label a particular object, stored information is retrieved and evaluated, and only when there is a satisfactory resemblance between new and stored information is an object accepted into the respective category (Balsalou, Huttenlocher, & Lamberts, 1993). Objects that are perceived as belonging to the same category are then assigned the same label (Mervis & Rosch, 1981). Researchers are intrigued as to how children learn to categorize and label objects in their environment, and how this process develops and gradually becomes more refined over time. Numeral classifier systems offer researchers an ideal opportunity to examine how children learn to categorize and label objects in their environment using a constrained framework or system.

Numeral classifiers (NumCl) are typically used in counting objects. An object is paired with a particular numeral classifier dependent on its features or characteristics; for example in Malay, a ball is round and small and thus takes the numeral classifier biji, which is used for any small, round object; e.g., 3 biji bola (three NumCl ball) is used to refer to ‘three balls.’ Speakers from numeral classifier languages need to learn how to categorize objects in their environment and pair them with the appropriate numeral classifier, using the language-specific classification system.

Malay, an Austronesian language, has a rich and complex numeral classifier system (Dirin, 2000; B. Othman, 2004), which enables us to investigate the development of categorization and labeling through numeral classifier categories. The current study aims to investigate the acquisition of Malay shape-based numeral classifiers through an elicited production task in children aged 6–9 years old. First, we review background information on the characteristics of numeral classifier systems, with a particular focus on the Malay numeral classifier system. Next, we review previous research conducted on the acquisition of numeral classifiers, prior to outlining the specific research aims and hypotheses of the current study.

NUMERAL CLASSIFIER SYSTEMS

Numeral classifiers are a syntactic-semantic category that is common in Sino-Tibetan, Atlantic-Congo, and Austronesian languages. Numeral classifier systems share some universal aspects in their classification of nouns by classifying objects based on primary parameters, i.e., distinctions between animate and inanimate objects and between human and nonhuman distinctions (Adams & Conklin, 1973; Allan, 1977). However, at a more specific level, substantial differences in numeral classifier languages exist and it is at this level that semantic features like rigidity and size (which Adams & Conklin [1973] described as secondary parameters due to dependence of these features on the primary parameters) are combined with the primary parameters to form numeral classifier classes (Uchida & Imai, 1999). Functional parameters (e.g., tools, footwear, written materials) form even more language-specific numeral classifiers because their usage reflects the interests of the specific culture (Adams & Conklin, 1973).
According to Yamamoto (2005), both cognition-based universal features and culture-based idiosyncratic features have to be considered in forming a complete model of numeral classifier systems. The categorization of numeral classifiers largely cuts across taxonomic categories based on the attributes of their referents (T'Scu, 1976), resulting in striking language partitioning differences (Yamamoto & Kei, 2000). For example, in Mandarin Chinese, snakes are classified as tiao (i.e., long, thin, curved objects) together with rivers and roads while monkeys are classified in the same category with human as ge. In contrast, animals in Japanese are never classified with either inanimate objects or with human beings. From the semantic, cognitive, and cultural perspectives, numeral classifiers function as a means to communicate a few especially important classes that objects fall into in the manner they are perceived by the speech community (Craig, 1986).

![Diagram of Malay numeral classifier system]

**Figure 1** The Malay numeral classifier system

Note: The shape-based numeral classifiers require mixed semantic criteria involving dimensionality and rigidity or dimensionality and size. The specific classifiers are more extensive than illustrated.
THE NUMERAL CLASSIFIER SYSTEM OF MALAY

The numeral classifier system of Malay is considered to be conceptually complex, as the selection of numeral classifiers in classifying Malay nouns has a high degree of arbitrariness (Dirin, 2000; B. Othman, 2004) (see Figure 1). Apart from the numeral classifiers for human (orang – literally ‘human’), animal (ekor – literally ‘tail’), and trees (batang – literally ‘stem’), which are relatively semantically transparent, the acquisition of other Malay numeral classifiers appears to be more complex. There are many exceptions to the rule in the selection of some numeral classifiers: e.g., the classifier buah, which literally means ‘fruit,’ is used for example with cars, but not fruit; the classifier biji, which literally means ‘seed,’ is for example used for cups and plates but not seeds, and the classifier kaki, which literally means ‘leg,’ is used for umbrellas, some plants, but not legs. This kind of arbitrariness of the majority of Malay numeral classifiers means that the Malay numeral classifier system is often regarded as loose and opaque and as ‘semantically non-transparent’ (Omar, 1972, p. 89).

Structurally, similar to Japanese and Chinese, Malay numeral classifiers are preceded by a numeral and followed by a head noun:

<table>
<thead>
<tr>
<th>Num</th>
<th>NumCl</th>
<th>N</th>
<th>Num</th>
<th>NumCl</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiga</td>
<td>orang</td>
<td>kanak-kanak</td>
<td>Empat</td>
<td>ekor</td>
<td>arnam</td>
</tr>
<tr>
<td>Three</td>
<td>NumCl</td>
<td>child</td>
<td>Four</td>
<td>NumCl</td>
<td>rabbit</td>
</tr>
<tr>
<td>‘three children’</td>
<td>‘four rabbits’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When numeral classifiers are used in Malay, the noun that forms the head of the noun phrase appears in a singular form as illustrated above. Situations where numbers are not used to indicate plurality, numeral classifiers are obligatorily deleted. Such instances occur usually when nouns are reduplicated (Omar, 1974):

- N-pl
  - Rumah-rumah
  - Buku-buku
  - ‘houses’
  - ‘books’

In isolated cases, instead of appearing to the left of the head noun, numeral classifiers may also appear to the right of the noun, following a numeral. Such a structure is usually found in itemizing different types of nouns together with their respective quantities:

<table>
<thead>
<tr>
<th>N</th>
<th>Num</th>
<th>NumCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>anak</td>
<td>dua</td>
<td>orang</td>
</tr>
<tr>
<td>‘Child’</td>
<td>‘two’</td>
<td>‘two’</td>
</tr>
<tr>
<td>N</td>
<td>Num</td>
<td>NumCl</td>
</tr>
<tr>
<td>Kerbau balar</td>
<td>tujuh</td>
<td>ekor</td>
</tr>
<tr>
<td>‘Buffalo albino’</td>
<td>‘seven’</td>
<td>‘seven’</td>
</tr>
</tbody>
</table>

Despite being syntactically different from the noun, what is most significant in the two examples above is that the choice of numeral classifiers is maintained and
consistent for each type of noun, i.e., the noun 'child' co-occurs with the classifier orang [animate: human], while the noun 'buffalo' co-occurs with the classifier ekor [animate: animal].

Imai and Gentner (1997) describe numeral classifiers as similar to modifiers that unitize uncountable objects for quantifying purposes (i.e., unifying modifiers) such as a spoonful of sugar, a bowl of rice, and a cup of coffee. However, this proposition does not hold for Malay numeral classifiers, as numeral classifiers are not used for quantifying mass nouns. Instead, numeral classifiers are used to qualify count nouns based on the permanent properties of the objects.

As presented in Fig. 1, the categorization of inanimate objects in Malay occurs in two ways – either based on shape or based on specific attributes of the objects. Specific numeral classifiers of Malay are relatively straightforward as they involve a one-to-one mapping between the object and the numeral classifier, e.g., all flowers take kunang; whereas Malay shape-based numeral classifiers are more complex, e.g., long objects can take either batang or utas. The shape-based numeral classifiers involve mixed semantic criteria comprising dimensionality and rigidity or dimensionality and size. Malay 1D and 2D numeral classifiers have two forms determined by the rigidity of the object (rigid and non-rigid), whereas Malay 3D numeral classifiers have four different forms depending on the size of the object (fine, small, medium, and big). In contrast to Chinese, Japanese, and Thai numeral classifier systems (Carpenter, 1991; Hu, 1993; Matsumoto, 1985), Malay does not have an unmarked general numeral classifier that allows the substitution of specific numeral classifiers within the same ontological boundaries without distorting the syntax and the semantics of the language.

REVIEW OF LITERATURE ON THE ACQUISITION OF NUMERAL CLASSIFIERS

Research on language acquisition has largely focused on the years up to the age of 5. However, language continues to develop throughout the school years (Crutchley, 2007; Nippold, 1988). Cross-linguistic studies on numeral classifier acquisition have found that the acquisition of numeral classifiers occurs relatively late in development (Carpenter, 1991; Gandour, Petty, Dardaranada, Dehongkit, & Mukhngoen, 1984; Mak, 1991; Ng, 1989; Sanches, 1977; Yamamoto, 2005).

Pinpointing the precise acquisition order of specific numeral classifiers is difficult because the internal structure of each category domain differs from language to language (Yamamoto, 2005). Previous studies (e.g., Burin, 1959; Erbaugh, 1984, 1986; Gandour et al., 1984) observed early distinctions made between animate and inanimate numeral classifier usage in children, which corresponded to preschool children's capacity to be able to readily discriminate or categorize animate and inanimate objects. Consequently, they concluded that cognitive-based categories are fundamental in the acquisition of numeral classifiers.

The picture becomes more complex when we consider the acquisition order of shape-based categories in numeral classifier systems. Erbaugh (1984, 1986), in her longitudinal and cross-sectional studies on Japanese numeral classifier acquisition,
found that 1D classifiers were produced prior to 2D classifier and that 3D classifiers were produced last. She attributed these findings to the fact that length is noticed and understood prior to flatness. However, Matsumoto (1985), in a study of 5- to 7-year-old Japanese children, found that acquisition of 3D classifiers occurred prior to 2D and 1D classifiers. Matsumoto's participants also showed the production of the classifier for tiny objects (termed by Yamamoto [2005] as 0D classifier) before producing the 3D classifier 3. Matsumoto's shape-based numeral classifier acquisition order is dissimilar to Loke's (1991) findings on Mandarin Chinese numeral classifier acquisition, as he found that 3D classifiers were produced after 1D classifiers but before 2D classifiers. However, a study conducted on the acquisition of Hokkien numeral classifiers demonstrated that dimensionality does not necessarily determine the order of acquisition. Ng (1989) proposed that rigidity determines shape-based classifier acquisition order, as her study showed the following acquisition sequence: [1D: -rigid] > [2D: -rigid] > [3D: spherical] > [2D: +rigid]. In Cantonese, Mak (1991) utilized an elicited production task with 4- to 8-year-olds, and found that shape-based numeral classifiers proceeded in the following order: [1D: -rigid] > [1D: +rigid], [2D: +rigid] > [3D: small] > [3D: lumpy] > [2D: -rigid]. The contradictory findings from these studies suggest that it is neither dimensionality nor rigidity that determines the order of shape-based numeral classifier acquisition across languages per se; instead, it is the relative complexity of the classifier's semantic features involving a combination of dimensionality and rigidity or dimensionality and size that affects the acquisition order of shape-based numeral classifiers (Yamamoto, 2005).

Semantic complexity or cognitive complexity refers to the cumulative number of semantic features in a particular semantic category. It has been argued that children first learn to categorize objects requiring less semantic features or contrasts and then proceed to categories involving more complex categorization processes (Clark, 1972, 1973), suggesting that multiple-dimension categorization rules are harder to acquire than single-dimension categorization rules (Madoox, 1992). Thus, based on this assumption, numeral classifiers involving more complex features are acquired later than those involving less complex features (Uchida & Imai, 1999).

Cross-linguistic studies on numeral classifier acquisition reveal that general numeral classifiers are usually acquired relatively early by children and are used by children as default classifiers before the mastery of more specific classifiers (Carpenter, 1991; Chien, Lust, & Chiang, 2003; Wei & Lee, 2001; Yamamoto & Keil, 2000). For example, in Japanese, young preschoolers (1.6–4.0) use the more general classifiers 3tu [Inanimate: generic] and 35 [Inanimate: 3D] in their production prior to the use of other classifiers (Matsumoto, 1985). Similarly, Thai children often use the general classifier 3ai as a 'default' or 'fall back' option in place of more semantically specific numeral classifiers and before learning the limitations of the semantic relationship between a head noun and its specific numeral classifier (Carpenter, 1991; Yamamoto, 2005). Relatively 'semantically heavy' or complex classifiers, such as configurational (shape-based) classifiers, have been found to be relatively difficult for children to acquire (Gandour et al., 1984). By using more general default classifiers as substitutes for more semantically specific or complex
numeral classifiers, children can reduce the semantic load by making classification less dependent on inherent semantic characteristics of referents.

Researchers have also found that substitution responses for shape-based numeral classifiers tend to be ‘same dimension’ rather than ‘different dimension’ substitution responses (Loke, 1991; Mak, 1991). In Cantonese, which has multiple shape-based numeral classifiers with complex combinations of features, the majority of children’s substitution errors were within the same dimensional domain. Similarly, Wei and Lee (2001) found that Chinese–English bilingual children often inappropriately overextended numeral classifier usage to referents in the same semantic category, especially using shape criteria, as shape appears to be more salient than function (Wei & Lee, 2001). Substitutions and errors may also suggest which numeral classifier terms children have available and give clues of the acquisition order of terms within a semantic domain (Bernstein Ratner, 2000).

ROLE OF INPUT IN LANGUAGE ACQUISITION

There appears to be a general consensus that children tend to learn words earlier when they are produced more frequently in speech directed to them than less frequently occurring words (Gallaway & Richards, 1994; Snow & Ferguson, 1977). For example, Tare, Shatz, and Gilberston (2008) examined the frequency and uses of subsets of color and number words in mothers’ speech to toddlers and of time words to preschoolers. Their findings support the argument that varied conversational input provides useful data for children to create early word–word mappings for nonobject terms. In another study, Goodman, Dale, and Li’s (2008) analysis of MacArthur–Bates Communicative Development Inventory norming data and for parental input frequency from the CHILDES database, showed that within each of the six lexical categories studied, the more frequent the word is heard by the child the earlier the word is produced. An important consideration, which has been much debated, is what is the best method to quantify the input children are receiving, as the input being measured may not be a true representation of what children are actually exposed to in their linguistic environment (Hoff-Ginsberg, 1992; Pine, 1992).

Some studies suggest that preschool children’s language development can be influenced by children's television programs (Ely & Gleason, 1995; Rice, 1990), storybook reading (Fisch, 2004; Snow, 1991; Snow, Tabors, Nicholson, & Kirkland, 1995), and noninteractive language such as overhead conversations (De Houwer, 1995). Input frequency has also been cited to be an important factor in the acquisition of Japanese numeral classifiers (Matsumoto, 1985; Yamamoto, 2005). Yamamoto (2005), for example, found that high frequency numeral classifiers in both speech and written texts such as -tsu, -ko, -ni, -hiki, and -dai ‘emerged maturationally’ earlier than lower frequency numeral classifiers.

In order to examine the role of input in acquisition, research on early language development has mostly focused on data from ‘dyadic adult–child play settings’ (Scott, 1988, p. 51). However, according to Nippold (1988), for later language development both spoken and written forms of communication are a significant...
source of stimulation that should be ‘scrutinized.’ This is because, in later language development, school-going children learn language both in informal settings (through indirect modeling and reinforcement) and formal instruction. As numeral classifier acquisition takes place at a later stage in language development (Carpenter, 1991; Gandour et al., 1984; Mak, 1991; Ng, 1989; Sanches, 1977; Yamamoto, 2005), input from various forms of discourse and not limited to those concerning mother–child interaction is likely to play a role in the production of Malay numeral classifiers in children.

THE MALAYSIAN EDUCATION SYSTEM

In the Malaysian education system, children begin primary school at the age of 7. Prior to this age, most children attend either private or public preschools as preparation for primary school. Children are only formally taught Malay numeral classifiers in Grade 1 at 7 years of age, and at this age only six numeral classifiers are introduced in their workbooks, i.e., orang [animate: human], ekor [animate: animal], batang [1D: +rigid], keping [2D: –rigid], buah [3D: big], and biji [3D: small] (Yusof, Jeon, & Sukor, 2004). In Grade 2 textbooks, three other numeral classifiers, namely helai [2D: –rigid], ketol [3D: medium], and katum [specific: flower], are introduced (A. J. Othman & Ismail, 2004). In Grade 3 textbooks, the following additional numeral classifiers are introduced: utas [1D: –rigid], butir [3D: fine], unat [1D: +rigid], pucuk [specific: weapon, letters], bentuk [specific: ring, hook], and kaki [specific: umbrella] (Jeon, Sukor, & Yusof, 2005).

THE CURRENT STUDY

The current study investigates the acquisition of eight of the most commonly used Malay shape-based numeral classifiers through an elicited production task. The aim is to examine the developmental patterns observed in the production of Malay shape-based numeral classifiers. Based on the literature reviewed, the following predictions were formulated on the order of acquisition of Malay numeral classifiers:

1. Previous research suggests that the acquisition order of numeral classifiers within a particular language is affected by the complexity of the semantic features of the individual numeral classifiers with more features acquired later than those with fewer features (Uchida & Imai, 1999). In relation to Malay shape-based numeral classifiers, it is predicted that shape-based classifiers with complex features such as the multiform 3D classifiers (four forms: fine, small, medium, and big) will be acquired later than those with less complex features, i.e., the 1D and 2D classifiers (two forms: +rigid and –rigid).

2. (a) Previous research indicates that more general numeral classifiers are usually used by children as default classifiers prior to mastery of more specific numeral classifiers and are often overused for references (Carpenter, 1991;
SALEHUDDIN & WINSKEL: INVESTIGATION INTO MALAY NUMERAL CLASSIFIER

Chien et al., 2003; Wei & Lee, 2001; Yamamoto & Kell, 2000). Even though Malay does not have a specific general classifier per se, we can still predict that more general classifiers will be used by Malay children as substitutes for other more specific classifiers, and give clues about the acquisition order. (b) Furthermore, as previous studies (e.g., Loke, 1991; Mak, 1991) have found that alternative or substitution responses for shape-based numeral classifiers tend to be ‘same dimension substitutions’ rather than ‘different dimension substitutions,’ we can expect Malay to adhere to a similar trend.

3. As previous research on numeral classifier acquisition proposes that input plays a role in the acquisition of terms within a semantic domain, we expect to find an association between input and classifier word frequencies (as estimated by numeral classifier word frequencies obtained from a 150,000-word corpus) and formal teaching with the production order of numeral classifiers in Malay.

METHOD: THE ELICITED PRODUCTION TASK

The study was designed to investigate children’s ability to produce eight of the most common shape-based numeral classifiers in Malay, namely batang [1D: +rigid], utas [1D: -rigid], keping [2D: +rigid], helai [2D: -rigid], busah [3D: big], ketul [3D: medium], buji [3D: small], and butir [3D: fine], through an enumeration task. Pictures were used as stimuli to elicit responses from the children.

Participants

One-hundred-and-forty children attending a preschool and a primary school in the same national school participated in the study. The children were all native speakers of Malay and spoke Malay as their first language. All the 6-year-olds were preschoolers while the 7-, 8-, and 9-year-olds were in their first, second, and third year of the Malaysian primary education system respectively. There were 31 6-year-olds, 36 7-year-olds, 41 8-year-olds, and 32 9-year-olds. There were also 20 adults participating in the experiment as a comparison group. The adults lived in the vicinity of the school and had a mixed educational background. All participants were from a middle SES. A description of the participants, including the gender ratio, is given in Table 1.

Stimuli

The stimuli consisted of pictures of objects that were selected based on earlier ratings by 30 adult Malay native speakers. Rather than using only a single stimulus for each numeral classifier category, a range of exemplars of each numeral classifier was selected as test stimuli to give a broader representation of each classifier category. Adults were asked to rate pictures of familiar, everyday objects from most typical to most atypical exemplars for each of the eight numeral classifier categories. An object rated as very typical was given a score of ‘5,’ whereas a very
Table 1: Description of participants in the study

<table>
<thead>
<tr>
<th>Age range</th>
<th>Mean age</th>
<th>No. of participants</th>
<th>No. of males</th>
<th>No. of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year-olds</td>
<td>6.8-6.7</td>
<td>6.18</td>
<td>31</td>
<td>14</td>
</tr>
<tr>
<td>7-year-olds</td>
<td>6.8-7.6</td>
<td>7.13</td>
<td>36</td>
<td>13</td>
</tr>
<tr>
<td>8-year-olds</td>
<td>7.9-8.8</td>
<td>8.25</td>
<td>41</td>
<td>19</td>
</tr>
<tr>
<td>9-year-olds</td>
<td>8.11-9.8</td>
<td>9.28</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td>Adults</td>
<td>17.3-77.8</td>
<td>48.07</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>160</td>
<td>63</td>
<td>97</td>
</tr>
</tbody>
</table>

Atypical exemplar was given a score of '1'. The responses given by the adults were averaged and subsequently all five pictures of objects ranked as very typical, typical, moderate, atypical, and very atypical exemplars for each of the shape-based numeral classifiers were used as test stimuli for the current study. All the items appeared in units of three so as to encourage the children to use a number, a numeral classifier, and a noun in naming the respective objects. If the items appeared once, naming the objects only by their respective names could be considered appropriate and if the items appeared twice, using the phrase sepasang ('a pair of') would be considered correct as well.

In addition, in order to gain an estimation of input frequency of the numeral classifiers in the children's environment, a corpus from a range of texts relevant to preschool and school-aged children was analyzed. The corpus consisted of children's television programs (25%), children's storybooks (25%), and folk stories and bedtime stories available online from the website of Dewan Bahasa dan Pustaka (50%), an official body in charge of promoting and preserving the Malay language. From this 150,000-word corpus the word frequencies listed in Table 2 were obtained. It can be seen that buah [3D: big] was the most frequently used shape-based numeral classifier followed by batang [1D: +rigid] and biji [3D: small], whereas butir [3D: fine] did not appear at all in the corpus (Table 2).

Procedure

Prior to the experimental session, a slide display of all 44 pictures with an audio presentation of the names of the respective objects in the picture display were shown to the children (Fig. 2). In this picture-familiarization session, children were asked to repeat the names of the objects after the audio presentation of the respective objects before proceeding to the next slide. This was to ensure that the children were familiar with the items presented to them. They could also ask for clarification at this point.

Four of the 44 pictures were used for practice trials (Table 3). Children were shown what to say when a prompt was given to them, e.g., *Tiga bentuk cincin* ('Three NumCl [specific: rings] rings') and *Tiga pasang cemilan mata* ('Three NumCl [pair] glasses') respectively. The numeral classifiers used in the practice trial were
not used in the experimental session. A prompt in the form of the expression 
Berapa? (‘How many?’) was given to the children if they merely mentioned the 
name of the object, e.g., Lelaki (‘Man’). The children were further prompted with 
Tiga apa lelaki? (‘Three what men?’) if they merely gave Tiga lelaki (‘three men’) 
as their answers, with apa deliberately stressed to signal to the children as well to 
emphasize that they needed to fill the gap between the numeral and the noun. 
Only when the researcher was convinced that the participants had understood the 
task did the experimental session commence.

In the experimental session, children were asked to count the items that were 
presented on the Microsoft PowerPoint slide presentation one by one. The correct 
expected response was to count the items using the correct numeral classifier e.g., 
Tiga batang pensil (‘three NumCl pensil’). All the objects required Malay shape-
based numeral classifiers in the counting activity. The pictures in the experiment
are described in Table 4. Pictures were presented in a fixed random order to all the children. Unlike in the practice trial, the children were not prompted or doubly prompted with ape in the experimental session. The experimental session lasted about 20 min for each child.

Table 3  Description of practice trial items

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Sample response</th>
<th>Classifier required</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 rings</td>
<td>3 bentuk circon</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 hooks</td>
<td>3 bentuk mata kail</td>
<td>bentuk</td>
</tr>
<tr>
<td>3 wings</td>
<td>3 pasang sayap</td>
<td>pasang</td>
</tr>
<tr>
<td>3 glasses</td>
<td>3 pasang cermin mata</td>
<td>pasang</td>
</tr>
</tbody>
</table>

Table 4  Description of test stimuli

<table>
<thead>
<tr>
<th>Classifier tested</th>
<th>Classifier type</th>
<th>Stimulus</th>
<th>Typicality ranking</th>
<th>Correct response</th>
</tr>
</thead>
<tbody>
<tr>
<td>batang</td>
<td>1D: +rigid</td>
<td>3 pencils</td>
<td>Very typical</td>
<td>3 batang pensil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 trees</td>
<td>Typical</td>
<td>3 batang pokok</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 brooms</td>
<td>Moderate</td>
<td>3 batang penyapu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 rivers</td>
<td>Atypical</td>
<td>3 batang eungai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 roads</td>
<td>Very atypical</td>
<td>3 batang jalan</td>
</tr>
<tr>
<td>utas</td>
<td>1D: -rigid</td>
<td>3 necklaces</td>
<td>Very typical</td>
<td>3 utas rantai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 ropes</td>
<td>Typical</td>
<td>3 utas tali</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 strands of thread</td>
<td>Moderate</td>
<td>3 utas benang</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 watches</td>
<td>Atypical</td>
<td>3 utas jam tangan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 chain links</td>
<td>Very atypical</td>
<td>3 utas rantai kunci</td>
</tr>
<tr>
<td>keping</td>
<td>2D: +rigid</td>
<td>3 photographs</td>
<td>Very typical</td>
<td>3 keping gambar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 wooden planks</td>
<td>Typical</td>
<td>3 keping papan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 pieces of biscuits</td>
<td>Moderate</td>
<td>3 keping baskut</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 CDs</td>
<td>Atypical</td>
<td>3 keping CD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 pieces of cokos</td>
<td>Very atypical</td>
<td>3 keping kek</td>
</tr>
<tr>
<td>helai</td>
<td>2D: -rigid</td>
<td>3 pieces of papers</td>
<td>Very typical</td>
<td>3 helai kartas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 shirts</td>
<td>Typical</td>
<td>3 helai baju</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 leaves</td>
<td>Moderate</td>
<td>3 helai daun</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 handkerchiefs</td>
<td>Atypical</td>
<td>3 helai saputangan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 pairs of pants</td>
<td>Very atypical</td>
<td>3 helai seluar</td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>Classifier tested</th>
<th>Classifier type</th>
<th>Stimulus</th>
<th>Typicality ranking</th>
<th>Correct response</th>
</tr>
</thead>
<tbody>
<tr>
<td>buah</td>
<td>BD: big</td>
<td>3 buses</td>
<td>Very typical</td>
<td>3 buah bas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 boxes</td>
<td>Typical</td>
<td>3 buah kotak</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 families</td>
<td>Moderate</td>
<td>3 buah kalauage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 planets</td>
<td>Atypical</td>
<td>3 buah planet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 robots</td>
<td>Very atypical</td>
<td>3 buah robot</td>
</tr>
<tr>
<td>ketul</td>
<td>BD: medium</td>
<td>3 stones</td>
<td>Very typical</td>
<td>3 ketul batu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 chunks of meat</td>
<td>Typical</td>
<td>3 ketul dagang</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 drumsticks</td>
<td>Moderate</td>
<td>3 ketul paha ayam</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 gold ingots</td>
<td>Atypical</td>
<td>3 ketul jongkong</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>emas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 chocolates</td>
<td>Very atypical</td>
<td>3 ketul cocklat</td>
</tr>
<tr>
<td>biji</td>
<td>BD: small</td>
<td>3 rambutans</td>
<td>Very typical</td>
<td>3 biji rambutan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 balls</td>
<td>Typical</td>
<td>3 biji bola</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 oranges</td>
<td>Moderate</td>
<td>3 biji cren</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 cups</td>
<td>Atypical</td>
<td>3 biji cawan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 plates</td>
<td>Very atypical</td>
<td>3 biji pinggan</td>
</tr>
<tr>
<td>butir</td>
<td>BD: fine</td>
<td>3 grains of rice</td>
<td>Very typical</td>
<td>3 butir heras</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 stars</td>
<td>Typical</td>
<td>3 butir bintang</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 precious stones</td>
<td>Moderate</td>
<td>3 butir batu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>permata</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 seeds</td>
<td>Atypical</td>
<td>3 butir biji benih</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 grains of sugar</td>
<td>Very atypical</td>
<td>3 butir gula pasar</td>
</tr>
</tbody>
</table>

Coding of responses

Each of the children's production responses was classified into either one of the two main categories of correct or incorrect response, and if incorrect, their responses were further classified into three subcategories:

1. Correct response: responses were labeled as 'correct' if they were as described in Table 4 (Column 5).
2. Incorrect response: responses were labeled as 'incorrect' if they were different from those described in Table 4. This type of response was further divided into three subcategories:
   a. No classifier response: when no numeral classifier was used in the response, e.g., instead of 3 batang pensil ('3 stem' or [BD: + rigid] pencils), 3 pensil ('3 pencils') was used.
   b. Alternative classifier response: when an alternative numeral classifier was used instead of the correct numeral classifier, e.g., for 3 batang pensil ('3 stem' or [BD: + rigid] pencils), *3 buah pensil ('3 [BD: big] pencils') was substituted.
Appendices

- 298 -

c. Other response: when another response besides an alternative numeral classifier response was used, i.e.,

- Utilizers: when a Malay expression to quantify mass nouns was used, e.g., 3 genggam gula pasir ('3 handfuls of sugar') for 3 butir gula pasir ('3 grains of sugar'). Although this is grammatically correct, it is not a numeral classifier.

- Collective nouns: when a Malay expression to quantify multiple count nouns via a larger object was used, e.g., 3 tangkai daun ('3 stalks of leaves') for 3 helai daun ('3 pieces of [2D: -rigid] leaves'). Although this is grammatically correct, it is not a numeral classifier.

- Measure words: when an international measuring unit was used, e.g., 3 kilogram gula pasir ('3 kilograms of sugar') for 3 butir gula pasir ('3 grains of sugar'). Although this is grammatically correct, it is not a numeral classifier.

- Content words: when a noun, a verb or an adjective was used, e.g., 3 tajam batu ('3 'sharp stones') for 3 ketul batu ('3 pieces of stone').

RESULTS

Only a small number of correct shape-based numeral classifiers (5.25%) were produced by the youngest age group, i.e., the 6-year-olds. As expected, the mean number of correct responses increases with age: the 7-year-olds gave 18.20% correct responses, the 8-year-olds 41.53%, and the 9-year-olds 51.25%. Despite showing the highest percentage of correct responses among the children, the responses given by the 9-year-olds were markedly different from the proportion of correct responses given by the adults (77.63%). Figure 3 shows that the 6-year-olds only produced three correct numeral classifiers, i.e., biji [3D: small], batang [3D: -rigid], and helai [2D: -rigid], whereas the 7-year-olds produced all but one of the shape-based numeral classifiers tested, i.e., butir [3D: fine]. The 8- and 9-year-olds produced all the eight tested numeral classifiers.

In order to compare the number of correct responses for the different numeral classifiers in the different age groups and across gender, a mixed repeated measures analysis of variance (ANOVA), with numeral classifier and typicality type as within-subject factors, and age group (6-year-olds, 7-year-olds, 8-year-olds, 9-year-olds) and gender as between-subject factors, was conducted. There was a significant main effect of numeral classifier, F(7, 924) = 90.62; p < 0.001, partial η² = 0.407, a significant effect of typicality, F(4, 528) = 107.61; p < 0.001, partial η² = 0.449, and a significant main effect of age group, F(3, 132) = 105.59; p < 0.001, partial η² = 0.706. There was no significant effect of gender, p > 0.2. There was a significant interaction effect between numeral classifier and age group, F(21, 924) = 12.07; p < 0.001, partial η² = 0.215. No other effects were significant. Tukey's post-hoc analysis at α² = 0.01 revealed that correct responses for the different age groups were all significantly different from each other.

The mean numbers of correct responses for the eight shape-based numeral classifiers for the children are listed in Table 5. This also gives an indication of the
SALEHLODIN & WINSKEL: INVESTIGATION INTO MALAY NUMERAL CLASSIFIER

Figure 3  Mean number of correct numeral classifier responses produced by the different age groups for each shape-based numeral classifier

Malay shape-based numeral classifiers acquisition order based on the production task. Planned contrasts revealed that all numeral classifier pairs except buah and biji were significantly different: helai and batang, t(139) = 4.35, p < 0.001; batang and buah, t(139) = 3.45, p = 0.001; biji and utas, t(139) = 4.50, p < 0.001; utas and keping, t(139) = 2.08, p < 0.05; keping and ketul, t(139) = 3.65, p < 0.001; and ketul and butir, t(139) = 3.84, p = 0.005. The order of shape-based numeral classifier production shows that while one of the 2D classifiers was produced the earliest (helai), the other 2D classifier (keping) was only produced after the production of two 1D classifiers and two 3D classifiers.

From Fig. 4, it can be seen that there is a general trend in the number of correct responses from very typical > typical > moderate > atypical, very atypical exemplars across all age groups, which corresponds to the adult rating data.

Table 5  Stage of Malay shape-based numeral classifier acquisition based on the mean number of correct production responses across all children’s age groups (standard deviations are in parentheses)

<table>
<thead>
<tr>
<th>Numeral classifier</th>
<th>Mean number of correct responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>helai [2D: -rigid]</td>
<td>2.81 (2.11)</td>
</tr>
<tr>
<td>betang [1D: +rigid]</td>
<td>2.24 (1.42)</td>
</tr>
<tr>
<td>buah [3D: big]</td>
<td>1.86 (1.62)</td>
</tr>
<tr>
<td>biji [3D: small]</td>
<td>1.84 (1.48)</td>
</tr>
<tr>
<td>utas [1D: -rigid]</td>
<td>1.20 (1.64)</td>
</tr>
<tr>
<td>keping [2D: +rigid]</td>
<td>0.96 (1.23)</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td>0.02 (1.19)</td>
</tr>
<tr>
<td>butir [3D: fine]</td>
<td>0.33 (0.76)</td>
</tr>
</tbody>
</table>
In relation to the incorrect responses, it can be seen from Fig. 5 that the youngest children tended to give a 'No classifier' response, whereas with age the children tended to give an 'Alternative classifier' response. In total, 21 numeral classifiers were used as alternative substitution responses. Sixteen 'other responses' given by the children included 10 unitizers, e.g., gelas ('a glass of'), cubit ('a pinch of'), bungkus ('a packet of'); four collective nouns, e.g., kumpulan ('a group of'); banis ('a row of'); and two measure words, e.g., batu ('a mile'), kilo ('a kilogram/kilometer'); while the remaining 28 responses were nouns and adjectives, e.g., bingkai ('frame'), lurus ('straight').

Figure 4  Mean number of correct responses by each age group based on typicality

Figure 5  Mean number of the four response types used by each age group. 'Others' includes 'unitizers,' 'collective nouns,' 'measure words,' and 'content words'
As can be seen from Table 6, the classifiers *buah* [3D: big] (446 occurrences) and *biji* [3D: small] (422 occurrences) were the most frequently used alternative classifiers. The classifiers *keping*, *helai*, *batang*, *ketul*, and *utas* were used as alternative classifiers to a lesser extent. Other numeral classifiers used as alternative classifiers included *pasang* [arrangement: pair] (89 instances), *orang* [animate: human] (55), and *ekor* [animate: animal] (36). The classifiers *butir* (292), *ketul* (274), followed by *keping* (221) had the most substitutions made for them.

The numeral classifiers *buah* [3D: big] and *biji* [3D: small] were used predominantly in place of 3D classifiers; *buah* was mainly used in place of other 3D classifiers (51%) rather than replacing 1D (29%) or 2D (15%) classifiers; and *biji* was used predominantly in place of other 3D classifiers (77%) in contrast to 1D (8%) and 2D (15%) classifiers. The classifier *helai* [2D: +rigid] was used as an alternative classifier predominantly in place of *utas* [1D: +rigid] (48%) and *keping* [2D: +rigid] (39%). The classifier *batang* [1D: +rigid] was mainly used as a substitution response for *keping* [2D: +rigid] (61%).

A Spearman’s rank correlation coefficient was conducted between the number of correct numeral classifier responses and the frequency of occurrence of classifiers in the 150,000-word corpus. A moderate positive relationship was found between the number of correct classifier responses and frequency of occurrence of classifiers in the corpus, $r_s = 0.89$, $p$ (one-tailed) < 0.05. This indicates that there is a positive relationship between the frequency of occurrence of numeral classifiers as reflected in the 150,000-word corpus, and the production of correct numeral classifier responses in the enumeration task.

**DISCUSSION**

Malay numeral classifier production appears to be a relatively prolonged developmental process. The percentage of correct production responses by the preschool children in this study, i.e., the 6-year-olds, was very low (5.25%). The Grade 1 primary children, the 7-year-olds, scored only 18.20% correct and even the oldest children, the 9-year-olds, produced only 51.25% correct responses. In addition, there was a marked difference between the performance of the oldest children tested, the 9-year-olds, and that of the adults. The 6-year-olds only produced three correct numeral classifiers, i.e., *biji* [3D: small], *batang* [3D: +rigid], and *helai* [2D: +rigid], whereas the 7-year-olds produced all but one of the shape-based numeral classifiers tested, i.e., *butir* [3D: fine], and the 8- and 9-year-olds produced all eight numeral classifiers tested. In general, the youngest children tended not to use a numeral classifier in their responses, whereas the older children either used the correct numeral classifier or substituted another ‘alternative’ classifier for the correct classifier. The older children also still had a tendency to not use a numeral classifier in their responses, although this declined with age. The adults, in contrast, gave either the correct classifier response or used an alternative classifier.

Substitution and error responses can give us clues about the acquisition order within a semantic domain (Bernstein Ratner, 2000). The 3D classifiers *buah* [3D: big] and *biji* [3D: small] were the most frequently used ‘alternative’ classifiers, whereas
<table>
<thead>
<tr>
<th>Correct classification</th>
<th>Alternative classification</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>buah [3D: big]</td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>by [3D: small]</td>
<td></td>
<td>68</td>
</tr>
<tr>
<td>keping [2D: rigid]</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>ketul [3D: medium]</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ketul [3D: small]</td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>keji [2D: split]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>lele [2D: split]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>mar [3D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [2D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: split]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: medium]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: long]</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>marang [1D: small]</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
the classifiers _butir_ [3D: fine] and _ketul_ [3D: medium] had the most substitutions made for them. These substitution responses roughly correspond to the production order found in the enumeration task, as both _buah_ and _biji_ were found to be produced earlier than _butir_ or _ketul_ (Table 5).

The production of correct responses in the current study did not follow the predicted order based on semantic complexity, i.e., that numeral classifiers with fewer features, 1D and 2D classifiers (two forms each), are acquired earlier than 3D classifiers with more features (four forms) (Clark, 1972, 1973; Maddox, 1992). The actual overall production order was more complex than this predicted order: _helai_ [2D: -rigid] > _batang_ [1D: +rigid] > _buah_ [3D: big]; _biji_ [3D: small] > _utas_ [1D: -rigid] > _keping_ [2D: +rigid] > _ketul_ [3D: medium] > _butir_ [3D: fine] (as illustrated in Table 5). The initial order that a 2D and a 1D numeral classifier were produced prior to a 3D classifier corresponds to the predicted order, i.e., _helai_ and _batang_ are produced prior to _buah_ and _biji_; however, the later order is not consistent with the prediction made, as the 3D classifiers _buah_ and _biji_ were produced prior to the 1D and 2D classifiers _utas_ and _keping_. Clearly, we need to examine additional contributing factors that affect production order.

Input frequency has been found to be an important factor in the acquisition of Japanese classifiers (Matsumoto, 1985; Yamamoto, 2005). The analysis of numeral classifier frequency from the Malay children’s corpus showed that _buah_ [3D: big], _batang_ [1D: +rigid], and _biji_ [3D: small] were the most frequently occurring shape-based numeral classifiers (Table 2). The relatively high frequency of the two 3D classifiers, _buah_ and _biji_, in comparison with the other shape-based numeral classifiers (Table 5) could explain the relatively early emergence of these 3D numeral classifiers in comparison with the 2D and 1D classifiers in the enumeration task.

As formal instruction can also play a role in later language development (Nippold, 1988), we can also expect the formal teaching of numeral classifiers, which commences in Grade 1 primary school, to have an impact on numeral classifier production responses. Hence we can expect school-aged children, the 7-year-olds, to be more proficient at producing numeral classifiers than the preschool children, i.e., the 6-year-olds, and expect this to be reflected in the higher rate of correct production responses of the primary school children. There was an improvement in correct production responses of the 7-year-olds in comparison to the 6-year-olds. Furthermore, the 6-year-olds produced only three correct numeral classifiers, whereas the 7-year-olds produced all but one of the classifiers. With age, there was a steady increase in performance in the correct production of numeral classifiers in the enumeration task. In general, the more frequently used numeral classifiers are taught earlier in primary school than the less frequently used classifiers. The significant increase in usage of the classifiers _batang_ and _buah_ between the preschool children and the Grade 1 children, as well as the significant increase in _utas_ and _butir_ between the Grade 2 and Grade 3 children, correspond with the fact that the classifiers _batang_ and _buah_ were introduced in Grade 1 workbooks while _utas_ and _butir_ were introduced in Grade 3 workbooks. In sum, it appears that the formal instruction of numeral classifiers plays an important role in the production of Malay shape-based numeral classifiers, in particular the less commonly used classifiers, e.g., _ketul_ and _butir._
Malay does not have a general numeral classifier that can acceptably be substituted semantically for other numeral classifiers, in contrast to for example the Thai classifier រស. However, the Malay children did tend to use buah [3D: big] and biji [3D: small] for this purpose, which suggests that these two numeral classifiers are relatively accessible to the younger children. Furthermore, these substitution responses correspond to the high frequency of these two numeral classifiers found in the Malay children's corpus, suggesting that input frequency does play a role in the production of numeral classifiers.

Children were also found to commonly substitute a 'within-the-same-dimension' numeral classifier or a classifier that shared one of the features with the correct classifier response. The classifiers buah [3D: big] and biji [3D: small] were frequently substituted for other 3D classifiers, which concurs with results found on Cantonese and Mandarin Chinese (Loke, 1991; Mak, 1991). In addition, children were also found to substitute an alternative classifier that varied in just one of the features, either dimensionality or rigidity. For example, the classifier helai [2D: +rigid] was used as an alternative classifier predominantly in place of utas [1D: +rigid] and kepang [2D: +rigid], and the classifier batang [1D: +rigid] was mainly used as a substitution response for kepang [2D: +rigid]. These substitution errors suggest that children are having difficulty in making finer categorization distinctions between 1D and 2D objects, rigid and nonrigid objects, as well as relative size judgments of 3D objects into fine, small, medium, or big categories.

An important consideration is that the categorization of some objects into particular numeral classifier categories is fairly straightforward, whereas the categorization of other objects is not always clear-cut as there is a degree of flexibility or choice in how some objects are classified. Items can be paired with different classifiers depending on their shape, appearance or presentation, as categorization is dependent on how the item is perceived by the speaker. For example, coklat ('chocolate') is paired with the classifier ketul [3D: medium] if it is perceived as chunky, but if it is perceived as small in size it may be collocated with the classifier biji [3D: small]. However, if the chocolate is long (finger-like), it will likely be used with batang [1D: rigid]. Hence, it can be seen that there can be quite a high degree of variation between how some items are sorted into different categories by different speakers.

The categorization of an object involves the retrieval and evaluation of stored information on related objects, and only when there is a satisfactory resemblance between new and stored information is an object accepted into the respective category (Barsalou et al., 1998). Some objects are readily categorized into a particular numeral classifier category, whereas other objects are more difficult to categorize as they can be placed into several different categories based on sometimes quite fine perceptual distinctions. As categorization within the Malay numeral classifier system involves building up a comprehensive knowledge of complex inherent semantic characteristics of the different numeral classifier categories, it takes an extended period of time for children to acquire. In conclusion, this study shows that the production pattern of Malay numeral classifiers is a complex and relatively prolonged process (which continues into the school years) with multiple factors affecting production order, including semantic complexity, input frequency, and the formal teaching of numeral classifiers.
REFERENCES


Appendices

SALEHUDDIN & WINSKEL: INVESTIGATION INTO MALAY NUMERAL CLASSIFIER


ADDRESS FOR CORRESPONDENCE
Khairiyati Salehuddin
MARCS, Auditory Laboratories, University of Western Sydney, Bankstown Campus,
Locked Bag 1797, Penrith South DC, NSW 1797, Australia
E: K.salehuddin@uws.edu.au

Downloaded from http://rsc.oxfordjournals.org at University of Western Sydney on July 13, 2006