RECOGNISING THE SOCIOLINGUISTIC REALITY OF SPOKEN INDONESIAN: A CORPUS AND USAGE ANALYSIS OF A MIDDLE DIGLOSSIC VARIANT

HALIM NATAPRAWIRA

BA, Dip Ed, MA

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of the Sunshine Coast

Faculty of Arts, Business and Law

Submitted 12 June 2017
Abstract

The purpose of this research was to contribute to the learning and teaching of Indonesian as a foreign language by defining Spoken Indonesian (SI), and incorporating it in a pedagogic design in order to increase Australian students’ proficiency in Indonesian as the target language.

Indonesian language is one of the Asian languages taught in Australian schools and universities. The Australian Government’s 2012 White Paper, Australia in the Asian Century, identifies the learning of Asian languages as a key strategic component in aligning Australia’s economic and security interests with the countries of the Asian region.

As a step in the direction outlined in the White Paper, this research undertakes a linguistic and sociolinguistic study of the everyday, informal, spoken Indonesian language. The sociolinguistic epistemology of this research uses the theory of diglossia. In classical diglossia theory, a diglossic language consists of a High (H) and a Low (L) variant. Framing the research methods within the diglossia concept allows for sociolinguistic and linguistic recognition of different language variants that exist within the universe of the Indonesian language (bahasa Indonesia). As such, a description and function of each language variant can be constructed. The H variant of bahasa Indonesia is Formal Indonesian (FI). FI is what is taught as the target language in Australia. The L diglossic variant is Colloquial Indonesian (CI). CI is little understood and is therefore generally not included in Indonesian language pedagogy.

The FI model that is taught as the target language follows a prescriptive language teaching model that is appropriate for formal language settings. A misalignment occurs when FI is also presented as an everyday informal language. The premise of this research is that, rather than FI being spoken as the everyday informal language, SI, a blend of FI and CI, is spoken as the everyday language. Based on this assertion, the ‘M hypothesis’ is constructed. The ‘M’ here refers to the Middle diglossic variant, a merging of the H and L diglossic variants.

This research uses a body of linguistic data (corpus) to assess the SI phenomenon. To ascertain that SI is the informal spoken language, the linguistic data of the corpus must contain the language features of CI. The grammar of FI is the standard Indonesian grammar presented in textbooks and dictionaries. At the writing of this thesis, there exists no prescribed grammar reference for CI, although various isolated treatises on different aspects of CI exist, most of them academic. Thus, this research sets out to provide a description of CI grammar by identifying linguistic features in the corpus that are non-FI,
using precursory information provided by previous academic research on CI. This description of CI linguistic features encompasses lexis, phonology, morphology and semantics.

The aim of this research was threefold:

1. To ascertain that Indonesian is a diglossic language.
2. To test the M hypothesis that CI is present in the corpus of informal language alongside FI, and to ascertain that SI is the everyday spoken language.
3. To form a body of information from which a reference resource can be developed, and to design a pedagogic model for the learning and teaching of SI.

The research found that CI linguistic features are prevalent and are a significant part of the spoken Indonesian language. This finding has two implications. First, it means that the teaching of Indonesian language needs to be redefined in terms of diglossia. Second, in practical terms, the diglossic nature of Indonesian exposes the need for a new approach to the pedagogic design of Indonesian language. This thesis proposes a notional–functional syllabus model in which unit models are arranged into topics. These topics can be organised to correspond with the formal and informal themes of Indonesian diglossia and thus the High and Low diglossic variants are represented in their respective sociolinguistic settings.
# Table of Contents

Abstract ii

Table of Contents iv

List of Figures vii

List of Tables vii

List of symbols and abbreviations viii

Statement of original authorship x

Acknowledgments xi

Chapter 1. A Research Paradigm for Spoken Indonesian 1

1. Introduction: The scope of this research 2
   1.1 Indonesian sociolinguistic research 3
   1.2 Terms and construct in this thesis 4
   1.3 Research aims and significance 5
   1.4 The research question: Framing the hypothesis 6
   1.5 The data 6
   1.6 Methodological approach and research methodology 7
   1.7 The practical application of research 8
   1.8 Organisation of the thesis 8
   1.9 Summary 10

Chapter 2. The Evolution of Spoken Indonesian 11

2. Introduction 12
   2.1 Diglossia 12
      2.1.1 The evolution of Indonesian diglossia 18
   2.2 A brief history of bahasa Indonesia 19
      2.2.1 Early origins 19
      2.2.2 Philology and etymological examples 21
      2.2.3 Classical Malay 24
      2.2.4 Colonial language planning 25
      2.2.5 Bahasa Indonesia and nationalism 27
      2.2.6 Post independence 30
      2.2.7 Government language planning 31
      2.2.8 Spelling systems using the Latin alphabet 35
   2.3 The sociolinguistic functions of Indonesian diglossic variants 38
      2.3.1 ‘Weak’ diglossia 39
Chapter 3. Methodology and Description of the Data: Qualitative and Quantitative Corpus Analysis of Spoken Indonesian

3. Introduction
3.1 The taxonomy, terminology and coding of Indonesian diglossia
3.1.1 Confused terminology
3.2 The SI spectrum
3.3 Data samples
3.4 Methods and sampling
3.5 Data analysis
3.5.1 Method 1 – Establishing whether CIAL is present in SI corpora
3.5.2 Method 2 – Qualitative method: IPA – identifying and isolating non-FIAL, and constructing CIAL
3.5.3 Method 3a – Quantitative method: CIAL:SI ratio
3.5.4 Method 3b – Quantitative method: CIAL usage across Indonesia
3.6 Conceptualising the interrelationship between FI and CI in SI
3.6.1 Register
3.7 The Indonesian phonemes
3.7.1 The IPA system applied to Indonesian
3.7.2 The Indonesian vowels
3.7.3 The Indonesian consonants
3.8 Conclusion

Chapter 4. The Linguistic Components and Sociolinguistic Nature of Spoken Indonesian: Findings and Discussion of Corpus Analysis

4. Introduction
4.1 Findings ad discussions of the corpus analysis
4.1.1 Qualitative results: CIAL in SI corpus and description with the International Phonetic Alphabet
4.1.2 Quantitative results: CIAL:SI
4.2 The sociolinguistic implications of the findings
4.3 Conclusion

Chapter 5. The Frequency and Demographic Range of Spoken Indonesian: Findings and Discussion
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>Introduction.</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>The demographic reach of CI: Questionnaire.</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Jakarta data: Analysis and discussion.</td>
<td>84</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Comments by participants from Jakarta.</td>
<td>87</td>
</tr>
<tr>
<td>5.3</td>
<td>Regional Indonesia data: Analysis and discussion.</td>
<td>89</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Comments by participants from regional Indonesia.</td>
<td>93</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusion.</td>
<td>96</td>
</tr>
<tr>
<td>6.</td>
<td>Teaching Spoken Indonesian – the Practical Application of this Research</td>
<td>98</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction.</td>
<td>99</td>
</tr>
<tr>
<td>6.2</td>
<td>Towards the teaching of SI.</td>
<td>99</td>
</tr>
<tr>
<td>6.3</td>
<td>The state of Indonesian language teaching in Australia.</td>
<td>100</td>
</tr>
<tr>
<td>6.4</td>
<td>Prescriptive grammar versus descriptive grammar.</td>
<td>101</td>
</tr>
<tr>
<td>6.5</td>
<td>A review of language content in Indonesian L2 textbooks.</td>
<td>104</td>
</tr>
<tr>
<td>6.6</td>
<td>Comparing the CI content in Indonesian language textbooks with corpus data.</td>
<td>107</td>
</tr>
<tr>
<td>6.6.1</td>
<td>The descriptive-pragmatic process: Defining mainstream SI.</td>
<td>109</td>
</tr>
<tr>
<td>6.7</td>
<td>Contextualising diglossia in the Notional-Functional Syllabus Model.</td>
<td>111</td>
</tr>
<tr>
<td>6.8</td>
<td>Conclusion.</td>
<td>114</td>
</tr>
<tr>
<td>7.</td>
<td>Conclusion.</td>
<td>115</td>
</tr>
<tr>
<td>7.1</td>
<td>Overview.</td>
<td>116</td>
</tr>
<tr>
<td>7.2</td>
<td>Problems encountered.</td>
<td>117</td>
</tr>
<tr>
<td>7.3</td>
<td>Future research.</td>
<td>119</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Final remarks.</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>References.</td>
<td>121</td>
</tr>
<tr>
<td>Appendix 1.</td>
<td>SI Corpus.</td>
<td>128</td>
</tr>
<tr>
<td>Appendix 2.</td>
<td>Semantics and Pragmatics of CI Lexis and Collocations.</td>
<td>168</td>
</tr>
<tr>
<td>Appendix 3.</td>
<td>Extended Discussions on CI Linguistic Features.</td>
<td>215</td>
</tr>
<tr>
<td>Appendix 4.</td>
<td>Some Issues of English L1 Interference in Learning Indonesian.</td>
<td>229</td>
</tr>
<tr>
<td>Appendix 5.</td>
<td>The Influences of Regional Languages on CI.</td>
<td>237</td>
</tr>
<tr>
<td>Appendix 6.</td>
<td>The Use of CI in News Reporting.</td>
<td>251</td>
</tr>
<tr>
<td>Appendix 7.</td>
<td>Research Instrument: Survey.</td>
<td>255</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. The synthesis and composition of Spoken Indonesian……………………………………..38
Figure 2. The research methodology………………………………………………………………….54
Figure 3. The merging of H and L into the M variant (ΔFI U ΔCI = ΔSI)…………………………….57
Figure 4. The formal-informal spectrum of FI–SI–CI……………………………………………….58
Figure 5. CIA:SI ratios for all data sets……………………………………………………………..76
Figure 6. The spread of data in the FI-SI-CI sociolinguistic spectrum…………………………..77
Figure 7. Indonesian diglossia: σδ as a function of Aλ……………………………………………….78
Figure 8. Traditional Fergussonian diglossia: Aλ as a function of σδ…………………………….79
Figure 9. The intervallic scale of the research instrument…………………………………………84
Figure 10. The descriptive-pragmatic process…………………………………………………….109

List of Tables

Table 1. Fergussonian Diglossia……………………………………………………………………13
Table 2. Comparison of Personal Pronouns Between Indonesian and Cham……………………..22
Table 3. Comparison of Possessive Use of Personal Pronouns in Three Languages…………….23
Table 4. Data Samples in the SI Corpora…………………………………………………………52
Table 5. Mixed Methods Application to SI…………………………………………………………53
Table 6. The Indonesian Vowels……………………………………………………………………61
Table 7. The Phonetic Table of Indonesian Vowels…………………………………………………62
Table 8. The Indonesian Consonants………………………………………………………………64
Table 9. The Phonetic Table of Indonesian Consonants…………………………………………65
Table 10. Correspondence Analysis of All Data Sets……………………………………………77
Table 11. Indonesian Diglossia……………………………………………………………………80
Table 12. FI and CI Usage by the Jakartan Participants……………………………………….85
Table 13. Summary Statistics on Language Use in Jakarta Area…………………………….86
Table 14. FI and CI Usage Among Participants From Regional Indonesia……………………89
Table 15. Summary Statistics on Language Use in Regional Areas……………………………92
Table 16. A Comparison Between CI Content in Textbooks and the Corpus……………….108
Table 17. The Prescriptive Sieve: Filtering Non-Mainstream Colloquialism……………….111
List of symbols and abbreviations

The Glossary of Indonesian Diglossia

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL (Aλ)</td>
<td>Aspect of Language / linguistic features</td>
</tr>
<tr>
<td>aux</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>AV</td>
<td>Audiovisual</td>
</tr>
<tr>
<td>Bahasa Baku</td>
<td>Lit. ‘Proper Language’. The term for the convention of Indonesian grammar that is officially considered correct by the Indonesian language authority.</td>
</tr>
<tr>
<td>Bahasa Gaul</td>
<td>‘Teen language’. A subset of CI.</td>
</tr>
<tr>
<td>Bahasa Sehari-hari</td>
<td>A term, albeit not official, that is used to refer to the informal everyday language (CI/SI).</td>
</tr>
<tr>
<td>BB</td>
<td>Billboard</td>
</tr>
<tr>
<td>BD</td>
<td>Bahasa Daerah (regional language).</td>
</tr>
<tr>
<td>BI</td>
<td>Bahasa Indonesia. Lit. ‘Indonesian Language’. The official name for the Indonesian language.</td>
</tr>
<tr>
<td>BIPA</td>
<td>Bahasa Indonesia untuk Penutur Asing (Indonesian for foreign learners).</td>
</tr>
<tr>
<td>CI</td>
<td>Colloquial Indonesian</td>
</tr>
<tr>
<td>CJI</td>
<td>Colloquial Jakartan Indonesian</td>
</tr>
<tr>
<td>CIAL</td>
<td>CI Aspect of Language / linguistic features of CI.</td>
</tr>
<tr>
<td>cop</td>
<td>Copula</td>
</tr>
<tr>
<td>DET</td>
<td>Determiner</td>
</tr>
<tr>
<td>Diglossia</td>
<td>a linguistic situation within a language were there exists a High variant and a Low variant.</td>
</tr>
<tr>
<td>DP</td>
<td>Discourse particle</td>
</tr>
<tr>
<td>D-P</td>
<td>Descriptive-pragmatic</td>
</tr>
<tr>
<td>Δ</td>
<td>Diglossic variant</td>
</tr>
<tr>
<td>FI</td>
<td>Formal Indonesian</td>
</tr>
<tr>
<td>FIAL</td>
<td>FI Aspect of Language; the linguistic features of FI.</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>H</td>
<td>High diglossic variant</td>
</tr>
<tr>
<td>IPA</td>
<td>International Phonetic Alphabet</td>
</tr>
<tr>
<td>(\sigma\delta)</td>
<td>Informal sociolinguistic domain</td>
</tr>
<tr>
<td>JABOTABEK</td>
<td>Jakarta Bogor Tanggerang Bekasi</td>
</tr>
<tr>
<td>KBBI</td>
<td>Kamus Besar Bahasa Indonesia (Official Indonesian Language Dictionary)</td>
</tr>
<tr>
<td>L</td>
<td>Low diglossic variant</td>
</tr>
<tr>
<td>Lit</td>
<td>Literature</td>
</tr>
<tr>
<td>L1</td>
<td>First language</td>
</tr>
<tr>
<td>L2</td>
<td>Second language learner / Second language learning</td>
</tr>
<tr>
<td>M</td>
<td>Middle diglossic variant</td>
</tr>
<tr>
<td>MM</td>
<td>Mixed Methods</td>
</tr>
<tr>
<td>neg</td>
<td>Negator</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>(\emptyset)</td>
<td>Null-Element / ellipsis</td>
</tr>
<tr>
<td>PIM</td>
<td>Pre-Indonesian Malay</td>
</tr>
<tr>
<td>prep</td>
<td>Preposition</td>
</tr>
<tr>
<td>pro</td>
<td>Pronoun</td>
</tr>
<tr>
<td>RCTI</td>
<td>Rajawali Citra Televisi</td>
</tr>
<tr>
<td>RI</td>
<td>Recording of interviews</td>
</tr>
<tr>
<td>SCTV</td>
<td>Surya Citra Televisi</td>
</tr>
<tr>
<td>SI</td>
<td>Spoken Indonesian</td>
</tr>
<tr>
<td>(\sigma\delta)</td>
<td>Sociolinguistic domain</td>
</tr>
<tr>
<td>TC</td>
<td>Target Culture</td>
</tr>
<tr>
<td>TL</td>
<td>Target Language</td>
</tr>
<tr>
<td>TPI</td>
<td>Televisi Pendidikan Indonesia</td>
</tr>
<tr>
<td>TVRI</td>
<td>Televisi Republik Indonesia</td>
</tr>
<tr>
<td>VP</td>
<td>Verb Phrase</td>
</tr>
<tr>
<td>(wh)-</td>
<td>Question word</td>
</tr>
<tr>
<td>X(^a)</td>
<td>The Indonesian base word, the uninflected form of a lexical item</td>
</tr>
</tbody>
</table>
Statement of original authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature:

Date: 12 June 2017

Statement of editorial assistance

Professional editors Dr Leigh Findlay (TrueNature Writing & Editing) and Glenine Hamlyn (Envision Editing) provided copyediting and proofreading services, according to the guidelines laid out in the university-endorsed national Guidelines for editing research theses.
Acknowledgements

First and foremost, my sincerest gratitude goes to my supervisors: Phillip Mahnken, Michael Carey and Richard Curtis. They have believed in my ideas, supported me in my research, and guided me to complete this thesis. My candidature has been part-time and the enthusiasm that I started with waned towards the end as emotionally and mentally I found myself depleted. This, together with the demands of fulltime work, created periods of stress, panic and a sense of helplessness. Thankfully, somebody in the faculty administration at the University of the Sunshine Coast recognised the situation and in their wisdom intervened and provided the support mechanism to get the thesis back on track. My supervisors also deserve great respect for their patience and tolerance in dealing with my penchant for complication, esoteric concepts and semiotic abstractions; they kindly steered me away from occult academia so that I could deconstruct my ideas, and deliver and communicate them more simply. Thus this thesis is the more logical, meaningful and substantial because of the contributions of my supervisors.

The next person to whom a special thank you is owed is Birgit Stein, my lifelong friend, who has been most supportive, understanding and accommodating to the needs and demands of this journey. Indeed, many amazing people have shown interest and goodwill in helping and supporting me in their various ways. They include family, friends, survey participants, colleagues, managers and employers. Some have volunteered to proofread, others offered help with formatting or allowed me leave so that I could work on this thesis. Trying to list every individual who contributed would be impossible; I am bound to omit names that are deserving of acknowledgement. Suffice to say I am grateful to everybody who has directly or indirectly, in one way or another, been a part of this journey.
Chapter 1

A Research Paradigm for Spoken Indonesian
1. Introduction: The scope of this research

This research has three objectives. First, it aims to establish the linguistic features of Colloquial Indonesian (CI). Second, the research aims to use the linguistic features of CI to provide a sociolinguistic description of the spoken informal Indonesian language, referred to in this thesis as Spoken Indonesian (SI). Third, the practical aim of this research is to design a pedagogic model for Indonesian language that incorporates CI.

Indonesian language is in a state of diglossia (Errington, 1986; Sneddon, 2002, 2003a). Ferguson (1959) defined diglossia as a sociolinguistic situation in which a High variant (H) and a Low variant (L) exist. In the diglossia of Indonesian, the H variant is the standard official language and bahasa Baku is the term used for its proper grammar. In this thesis the term Formal Indonesian (FI) is used for the H standard Indonesian. The L variant is the colloquial spoken language (CI). The key hypothesis tested in this research was that CI and FI have merged to form an everyday, informal spoken language. This merging results in a Middle (M) diglossic variant, SI. There are no official terms for this spoken language, but it is sometimes referred to as bahasa sehari-hari (everyday language).

The practical objective of this research was to create a pedagogy for Indonesian language that incorporates CI as the main component of SI in learning materials for foreign learners of Indonesian. Sneddon (2003b) has stated that many second language L2 learners of Indonesian, specifically Australian students, often find themselves ill-equipped to converse in Indonesian despite having learnt the language for many years. Sneddon attributed this to the fact that, for most part, only the H variant FI is taught in schools and universities.

The research data consists of corpora gathered from authentic everyday situations, particularly informal conversations. Contrary to the findings of this research, pedagogies for Indonesian as a foreign language often portray FI as the language that is spoken across all sociolinguistic domains, including informal conversational settings (Sneddon, 2001, 2006; Djenar 2006). Based on the linguistic findings of this research, this thesis demonstrates why and how CI should be incorporated in pedagogies for Indonesian language because CI is an important element of Spoken Indonesian.
1.1 Indonesian sociolinguistic research

The study of the sociolinguistics of Indonesia is a complex area that involves linguistic interaction between Indonesian, as the national lingua franca, and numerous regional dialects and languages. The focus of this thesis is the sociolinguistic variety within Indonesian itself. This focus entails exploring the historical linguistic development that has led to the current linguistic situation of Spoken Indonesian, including the influence of regional and other languages (Goebel, 2013).

Although a second-language lingua franca for the majority of Indonesians scattered throughout Indonesia’s vast archipelago, Indonesian is the mother tongue for most of the approximately 14 million inhabitants of greater Jakarta, Indonesia’s capital. Errington (1986) used the diglossia paradigm to describe the Jakartan Indonesian (JI) dialect as an L. Sneddon (2006) referred to Jakartan Indonesian as Colloquial Jakartan Indonesian (CJI) and also described it as the L diglossic variant. More recent studies use the term Colloquial Indonesian (Djenar & Ewing, 2015) instead of JI or CJI, which is the term used in this thesis. Although this L variant originated in Jakarta, its scope and usage has extended far beyond the capital and thus warrants the name Colloquial Indonesian (CI).

The sociolinguistic theory of diglossia defines an H diglossic variant and an L diglossic variant. Sneddon (2003a) identified FI as the H diglossic variant and CI as the L diglossic variant. This thesis continues the study of Indonesian diglossia and uses it as the epistemological framework for the research. Using the conceptual framework of diglossia, this thesis qualifies the informal spoken language CI as a legitimate language variant that allows for linguistic description and categorisation in juxtaposition to FI. Sneddon (2003b) also raised the notion that FI and CI could merge into an M diglossic variant. This postulation framed the M hypothesis of this research, which was to ascertain by means of qualitative and quantitative analysis if such an M diglossic variant (SI) has indeed emerged.

Currently the general approach towards CI by educators and language authorities is that it is a linguistic variant that does not warrant inclusion in pedagogy. Nonetheless, study of CI has increased, primarily by academics based outside Indonesia. Djenar (2006, 2008) and Djenar and Ewing (2015) have written on aspects of CI and its uses in youth literature. Similarly, Smith-Hefner (2007) has written on sub-variants of CI that are used by youth and the middle class. Boellstorff (2004) has provided descriptions of CI features and so has Kushartanti (2014) who has also performed extensive researched on CI language acquisition by Indonesian children. Manns (2014) completed an observational study on the use of CI in youth radio programs outside Jakarta.
SI contains both FI and CI. FI provides much of the modern and technical vocabulary while CI is the spoken component of SI. The natural spoken element of SI gives the language its conversational quality. Therefore, a linguistic description of SI necessitates a description of CI.

The ‘standard’ Indonesian language, known in the native vernacular as bahasa Indonesia, differs significantly from the language that is actually spoken in everyday, informal interaction. CI is the ‘non-standard’ language in the official sense, but in terms of informal, everyday language, the linguistic features of CI can be regarded as ‘mainstream’ because CI in an integral part of SI.

The significant difference between FI and CI is more than a difference of register in terms of colloquial choice of vocabulary and expressions. There are fundamental differences in the grammar and in the aspects of language (linguistic features): the lexis, phonology, morphology and semantics. Even when the root of a word – more commonly termed in Indonesian language grammar as base word – is shared between FI and CI, the differences in aspects of language can realise very distinct lexical forms. For example, the base word *dengar* (to listen/to hear) as an active verb inflects as *mendengarkan* in FI but as *ngedengerin* in CI. Expressing degrees, such as ‘too big’, is *terlalu besar* in FI but is *kegedean* in CI. The base word *buat* (the verb ‘to make’) in FI has a different word class function in CI (the preposition ‘for’). These differences in meaning would be difficult to understand for students who have only learnt FI. Moreover, educated Indonesians switch between codes seamlessly as the social situation requires. Indonesians go about their daily lives speaking SI, but when they find themselves in a formal situation, such as when talking to authority, they can instantaneously switch to FI.

1.2 Terms and constructs in this thesis

The paradigm of this thesis’s discourse is the Indonesian sociolinguistic diglossia phenomenon. Three sets of diglossic variants form the dialectics of discussion:

- Formal Indonesian (FI), the H diglossic variant (bahasa Baku)
- Spoken Indonesian (SI), the M diglossic variant
- Colloquial Indonesian (CI), the L diglossic variant.

FI is the formal–official language, while SI is the everyday spoken language. In the formal–informal spectrum, CI is the most informal–colloquial form. CI is a direct linguistic descendant from Betawi.
Malay, the native Malay dialect originally spoken by the Betawi ethnic group of Jakarta. This dialect is no longer widely spoken, although it is spoken in certain pockets of Betawi communities in and around Jakarta. While the prospect lingers that the Betawi Malay dialect may diminish in usage and possibly become extinct, its linguistic features – the lexicon, phonology, morphology, speech acts and semantic properties – has survived and is continuing as a part of CI. Traditional Fergussonian diglossia (two languages) consists of H and L language variants. The emergence of a third middle M variant is a phenomenon when a merging of H and L variants occurs. The term triglossia has not become a common conventional term in theoretical discussions; in the case of the M variant as a merging of H and L variants, the M variant is not an independent variable but rather a dependent variable conditional on the existence of the H and L variants. Moreover, data in this thesis suggests that a pure form of CI no longer exists; all informal conversation contains FI. FI provides a modern and technical lexis that cannot be divorced from conversation about technology and modern abstract concepts that are part of everyday life. However, FI is still used in formal domains without any incursions by CI. In the informal domains, the linguistic features of both FI and CI are integral to SI. Thus Indonesian diglossia in practical terms is still a diglossia consisting of FI and SI. Nevertheless, in this thesis CI is categorised as an independent variant in order to isolate and identify its features. In reality, the Indonesian linguistic landscape is beyond diglossia or even triglossia – it is ‘multiglossic’: the ‘Indonesian diglossia’ is interwoven with regional languages, many of which are diglossic, triglossic or multiglossic themselves.

1.3 Research aims and significance
The significance of this research is that it works towards developing a comprehensive linguistic description of CI, as well as designing a pedagogical model for delivering the teaching of SI.

Once CI is understood in terms of its features, structure and grammar, then SI can be included in the pedagogies for Indonesian language. Consequently, teachers of Indonesian language will be able to use authentic, contemporary Indonesian resources and incorporate authentic spoken Indonesian language into the classroom.

The aims of this research were threefold:

1. To study the Indonesian L diglossic variant CI and, as such, to provide a description of the linguistic features of CI.
2. To use a mixed methods approach to analyse a corpus of authentic, informal spoken language to test the M hypothesis, in order to establish if an M diglossic variant (SI) exists.

3. To provide a pedagogic model for the teaching of CI linguistic features in the context of the SI diglossia.

Aim 2, testing the M hypothesis, validates the purpose of researching CI linguistic features. If the hypothesis is correct, then CI linguistic features form an integral part of SI and function as the informal spoken language, thus supporting the case for including CI in Indonesian language pedagogy to enable authentic language-teaching practice.

The research reported in this thesis continues the tradition of studying the Indonesian language within the sociolinguistic diglossic framework and explores how spoken Indonesian has evolved into its current form. Researchers of Indonesian sociolinguistics have identified H and L diglossic Indonesian variants (Djenar, 2006; Errington, 1986; Sneddon, 2003a), and have raised the prospect that a merging of the H and L variants into an M variant might take place (the M hypothesis) (Sneddon, 2003b).

1.4 The research question: Framing the hypothesis

The central questions driving this research concern the diglossic situation of the Indonesian language. If Indonesian is a diglossic language, then the corpus data will show that an informal, everyday spoken language exists, which differs linguistically and grammatically to formal–standard Indonesian. The premise of this research is that the FI has merged with CI into a middle diglossic variant M (SI) that functions as an everyday spoken language. The focal point of this research is testing the M hypothesis. If such an M spoken diglossic variant exists, then the research question is ‘Are CI linguistic features an integral part of everyday spoken Indonesian (SI)?’ In terms of an alternative hypothesis, the question is framed as $H_1$: CIAL:SI >0.

1.5 The data

The main linguistic data comprises a corpus ($n = 14000+$ words) from a variety of data sets ($n = 48$) from everyday, informal settings. The corpus of informal spoken language is drawn from consists of dialogues of everyday social interactions that include recorded interviews, contemporary literature, advertisements, TV shows and films from the 1970s to 2010s.
A second set of data was generated from a research instrument (survey) to collect data about the demographic use of the CI language. The survey was distributed to 98 native Indonesian speakers and contained multiple questions about the use of FI and CI phrases. The participants were given everyday informal scenarios and asked to select which phrase they considered appropriate to use in particular interactions. The demographic survey also collected data about age, gender, education, profession and place of residence.

1.6 Methodological approach and research methodology

First, a qualitative description of the corpus was performed using the International Phonetic Alphabet (IPA) to analyse the lexis, phonology, morphology and semantic properties of CI lexis and speech acts. Next, a qualitative analysis was performed to isolate all the lexicon and linguistic features in the SI corpus that were not part of standard bahasa Baku or FI. These isolated items were identified as the linguistic features that comprise the grammar and lexicon of CI. This analysis produced a CI data set that was coded as the Colloquial Indonesian Aspect of Language (CIAL).

Next, the results of the above analysis – the CI linguistic features in the corpus – were quantitatively analysed. The CIAL count from the corpus has a numeric value, that is, the number of individual CIAL items in each data set. Every CIAL item is counted across all data sets and measured as a ratio of CIAL:SI. The presence of CIAL in the corpus was used to test the alternative hypothesis that CI is present in the SI diglossic variant $H_1$: CIAL:SI > 0. Finding both FI and CI linguistic elements in the corpus supports the proposition (M hypothesis) that they combine to form an M diglossic variant (SI). The second data type generated from the survey was also measured in ratio terms. In the survey, the participants were given everyday informal scenarios and asked to indicate if they would use an FI or a CI phrase. The result is the FI:CI ratio of usage by the individual. Analysis of the accompanying demographic data revealed the scope and geographic spread of CI in Indonesia. Because the use of CI originated in Jakarta (Oetomo, 1990), a high score for CI use was expected among participants from Jakarta and surroundings. Of interest are data gathered from participants outside the Jakarta area. For example, has the use of CI, which features strongly in television and popular-culture media (Turner & Wong, 2010), spread to regional cities? This research presumed that cosmopolitan regional cities and those where bahasa Indonesia is widely spoken would have a high usage of CI, and that CI would be used less frequently in remote regional areas where regional languages are the everyday spoken language.
1.7 The practical application of the research

Most foreign students of Indonesian, despite spending many years learning the language, have problems in understanding and communicating in everyday spoken Indonesian (Sneddon, 2003b), primarily because Indonesian is a diglossic language in which the grammar and vocabulary sets differ between the H diglossic variant and the L diglossic variant.

Traditionally, textbooks for learning the Indonesian language, and other learning resources, have focused on the H variant, FI (Djenar, 2006; Sneddon, 2006). This research tested the assumption that FI is the everyday spoken language; the thesis offers a counterview that the inclusion of only FI in pedagogies does not reflect the spoken language in everyday interactions. Sneddon (2003b) noted that this problem is partly due to a lack of understanding of the diglossic situation and hence the absence of resources for learning spoken Indonesian. The traditional perception of educators that teaching the L language is not appropriate is another contributing factor.

The recommendations in the conclusion of this thesis include a pedagogical model in which Indonesian language learning is contextualised into topics that align with Indonesian diglossia; hence, formal and informal spoken Indonesian would be taught in their respective sociolinguistic settings. Further, this research provides a basis for future projects, including the production of learning materials related to the spoken Indonesian language that will benefit foreign students of the language. To this end, this thesis provides a collection of the linguistic features of CI, the semantic and pragmatic values of CI, the CI lexis and CI collocations (Appendix 2).

1.8 Organisation of the thesis

Chapter 2 of this thesis discusses my epistemology and the sociolinguistic theoretical framework of diglossia. This chapter explores diglossia as a historical and global phenomenon, how diglossia differs from dialect, and how the Indonesian language can be classified as a diglossic language. Chapter 2 then investigates the historical origins of informal and formal Indonesian. Standard formal Indonesian is derived from the royal court Malay language of Riau in Sumatra. This Malay variant was well established in written literature. This ‘high’ version of Malay was subsequently chosen by the colonial Dutch rulers to serve as an administrative language. The Dutch educators developed the first system of grammar for this early form of Indonesian. Sneddon (2003b) noted that the design of the Dutch colonial language curiously resulted in ‘European’ structures. The other more commonly spoken dialects of Malay, known at the time as ‘Market Malay’, were eschewed by the colonial administration. Herein began the prescribed form of Indonesian alongside the evolution of a colloquial variety of the Indonesian language, which led to the current diglossic situation.
A linguistis descendant of Betawi-Market Malay, CI has always been a natural spoken language and prevails to be one. As an organic and natural spoken language, CI contains spoken expressions and pragmatics that are not found in FI. Further, CI is a ‘mother tongue’, primarily for those living in Jakarta. People who speak Indonesian as their first language (as opposed to a regional dialect or language) acquire CI at home, while FI is learned in school (Smith-Hefner, 2007). This chapter also discusses how SI became the language of popular culture and how it spread from Jakarta to other parts of Indonesia through the media and the internet.

Chapter 3 establishes the theoretical framework that underlies the mixed method approach. The chapter outlines the IPA for Indonesian and how different diglossic variants are coded and categorised within the theoretical framework. Applying a theoretical framework draws a clear distinction between the diglossic variants and enables a visual conceptualisation of how the variants interact. This chapter also clarifies the differences between CI and other language variant subsets, such as bahasa gaul (teen speak).

Chapter 4 presents the findings from the main corpus. The chapter provides a description of the CI linguistic features using IPA. The M hypothesis is tested and the spread and amount of CI linguistic features across the data sets indicate the nature of SI in terms of CI composition.

Chapter 5 presents the findings from the second data set generated from the survey questionnaire. The demographic spread and scope of CI indicates its usage as an informal language.

Chapter 6 discusses the conclusion that is extrapolated from the findings of the data analysis to accept or reject the M hypothesis. Recommendations for the practical application of this research are also made. The CI content of spoken Indonesian derived from the findings of this research is compared with that of Indonesian language textbooks. A pedagogical model that incorporates CI linguistic features is also recommended for the teaching of SI. The pedagogic model is framed around the sociolinguistic settings of Indonesian diglossia. In this model, FI is taught in the context of a formal setting, while CI is taught in the context of an informal setting.
1.9 Summary

The research presented in this thesis explores the informal Indonesian language in the context of Indonesian diglossia. The following points summarise the research objectives and related concepts in this thesis:

- To test the M hypothesis: Are FI and CI both present within SI?
- To determine if SI is the everyday spoken language.
- To define the linguistic features of CI, that is, the colloquial component of SI.
- To describe the linguistic features, patterns and semantics of CI.
- To design a pedagogic model for the Indonesian language that incorporates the teaching of CI.
Chapter 2

The Evolution of Spoken Indonesian
2. Introduction

This chapter discusses the sociolinguistic theory of diglossia and the historical development of Indonesian diglossia from Malay. Conceptualising Indonesian within the diglossia framework is fundamental to the systemic analysis of this research because it allows for the categorisation of the FI-SI-CI linguistic variants.

*Bahasa Indonesia* belongs to the Austronesian Malay language family. Variations of Malay existed in the historical development of the language. The term *bahasa Indonesia* was proclaimed in 1928 by the young nationalists as the future national language of independent Indonesia (Foulcher, 2000). Nevertheless, the emergence of bahasa Indonesia in its modern form is due also to language planning by the Dutch colonials. The Dutch were the first to collate the patterns of *Classical Malay*, which was the beginning of the standardisation of Indonesian grammar. The Dutch selected the Classical Malay from Riau over other Malay dialects because this dialect was deemed more prestigious than the other variants. The literature of Riau Malay forms the majority of the collection of Classical Malay works.

The other variants of Malay in the East Indies at the time were mostly spoken variants of ‘Market Malay’. Today’s diglossia in Indonesia is a direct product of the language selection that started during this Dutch colonial era (Sneddon, 2003b). Since independence, successive Indonesian governments have continued to oversee the development of standard Indonesian language (H), primarily through the national Pusat Bahasa (Language Centre).

The object of this thesis is to prove that Indonesian is in a state of diglossia and to challenge the notion that standard Indonesian (FI), the H variant, is the natural spoken language. In a diglossic situation, the L (or M) variant is the vernacular spoken language. Despite decades of official endeavour to promote standard Indonesian (FI) as the proper and the only Indonesian variant that ought to be spoken, the informal language CI has persisted and is inextricably imbedded in the spoken language. Unlike CI, the most likely reason for the failure of FI to be accepted as the exclusive everyday informal language is because FI is not naturally an everyday informal language.

2.1 Diglossia

‘Every living language displays a greater or lesser degree of diglossia between its spoken and written use’ (Lubliner, 2002, p. 2).

Fergusson (1959) defined diglossia as a language situation in which a High (H) diglossic variant and a Low (L) diglossic variant exist. He identified this phenomenon within certain world languages such as Arabic, Greek and German. Similar to dialects, H and L diglossic variants usually come from the same language family tree; they share common linguistic roots and share vocabulary and linguistic
features, albeit with variations in phonology, lexis and morphology. Fergusson (2003, p. 347) allocated parameters for the predominant use of either the H or the L language variant according to their respective social contexts. Fergusson’s (2003, p. 347) observation of diglossia in Table 1 shows that the L variant is the predominant language used in situations that are informal or involve close personal relationships, with the exception of ‘Personal letter’. In contrast, the H variant is the language of government and education.

Table 1
Fergussonian Diglossia

<table>
<thead>
<tr>
<th>Type of communication</th>
<th>H</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sermon in church or mosque</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Instruction to servant, subordinates</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Personal letter</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Political speech</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>University lecture</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Conversation with friends, family, colleagues</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>News broadcast</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Radio (soap opera)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Newspaper editorial, news story</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Caption on political cartoon</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Poetry</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Folk literature</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

*Note. Based on Fergusson (2003, p. 347)*

The sociolinguistic term diglossia is derived from the French *diglossia*, which is an adaptation from the introduction by the Greek–French writer Jean Psichari (Ioannis Psikharis) of the Modern Greek διγλωσσία. This word simply means ‘bilingualism’ and is typically defined in Modern Greek dictionaries as the use of two languages or two varieties of the same language by a community (Lubliner, 2002).

Errington (1986, p. 4) also proposed a diglossic identity for Indonesian in order to ‘develop better generalizations about the development of “modern” Indonesian. I argue that it is useful to introduce a well-known sociolinguistic concept, diglossia, to describe the Jakartan (L) situation.’

Errington specifically refers to the evolution of diglossia in Indonesia, and how it shares with other developing countries the need of language evolution to meet the demands of the modern world. Modern Indonesian is a product of the new national Indonesian culture that emerged in the 20th century and that continues to undergo an evolution in social norms, functions and activities. The
application of modern Indonesian diglossia in different social settings plays an integral part in fulfilling the communication needs of various administrative, religious and cultural activities. According to Errington, the dichotomy of H and L Indonesian is the result of ‘the elaboration and spread of standard Indonesian through national institutions which have developed a diglossic situation in which L and H Indonesian are functionally complementary codes’ (Errington, 1986, p. 13). Errington argues that the growth of a body of prescriptive writings on Indonesian, supported by the government and disseminated through schools, literature and the media, has fostered the differentiated social situations that are typical in diglossic speech communities.

Identifying Indonesian as a diglossic language where sociolinguistic boundaries are defined allows linguistic descriptions of CI. Formally acknowledging Indonesian as a diglossic language also promotes the status of the L spoken variant (CI) as a legitimate language, rather than dismissing it as incorrect, inappropriate or inferior. Within the sociolinguistic diglossia theory, the spoken language is not classified as a low-status vernacular or lesser form of the language but as a legitimate entity in its own right. Removing the negative labelling of CI also allows the L variant to be taught to foreign learners of the Indonesian language.

Fergusson (1959) recognised that the H variant is the official language endorsed by governments, taught at school, represented in the media, and used in official publications, while the L variant is the informal language of daily interaction. The L variant is the language that has evolved organically among the speech community and its development is free from direct official control. Furthermore, Fergusson elaborated that the H variant is not regularly used by anyone in a speech community as a medium of ordinary conversation, and any attempt to do so is felt to be either pedantic or artificial. Native speakers of the H form do not really exist per se because it is learnt through formal schooling.

Fergusson made the distinction between diglossia, as he defined it, and the more common ‘standard-with-dialect’ dichotomy, the difference being that, in the latter situation, some people actually speak the standard language, whereas in a diglossic ecology no one speaks the H variant colloquially. In other words, both dialect and the standard language are informal spoken variants. A dialect is used in intra-ethnic communication but the use of L is not limited within a particular ethnic setting. H on the other hand is never used as an informal language. Lubliner (2002) used a Greek example, where *katharevousa* (L) is the natural medium of expression and no one really speaks standard Greek as taught in school. The grammatical and syntactic rules of the standard language are used in formal expository writing. The correlation in the Indonesian context is that CI (L) is the natural medium of expression and FI is the standard language.
Fergusson (1959) stated that an average speaker in a diglossic speech community is at home in L to a degree he almost never achieved in H. The grammatical structure of L is naturally learned without explicit discussion of grammatical concepts, while the grammar of H is learned in educational institutions in terms of prescriptive rules and norms. In other words, speech is acquired through daily spoken interaction while literacy must be learned through education. An H diglossic variant usually acquired the prestigious ‘High’ status because it has developed into the language of literacy – the language in which scriptures, legal and governance documents, and classical literature were written. The L variant on the other hand, seldom or sometimes not at all served as a written language. It is mainly only in modern contemporary literature that the L variant is written in publications such as novels and comics – and often to portray direct speech quote.

Traditionally, a grammatical study of the H form of the language exists. Language publications concentrate on grammars, dictionaries, treatises on pronunciation, writing style, and so forth. Norms for pronunciation, grammar, and vocabulary use are well established and allow variation only within strict boundaries. The orthography is also well established and out-of-norm variations are deemed incorrect. ‘By contrast, descriptive and normative studies of the L form are either non-existent or relatively recent and not comprehensive. Often these have been carried out firstly by scholars outside the speech community and are written in other languages’ (Fergusson, 2003, p. 350). Fergusson’s observation of diglossia in non-Indonesian languages accurately describes the linguistic situation in Indonesia. Sneddon (2003b, p.122) claims that

no study of the informal language has been undertaken by the Indonesian (language) authority ... there has been very little research on colloquial language by non-Indonesian linguists ...

They differ to such an extent that an attempt to account for both by a single unified approach would be extremely complex, if not impossible.

FI has been given privilege over nonstandard variants and has been the only variant given any grammatical treatment by the education authorities in Indonesia (and elsewhere). Recently, however, the spoken variants have drawn the attention of academics, primarily foreign scholars, who have given linguistic and grammatical descriptions. CI, despite its use every day by Indonesians, has been deemed as not ‘proper’ Indonesian (Sneddon, 2003b). However, some Indonesian linguists have acknowledged the prevalence of CI in informal settings and have called for FI and CI to be considered parts of the same language (Kridalaksana, 1978).

The nature of FI as an officially constructed language is reflected in its intonation. FI has been described as a language that is unnaturally stilted; its intonation is impersonal and neutral, which
establishes psychological distance between speakers (Sneddon, 2003b). Yet, native Indonesian language teachers consider FI as the only variant worthy of being taught to foreigners; the common belief is that it is better for a foreign student to use FI than to use a substandard language. As a result many foreign students may be uninformed about the common use of the colloquial language. Sneddon (2003b) added that a lack of both understanding of Indonesian diglossia and command of CI among teachers can result in SI not being conveyed to students. For students to gain a complete understanding of the Indonesian language, CI needs to be taught alongside FI.

Fergusson (2003) noted the importance of using the right variant. He explained that outsiders become an object of ridicule when they learn to speak fluent, accurate L but then use it in a formal speech. Conversely, members of the speech community would likely be an object of ridicule if they use the H form in a conversational situation or during an informal activity like shopping. Fergusson (1959) used the case of Arabic to demonstrate the usage of H and L variants. He explained that it is typical to have someone read aloud from a newspaper written in the H form and then proceed to discuss the contents in the L form, or to listen to a formal speech in H form and then discuss it with the speaker in L form. Formal university lectures are typically given in H form, but tutorial discussions may be largely conducted in L form. Often a considerable part of the teacher’s time is taken up with explaining in L form. If a non-speaker of Arabic asked an educated Arab for help in learning to speak Arabic, the Arab would normally try to teach the student H forms, insisting that such forms are the only ones used. Often, educated Arabs will maintain that they never use L forms at all, despite direct observation showing that they use it constantly in ordinary conversation. These dynamics in the usage and perception of H and L forms in Arabic, as described by Fergusson, also apply in the Indonesian case.

Ferguson’s initial study on diglossia was affirmed by other sociolinguists such as Harold Schiffman (1999). Schiffman, in his article ‘Diglossia as a Sociolinguistic Situation’ expounded the important conditions that engender a diglossic language. He stated that there is always a clear separation in the function of the H and L variants, and each has its appropriate function within its own sociolinguistic domains. H and L variants are used for different purposes and native speakers of the community would find it odd, ludicrous, or even outrageous if anyone used the H variant in the L domain or vice versa.

The H variant has greater prestige than the L, being perceived as the language of ‘great’ literature, canonical religious texts, ancient poetry, public speaking, and of pomp. Hence literary heritage is supposedly in the domain of H. Written L belongs to the domain of ‘common’ folk poetry and popular genres. Yet native speakers are born into the L domain. L is the natural language, the variant that is
acquired first; it is the mother tongue and the language of home. The H variant is acquired through schooling. Many ‘native scholars acknowledge only the H variety as the language’ (Schiffman, 1999, p. 2). The H variant is strictly standardised, with entries in grammars, dictionaries and canonical texts. The L variant is rarely standardised in the traditional sense; if formal grammar descriptions exist, they are usually written by outsiders.

Schiffman (1999) observed that diglossias are generally stable, some persisting for centuries or even millennia. Occasionally, over a long period, L variants displace the H variant, or a merging of the two can also occur. As will be elaborated throughout this thesis, this case of H–L merging has taken place within Indonesian diglossia in the informal spoken domain.

A European example of M diglossic languages that emerged from H and L variants are the Romance languages that merged Latin (H) with local vernaculars (L) in medieval Europe. The post-Roman and medieval populations of today’s France, Italy, Spain and Portugal, whose local (Gaul, Italic, Iberian) cultures were Romanised and linguistically Latinised, began speaking forms of Latin that were superimposed with existing local linguistic vernaculars. In time, these local forms merged with Latin grammatical concepts and evolved into the modern Romance languages such as French, Spanish, Italian and Portuguese (Haugen, 2003).

Often the grammar of the H variant is regarded as more complex – implying greater linguistic sophistication – than the grammar of the L variety (Fergusson, 2003). In modern objective linguistic study, however, this view is not always held. Traditionally, H language is regarded as having more complex tense systems, gender systems, agreement and syntax than those of the L variety. Many of the lexical items are often shared, but generally there is differentiation; the H form has vocabulary that the L form lacks, while the L form is rich in idioms, expressions and interjections that are not appropriate in the H form’s domain. This differentiation is also found in the FI–CI juxtaposition in Indonesian, as will be demonstrated in this thesis.

Schiffman (1999) further comments that, in most cases, two systems of phonology are discerned. In one, the H and L variants share the same phonological features; in the other, the H variant may have different morpho-phonemics, or the H variant is a subset of the L variant inventory. However, speakers often fail to differentiate between the two systems. H and L variants may borrow elements from each other in particular items. As we will see in Chapter 4, the case of Indonesian diglossia encompasses different degrees of the above phonological variations between the H and L variants. To provide a brief example: phonological realisation of the H word hitam (black) in CI can be hitam, itam or item.
2.1.1 The evolution of Indonesian diglossia

Diglossia in Indonesia is diverse and complex. In the Indonesian linguistic landscape, a dialect or regional language is often the third language used by the speech community alongside the standard H and L diglossic variants. Moreover, often a regional language – such as Javanese, Sundanese or Balinese – has its own diglossic dynamics involving complex language hierarchies. An example is the provincial cultural and educational centre, Yogyakarta, whose inhabitants – apart from Indonesia’s largest ethnic group, the Javanese – include many tertiary students from throughout Indonesia. At a university campus where interethnic interactions between young people outside class are common, CI (L) would be the language used. When in lectures, FI (H) is the norm, as a sign of respect acknowledging the social status of the lecturer. When Javanese individuals converse, they would use ngoko (the Javanese L variant) if neighbourhood friends and kromo (the Javanese H variant) when talking to parents or grandparents at home. Thus, within one day, a contemporary Javanese person would often speak and mix four language variants (FI, CI, kromo, ngoko).

The diglossic nature of many Indonesian regional languages has played a part in reinforcing diglossia in Indonesian. The speakers of these regional diglossic languages, notably Javanese and Sundanese, have influenced Indonesian diglossia (Turner & Wong, 2010). The phonological and morphological aspects of CI are heavily influenced by these regional languages (Rafferty, 1982; Prawirasumantri, Husen & Sjamsuri, 1979; Sugiarto, 2007) (See Appendix 5). Much of the early linguistic transfer from regional languages into CI has been via Betawi Malay. The development of CI from a localised Betawi Malay dialect into a national spoken language is inextricably connected to Jakarta’s becoming the capital of independent Indonesia (Grijns, 1991). Had another city become the capital – and hence the economic centre and, more importantly, the hub of modernity, including popular culture – perhaps CI would never have been more than a regional dialect. Jakarta as the capital allowed for the eventual language fusion of the official language FI with CI. ‘In recent decades, a variety of informal Indonesian has emerged in the speech of the Jakartan middle class, which is developing into a standard colloquial variety of the language’ (Sneddon, 2003b, p. 11).

After independence, Indonesians from all parts of the country converged in Jakarta for reasons of politics, administration, business and education, among others. Inter-ethnic marriages between newcomers to Jakarta were common and, to a lesser degree, with the indigenous Betawi people. This interethnic melting pot created a new generation of Jakartans who began fusing CI Betawi linguistic features with FI, conceiving a new hybrid linguistic phenomenon: SI (Sneddon, 2003b). Kushartanti (2014), in her research of language acquisition by children in Jakarta, reinforced the general understanding amongst academics that Indonesians who are born and raised in Jakarta grow up speaking SI, which is imbued with strong CI linguistic features, as their mother tongue. Some examples of CI phrases are Iya! Yuk kita bikin sesuatu yang asyik! (FI: Iya! Ayo kita melakukan
The grammar of bahasa Baku (the official Indonesian term for FI) is still the official standard language, learnt in formal schooling, while spoken Indonesian is learnt outside school through daily interactions (Smith-Hefner, 2007). While SI was initially the mother tongue of the Jakartan speech community and its surrounding communities, the prominence of Jakarta as the capital city and as an exporter of culture and language through the media to the rest of Indonesia has resulted in SI becoming the common spoken Indonesian language in other parts of Indonesia (Manns, 2014). Manns collected samples of the use of non-formal colloquial Indonesian outside the capital and noted how radio stations in regional cities such as Bandung, Denpasar and Padang introduced less formal colloquial language (CI). Ewing (2005) also notes that young speakers in Javanese cities use non-formal colloquial Indonesian as an ‘in-group code’ and to project youth identity, especially during interethnic interactions.

Before further elaborating on the interrelationship of CI and FI, the next section discusses the origins and historical development of Malay and how it became a major language in the region.

2.2 A brief history of bahasa Indonesia

2.2.1 Early origins

Malay belongs to the Austronesian language family, which ‘covers a vast geographical area from Madagascar to Easter Island and from Taiwan to Hawaii. It is also known as the Malayo-Polynesian language group and some 1200 languages belong to this linguistic family’ (Crystal, 2006). The speakers of this language family are thought to have originated in southern China (the Yellow River valley) and migrated via Taiwan into the islands of the Philippines about 2500 BC, and then into Indonesia and out to the Pacific about 1000 BC.

Hervas was the first linguist–anthropologist who suggested an Austronesian language family. He was a pioneer in the study of world language families and identified some of the major language groups, including the Indo-European and the Austronesian language families. Hervas studied languages from around the world and compared their linguistic features. Subsequent linguists and philologists built on Hervas’s original findings. These early 19th-century European anthropologists and linguists diligently researched and classified the branches of the Austronesian language, upon which much of today’s knowledge on the Austronesian language family is based.
Another important academic was Wilhelm von Humboldt who, through his work *Malaischer Sprachstamm (Origins of Malaysian Language)*, was at the forefront of Austronesian language philology. He used comparative linguistic data to provide compelling evidence of a connection between the languages in Indonesia and Pacific Polynesia; he identified this language family as the *Malay-Polynesia* language group (Slametmuljana, 1982). This study concluded that there is an older, ancient form of Malay-Polynesian language and that its language descendants today are spoken in an area that includes Madagascar, parts of mainland Southeast Asia, the Indonesian archipelago and the Polynesian islands of the Pacific.¹

In 1848, another anthropologist–linguist, Crawfurd, made comparisons between Malagasy, Maori, Tagalog, Malay and Javanese. From 8000 Malagasy words, he only found 140 words that can be traced back to Malay or Javanese, from 4560 Maori words he found 103 words that correspond to Malay or Javanese, and from 9000 Tagalog words he found only 300 items that can be paired with Malay or Javanese. He attributed these lexical relationships to the influence of Malay upon the other languages in the region through migration, and not to a common language stem. He found higher correlations among the languages in the Indonesian archipelago. He studied the regional Indonesian languages of Madura, Lampung, Bali and Bugis, and found that each of the vocabularies studied showed about 60% affinity with Malay. Crawfurd himself believed that the people of Indonesia did not migrate from mainland Asia but were indigenous to the islands (Slametmuljana, 1982).

The generally accepted theory today is that the Malay culture and language dispersed to Indonesia through a migration from Asia. J. R. Logan, a 19th-century anthropologist who studied the cultures of Southeast Asia, supported the Asian-migration theory. In 1848 he found similarities among the customs of some of the tribes in Sumatra and Borneo and those of the Naga tribe from Assam in mainland Southeast Asia. Logan was also the first to use the name *Indonesia* (Greek for ‘Islands of India’). This name was later popularised and used by Dutch academics (Slametmuljana, 1982).

The investigations of the pioneering anthropologists and linguists of the 19th century gave rise to questions about whether similarities amongst the Austronesian languages were primarily due to language evolution from a common older language or to language exchanges through the interaction of

¹In *Voyage Round the World*, a study of world languages in 1776, philologist Foster surmises that the similarities between the Polynesian languages and the Malay languages suggest the existence of an older language that was the precursor to these modern languages. English academic Marsden came to a similar conclusion in his 1843 book *On the Polynesian or East Insular Languages*. Marsden coined the word *Polynesia*, which has become the accepted term, whereas Foster used the descriptor *South Sea Islands*. 
of peoples. Most modern philologists accept the explanation of a common language ancestry. However, this explanation does not preclude incidences of word borrowing through interactions.

The Malay language has been the lingua franca in the Southeast Asian archipelago and peninsula for many centuries, its spread catalysed by the Buddhist Empire of Sriwijaya. The Sriwijaya polity of central Sumatra began in the 3rd century AD and ended in the 10th century. Its territories or vassal states spanned most of today’s Indonesia and Malaysia. The political influence of Sriwijaya reached areas in the Philippines, Thailand and Cambodia, which may be important when considering possible language ‘export’ from the Malay islands to mainland Southeast Asia.

2.2.2 Philology and etymological examples

Identifying ‘original’ words is fundamental to the study of the Austronesian language family (and to any language family). Many of the Asiatic Austronesian languages have had strong linguistic input from Sanskrit or Arabic, or both. Words that cannot be traced to either Indian or Arabic origins, such as aku (I), hidup (life) and mati (die), and that have etymological ‘equivalents’ in other Austronesian languages, can be considered original. However, loanwords introduced from Sanskrit, which began around the first centuries AD (e.g. saya meaning ‘I’), and from Arabic, which began around the 9th century AD (e.g. wafat meaning ‘die’), are excluded from Austronesian etymological studies (Slametmuljana, 1982) but are still considered an integral part of the study of the philology and etymology of Indonesian.

The belief in the Asian origins of the Malay–Polynesian peoples has partly been founded on linguistic similarities between the languages of the islands with those from parts of mainland Asia. Schmidt (1906), in his book *Die Mon-Khmer Völker* identified that the mainland Southeast Asian language Cham displays very strong linguistic similarities and is grammatically related to the Malay–Polynesian languages. He studied the similarities among the Cham, Jarai and Rhade languages and the Malay–Polynesian languages.

A study by Hamy (1877, cited in Slametmuljana, 1982) analysed some of the languages from the Cochin mountain region and Indo-China such as Cham and Jarai and described them as Continental Malay. He believed that words from Continental Malay such as apoi (Malay: api meaning ‘fire’) and ayar (Malay: air meaning ‘water’) are not borrowed Malay words but share ancestry in a common, ancient Austronesian language. Many of the lexical features of Cham are found in the vocabulary of the Malay languages, and even in the languages of Pacific Polynesia today. The Malay word mati (‘dead’ or ‘died’) in Cham language has the form matai. In the ancient Sriwijaya language of Sumatra, it retains the form matai. Some of the Dayak languages of Borneo have the forms matai,
pampatai, matei or natei. Amongst the Polynesian languages of New Zealand (Maori), Samoa, Tahiti, Hawaii, Tonga and Marquesas the form is mate. The Malay word hidup (live; to be alive) in Cham takes the form hudip and hedip in Jarai. The Javanese word asu (dog) takes the form asau in Cham and Jarai and aso in Tagalog. In Cham, the Malay word laki (male) takes the form lakei and in Jarai rekei, which might be related to the Javanese word rek (‘mate’, a greeting to males). The Polynesian forms of fafine (Samoa), wahine (Hawaii) and fifine (Tonga) have also been linked to the form binai/bini. Traces of Continental Malay are also found in the Khmer language. Khmer words such as bong (Malay: abang meaning ‘older brother’), meas (Malay: emas meaning ‘gold’), preak (Malay: perak meaning ‘silver’) and kardus (Malay: kardus/kertas meaning ‘paper’) bear marks of Austronesian lexical footprints (Slametmuljana, 1982).

Sriwijaya inscriptions provide early samples of active and passive morphological markings. The prefixes ma and ni are found in the predicate forms mangujari and niujari (the modern Malay root word is ajar meaning ‘teach/learn’). The Sriwijayan word niwunuh has the modern Malay equivalent of dibunuh (killed). During the classical Malay period, the ma and ni morphs evolved into the me- and di- affixes. In passive predicate form, the first person pronoun marking in Sriwijayan Malay was attached as suffix. Hence the modern Malay equivalent of the Sriwijayan word nigelarku is kugelari (that I wore; that I bestowed; that I gave title) (Slametmuljana, 1982).

The resemblance in form and syntax of personal pronouns extends beyond the comparison between Sriwijayan and modern Malay; some interesting similarities are found between modern Indonesian and Cham in the personal pronoun forms. Consider the following examples from Slametmuljana (1982, p. 66) in Table 2:

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Cham</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>aku beli</td>
<td>kau blei</td>
<td>I bought</td>
</tr>
<tr>
<td>kau beli</td>
<td>heu blei</td>
<td>you bought</td>
</tr>
<tr>
<td>dibeli</td>
<td>nhublei</td>
<td>he/she/was bought</td>
</tr>
<tr>
<td>kita beli</td>
<td>gita blei</td>
<td>we bought</td>
</tr>
</tbody>
</table>

The phonetic and phonological relationships such as those between aku and kau, kau and heu, and kita and gita give foundation to the philological linguistic studies of language families. For the third person dibeli–nhubeli, other philological developments require investigation. For example, the Cham
personal pronoun *nhu* bears closer resemblance to the Sriwijayan morphological prefix *ni*, which later became the modern Malay *di-*, as discussed earlier. A possible shift in word class might have occurred from the third-person pronoun *dia*, which shortened into a passive third-person morphological prefix *di-* (e.g. *dibelinya* meaning ‘bought by him/her’). However, this idea conflicts with the Sriwijayan usage of the two affixes *ni-gelar-ku* – unless by this time the morph *ni* had already transformed into a passive prefix and its use was no longer specific to the third person (Slametmuljana, 1982).

Lexical input forms most of the linguistic contribution from the Indian subcontinent, but some similarities in the possessive usage of personal pronouns might also indicate structural influence (Table 3):

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Mundari</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>rumahku</td>
<td>omah-ku</td>
<td>my house</td>
</tr>
<tr>
<td>rumahmu</td>
<td>omah-mu</td>
<td>your house</td>
</tr>
<tr>
<td>rumahnya</td>
<td>omah-e (deweke)</td>
<td>his/her/the house</td>
</tr>
</tbody>
</table>

The lexical and morphological similarities in Table 3 illustrate linguistic ties. The Mundari form *deweke* has not transferred to Indonesian but still exists in Javanese (he/she/one’s self) (Slametmuljana, 1982, p. 73).

There are phonetic and functional similarities between the Indonesian *pe-* nominal marking and the *pan-* nominal marking of the Palaung language from the Mon-Khmer language branch. In Indonesian, the attachment of the *pe-* prefix to a root word (commonly called *base word* in Indonesian grammar), marks a nominalisation of the word. For example, if *pe-* is attached to the word *main* (play) to form the word *pemain*, the word becomes a noun meaning ‘the person who plays’. An identical nominalisation process occurs in Palaung: *dah* (to say) becomes *pandah* (a word); *deh* (to give) becomes *pandeh* (a gift); *rah* (to love) becomes *panrah* (the love); *yar* (to comb) becomes *panyar* (a comb); and *gop* (to cut) becomes *pangop* (scissors) (Slametmuljana, 1982). Note the difference in the type of noun between the Indonesian and Palaung examples. In the former, the *pe-* prefix forms an agent noun while in Palaung the *pa-* forms a concrete or an abstract noun.

The lexical and linguistic similarities among the Austronesian languages, encompassing a geographic span from Madagascar to the Pacific Islands, undoubtedly provide evidence of common linguistic etymology. Other historical variables, such as Sriwijaya’s exportation of *Ancient Malay* to other
Southeast Asian cultures, may also partly explain similarities among the languages in the region. The spread of Ancient Malay by the Sriwijayan Empire would have elevated the status of this language to the language of an empire and possibly set the foundation for Malay’s development into a lingua franca. The next milestone in the Malay language’s prominence was the period of Classical Malay, during which other variants of Malay were also spoken. The diglossic situation in Indonesia today can be traced back to the diversity of Malay variants centuries ago.

### 2.2.3 Classical Malay

Today the Malay language is the national language of Indonesia (bahasa Indonesia), Malaysia (bahasa Malaysia), Brunei and Singapore (bahasa Melayu). Modern Indonesian is a development from the Riau Royal Malay language of Sumatra, one of the more ‘prestigious’ Malay dialects that existed in the regions of today’s Indonesia and Malaysia. Scholars place the period of Classical Malay from the 14th century to the 18th century. By the 19th century, many dialects of Malay existed in the region, but the dialect regarded as the most scholarly was the Malay language used by the Riau Sultanate, which provided the basis for the High form of Formal Indonesian, that is, the standard language bahasa Baku. This Riau version of Malay is still regarded as a refined version of Classical Malay, alongside the very similar Johor Malay across the Malacca Strait, is a linguistic descendant of Ancient Malay, the language of the Sriwijaya Empire. Ancient Malay, besides being the language of an empire, was also a language of Buddhist literature and culture; Classical Malay became the language of Islamic literature and scholarship. By the time of the arrival of the Europeans to the Indonesia–Malaysia region, the Malay language was an integral part of the communication among the different linguistic cultural groups. Many variants of Malay existed in different regions, but mostly speakers were understood by each other. A Dutch sailor travelling in the late 16th century wrote that ‘anyone who does not know Malay in the Indies will not get anywhere’ (Sneddon 2003b, p. 61).

Classical Malay also played an important part in the Islamic proselytising of the Malay world, being the medium by which the message of Islam was disseminated in the region and thus further legitimising Malay’s linguistic status and entrenching its position as the lingua franca.

In the minds of Indonesian nationalist historians, the existence of modern Indonesia is a natural progression from the ancient Buddhist and Hindu empires of Sriwijaya (7th–13th century) and Majapahit (13th–15th century) because the combined area controlled by these empires partially encompassed modern Indonesia. The tenets of modern Indonesian nationhood, the Garuda Pancasila, (Five Principles of Garuda) and the national motto, Bhinneka Tunggal Ika (often translated as ‘Unity in Diversity’; a literal translation is closer to ‘Diverse (but) One (in) Essence’), are taken from ancient texts from those eras. The Pancasila are the five tenets of ‘righteous living’ in Buddhism and the
Garuda is the mythical bird that the god Vishnu rode. Bhinneka Tunggal Ika was a Sanskrit phrase that described the spiritual oneness of Javanese Buddhism and Shivaism, although their worldly religious–cultural forms differed. These concepts have been recontextualised by the fathers of modern Indonesian nationalism to encompass all of the ethnic and religious diversities found in the archipelago.

Under the Majapahit Empire, the term *Nusantara* (Islands in the Midst) was used to refer to the Malay Archipelago. This term is still widely used in Indonesia today. The political geography of Nusantara during the ancient kingdoms differed from the territory of the modern Indonesian Republic. Nonetheless, Indonesian nationalism often makes links to the ‘golden era’ of the Majapahit Empire and creates a sense of continuity in national identity (Ricklefs, 1993).

European colonisation contributed to setting the direction of the region’s political and linguistic landscape. Indonesia’s current territory was defined as a by-product of Dutch colonisation. Linguistically, the Dutch played a central role in determining the fate of the Indonesian language. ‘With the adoption of Malay by the Dutch as a language of the colonial administration, it was progressively standardised and disseminated among the educated native elite of the Dutch East Indies’ (Errington, 1986, p. 8).

### 2.2.4 Colonial language planning

European colonisation was instrumental in the Indonesian–Malaysian language split. In the treaty of London in 1824, the British and Dutch governments divided the Malay-speaking world into two colonies. The British claimed the Malay Peninsula while the Dutch claimed the islands south of the Malacca Strait. Such a divide did not previously exist. The Malay language homeland included both Sumatra and the Malay Peninsula, which were often part of the same polity since the times of Sriwijaya.

The Dutch colonial administration devised a ‘formal’ variant of Malay that served as an administrative language of the East Indies. This institutionalised Malay, already an unofficial lingua franca. In the early 20th century, the Indonesian nationalists renamed this official version of Malay *bahasa Indonesia*, which became official standard Indonesian, *bahasa Baku* (FI), after independence.

Besides Malay, the Dutch considered Javanese for the official language, but they abandoned it because of the intricate Javanese hierarchical system and its the linguistic complexities. In the mid 19th century, the Dutch decided to permanently institutionalise Malay, which became known as
dienstmaleisch (service Malay) (Errington, 1986). In 1865, Malay was adopted as the second official language of the colonial administration alongside Dutch. Most interactions with indigenous Indonesians were conducted in Malay from this point (Abas, 1987).

Malay can be viewed as already in a state of diglossia when the Dutch arrived in the region: classical Malay from Riau existed alongside variants of Market Malay; perhaps the Dutch administrative adoption of Classical Riau Malay followed an existing trend. The Riau Malay variant was well recorded – much literature, both folklore and Islamic literature, and correspondences between local rulers were written in Classical Malay. In contrast, little written evidence exists of the commonly spoken Market Malay variants. Thus the Dutch language planning probably reflects the existing linguistic situation.

Malay-language education would use High Malay … the style that had evolved in Classical Malay literature. This was, to a large extent, an artificial construct, developed over many centuries, which had diverged significantly from the various vernacular forms of the language spoken by the common people in their daily lives. (Sneddon, 2003b, p. 9)

In the late 19th and early 20th centuries, the Dutch education administrator in the Indies, Ophuijsen, standardised Malay grammar and spelling using the Latin alphabet and basing it on the court language of Riau (Sneddon, 2003b). This standardisation laid the foundation of today’s diglossia between Formal and Spoken Indonesian, because the Riau court language was primarily a language of literature and court ceremony, while the numerous spoken variants of Malay at that time were known as bahasa Melayu Pasar (Market Malay). This was a pivotal time in the history of the Indonesian language because it was the inception of bahasa Baku as an official language.

Sneddon (2003b) noted that the choice of Riau Malay deviated from the common perception among Dutch linguists of the day. Despite Ophuijsen’s emphasising the importance of oral traditions and spoken forms of language, standard Malay was based on the written forms of Malay found in Riau manuscripts rather than on the language that was spoken by local populations.

The grammatical structure of today’s FI is significantly different from its precursor language of Classical Riau Malay. The Dutch produced a language structure grounded in European syntax (e.g. subject-focused sentence heads) and translation that are alien to daily speech. FI did not evolve from everyday communal interactions of ordinary life situations but from colonial language planning (Oetomo, 1991; Sneddon, 2003b). While the high Riau Malay was promulgated by the colonial administration, the spoken varieties of Malay continued to develop organically at the communal level.
as everyday informal languages. ‘By independence the two styles, formal language and informal language had become quite distinct, a situation that continues to the present’ (Sneddon, 2003b, p. 9).

Another milestone for the institutionalisation of Malay in the East Indies was its inclusion in the colonial education system. The ethical policy that was initiated from the central government in Holland spurred the colonial government of the East Indies to provide education to the indigenous population. At first only the children of the aristocracy had access to the Dutch schools. In 1891 entry to the Dutch medium–lower schools system was made affordable and was opened up to selected children of lower-income Indonesian families. Malay began to be taught alongside Dutch at this level of schooling, although the Dutch language was still the main language of education for secondary and tertiary levels. Access to middle schooling allowed higher-education opportunities for students from a wider socioeconomic background (Ricklefs, 1993). In reality, however, only a select few school-age children and youth had such opportunities. In 1930, only 8 per cent of school-age youth attended some form of schooling (Dardjowidjojo, 1998). Although limited, dissemination of Malay among the indigenous Indonesian population had begun. The educated local population provided the administration of the Dutch East Indies with a civil-service workforce: just before World War II began, the indigenous population made up about 90 per cent of the workforce of the colonial administration (Cribb & Brown, 1995). However, the opportunities for indigenous Indonesians were still restricted. Educated locals with university degrees and fluency in Dutch were unable to access senior positions, which were occupied by Dutch nationals (Cribb & Brown, 1995). Colonial inequalities promoted an emancipation cause among the Indonesian youth, which gave rise to Indonesian nationalism.

2.2.5 Bahasa Indonesia and nationalism
The emergence of Indonesian nationalism in the early 20th century was the next important step in the ascent of the Malay language to the official language of independent Indonesia. On 28 October 1928 at an independence congress of young nationalists, a pledge, Sumpah Pemuda (One land, one people, one language) adopted bahasa Indonesia (Errington, 1986) as the language of unity for the future independent Indonesia (Foulcher, 2000). This ‘Youth Pledge’ articulated three important personal beliefs:

- That those present and also all indigenous peoples in ‘Indonesia’ share a common homeland: ‘We the sons and daughters of Indonesia acknowledge that we have one birthplace, the Land of Indonesia (Tanah Air Indonesia).’
- That all indigenous peoples of Indonesia belong to a single people: ‘We the sons and daughters of Indonesia acknowledge that we belong to one people, the People of Indonesia (Bangsa Indonesia).’
That a language of unity exists among the Indonesian nation, identified as ‘Indonesian’: ‘We the sons and daughters of Indonesia uphold the language of unity, the Language of Indonesia (Bahasa Indonesia).’

The pledge ‘One nation, one people, one language’ became the invocation of the independence movement (Ricklefs, 1993).

Once chosen as the future national language in the Sumpah Pemuda, bahasa Indonesia became the medium that carried the independence momentum across the Dutch East Indies. A national consciousness was forming and bahasa Indonesia was the quintessence of Indonesian nationalism. Bahasa Indonesia was used by the nationalists as the unifying element of an archipelago that was ethnically, culturally and religiously diverse. Any notion of nationhood as a continuum of Nusantara (the homeland archipelago) that existed had to be rejuvenated, recreated and reinforced. The early nationalist leaders dedicated their lives to spreading the idea of an independent, new, modern nation of Indonesia. The passion for independence rapidly gained popularity and unified the different ethnic and ideological groups in the fight against the common colonial enemies. However, after the colonial masters were routed, unity became elusive and the inherent diversities proved potentially divisive. Nation-building was not going to be easy! Many conflicts erupted between the (secular–nationalist) government of the new Republic of Indonesia and factions from opposing political spectrums. Despite the political turbulence in the young Indonesia, bahasa Indonesia’s status as national language was never in doubt. Bahasa Indonesia was the language of all the major political factions and their political propaganda, whether nationalism, Islam or communism. For example, a number of nationalist organisations such as Budi Utomo (Beautiful Endeavour, formed in 1908), Sarekat Islam (Islamic Guild, established in 1912) and the Communist Party of Indonesia (formed in 1920) used Malay as the language of their political activities. In 1927, the future first president of Indonesia, Sukarno, created a united political front called Permufakatan Perhimpunan-Perhimpunan Politik Kebangsaan (Indonesia Federation of Indonesian Nationalist Political Movements) to engage all the various nationalist groups (Abas, 1987). This important development consolidated the nationalist consciousness, which culminated in the Sumpah Pemuda event a year later.

The Peranakan or Chinese descendants in Java also contributed significantly to the early promulgation of Indonesian literature. Many of the publishing houses in the early 20th century were owned by Chinese Indonesians. Prior to the nationalisation of Malay into bahasa Indonesia, the Peranakan community were known to have spoken a variant of ‘Chinese Malay’. The Chinese Indonesian presses began in the 1850s and were an important part of the development of modern bahasa Indonesia because there were not many other early indigenous publications in Malay at the
time. Many of today’s idioms and language structures can be attributed to the contribution of the early Chinese Indonesian publications in Malay. Oetomo (1991) states that written evidence from printed materials from that time suggests that the pre-Indonesian Malay (PIM) that was spoken by the *Peranakan* community, which was heavily influenced by Javanese, have many resemblances with the linguistic features of today’s CI (elision, allomorphy and phonology, assimilation, syntax, lexis). Salmon (1981) in her book *Literature in Malay by the Chinese of Indonesia: A Provisional Annotated Bibliography*, has compiled a seminal catalogue of these Chinese Malay works from the 1880s to the 1930s.

The socio-political importance of the Indonesian language further increased when the Japanese took over the East Indies in World War II. The Japanese overran the Dutch East Indies in 1942 and ruled the East Indies until 1945. The Japanese promised independence and established an independence committee (Panitia Pesiapan Kemerdekaan Indonesia – The Preparatory Committee for Indonesian Independence), giving Sukarno and the nationalist leaders a forum through which to plan for eventual independence. Furthermore, the Japanese prohibited use of the Dutch language in any capacity, public and private. This edict required replacing Dutch with bahasa Indonesia in areas that were previously and exclusively the domain of the Dutch language. During this period Indonesian was promulgated replacing Dutch as the language of the press, literature and education. The poet and writer Alisjabana became the head of the language commission which created new terms and regulated the grammar. New Malay words were coined for technical concepts and new textbooks were written in Malay for use in higher education. All of these factors further institutionalised bahasa Indonesia as the future language of Indonesia (Ricklefs, 1993).

The choice to base the language of national unity on Malay was logical. As noted above, Malay had been widely used in much of the archipelago as a lingua franca for trade and diplomacy since the ancient Hindu–Buddhist era and then became the language of Islamic proselytising and later on chosen by the Dutch as the second official language in administration and lower-level education. The other candidate language was Javanese. But the Javanese people were already the dominant ethnic group, comprising 45 per cent of the indigenous population. Such a choice could create future resistance because of the perception of ‘Javanisation’ by other ethnic groups. Malay was the ‘neutral language’ and already was widely used in many regions such as *Betawi, Banjarmasin, Manado and Ambon*, and could therefore foster a sense of collective ownership. These spoken varieties of Malay all contributed to CI becoming a national spoken language.
2.2.6 Post independence

In the period following Dutch concession of Indonesian independence, the government of the new Republic of Indonesia implemented the teaching of the Indonesian language in schools. Indonesian was used at primary and secondary levels. However, Indonesian did not completely replace regional languages. In regional areas the use of the local mother tongue as the medium of instruction was permitted for the first 3 years of primary schooling. This approach aligned with the national philosophy of celebrating diversity (kebhinnekaan); the conservation of regional culture and languages is supported by the constitution of 1945, which states that the regional and local languages of Indonesia are to be protected and fostered as assets of the nation (Dardjowidjojo, 1998).

This post-independence period is of particular interest in this thesis because it is when Indonesia’s FI and CI language variants began to merge. The choice of Jakarta as the capital city of independent Indonesia was crucial in this regard: ‘Today, the prominence of Jakartan-Malay is measured by its sociolinguistic position in so far that it has become the daily spoken language (Grijns, 1991, p. 124).

The local dialect of the indigenous Betawi Jakartans was also a Malay variant which contains linguistic features from Javanese, Sundanese and Balinese. Within a relatively short time after independence, Betawi Malay influenced the spoken Indonesian language in Jakarta from which CI evolved. Furthermore, ‘The Jakartan accent has been a big contributor to the alternative linguistic forms of bahasa Indonesia, even penetrating into the domain of bahasa Baku’ (Grijns, 1991, p. 106). The multicultural and multiethnic nature of the Betawi culture was an ideal soil in which a national Indonesian identity could develop. Jakarta was an ethnically ‘neutral’ place and the local linguistic environment was already Malay.

In the 4th century, Jakarta was part of the ancient Hindu kingdom Tarumanagara, and later was part of the Buddhist kingdom Sriwijaya (8th–12th centuries). In the 13th century, Jakarta became a port city of the Sundanese Hindu Padjajaran kingdom, bearing the name Sunda Kelapa. In the 16th century, it was invaded by the Muslim sultanates of Demak and Banten and was renamed Jayakarta (from which the modern name is derived). There are insufficient records to establish exactly when Malay began to be spoken in Jayakarta. However, when the Dutch arrived, Jayakarta was clearly a cosmopolitan port city where traders from different parts of the East Indies converged and Malay was commonly spoken. When the Dutch made Jayakarta its colonial capital, they renamed it Batavia. The name of the local Betawi people is believed to be derived from the name Batavia. The intermarrying of different ethnic groups continued in Batavia and its melting-pot culture gave the local Betawi people their identity.
The intermarriage of the various Malay and non-Malay peoples (e.g. Chinese, Arabic and European) is reflected in the many linguistic influences on the development of the Betawi Malay dialect, which include Sundanese, Javanese, Balinese, Portuguese, Dutch and Chinese (Goebel, 2013; Sneddon, 2003b). The Malay that the Betawi people spoke became known as bahasa Betawi. After independence, bahasa Betawi spread among the rest of the growing non-Betawi population; the original Betawi population, however, became a minority ethnic group in their home town. A non-Betawi but Jakartan-born population emerged from the intermarriage of post-independence emigrants. The Indonesian language spoken by these new Jakartans is a blend of standard Indonesian FI and the linguistic features of Betawi Malay CI, that is, SI—the focal language of this thesis. Sneddon (2006) referred to the Indonesian spoken in Jakarta as Colloquial Jakartan Indonesian (CJI). This term suggests that Jakartan Indonesian does not contain FI elements, whereas this thesis posits that it contains both FI and CI elements. The scope of the language examined in this thesis extends beyond Jakarta and is therefore referred to as Spoken Indonesian.

2.2.7 Government language planning

When the independence war ended in 1949, the government continued language development through the official publishing house Balai Pustaka. The task began to promote and spread Indonesian through the press and the education system. The national language institute under the Ministry of Education, Instruction and Culture was located at the University of Indonesia and its purpose was to develop Indonesian to meet modern demands. The Old Order Sukarno era was stricken with political strife and upheaval. Unity and political stability were achieved only under the New Order (Suharto Government, 1967–98) though Suharto’s authoritarian-style government. This was also the period when bahasa Indonesia matured into its current form. National education and official media consolidated bahasa Baku as the national language, while the informal spoken language that had been taking shape in the national capital had also begun to spread through popular media and literature.

Language planning by the Indonesian state has been concerned primarily with FI for use in government, formal business, the law, the press and education. FI is taught at school but children have very little contact with it before then. For children who are born into Indonesian-speaking families (as opposed to regional languages or dialects), CI is the mother tongue, the language of home and of everyday communication (Smith-Hefner, 2007).

In post-independence Indonesia, vigorous propagation of and language planning for bahasa Indonesia was a priority of the national Indonesian Government. Modern nation building has universal aspects in a national language-planning program: there is *the Participatory Function*, which is the function of
language that facilitates participation in world-wide cultural developments. Then there is the language spread, which is usually motivated by pragmatic or political considerations. Ultimately, the purpose of adopting an official or national language is to facilitate the functioning of government and institutions and contribute to the political unification of the country’ (Nahir, 2003, p. 431).

Indonesian language planning conforms to a worldwide phenomenon as described by Nahir, in which language planning is a deliberate, organised effort by official institutions to affect the linguistic and sociolinguistic status development of the language ‘through a process of selection, codification, implementation, and elaboration’ (Nahir, 2003, p. 423). Often, ‘purification campaigns’ are part of language planning. Nahir identifies two types of language purification. The first is Internal Purification, which protects the accepted standard code against deviations that can occur in the form of non-standard (non-official) or ‘incorrect’ usage. This theme certainly recurs in bahasa Indonesia: bahasa Baku is constantly promoted by the Indonesian language-planning authority as the language that is ‘good and proper’ (baik dan benar) and is represented in government television and newspapers.

Nahir’s second form of purification is the External Purification, which tries protect a language against incursions from foreign languages. This form also features in Indonesian language planning: the official Indonesian Language Centre (Pusat Bahasa) constantly attempts to derive words from other Austronesian languages or to revive ancient Sanskrit words (or invent derivations of such words) to function as a modern socio-economic, technological and scientific lexicon.

While the Old Order regime of Indonesia’s first president, Sukarno, conceptualised and instigated Indonesian nationhood, the New Order of President Suharto consolidated and then systematically implemented the ideology and precepts that Sukarno’s generation devised. The systemised inculcation of bahasa Indonesia was executed hand-in-hand with the inculcation of the state’s ideology, Pancasila (the Five Principles). The New Order regime built schools in every village, through which nationalism and the Indonesian language were disseminated. Further, in the 1970s Lembaga Bahasa Nasional (the National Language Institute), via the national television channel TVRI, began broadcasting a daily half-hour program on ‘How to speak good and proper Indonesian’ (bahasa Indonesia yang baik dan benar). Propagation of Indonesian nationalism increased when Suharto launched Indonesia’s own Papala satellite that enabled the state broadcasting to penetrate into the remote corners of Indonesia. In 1978, the government’s Pusat Bahasa (Language Centre) held the Third Indonesian National Language Congress, which became a five-yearly institution.
These combined efforts of the New Order have produced impressive results and can be considered one of the most successful state-driven language-planning projects in modern history. In 1971, only 41 per cent of Indonesians were counted as native Indonesian speakers. This figure rose to 83 per cent in 1990, and by 2003 was approaching 100 per cent (Sneddon, 2003b).

However, the Indonesian public and media have no sense of obligation to conform to the linguistic demands of the language planning authority (Sneddon, 2003b); hence the development of the Indonesian language is as much propelled by popular usage as by language prescription. In effect, the ‘anglicisation’ of the Indonesian language continues unabated, especially in the case of modern lexical items that have no ‘natural’ native equivalents. Attempts at creating new words by official language planners are often subject to public mockery. Some of the Language Centre’s neologism struggles to compete with English-origin words as common usage. For example, the language authority Pusat Bahasa tried re-contextualise Javanese words for the cyber-world: *unduh* is to ‘download’ and *unggah* is to ‘upload’, while *pranala* is an ‘internet link’. Yet in daily speech these terms seem cumbersome; Indonesians more commonly use the original English terms because they sound contextually more natural to the technology. Many new anglicisms have become part of common usage, including non-technical anglicised phrases: *lagi ngga mood* is ‘not in the mood’, *ngga ngefek* means something ‘has no effect’, and *ngecek* is ‘to check’.

Inevitably in the modern linguistic landscape, anglicism has become common in both informal and formal spoken situations. Nonetheless, the language authorities continue their efforts to encourage the use of Indonesian words instead of foreign words. Untung Yuwono (2014), a linguistic academic from the University of Indonesia, reiterated the official stance and encouraged Indonesians to use the Indonesian vocabulary provided by Pusat Bahasa rather than foreign equivalents. He exhorts that the foreign words should only be used when there is no contextual equivalent in the Indonesian dictionary.

Formal Indonesian is widely accepted as an official language that has a rightful place in its appropriate sociolinguistic domain. Conversing educators and academics will choose language that is more inclined towards FI, and data in this study supports this observation. FI has been integral in Indonesian nation building as the language of administration and education: ‘Official and national languages have to have a high level of standardisation as a matter of cultural necessity. This means primarily that these languages should have the structural properties of a standard language to the highest possible extent, both that of flexible stability and that of intellectualisation’ (Nahir, 2003, p. 431). The intellectualisation of FI has transformed it into a modern language that can function as an academic, technical and administrative language.
FI is still the only official Indonesian language; uncertainty over the linguistic paradigms of CI is one reason for FI’s official standing. The CI component of SI does not adhere to the official spelling standard, Ejaan Yang Disempurnakan (EYD). Currently there seems little scope for discussion among the language authorities to consider the CI linguistic features in SI as a legitimate spoken language. As mentioned previously, the juxtaposition of language variants into a sociolinguistic hierarchy is imbedded in many Indonesian cultures. The conservative stance of the official Indonesian Government merely reflects the conservative Javanese linguistic mindset. The Javanese are the politically dominant ethnic group in Indonesia. Javanese values permeate the Indonesian culture, including in language planning. Javanese itself has a linguistic culture that stratifies language. In Javanese there are low (ngoko) and high (krama) variants of the language. Ingrained in this traditional linguistic mindset is the notion of a low and a high language, which has parallels in European languages – official standard versus vernacular – and exemplifies diglossia (H variant versus L variant). The status quo among the language authorities is that variants other than FI are perceived as low and hence not worth consideration. Fergusson made an observation on the traditional perception of the H and L variants in diglossic languages:

The proponents of H argue that H must be adopted because it connects community with its glorious past or with the world community and because it is a naturally unifying factor as opposed to the divisive nature of the L dialects. In addition to these two fundamentally sound arguments are usually pleas based on the beliefs of the community in the superiority of H: that it is more beautiful, more expressive, more logical, that it has divine sanction (as in the case of Arabic), or whatever their specific may be. When the arguments of the superiority of H are examined objectively their validity is often quite limited, but their importance is still very great because they reflect widely held attitudes within the community. Meanwhile, the proponents of L argue that some variety of L must be adopted because it is closer to the real thinking and feeling of the people … and it is a more effective instrument of communication at all levels. (Fergusson, 2003, p. 356)

FI, associated with official, prestigious institutions and with literacy and education, is clearly viewed as a socially ‘high’ language in relation to which the ‘nonstandard’ (for some, ‘substandard’) dialects of traditional and ordinary Jakartanese are ‘low’, non-prestigious variants. Stylistic elaboration in Indonesian from Javanese influence is making Jakartanese something akin to low Javanese ngoko in relation to standard Indonesian, which is akin to high Javanese krama.

A significant factor in the sociolinguistic Javanisation of the Indonesian language is that the Javanese ethnic group comprises a substantial portion of the population of the Indonesian capital. Errington
(1986, p. 8) states that the Javanese cultural influences on the Indonesian language ‘have been due not just to their political and cultural importance in the nation, but to their numbers in the city.’ He observed that the transfer of the ngoko–kromo hierarchy is reflected in the way spoken Indonesian contrasts with standard–official Indonesian. It is ‘akin to the Javanese system of linguistic politesse to contrast the social significances of their uses’ (Errington, 1986, p. 8).

2.2.8 Spelling systems using the Latin alphabet

The Latin alphabet is the ‘system of signs’ used in the orthography of Modern Spoken Indonesian, which has significant repercussions. Firstly, the Latin alphabet, the most common writing system of western European languages, is associated with modernity and secularism. This factor has allowed the Indonesian language in general to be more aligned with secularism, modernity and academic learning. A language that is grounded in an orthography of religious origins, such as Arabic, will always be in danger of maintaining some bias towards the philosophy from which the writing system originated. This is not a hard-and-fast rule; it can be argued the Latin alphabet was once the language of Christian liturgy. In the post-renaissance, secular modern world, the Latin alphabet has been divorced from the exclusivity of Christian thinking and in subsequent centuries has become the writing system of most western European languages.

As a ‘meta orthography of learning’, the Latin alphabet also allows for easier and more direct transfer of knowledge – particularly in the modern context – from the English language body of knowledge to the collective Indonesian epistemologies. Therefore, when studying Indonesian linguistics, given that the study of modern linguistics itself is a discipline pioneered by Western academia, presenting linguistic analysis of the Indonesian language in the common meta-orthography of learning is easier.

SI, more so than FI (which at times is imbued with nationalistic/patriotic rhetoric), is a language with secular connotation and modernity. No specific religious or ideological contexts are associated with SI. Thus the Latin alphabet caters for the fluidity and flexibility of SI as a modern language that readily absorbs new lexis and concepts of technology and the modern world.

In its history the Malay language had used two main writing systems before the Latin letters became the modern orthographic system in both Indonesia and Malaysia. Orthographic engravings (prasasti) from the ancient Hindu–Buddhist period (2nd century AD to 16th century AD) used the imported Indian Pallawa and Kawi writing systems. The Hindu–Buddhist culture allowed for the absorption of an extensive array of Sanskrit loan words into Ancient Malay. By the 16th century, Islam had replaced Hinduism and Buddhism as the major religion. A modified form of Arabic calligraphy (Jawi)
became the orthographic medium of Classical Malay. The Dutch deciphered Riau Malay texts into Latin script from the Jawi script. The coming of Islam did not alter the role of Malay as a lingua franca in the previously Hindu–Buddhist environment. Islamic proselytising simply applied the existing lingua franca to its own cause. Malay became associated with the language of Islamisation. During this period Arabic loan words were absorbed into Classical Malay, which became the language of the court, regional correspondence and diplomacy (Moeliono, 1986). Today, both Sanskrit and Arabic loan words have become a part of standard Indonesian vocabulary.

Both the Portuguese and the Dutch introduced the Latin writing system to Malay. During colonial times Jawi continued to be used especially in Islamic literature. Today in Malaysia Jawi is still an alternative writing system. The more secularly inclined Indonesian government has opted to use the Latin alphabet as the main national alphabet. Considering the religious diversity of the indigenous Indonesian population, the Latin letters are religiously neutral and the alphabet is not automatically associated with Christianity and is a unifying linguistic medium. To revert to a Sanskrit-based writing system would be intolerable for strict Muslims; they would associated such writing with non-Islamic belief systems. In contrast, to impose the Arabic-derived Jawi would alienate the non-Muslim populations. Sukarno and the other secular nationalists were adamant that Indonesia was to be a secular country (secular–theistic) where the major religions have equal standing before the constitution. Thus the Latin alphabet was the most religiously neutral choice, representing both secularism and modernity.

Most Indonesians today are well versed only in the Latin alphabet as an orthographic system; only religious scholars and philology academics are able to read the older systems. Thus, when Spoken Indonesian is represented in writing today – in literature, magazines, song lyrics and social media – the Latin letters have become the natural orthography of SI. The syllabic constructions of CI do not deviate from the paradigm of the ‘Perfected Spelling’ (Ejaan Yang Disempurnakan) of FI. Because no official standard convention for the spelling of CI exists, in stylistic writing its spelling is quite liberal, including its use of older spelling systems. For example, the personal pronoun elu (you; from Hokkien) is sometimes spelled loe. Discourse particles (DP) such as sih (+affirmative) is written as se, and the negator enggak written as engga, nggak, gak, or ga. Expressive collocations such as ayo loh! (DP +exclamatory) and mau aja (to agree) are spelled hayo lho! and mo aja respectively.

The spelling system of Indonesian–Malay using the Latin alphabet has had three main developmental phases. Grijns (1991) provided an account of the evolution of the spelling system of the Indonesian–Malay Latin alphabet from its beginnings to its current form.
The Van Ophuijsen spelling was the first official spelling of Indonesian Malay instituted by the Dutch administration. This spelling system was officially in use from 1901 until 1947. Under this system the orthographic form ‘e’ (italicised) stood for the central-unrounded [ə] schwa phone, while the normal letter ‘e’ - and before the revolution war also ‘é’ - represented the front-unrounded [e] phone. Another characteristic feature was the use of the letters ‘oe’ for the vowel ‘u’ or the phone [u].

The Soewandi spelling, also known as Ejaan Republik (the Republic’s Spelling) was instituted in 1947 as a simplification initiative of the Latin spelling system of Indonesian. The old spelling distinction between the /ə/ schwa and the /e/ phonemes was abolished and both became represented by the letter ‘e’. The ‘oe’ spelling form was also simplified to the letter ‘u’.

A cooperative agreement in language development with pre-Malaysia Malaya was signed in 1959 to form a joint spelling system, the Melindo (Malay–Indonesia) spelling system. Cooperation with Malaya broke down during Sukarno’s confrontation with the Federation of Malaysia in 1963. Suharto normalised relations with Malaysia in 1966 and language planning with Malaysia was resumed, culminating in Ejaan Yang Disempurnakan (EYD) (The Perfected Spelling), which was introduced in 1972 and revised in 1975.

The EYD deliberations of the Education Department produced guidelines for Indonesian word formation that served as a foundation for spelling all Indonesian words, including loanwords that have become part of common speech. Every Indonesian syllable contains a vowel (V). This vowel may be preceded or followed by a consonant (C). There are eleven ways to construct an Indonesian syllable:

V → a-nak, i-tu, ba-u
VC → ar-ti, ma-in, om-bak
CV → ra-kit, ma-in, i-bu
CVC → pin-tu, hi-lang, ma-kan
CCV → pra-ja, sas-tra, in-fra
CCVC → blok, trak-tor, prak-tis (Dutch/English loaned words)
VCC → eks, ons
CVCC → teks, pers, kon-teks (Dutch/English loaned words)
CCVCC → kom-pleks (Dutch/English loaned words)
CCCV → stra-tek, in-stru-men (Dutch/English loaned words)
CCCVC → struk-tur (Dutch/English loaned words)

These syllabic and vowel constructions apply to all Indonesian diglossic variants and are consistent in CI even when lexical, phonological or morphological changes take place.
2.3 The sociolinguistic functions of Indonesian diglossic variants

A concise review follows on the diglossic variants of Indonesian that are posited in this thesis. Indonesian diglossia consists of Formal Indonesian (the H variant), Spoken Indonesian (the M variant), and Colloquial Indonesian (the L variant). Sociolinguistic studies on Indonesian by Sneddon (2001, 2002, 2003b, 2006) have suggested that an M variant has emerged from a blending of the H and L variants. Historical linguistic studies on Indonesian also suggest that this blending is primarily a phenomenon of post-independent Indonesia, a process that started in the capital Jakarta (Sneddon, 2003). Figure 1 illustrates the process of synthesis of FI and CI and the resulting SI, including the domains of each variant.

Indonesians, especially educated Indonesians, would intuitively traverse between any of these three categories as the social setting demands. SI is the language of popular culture (e.g. TV, radio, movies, comedies, contemporary youth magazines) and has spread to all corners of Indonesia through both the national media and the transmigration program that scatters Indonesian speakers to traditionally non-Indonesian speaking areas.

CI and SI occupy a different sociolinguistic domain from that of FI, which results in different phonological and prosodic realisations. The main difference in the intonation of CI and FI is the longer vowels in the tonic syllable of CI. The prosodic features – features of intonation and rhythm –

Figure 1. The synthesis and composition of Spoken Indonesian.
in CI are important pragmatic and semantic markers and reflect the human emotions intrinsic to social interaction (Van Heuven & Van Zanten, 2007). The intonation of SI naturally expresses humour, desires, opinions, attitudes, sadness, anger, etc, whereas FI is primarily a tool that conveys non-personal information. FI is mainly composed in writing. When spoken or read to an audience, its prosody is almost monotone compared to that of CI and it does not adequately express the spontaneous human emotions that are required in casual, everyday conversations. Observers of the Indonesian language who do not understand the distinction between FI and CI and mistake Indonesian as exclusively FI have described the language as a “soulless” alienating language that is curiously impersonal, neutral and sets up psychological barriers between its speakers’ (Sneddon 2003b, p. 18). Indonesians who can speak FI use it only in formal settings. Consistent with Fergusson’s (2003) observation of diglossic speech communities, Indonesians perceive FI as irritating when used in informal settings. ‘No segment of the speech community in diglossia regularly uses H as a medium of ordinary conversation, and any attempt to do so is felt to be either pedantic or artificial’ (Fergusson, 2003, p. 354). In contrast with the emotionless language of FI, CI is lively and dynamic, and rich in humour (Sneddon 2003b).

2.3.1 ‘Weak’ diglossia

There are several types of linguistic relationships between diglossic variants. This section considers the inter-language (or intra-language in the context of this study) relationship using the criteria of Paulston and Tucker (2003), which classify languages in terms of their common features: (1) genetic classification (a common linguistic ancestor), (2) areal classification (spoken in a common area), (3) typological classification (common structures), and (4) functional classification (common social use).

FI and CI share commonalities under categories 1, 2 and 3. In Category 4, they diverge as explained in the previous section, although overlaps and linguistic blending occur in usage by native speakers as they consciously or subconsciously (intuitively) switch registers.

This thesis concludes that the nature of the Formal Indonesian–Spoken–Colloquial Indonesian diglossia in the Indonesian language is a weak diglossia. Lubliner (2002, p. 6) first posited the notion of strong and weak diglossias:

Strong diglossia is what has up to now been called diglossia *tout court*: a situation where dialect and standard are different enough for their users to be fully aware of their being two distinct codes, with different names for them. Weak diglossia, by contrast, occurs when the dialect is what I have called a parastandard, and in this case speakers usually think only in terms of ‘the language’ (‘English,’ ‘French’ or whatever) which may be used more or less ‘correctly’ (the model of ‘correctness’ being of course the standard), and discrepancies are
perceived as ‘faults,’ or at the very least as ‘colloquialisms,’ but not as indications of distinct language varieties.

In accordance with Lubliner’s definition, Indonesian diglossia is weak, because the spoken variants are deemed incorrect by the traditional grammarians, as opposed to their being different dialects or languages. The common genetic Malay heritage of Indonesian diglossia and common features of syntax, morphology and phonology provide more evidence of that Indonesian diglossia is a weak diglossia. Mostly the variations in lexicon, syntax and morphology are ‘surface’ differences. The emergence of an M variant from the H and L variants also indicates Indonesian diglossia is weak, suggesting enough common linguistic foundation to allow for such a merging.

The merging of the H and L variants into an M variant is not unique to Indonesian. Fergusson (1959) noted from his studies of diglossia in world languages that a kind of standard L may arise. Speakers of other dialects imitate this variant, which tends to spread although it remains limited to the functions for which the L variant is appropriate. The communicative tensions that arise in the diglossia situation may be resolved by the use of intermediate forms of the language and repeated borrowing of vocabulary items from the H variant by the L variant. Fergusson’s observation is relevant to the Indonesian diglossia: SI fits the intermediate form described by Fergusson. In the case of SI, CI linguistic features that form part of SI – or are even ‘absorbed up’ into FI – are often not acknowledged as originating in CI. Once CI lexical items are in common use, they are assumed to be a natural part of the bahasa Indonesia (BI) lexicon. CI (Betawi) lexical items such as bakal (akanE1: will.Future aux.), ketimbang (dari padaE1: rather than), buat (untukE1: for) and biar (supayaE1: so that) have become common in formal journalism. These items have not entered the most formal FI domains such as state speeches, but their prevalence in other formal FI domains has seemingly promoted their linguistic status from a once-incorrect form of Indonesian into standard Indonesian. This upward movement of CI lexical items is primarily noticeable since the collapse of the New Order regime, after which journalists and writers in general had more freedom to employ the natural language.

As discussed earlier, many modern languages are a product of H–L diglossic merging. Fergusson (1959) made the point that modern Romance languages were a product of historical diglossias where linguistic convergences resulted from the borrowing of lexical items from the H Latin variant by the L lingua vulgate. If the Romance languages have resulted from merging of the H and L variants into an M variant, with the consequent disappearance of diglossia in those speech communities, this phenomenon may be a future possibility for the Indonesian language (Sneddon, 2003, p. 215):

One major question is the future of the diglossic state of the language. While diglossia can remain stable for centuries, there are cases where there is a drift away from the rigid division
between high and low varieties of a language, as has occurred in Greece in recent decades. Colloquial Jakartan Indonesian is acquiring the status of prestige variety of informal Indonesian, a trend that has continued into the new century. The breakdown in the previous fairly strict separation of High and Low Indonesian could eventually lead to the disappearance of diglossia, although any such trend is certainly at a very early stage.

For Indonesian, and for all world languages, modern universal media is a key catalyst that determines the fate of a language. Whereas in the past diglossias might have been stable for centuries and language change was a slow progression, the scope of modern media and international mobility have accelerated this process. Although the media has been used to advance the cause of FI, SI has also benefited from media advances. Globally, in countries in which authorities impose no linguistic restrictions on platforms such as popular mass media, social media and mobile technology, spoken diglossic variants have gained wider exposure: ‘In the twentieth century the process has been greatly accelerated by the mass media, which carry the parastandard (L/M) of the cultural centres to once isolated areas where formerly the standard (H) was learned only in school’ (Lubliner, 2002, p. 7). This transference could also accelerate the breakdown of diglossia as the ‘massive exposure to parastandard changes the perception of the “language” from that of a rigid, school-taught and rule-bound medium to that of a living, spoken one – since people don’t consciously differentiate between standard (H) and parastandard (L/M)’ (Lubliner, 2002, p. 7).

2.4 SI in the media
In 1901 Charles van Ophuijsen proposed a Romanised spelling system for Malay, including a grammatical description. This became known as the Ophuijsen spelling system and was adopted by the Dutch colonial government. In 1908 the Commissie voor de Inlandsche School en Volkslectuur (Commission for the Literature of Native Schools and Popular Literature) was established to produce publications and literature in Malay. In 1917 the Commission was renamed the Balai Pustaka (Literature Office) and published folklore and classical literature from throughout the Indies language. Equally important for the development of Malay in the Indies, the Balai Pustaka began publishing new and contemporary works in Malay, as well as translating well-known Western novels by authors such as Mark Twain, Jules Verne, and Rudyard Kipling into Malay (Abas, 1987). The Balai Pustaka also established libraries and became an important institution that spread modern Indonesian literature (Ricklefs, 1993).

In 1947 the National Language Council was established. The committee council formulated three language planning objectives: (1) coining new terminology pertaining to the development of the
modern world in areas such as science, technology and economics, (2) writing a grammar, and (3) coining daily words. In 1949, Alijahbana, a leading linguist of the new republic, published a grammatical description of Indonesian and established the direction of modern standard Indonesian grammar (FI) (Abas, 1987).

In the period following independence, FI was in its nascent stage and an H–L tension apparently had developed. CI would still have been contained and limited within the realm of Betawi Malay. In these early days of the new Indonesia, the official effort was focused on producing and promoting literature in the new national language. The ‘internal’ threat from CI had not become an issue. Not until the New Order era was there campaign to speak ‘good and proper’ Indonesian. What had been constructed to be the standard Indonesian language was not adhered to in everyday interactions. While the New Order government continued the efforts to propagate FI, it also made explicit exhortation against ‘improper’ language use. Under the Indonesian Minister of Information Decree No. 111/1990, all national television institutions were charged with the mission to support with full regard bahasa Indonesia yang baik dan benar (Indonesian language that is good and correct). The literature during the New Order was defined by distinctions between four categories of publications (Heryanto, 2008, p. 96): ‘1. Those accepted as “high” literature; 2. Those banned on the basis of political content; 3. Those pilloried by the guardians of literary aesthetics, often involving popular fiction, pop novels and comics; 4. Those regarded as non-Indonesian literature.’

During the New Order period, despite official directives for editors and publishers to conform to the grammar of FI, an unregulated publishing industry churned out bacaan liar (literally ‘wild reading materials’, semantically ‘illegal reading materials’) that contained CI text and were deemed inappropriate by Balai Pustaka.

Radio, film and television had more freedom to use CI. In the 1970s CI–SI was already the natural language amongst the Jakartan population. The local Betawi people, from whom CI was inherited, had by then already influenced the speech habits of the rest of the Jakartan population. The demographic phenomenon that took place in Jakarta was a crucial factor in the development of SI. This demographic phenomenon is the new generation of (non-Betawi) Jakartans who have emerged in modern (post-independent) times.

The media of the 1970s, notably radio, film and television, played a central role in the dissemination of CI via the media tentacles that stretched out from the capital to other parts of the country. Popular Betawi comedians and entertainers such as Benjamin S., an iconic pioneer of popular Indonesian culture, spoke and joked in CI. Benjamin S. and his productions were aired to the rest of the nation.
This honest depiction of language that begun in the 1970s celebrating Betawi culture continued into the 1980s with comedian–actors such as Warkop Prambors and popular films like *Catatan Si Boy* (Heryanto, 2008).

The status of Betawi Malay as a regional dialect afforded CI free expression. It was difficult to curb CI language depiction in the context of Betawi culture because the New Order authorities, while set to inculcate the use of FI, simultaneously encouraged celebration of regional languages and cultures (*bahasa dan budaya daerah*). Thus CI expressed in media in the context of Betawi culture was acceptable as an expression of regional culture.

The popularity of Betawi television programs was a loophole in the media control that challenged the state’s media definition of the language of popular culture. Into the 1990s, CI representation in the audiovisual media and popular fiction continued to undermine the New Order state’s capacity to circumscribe other varieties of the Indonesian language. Sen and Hill (2000) stated that during this time the emergence of SI was a manifestation of, and added to, hybrid local cultures; in a globalising world global symbolism expressed the mood of class and generation.

In the 1990s television media in Indonesia was privatised. The appearance of privately owned television channels provided a platform for even greater exposure of CI. For decades following independence, the state-owned television channel TVRI (Televisi Republik Indonesia) was the sole channel that broadcast nationally. The first television broadcast was on 17 August 1962, in tandem with the commemoration of Indonesian independence. TVRI was the only source of TV information until 1989, when licenses for private broadcast stations were released. In 1989 Rajawali Citra Televisi (RCTI) broadcast for the first time. Initially RCTI could only be received in Jakarta and surrounding districts (*JABOTABEK* – Jakarta Bogor Tangerang Bekasi). Later, other regions that could use decoders were also able to broadcast the new TV channel. In 1990 another private channel, Surya Citra Televisi (SCTV), also started up and, in 1991, Televisi Pendidikan Indonesia (TPI). RCTI and SCTV were not nationally broadcast until 1993. In 1994 ANTeve and Indosiar followed suit. Before long there were 11 nationally broadcast television stations. Many more have appeared since, such as Trans TV and TV7 (Susilo, 2010).

Local movies and television shows flourished with the emergence of these new private channels. Many Jakartan and Betawi films were produced and nationally broadcast during the 1990s, continuing into the new millennia. Films such as the 1994 *Si Doel Anak Sekolahan* (*Doel the Student Child*) were screened on RCTI; the film topped the ratings for the year, exceeding all previous TV ratings in Indonesia. *Si Doel* was set in an old Betawi kampong. By that time, the Betawi Malay dialect was
well understood in Jakarta and other large cities beyond the boundaries of the small Betawi community (Sen & Hill, 2000). Further, SI was realistically depicted as the standard spoken language of Jakartan in the 2001 film *Ada Apa Dengan Cinta?* (What’s up with Love?) (Heryanto, 2008). This was the first domestic film to attract huge blockbuster audience numbers across the country.

In tandem with the emerging new structure of the media economy, Indonesia also underwent political changes in the late 1990s. A new socio-political environment was dawning as the New Order era was closing. Students led mass protests against the New Order regime that, combined with a severe economic crisis, resulted in Suharto abdicating after a 32-year presidential tenure. Thus ended the New Order Government and Indonesia entered a new era known as *Masa Reformasi* (the Reformation Era). The demise of the old government ended the state’s authoritarian control on the media. A speedy cultural transformation was brought about by the free media of the Reformation Era. Free expression of CI in national television and films was crucial in CI’s spread and its adoption as the national spoken language (Susilo, 2010).

Moreover, Jackson (2005) noted that CI was also the semiotics in popular dissent during the revolutionary pro-democracy movement against the New Order. Expressing anti-establishment political opinions in the media through CI was a sign of solidarity with the protesting masses. During the revolution against the New Order, CI was identified as the *bahasa rakyat* (the people’s language). Jackson provided a collection of evidence of CI use in political cartoons and student magazines that were protesting against the New Order.

The subsequent governments of the Reformation Era embarked on the freeing up and opening of the media, which has created a new journalistic environment in print, electronic and TV mediums. The liberalisation of media laws from 1998 onwards saw a decade of media expansion. Print media tripled, radio stations and national television networks doubled and over 200 local television networks sprung up (Aspinal & Mietzner, 2010). Many new forms of television shows, such as talk shows and soap operas, were introduced to the national viewing community. The journalistic world was quick to capitalise on this new-found freedom. Within 10 years the unthinkable took place. Politicians at the highest level were the subject of ridicule on national television. A current affairs series that deserves mentioning is *Newsdotcom*, a satirical comedy that impersonated Indonesia’s presidents. In this show, comedians impersonated past and present presidents by turning the personality quirks of the presidents into a subject of humour. Notable too was the use of CI in the impersonations of the presidents. The personality Gus Pur, impersonated Indonesia’s fourth president, Gus Dur, who had a down-to-earth attitude and liked to use CI expressions such as *Buat saya sih gampang aja* (For me it’s just simple) and *Gitu aja kok repot* (Why fuss over just that?) (Heryanto, 2008).
The Indonesian Lawyer Club is another program depicting SI at work. In this program notable lawyers discuss political matters on national television. This show is a prime example of the use of elements of CI by educated, high-profile professionals. Guests on this show include state politicians such as the current President Jokowi.

In literature, CI was already beginning to be incorporated in the 1990s. Novels that passed official censorship before the 1990s still portrayed bahasa Baku (FI) as the everyday spoken language. In 1995 a novel containing CI language passed censorship. One of the largest publishers in Indonesia, Gramedia, published a series of Lupus novels under the category ‘Fiction for adolescents and children’. The Lupus series, which became a national best seller, contained a CI language variant that depicted the slang of Jakarta’s urban youth. Some of its linguistic humour used wordplay between English and Indonesian: Ayam sorry, Ayam sorry, Mom! Gout lagi busy sekalle! ‘Ayam’, which sounds like the English ‘I am’, means ‘chicken’ in Indonesian (Sen & Hill, 2000).

Djenar (2012) observed how CI gained greater representation in teenage literature in the period after the New Order Reformation. This period represented greater opportunities for openness. Journalists, writers and performers were able to express themselves more freely; in this era there has been greater representation of CI in media and literature. Novels for young people became more realistic by portraying CI as the form of everyday speech (Djenar, 2012). Television and entertainment programs were more prepared to use CI, while newspapers and magazines were more inclined to quote people speaking in CI (as opposed to retranslating speech into FI, previously the common prescriptive journalistic practice) (Sneddon, 2003b). The liberation of CI in the media drew concern from traditional prescriptive language figures; according to them, the free use of CI in dialogue and narration transgresses the ‘sensibilities of writing conventions’. Language authorities reacted to ‘teenlit’ They felt that the amount of colloquialism was excessive; teenlit ignored the fiction writing convention, which dictates that narration should be written in FI and that CI should be reserved for the dialogue; and the language is unstructured, containing mixing of standard and informal Indonesian (SI) (Djenar, 2012).

Such criticism, which was motivated by conservative, authoritative language-planning priorities missed the critical point that the language represented in teenlit simply reflects the linguistic reality of everyday speech, including that the fact that speakers freely mix FI and CI as required. Today, Indonesian television of the new millennium is saturated with CI and SI. All types of popular television shows – from soap operas, comedies and talk shows to cooking shows, and many more – are rich with, if not dominated by, the CI–SI language.

Lastly, the other modern media platform that has played an important role in the liberation of CI was the internet. Sen and Hill (2000) noted how at the turn of the new millennium the very freedom of the
internet became a constant reminder of the absence of openness and freedom in the other state-controlled media. Internet technologies had not been anticipated in the mechanism of state control (Sen & Hill, 2000). The internet and social media today major platforms for communication. Anyone who engages in Indonesian social media or blogs knows that CI–SI is the spoken language of contemporary Indonesia. In the past, language reality, or the language that people spoke, could be suppressed by the authorities’ control of education institutions, media and publications. The internet changed this vertical linguistic control. With the internet, and especially through social media, people are beyond the reach of government control and are free to use CI language.

2.5 Conclusion

This chapter has discussed the sociolinguistic theory of diglossia and its application to the current FI–CI linguistic juxtaposition in Indonesian. In the narrowest of definitions, the FI variant is the formal language used in formal situations, while CI is the colloquial variant. A review of the history of the Indonesian language gives evidence of the past existence of diglossia within the Malay language, which forms the basis of modern Indonesian language. FI, the modern H diglossic variant is a continuum of Royal Riau Malay and was grammatically constructed by the Dutch colonial authorities; CI is the variant inherited from the Market Malay language of the Betawi people of Jakarta. Political–historical circumstances have institutionalised a High variant, which, by default, relegated other variants to an L status. This chapter also discussed the importance of Jakarta in post-independent Indonesia as a hub of popular Indonesian culture and as a natural organic linguistic laboratory creating a new spoken national language in which the FI and CI diglossic variants merged to form the intermediary M diglossic variant, SI.

Political and media democratisation and television privatisation in the late 1990s, together with the internet age, have allowed free expression of CI. SI has become the standard spoken language, which is evident in contemporary films, popular TV shows and social media. Although it is natural to speak in CI, certain modern technical–abstract vocabularies need to be borrowed from FI. In this regard, CI and FI complement each other. The marriage of CI and FI was natural selection; one providing the speech acts, intonation and pragmatics of a natural spoken language while the other providing the terminologies for communicating in the modern world. Since Dutch colonial efforts, later continued by the Indonesian State, FI has undergone almost a century of language planning with a view to modernisation: a language has been designed that can facilitate universal Indonesian education and can function as a technical modern language (Sneddon, 2003b). FI has accomplished this goal. And for the purpose of everyday communication, the CI component in SI provides the dynamics of a living spoken language, ‘smoothing out’ and ‘sweetening’ the otherwise stilted and awkward prosody of FI.
The next chapter will lay out the study design and methodology employed in this research to test the hypothesis: that SI is the common everyday language. The gap in the existing literature addressed by this research is the paucity of empirical studies that describe the linguistic composition of SI. The remainder of this thesis uses linguistic data to identify and analyse the linguistic features of CI and to discuss how FL (the H variant) and CI (the L variant) have mixed to become SI (the M variant), the common informal spoken Indonesian language.
Chapter 3

Methodology and Description of the Data:
Qualitative and Quantitative Corpus Analysis of Spoken Indonesian
3. Introduction

This chapter outlines the methodology used to determine how SI will be described and measured in the thesis as a language variant that contains linguistic features of both FI and CI. The linguistic features of CI define SI as a spoken language. CI naturally evolved as a spoken language and thus possesses all the qualities of a spoken language in terms of expressions, semantics and pragmatics. On the other hand, as discussed in the Literature Review Chapter, FI is a formally prescribed language and may be perceived as unnatural, stilted and impersonal for the purpose of everyday informal expression. Nonetheless, the linguistic features of FI, particularly the lexicon of FI, equip SI to be a functional spoken language of the modern world, complementing CI where it lacks terminologies of modern abstract concepts such as technology, science, law, economics and politics. Before discussing the methodology of this research, it is important to explain the construction of the taxonomy, terminology and coding used in the methodology.

3.1 The taxonomy, terminology and coding of Indonesian diglossia

The FI–SI-CI diglossic taxonomy in this thesis corresponds to Sneddon’s (2003b) H diglossic variety, M variety and L variety. The FI–SI-CI coding allows for categorisation and hence the establishment of axioms for qualitative and quantitative linguistic analysis that clearly define the linguistic features of each variant. Subsequently it is possible to ascertain the composition of SI. All terms that refer to the formal standard Indonesian language are coded as FI (Formal Indonesian, i.e., the H diglossic variant also known as bahasa Baku or standard formal Indonesian). For the purpose of linguistic classification, colloquial varieties belong to the subset code of CI (Colloquial Indonesian, i.e., the L diglossic variant also known as bahasa Sehari-hari or informal everyday Indonesian).

While Indonesians certainly recognise the differences between FI and CI forms, according to Sneddon (2001) there is no clear dividing line between them, and intermediate forms are associated with semi-formal situations. Similarly, Djenar and Ewing (2015) observed that in practice there is no clear distinction between the use of FI as the formal language and CI as the informal language. In this thesis the linguistic grey zone described by Sneddon and Djenar is considered a continuum of SI. This study explores the possibilities of SI, how it works as a categorical variable containing a subset of linguistic components, and how it functions sociolinguistically.

3.1.1 Confused terminology

First it is important to clarify the terminology in relation to CI because there is no single consensus in the use of terms. Often CI and SI are confused with each other and with other sub-variants. Sneddon
In this thesis it is assumed that all colloquial Indonesian variants such as those listed above are but sub-variants of, or at times merely different terms for CI. The most common recent confusion amongst students of Indonesian language is that bahasa gaul (social language) is the same as CI. Djenar, a prominent researcher of bahasa gaul, has written on the prevalence of bahasa gaul among youth (Djenar, 2006, 2008; Djenar & Ewing, 2015). She claims that bahasa gaul is a social language that contains lexis pertaining to the semiotics of youth culture. Bahasa gaul has gained national exposure through the media. It adapts Jakarta dialectal and foreign borrowings, invents new terms and acronyms, and does not violate the grammatical patterns of informal spoken Indonesian (i.e., CI) (Smith-Hefner 2007). As Smith-Hefner stated above, bahasa gaul functions within the linguistic parameters of CI, with additional fad-words. It is constantly changing as new words or expressions become popular and older ones fall out of use. Given that bahasa gaul does not command its own linguistic features, apart from additional lexicon, but functions within the linguistic paradigms of CI, this thesis classifies bahasa gaul as a subset of CI.

At the lexis level, the position of bahasa gaul as a subset of CI does not lessen its contribution to SI. Some of its lexis and expressions are not a mere transient fad and do find their way into common use. A number of bahasa gaul words have been absorbed into SI and have are commonly used by the general population.

3.2 The SI spectrum
The Indonesian sociolinguistic language phenomenon represents a continuum due to the malleability of the use of language variants over different sociolinguistic domains (Sneddon, 2006). Previous research on Indonesian diglossia (Sneddon, 2006; Djenar & Ewing, 2015) has suggested that in the reality of daily informal speech, there are no clear boundaries separating FI, SI and CI in their usage. As Sneddon (2001) has stated, speakers may make their colloquial speech somewhat more formal by incorporating some features of formal language; thus characteristics of FI are not excluded from
informal conversation. Conversely, formal language varies in usage depending on the social context. A politician may use less formal language in an unprepared speech to demonstrate his populist intentions when trying to connect with the masses. Notwithstanding the malleable use of FI and CI across different sociolinguistic domains (i.e., social situations), the FI–SI–CI categorisation employed in this thesis is a system that allows for the study, dissection and analysis of linguistic patterns and features (phonology, morphology and lexica) of each variant. Due to the malleability of language variants in use across different sociolinguistic domains (Sneddon, 2006; Djenar & Ewing, 2015), the FI–SI–CI labeling fits into a continuous spectrum of formal-informal social domains. Subsets of CI such as bahasa gaul are located at the far right of the formal-informal continuum while usage of pure FI is positioned in the far left spectrum. Section 3.7 of this chapter further discusses these linguistic and sociolinguistic theoretical frameworks, which underpin the research conducted for this thesis.

It is important to remember that the distinction between CI and SI derives from the fact that CI linguistic features are an independent linguistic phenomenon, pre-existing to SI in Betawi Malay, whereas SI is a modern derivative of both CI and FI. SI has no inherent linguistic elements of its own; it projects only those of CI and FI. One objective of the research was to identify the presence of CI grammar and linguistic features in SI, thus defining SI’s function as an informal language variant.

3.3 Data samples
The linguistic corpus collected for this research represents natural language taken from the speech community under study, and it captures a subset of real-world language data. Only corpora that reflected the informal everyday Indonesian language were purposively selected.

An array of different types of written and audio-visual corpora was identified. Samples were obtained from recorded interviews with native Indonesian speakers compiled by Sneddon (2006), as well as from written media, internet content, billboard advertisements and audio-visual media such as television programs and films. Table 4 shows the number of data samples and the number of words each sample contained.
Table 4
Data Samples in the SI Corpora

<table>
<thead>
<tr>
<th>Data sample categories</th>
<th>N</th>
<th>SI</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorded interviews and conversations</td>
<td>6</td>
<td>6408</td>
<td>2130</td>
</tr>
<tr>
<td>Contemporary literature</td>
<td>16</td>
<td>4603</td>
<td>1298</td>
</tr>
<tr>
<td>Audio-visual media</td>
<td>14</td>
<td>3626</td>
<td>1745</td>
</tr>
<tr>
<td>Billboard advertisements</td>
<td>12</td>
<td>74</td>
<td>27</td>
</tr>
</tbody>
</table>

\[ \Sigma n = 48 \quad \Sigma SI = 14711 \quad \Sigma CI = 5200 \]

3.4 Methods and sampling

A Mixed Methods (MM) research design (or methodology) was applied to the data analysis. Using MM, the researcher collects, analyses and mixes (integrates or connects) both quantitative and qualitative data in a single study (Johnson, Onwuegbuzie & Turner, 2007).

The corpora (linguistic data) for this study were sampled purposively. The sampling was purposive in so far as only corpora that represent the informal spoken language matched the purpose of this research. Examples of purposive corpora for this research include everyday conversations between friends and family, social language of young people, and colloquial language that expresses familiarity between speaker and audience. Teddie and Yu (2007) defined purposive sampling as the technique of selecting units (e.g., individuals, groups of individuals, institutions) based on a specific purpose associated with answering the research study’s questions. Purposive sampling does not achieve representativeness of a broader group and/or set up comparisons among different types of cases. It targets instances that are representative or typical of a particular type of case, according to a particular dimension of interest. Purposive sampling aims to generate representative cases, which in this research were samples of everyday informal language.

The research methodology was grounded in positivism: analysing knowledge that is grounded on the empirical evidence in logical and mathematical terms (Morgan, 2007). Table 5 is an adaptation from Teddie and Yu (2007) that summarises the MM design in the context of this research.
Table 5

Mixed Method Application to SI

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Purposive sampling</th>
<th>Probability sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of analysis</td>
<td>Qualitative sampling: identifying and describing CIAL in the samples, using IPA conventions</td>
<td>Quantitative sampling: counting the amount of CI lexis in a sample and analysing it using descriptive statistics</td>
</tr>
<tr>
<td>Reasoning</td>
<td>Inductive: using sociolinguistic theory of diglossia as epistemological basis</td>
<td>Inductive: using numeric data to ascertain prevalence of CIAL in data samples</td>
</tr>
<tr>
<td>Paradigm</td>
<td>Positivist: using CIAL data to test the M hypothesis</td>
<td>Positivist: using CIAL data to test the M hypothesis</td>
</tr>
<tr>
<td>Purpose of sampling</td>
<td>To generate sample that addresses research question (Is Indonesian a diglossic language?)</td>
<td>To generate sample that addresses research question (Is Indonesian a diglossic language?)</td>
</tr>
<tr>
<td>Rationale for selecting cases</td>
<td>Cluster sampling: categories of secondary sources selected (audio visual media and literature representative of M sociolinguistic domains)</td>
<td>Cluster sampling and random sampling: secondary sources selected (audio visual media and literature) and survey (questionnaire) used to obtain primary data</td>
</tr>
<tr>
<td>Sample size</td>
<td>&gt;14000 words of primary data</td>
<td>&gt;14000 words of primary data</td>
</tr>
<tr>
<td>Time of sample selection</td>
<td>During the study</td>
<td>During the study</td>
</tr>
<tr>
<td>Method of sample selection</td>
<td>Uses criterion of M diglossic domain</td>
<td>Uses criterion of M diglossic domain</td>
</tr>
<tr>
<td>Form of data generated</td>
<td>CIAL</td>
<td>Numeric data gauging extent of CIAL in SI</td>
</tr>
</tbody>
</table>

Note: IPA = International Phonetic Alphabet, CIAL = Colloquial Indonesian Aspect of Language, SI = Spoken Indonesian.

3.5 Data analysis

Figure 2 maps out the methodology sequence and qualitative and quantitative interrelationship of this research. The data in the form of linguistic features, here forth referred to as Aspect of Language.
(AL), from the SI corpus first underwent qualitative analyses in Methods 1 and 2, followed by quantitative analysis in Methods 3a and 3b.

The first method of corpus analysis involved identifying the presence of CI lexicon in the corpora and qualitatively describing the phonological features of CI using the International Phonetic Alphabet (International Phonetic Association [IPA], 1999) in an impressionistic auditory analysis at a broad transcription, or phonemic, level. Following that the number of CI linguistic features, here forth CI Aspect of Language (CIAL), in SI was quantified to ascertain its composition as an M diglossic variant.

**Figure 2. The research methodology**

**3.5.1 Method 1 – Establishing whether CIAL is present in SI corpora**

To affirm that CIAL is present in the informal spoken language, it needed to be ascertained that CIAL is an element in the informal social context. The methodology for investigating whether CIAL is present in used qualitative analysis by identifying aspects of language that do not belong to FI. The

3.5.2 Method 2 – Qualitative method: IPA – identifying and isolating non-FIAL, and constructing CIAL

If the corpora data sets indeed showed the presence of CIAL, these AL would be analysed, expounded and illustrated by establishing the patterns and paradigms of CIAL, applying the second qualitative method by using the IPA system. In an earlier study on variations in Betawi Malay, Grijns (1981) used IPA to discuss some of the features of CI. The morphological analysis followed the common system used to describe affixes in Indonesian, as employed by Boellstorff (2002).

The process of constructing CIAL from the SI corpus involved a method of differentiation. The process of identifying CIAL simply meant highlighting all AL that were not FIAL and recording their phonology, morphology and lexicology. The aim of this method was to synthesise, unify and complement various descriptions of CIAL provided by previous researchers. This research pre-empted what CIAL might look like, and a guideline was devised as a tool to identify non-FIAL. The guideline included several indicators:

1. Syntactic ellipsis as a common feature in daily speech (Sneddon 2006)
2. Morphological divergences from FIAL (Fan 1990; Kushartanti 2014)
3. The phonological divergences from FIAL (Kushartanti 2014)
4. Elisions and allomorphy (Sneddon 2006; Kushartanti 2014)
5. Alternative lexicon not present in FI (Sneddon 2006; Djenar & Ewing, 2015)
6. Variation in semantic properties that are outside the paradigms of FI grammar (Sneddon 2006; Djenar 2008).

3.5.3 Method 3a – Quantitative: CIAL:SI ratio

The next objective of this research was to count the proportion of CIAL in the SI corpora to measure the linguistic nature of SI and determine the CIAL composition. This tested the M hypothesis that an M diglossic variant exists.

Quantitative methods were applied to test the alternative hypothesis that the CIAL ratio in the SI corpora is greater than zero; H1: CIAL:SI > 0. Descriptive statistical analysis of CIAL/SI consisted of
the mean and range across the data sets. The remaining non-CI lexis in the SI corpus are FI lexicon and the uninflected base word $X^o$ universal to both FI and CI.

3.5.4 Method 3b – Quantitative: CIAL usage across Indonesia

Method 3b involved collecting demographic data on CIAL usage by means of a survey. The main variable from these data of particular interest to this research was the use of CIAL outside Jakarta. The variables of age and occupation also indicated the scope of CIAL usage.

The questionnaire data were collected from 98 native Indonesian speakers from 35 different regions beyond, and including, Jakarta to determine FI versus CI language usage in everyday settings. The questionnaire consisted of 18 sets of questions from everyday informal contexts and gave the participants the option to indicate whether they would use FI or CI phrases in particular informal situations. The participants remained anonymous; the questionnaire gathered information on their age, occupation, and level of education, and the city/region in which they lived and worked. The questionnaire also gathered information about whether participants spoke bahasa Indonesia (BI) (Indonesian language), bahasa Daerah (BD) (regional language) or both (BI&BD) at home, socially and at work. The prevalence of CIAL usage by participants in their regional settings was measured as a percentage to indicate how much CIAL is used in spoken informal Indonesian outside the capital Jakarta.

To summarise, Method 2 comprised the qualitative component of the research methodology, and Methods 3a and 3b were the quantitative methods. The qualitative methods produced results that were subsequently analysed quantitatively. To cater for the quantitative conceptualisation and analysis, the qualitative results were assigned codes and symbols. These are discussed in Section 3.6. The quantitative results tested the M hypothesis that an M diglossic variant has emerged as the informal spoken language.

3.6 Conceptualising the interrelationship between FI and CI in SI

This section explains the theoretical interrelationship between FI, SI and CI. As noted in Section 3.1.1, the classification of FI, SI and CI into categorical variables allows for allocation of their respective linguistic elements into groups.

Each diglossic category contains its own linguistic features or Aspect of Language (AL). AL includes linguistic features such as phonological realisations, morphology and lexis. The FI Aspect of Language (FIAL) contains FI semantic properties, FI phonological realisations, FI morphological
markings and FI lexis. Conversely, the CI Aspect of Language (CIAL) contains CI phonological realisations, CI morphological markings and CI lexis. The coding $X^n$ denotes the standard base word (root word/infinitive) that it is present in both diglossic variants, although usage and semantic properties might differ.

FIAL is the established grammar of bahasa Baku that is taught in language textbooks and used in grammar references and dictionaries (Sneddon 1996, 2000; Quinn 2001). FIAL represents the official standard ‘Perfected Spelling and Pronunciation’ ([*Ejaan Yang Disempurnakan*] [EYD]) of Indonesian words (as opposed to regional and dialectal variances) prescribed by the official Indonesian language authority *Pusat Bahasa*. The morphologies are the standard affixes, such as *ber- $X^n$, me- $X^n$, me- $X^n$-kan*, *pe- $X^n*$, *pe- $X^n$-*, and *$X^n$-an*.

The base word or infinitive $X^n$ belongs to both FI and CI. It can be posited that many $X^n$ lexicon are more common to CI than FI. To give an example, the use of $X^n$ such as *beli* (to buy), *punya* (to have) and *masak* (to cook) is more common in everyday speech than the respective equivalent FIAL realisations of *membeli*, *mempunyai* and *memasak*. Nonetheless, in this study the $X^n$ has not been treated as non-FIAL. If it had been, the CIAL:SI would be much higher than the current results indicate. Historically, $X^n$ is most probably a linguistic phenomenon predating both FI and CI; therefore, in this research $X^n$ was treated as a set of lexicon universal to both FI and CI and categorically independent of the FI–CI dichotomy.

While the FI–CI dichotomy fulfils different sociolinguistic functions and in fact defines different registers, the two variants complement each other: thus the merging of the two in an M diglossic variant. The informal sociolinguistic domain, where there is a linguistic merging of FI diglossic variant ($\Delta$FI) with CI diglossic variant ($\Delta$CI) ($\Delta$FI $\cup$ $\Delta$CI), produces the M SI diglossic variant ($\Delta$SI) category that contains AL of both H & L variants (Figure 3).

![Figure 3](image)

*Figure 3.* The merging of H and L into the M variant ($\Delta$FI $\cup$ $\Delta$CI = $\Delta$SI)
The malleability of the use of language variants across different sociolinguistic domains creates a FI–SI–CI spectrum. Figure 3 shows the formal–informal spectrum. Usage of a pure form of FI falls at the formal end, and usage of pure CI falls at the informal end. The amount of CIAL determines the degree of formality or informality of SI in the sociolinguistic spectrum.

![FI--SI--------------------SI--CI](image)

*Figure 4. The formal-informal spectrum of FI–SI–CI*

### 3.6.1 Register

The understanding of the term *register* in this research aligns with Halliday’s definition, which states that register forms part of the general principles governing the use of language variation, and that its study aims to understand the situational factors that determine linguistic features (Lukin, Moore, Herke, Wegener, & Wu, 2011).

Halliday (1989) described register as ‘a variety of language, corresponding to a variety of situation’. Included in Halliday’s discourse on the *situation* of a register are the aspects of ‘field’, ‘tenor’ and ‘mode’, according to which field denotes the subject matter or the content discussed, tenor the relationship between the speakers and the purpose of the communication, and mode the channel of communication (e.g. direct speech, written language, video conference).

Furthermore, Halliday’s perspective on register is that it is located in the ‘semiotic spaces’ of field, tenor and mode. In other words, Halliday regards register as a set of semantic features that vary according to situation type and are activated by the contextual parameters of field, tenor and mode. Thus in Halliday’s definition of register, the wording or the lexico-grammar is the interface for situation and semantics (context → meaning → wording).

In the CI context of this research, Halliday’s understanding of the function of register would translate as:

informal sociolinguistic domain → CI semantics → CI lexico-grammar.

Thus, the register of CI lexico-grammar functions as the interface between informal sociolinguistic situations and the semiotic spaces of field, tenor and mode. The strength of engaging with Halliday’s
perspective on register is that it potentially provides a theoretical paradigm to zoom into the ‘semiotic spaces’ of the SI diglossic variant.

Nonetheless, for the purpose of this research, the use of the word ‘register’ in this thesis is for the most part limited to the context of Indonesian diglossic language variations; it primarily refers to the language of formal sociolinguistic situations and informal sociolinguistic situations, as outlined by Fergusson (Table 1). A task for future research could be to define the semiotic spaces (field, mode, and tenor) of each of Fergusson’s H–L diglossic situations for the Indonesian sociolinguistic context. In the research for this thesis the application of Halliday’s register discourse to the context of Indonesian diglossia was limited to Sneddon’s binary FI and CI categories in order to facilitate qualitative, quantitative, semantic and pragmatic comparisons between these two diglossic variants.

To give an example: if the sociolinguistic situation to be studied were a self-empowerment workshop, questions about the field may include whether the language was expositional or suggestive, while discussion on the tenor could explore the agentive role, social hierarchy and social distance of the individuals engaged in communication. The analysis of mode would explore the communication medium used in the interaction. While in the main discussion on SI the use of the term register is confined to ‘formal language’ and ‘informal language’, Appendix 2 elaborates on the semantic properties of CI lexicon and speech acts. This appendix contains a list of CI lexis and collocations, and the semiotic space it covers is the primary field’ in which the exposition and pragmatics of CI lexis, collocations and speech acts are discussed.

3.7 The Indonesian phonemes
The qualitative description of CIAL in Method 2 uses the system of phonetic symbols developed by the International Phonetic Association (IPA). The IPA symbols applied to the interpretation of CIAL are outlined in this section. These descriptions have been collated from existing research conducted by Sneddon (2006) and Kushartanti (2014), and from the IPA for Indonesian published by the International Phonetic Association (1999).

3.7.1 The IPA system applied to Indonesian
The IPA undertakes periodic revisions as new knowledge of spoken languages is acquired. It is the most comprehensive system used by modern linguists, as it represents the sounds of all known world languages (IPA, 1999). The IPA acknowledges the limitations of any system of symbols that tries to
capture sound: ‘Whilst in practice the distinction between what is linguistically relevant and what is
not may not always be clear cut, the principle of representing only what is linguistically relevant has
guided the provision of the symbols. The IPA does not provide symbols to indicate information such
as “spoken rapidly by a deep, hoarse, male voice”’ (IPA, 1999, p. 4).’

Selecting IPA symbols to represent accent and intonation within a description of Indonesian
phonology is, therefore, challenging. An oversimplification can produce an inaccurate phonological
representation of Indonesian, especially in the dynamics between FI and CI. If a standard
phonological formula is attempted based on an FI assumption, a subjective prescriptive transcription
will take place at the expense of a descriptive approach based on objective observation of the actual
spoken form of SI. The relative nature of the pronunciation of vowels is a reflection of the
interchangeability between phones. Where the /ɪ/ phone might be considered standard in words such
as ikut (to go along) or mengikuti (to follow somebody), in daily speech it might be realised as the [i].
Realisations of vowel-phonemes are mostly the same between FI and CI, but interchangeability
between vowel-phonemes does occur and is realised as various adjacent vowels between FI and CI.
This applies with the phones [i] and [ɪ]; [o] and [ɔ]; [a] and [ɑ]; [ɛ] and [ə]; [u] and [ʊ]. Moreover,
vowel phonetic realisations such as those from FI /a/ to CI [ɑ] or FI /ɛ/ to CI [ɛ] are not a matter of
uniform grammatical shifts from FI to CI but just a matter of self-expression and emphasis in the
pragmatics of daily speech.

The interchangeability of phones means that the orthography of Indonesian only has relative value. It
is often said that Indonesian is a phonetic language, meaning that each sound pronounced in the
language is consistently represented by the orthography (Clark & Yallop, 1995). However, this is not
entirely accurate. In Indonesian writing the letter ‘e’ represents the phonemes /ɛ/, /ɛ/ and /ə/; ‘o’
represents /ɔ/ and /ɔ/; ‘a’ represents /a/ and /ɑ/; ‘i’ represents /i/ and /ɪ/; and ‘u’ represents /u/ and /ʊ/.

While many words are sufficiently represented in phonetic terms by Roman orthography, it is often
the case that there is no correct or incorrect application of alternative phones in the pronunciation of
words, as it is often a matter of expression, register and intonation. Some tentative rules can be
derived to describe how phones may interchange between FI and CI in different syllable contexts. For
example, the letter ‘o’ is mostly the /o/ phoneme when it is the first letter in a word (e.g. [orang]
[person]); however, in the context between first-syllable consonants it is produced as the [ɔ] phone
(e.g. [kɔpi] [coffee]). In the case of the ‘i’ consonant, the /ɪ/ phoneme becomes [i] when it is in the
final position ([tapɪ] → [tapi]). More research needs to be done to record, identify and verify any
uniform syntactical and pronunciation patterns to the phonological changes between FI and CI, but
such a thorough undertaking is outside the scope of this thesis. For the purpose of this thesis, it is
deemed sufficient to present an awareness of the phonological variation between FI and CI. The only pronunciation variant that will be discussed in this thesis in later chapters will be the reduction of the phoneme /a/ realised as [ə], as in the case of dalam (FI) and [daləm] (inside) (CI).

The phonemes and phones discussed in this section do not take into account dialectal accents where regional-specific pronunciation of vowels might occur. For example, in the accent of a Batak person from North Sumatra, the word ‘lebaran’² would be pronounced [lebaran] while in standard FI it is pronounced [ləbaran]. Table 6 provides a list of Indonesian vowels with examples in FI (bahasa Baku) and CI.

Table 6

*The Indonesian Vowels*

<table>
<thead>
<tr>
<th>Latin alphabet</th>
<th>International phonetic alphabet</th>
<th>Examples in FI</th>
<th>Examples in CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a</td>
<td>aku, kasih,</td>
<td>kasih, masih</td>
</tr>
<tr>
<td></td>
<td></td>
<td>masih, makan</td>
<td></td>
</tr>
<tr>
<td>α</td>
<td></td>
<td>waduh!</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>ε</td>
<td>empat, sempat, t</td>
<td>empat, sempat, t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>empat</td>
<td>tempat</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
<td>Prefix: par-, tar-, mə-, bəsar, angkau</td>
<td>bəsar</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td>enak, sate, lempar</td>
<td>enak, sate, lempar</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>tidur, bisa</td>
<td>tidur, bisa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>lagi, tidur</td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>o</td>
<td>orang</td>
<td>ompong</td>
</tr>
<tr>
<td>o</td>
<td></td>
<td>kəpɪ, təpɪ,</td>
<td>səpir</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>totup</td>
<td>totup</td>
</tr>
<tr>
<td>u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td></td>
<td>bukan</td>
<td>ampon</td>
</tr>
</tbody>
</table>

---

² A Muslim celebration
3.7.2 The Indonesian vowels

Vowel categorisation depends mainly on the position of the tongue: the height of the tongue in the mouth (high, high-mid, low-mid or low) and the degree of fronting or retraction of the tongue (forward, centre or back). Lip position (unrounded spread/neutral, or rounded) also plays a role. These articulatory postures produce resonant acoustic frequencies called formant frequencies. The first two formant frequencies (F1 and F2) correlate with tongue fronting and backing, and tongue height, respectively. The third formant (F3) provides information about the presence of lip rounding (IPA, 1999).

Table 7.
*The Phonetic Table of Indonesian Vowels*

<table>
<thead>
<tr>
<th></th>
<th>Unrounded</th>
<th>Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
<td>Centre</td>
</tr>
<tr>
<td>Close</td>
<td>i / ɪ</td>
<td>u / ʊ</td>
</tr>
<tr>
<td>Close-Mid</td>
<td>e</td>
<td>ə</td>
</tr>
<tr>
<td>Low-Mid</td>
<td>ɛ</td>
<td>ɔ</td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

3.7.3 The Indonesian consonants

Most Indonesian words that begin with the /kh/ [x] and /sy/ [ʃ] phonemes are of Arabic origin. These two phonemes actually correspond to the Jawi (Arabic-Malay) script خ (khā) and ش (shīn) respectively: evidence of an Arabic linguistic contribution and a remnant from the period when Malay was written using the Jawi script. Many learned Indonesian Muslims today are still Arabic-literate. Together with the abundance of Arabic lexicon in the Indonesian language, this might explain the persistence, and possibly the permanent incorporation, of Arabic phonemes into Indonesian. Words of Arabic origin such as *bersukur*, pronounced [bəɾsukur], is also seen spelled as *bersyukur* and pronounced [bəɾʃukur] or [bəɾʃukur].
The Latin script does not represent the subtleties of the Jawi script; hence the consonant ‘h’ represents both the voiceless ح (ḥā) and the glottal ہ (hā). Only students of the madrasah\(^3\) might make that distinction of relevant Arabic-origin words containing these phonological shades.

The ‘f’, ‘q’, ‘v’, ‘x’ & ‘z’ consonants are also introduced (Ruijgrok, 2008), with ‘f’, ‘q’ and ‘z’ coming from Arabic and ‘x’ & ‘v’ from Dutch. Therefore, words containing these consonants are of foreign origin. Words containing /x/ are mostly names of foreign origin (e.g. from Mexico).

Under the Dutch, Malay was written using the Roman alphabet. Some Islamic texts continued to be written in the Jawi script by Islamic scholars during the colonial era. With the implementation of national education using the Latin script after 1949, the use of the Jawi script died out. Because /f/ is a phoneme that is traditionally foreign, historically many native speakers pronounce it as [p]: /fota/ (photo) is often pronounced [poto]. This speech characteristic is becoming less noticeable, however; the ‘f’ consonant has become a common phoneme. The consonant ‘v’ represents both the phonemes /v/ and /f/. Some native speakers pronounce /v/ as [f]. It must be stressed that the level of education of the native speaker often determines a person’s speech-production choices. With increasing study of English as a second language, many educated (notably middle-class urban) Indonesians have acquired English phones such as [θ] (thing) and [ð] (thus) (Ruijgrok, 2008). Nevertheless, this would be the exception rather than the rule. The phonemes of the consonant ‘k’ are /k/ when followed by vowel phonemes and the glottal stop [ʔ] when followed by consonant phonemes or occurring as the end of a word. Table 8 provides a list of Indonesian consonants, with examples in FI (bahasa Baku) and CI, and Table 9 classifies the consonants according to IPA.

\(^3\) Islamic school
### Table 8

*The Indonesian Consonants*

<table>
<thead>
<tr>
<th>Consonants/Phone</th>
<th>International phonetic alphabet</th>
<th>FI examples</th>
<th>CI examples</th>
<th>Examples in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b</td>
<td>bagus, benar</td>
<td>bagus, benar</td>
<td>big, blue</td>
</tr>
<tr>
<td>c</td>
<td>ʧ</td>
<td>ʧantik, (cantik)</td>
<td>ʧari, (cari)</td>
<td>chuck</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>dia, danau</td>
<td>dia, danau</td>
<td>door, duck</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>firdaus, fantasi, foto</td>
<td>firdaus, fantasi, foto</td>
<td>fiction, fun</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
<td>gugur</td>
<td>gede</td>
<td>gear, gap</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
<td>hukum</td>
<td>hati</td>
<td>hear</td>
</tr>
<tr>
<td>j</td>
<td>ʤ</td>
<td>ʤantung</td>
<td>ʤadi</td>
<td>job</td>
</tr>
<tr>
<td>k</td>
<td>k</td>
<td>kandang</td>
<td>kasih</td>
<td>cup</td>
</tr>
<tr>
<td>?</td>
<td>bapa?</td>
<td>ngga?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kh</td>
<td>x</td>
<td>ikhlas</td>
<td>khas</td>
<td>loch</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>lebih, lagi</td>
<td>lebih, lagi</td>
<td>leaf</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>masuk</td>
<td>main</td>
<td>more</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>naik</td>
<td>naik</td>
<td>near</td>
</tr>
<tr>
<td>ng</td>
<td>ñ</td>
<td>bangga</td>
<td>ɲapain</td>
<td>siñer</td>
</tr>
<tr>
<td>ny</td>
<td>ñ</td>
<td>ſaľa</td>
<td>ɲari</td>
<td>capon</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>Pakai</td>
<td>pake</td>
<td>paw</td>
</tr>
<tr>
<td>q</td>
<td>k</td>
<td>Qur’an</td>
<td>Qur’an</td>
<td>queen</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>Rasa</td>
<td>rasa</td>
<td>rake</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>Suka</td>
<td>suka</td>
<td>sink</td>
</tr>
<tr>
<td>sy</td>
<td>ſ</td>
<td>inřaf</td>
<td>inřaf</td>
<td>ſip</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>Tahan</td>
<td>tahan</td>
<td>tin</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
<td>volkano</td>
<td>variasi</td>
<td>vast</td>
</tr>
</tbody>
</table>
Table 9

*The Phonetic Table of Indonesian Consonants*

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Palatal-alveolar</th>
<th>Velar</th>
<th>Glottal</th>
<th>Post-alveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops - affricates</strong></td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td>g</td>
<td>ʧ / ʤ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flap/trill</strong></td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td>f</td>
<td>s / z</td>
<td>x</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td>m</td>
<td>n</td>
<td>ɲ</td>
<td>ɲ̊</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td>j</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.8 Conclusion
This chapter has presented the methodology and methods adopted in the research for this thesis, and outlines how MI is defined and measured as a language variant that contains both linguistic features of FI and CI. One of the objectives of this research was to test the M hypothesis. To this end, CIAL had to be identified and defined, and the prevalence of CIAL alongside FI needed to be measured in
the SI corpus. A qualitative method using IPA symbols was used to describe the phonological and morphological features of CI (CIAL), which in turn were quantitatively measured against the SI corpus to gauge how prevalent CIAL is in SI.

One area not covered by the research design was the intonation of CI. This research focused only on a word-level analysis of CI phonology, morphology and lexicon, omitting intonation, which is an equally important linguistic aspect of the semantic and pragmatic quality of CI as a natural spoken language. Such an undertaking could be the subject of a separate research project. In the next chapter the research findings are presented. These include an identification of the CI linguistic features within the corpus using the IPA system (research methods 1 and 2). Next, the quantitative result (CI:SI ratio) is given, which indicates the proportion of CI language in informal interactions (Method 3a). In Chapter 5, the results from survey instrument are presented. These results provide some indication of the use of CI and FI in informal settings from a population sample of native Indonesian speakers (Method 3b).
Chapter 4

The Linguistic Components and Sociolinguistic Nature of Spoken Indonesian: Findings and Discussion of Corpus Analysis
4. Introduction

This chapter presents and discusses the findings from the corpus analysis. The purpose is to test the M hypothesis that an SI exists, a middle Spoken Indonesian diglossic variant that is spoken in everyday informal situations. A language is defined as a middle diglossic variant when it consists of elements of both the high diglossic variant (in this case, FI) and the low diglossic variant (in this case, CI).

The premise of this thesis is that FI in its unadulterated form is not the informal spoken language and that the informal spoken language contains Colloquial Indonesian Aspect of Language (CIAL). To establish this, the research for this thesis sought to identify linguistic features AL (Aspect of Language) that are non-FI (non-FIA) in the everyday language corpus (SI), expressed in positive terms as CIAL. The lexicon and the phonological and morphological patterns of CIAL were then described and defined using the International Phonetic Alphabet. CIAL was also measured quantitatively as a proportion of the SI corpus (CIAL:SI) to gain an indication of the composition of SI in everyday informal speech. The 48 data sets for this study form a corpus of more than 14,000 words obtained from authentic written and media sources (contemporary literature, billboard advertisements, transcripts of interviews and audiovisual media) that represent informal, everyday situations.

4.1 Findings and discussion of the corpus analysis

4.1.1 Qualitative results: CIAL in SI corpus and description with the International Phonetic Alphabet

Data analysis showed that non-FIAL AL features are present in the SI corpus. These AL features also have sub-components. The non-FIAL features are: non-FIAL lexicon, non-FIAL morphological features, non-FIAL null parameter / ellipsis $\emptyset$, non-FIAL elisions, non-FIAL phonological realisations, and non-FIAL semantic properties. These non-FIAL features have now been positively constructed as CIAL.

The fusion of CIAL and FIAL in the SI corpora supports Sneddon’s (2006) hypothesis of the existence of an M diglossic variant in spoken Indonesian. Concurrently the notion that a pure form of FI is an informal spoken language can be rejected. CIAL can be positively verified as an integral part of the everyday informal language.

The next qualitative method involved identifying and isolating every non-FIAL item, analysing the structures and describing them using the IPA. The result constitutes the paradigms of CIAL that demarcate the diglossic boundary between SI and FI. These CIAL consist of CI lexicon, CI
morphological features, null parameter/ellipsis $\emptyset$, elisions, CI phonological realisations and CIA L semantic properties. The list below is a summary of CIA L that were identified in this research:

I. Word class ellipsis/null elements ($\emptyset$) in the syntax of daily speech

There are three notable common null elements in spoken Indonesian syntax:

a. the personal pronoun ellipsis in structures such as:

CI: $\emptyset$ Mau $\emptyset$ ke mana?

$\emptyset$-pro aux-mau $\emptyset$- verb-X$^n$ prep-ke wh-mana?

You want to go to where?

FI: Anda mau pergi ke mana?

‘Where are you going?’

CI: $\emptyset$ Enggak mau.

$\emptyset$-pro neg-Enggak aux-mau

I not want

FI: Saya tidak mau

‘I don’t want to.’

b. the adalah copula ellipsis in nominative structures such as:

CI: Bapak $\emptyset$ kepala desanya di sini.

pro-Bapak $\emptyset$- cop NP-kepala desaDET-nya prep-di NP-sini

Mister is head village-the in here

FI: Bapak adalah kepala desanya di sini

‘He is the village head here.’

c. the common null element parameter of the predicate pergi in phrases such as:

CI: $\emptyset$ lagi $\emptyset$ ke mana?

$\emptyset$-pro aux-lagi $\emptyset$-verb-X$^n$ prep-ke wh-mana?
II. Morphological features

a. ‘ng’ (/ŋ/) -X^n

This can be regarded as an allomorphy of the active me- prefix or as an active predicate morphological marker in its own right. This /ŋ/ phonemic form is an allomorphic speech habit and is a linguistic injection from the regional languages. The example ngopi also demonstrates the predication of a NOUN X^n that does not occur in FI:

CI:     ngopi [FI: minum kopi]

ŋ-X^n-(k)-opi

to drink coffee

CI:     ngirim [FI: mengirim]

ŋ-X^n-(k)-irim

to send

b. X^n-’in’ (/in/)

This morph replaces both FI’s predicate suffixes, ‘-kan’ and ‘-i’. It encompasses all the grammatical functions that these FI suffixes impart (accusative, dative-benefactive, dative-causative):

CI:     bikinin [FI: buatkan]

X^n-bikin-in [+benefactive]

to make something for somebody
CI:  *benerin*  
[FI:  *benarkan*]

\(X^n-bən-ə-r-in\)  [+causative]

to correct something

c. ‘-ny’ (/ŋ/) - \(X^n\)

Like the /ŋ/ phoneme, /ŋ/ is an allomorphic active prefix of *me* (or a proper morph) that operates on base words with first letters ‘c’ and ‘s’. Some examples include:

CI:  *nyuci*  
[FI:  *mencuci*]

\(ŋ-X^n-(ŋ)-uci\)

to wash; washing

CI:  *nyebar*  
[FI:  *menyebar*]

\(ŋ-X^n-(s)-əbar\)

to spread; spreading

d. ‘ng’ (/ŋ/) - \(X^n\) ‘-in’ (/–In/)

This is the active form of 1.3b. It is the CIAL variation of FI’s *me- X^n–kan* and *me-X^n–i*. The example *ngapain* is a predication of WH- lexical item *apa* and has two semantic values:

CI:  *ngapaiin*  
[FI:  *sedang apa? untuk apa?*]

\(ŋ-wh-apa-in- CONT-TENSE \ [+interrogative]\)

‘What are you doing?’ ‘What for?’

CI:  *ngebeliin*  
[FI:  *membelikan*]

\(ŋə-X^n-bəli-in\)  [+dative + benefactive]

to buy something for somebody
e. ‘-ny’ (ɲ/)-Xⁿ-‘-in’ (/–In/)

This is the active form of 1.3c. It is the CIAL variation of FI’s me-Xⁿ–kan and me-Xⁿ–I for base words with first letters ‘c’ and ‘s’. Some examples are:

CI: nyediain  
    [FI: menyediakan]
    ɲ-Xⁿ-(s)-ədia-in  
    [+dative + benefactive]
    to prepare something for somebody

CI: nyariin  
    [FI: mencarikan]
    ɲ-Xⁿ-(ʧ)-ari-in  
    [+dative + benefactive]
    to look for something for somebody

f. ‘ke-’ (/kə-/) Xⁿ-‘-an’ /-ən/

These are the alternative CI [+excessive] adverbial marker to FI’s adverb terlalu. Examples include:

CI: kegedean  
    [FI: terlalu besar]
    kə -Xⁿ-gəde-an  
    too large

g. Xⁿ-‘-an’ /-ən/

This affixation is a CI alternative to the FI adverb lebih [+comparative]:

CI: bagusan  
    [FI: lebih bagus]
    Xⁿ-bagus-an  
    nicer, better
III. Elisions, allomorphy and phonological variations to FI

a. Elision of first letters ‘s’ and ‘h’ in some common words. Some examples include:

CI:  *ama*  

\( X^n - (s) - ama \)  

with

CI:  *abis*  

\( X^n - (h) - abis \)  

finished

b. Elision of prefix me- (or /m/-\( X^n \) allomorphy) in active verbs with first letter ‘p’.

Examples:

CI:  *make*  

\( X^n - (p) - m - ak - e \)  

to use

c. Phonetic realisation [e], [ə] or [ɛ] – in place of second syllable ‘a’ vowel in the /a/ phoneme:

CI:  *item*  

\( X^n - (h) - it - e - m \)  

black

d. Phonetic realisation [e], [ə] or [ɛ] - in place of second syllable ‘a’ vowel in place of the /ai/ diphthong:

CI:  *make*  

\( X^n - (p) - m - ak - e \)  

to use
e. The [o] phone substitutes for the ‘u’ vowel:

- **CI:** sorga  
  - [FI: surga]
  
  - *s-o-r*ga
  
  - heaven

f. The [o] phone substitutes for the /au/ diphthong:

- **CI:** ijo  
  - [FI: hijau]
  
  - X\textsuperscript{o}-(h)-ij-o
  
  - green

IV. An existing array of alternative lexical features to FI, which are often preferred in speech rather than the ΔFI variants. Some examples include:

- **[CI: enggak; FI: tidak]**  
  - no, do not

- **[CI: cuma; FI: hanya]**  
  - only

- **[CI: pake VP segala?; FI: kenapa harus VP?]**  
  - why VP

- **[CI: mendingan; FI: lebih baik]**  
  - it is better to …

- **[CI: pengen; FI: ingin, mau]**  
  - to want

V. The frequent use of discourse particles that are absent in FI:

- **kok**  
  - [+interrogative]

- **deh**  
  - [+agreement]

- **sih**  
  - [+affirmative]

- **dong**  
  - [+ request +affirmative]

- **loh**  
  - [+interrogative]

- **mah**  
  - [+declarative]

- **nah**  
  - [+affirmative]
VI. The common use of constructions employing tag questions:

CI:  

<table>
<thead>
<tr>
<th>Word</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagus</td>
<td>nggak?</td>
</tr>
</tbody>
</table>

adj-bagus  neg/tag-nggak?

adj-good  neg/tag-not?

FI:  

<table>
<thead>
<tr>
<th>Word</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagus</td>
<td>atau tidak?</td>
</tr>
</tbody>
</table>

‘Is it good?’

CI:  

<table>
<thead>
<tr>
<th>Word</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucu</td>
<td>kan?</td>
</tr>
</tbody>
</table>

adj-lucu  tag-kan?

adj-funny  tag-kan?

FI:  

<table>
<thead>
<tr>
<th>Word</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucu</td>
<td>benar?</td>
</tr>
</tbody>
</table>

‘Funny, wasn’t it?’

VII. Variation in semantic properties of Indonesian lexica that are not traditionally recognised in prescriptive FI grammar. Some examples are:

[X^n jalan] CI [+V] (to go)  
FI [+N] (street)  [+V] (to walk)]

[X^n buat] CI [+prep] (for)  
FI [+V] (to make)]

[X^n biar] CI [+CP] (so that)  
FI [+V] (to let be)]

[X^n mau] CI [+aux +tense] (will)  
FI [+aux +modal (to want)]

[X^n suka] CI [+aux +tense] (habitually do …)  
FI [+aux +modal] (to like)]

[X^n pada] CI [+pronominal plural marker]  
FI [+prep] (on, at)]

4.1.2 Quantitative results: CIAL:SI

The quantitative method involved counting every lexical item with CIAL markings in each of the data samples in the corpus and statistically analysing these in terms of the CIAL:SI corpora ratio. The alternative hypothesis is constructed to measure a CIAL:SI ratio greater than zero. That is, \( H_1 \) is represented by CIAL:SI >0. SPSS produced a mean CIAL:SI ratio of .39, which means that \( H_1 \)
(CIAL:SI >0) can be accepted and that CIAL form a significant part of the spoken language in. Figure 5 illustrates the spread of the data samples, each as a CIAL:SI ratio. It visually displays the ratios for all corpora in their data-set categories. The majority of data samples had CIAL:SI ratios below .40, and the majority of data samples with CIAL:SI ratios above .40 were in the audiovisual and billboard categories (Figure 5).

Figure 5. CIAL:SI ratios for all data sets. Data-set categories are audiovisual (AV), billboard (BB), literature (LIT) and recordings of interviews (RI).

The next analysis correlated the dimensions of the CIAL:SI ratios of the data sets (Table 10) with the sociolinguistic spectrum of the FI–SI–CI continuum (Figure 6). The dimensions of the correspondence analysis were translated as intervallic variables in the informal spectrum to show the spread of the data samples. The 0 on the far left of the spectrum represents zero presence of CIAL, while the 1 on the far right of the spectrum represents usage containing exclusively CIAL. The bottom indicator marks the percentage the dimensions occupy as data sets from the corpus. Figure 6 represents the quantitative findings of the research located along the sociolinguistic continuum of SI (Sneddon 2006, Djenar & Ewing, 2015).
Table 10.

**Correspondence Analysis of all Data Sets**

<table>
<thead>
<tr>
<th>Data set (n)</th>
<th>1 (&lt;.2)</th>
<th>2 (.2 to .29)</th>
<th>3 (.3 to .49)</th>
<th>4 (.5 to .7)</th>
<th>5 (&gt; .7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB (12)</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LIT (16)</td>
<td>7</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AV (14)</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>RI (6)</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>% of corpus</td>
<td>0.02</td>
<td>0.21</td>
<td>0.56</td>
<td>0.17</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Note. BB = billboard; LIT = literature; AV = audiovisual; RI = recordings of interviews.*

![FI ---- SI --------- SI -------- SI ---- SI ---- CI](image)

**Figure 6.** The spread of data on the FI–SI–CI sociolinguistic spectrum.

Figure 6 shows that none of the corpora fell at the extreme ends of the intervallic scale (0 or 1), indicating that neither FI nor CI in their pure forms are used as informal spoken languages. The shaded range covering dimensions <.2 to >.7 contains most of the corpora data samples, with one RI data sample falling in the <.2 dimension and one AV and one BB sample falling in the >.7 dimension (Table 10). Data sets in the dimension of .3 to .49 form the largest proportion (56%) of the corpus (Figure 6), suggesting that a sociolinguistic spectrum with a CIAL:SI ratio from .3 to .49 is the most commonly encountered semiformal sociolinguistic situation.

4.2 **The sociolinguistic implications of the findings**

This section illustrates conceptual models of the dynamics of Indonesian diglossia. These abstractions are extrapolations inferred from the data and findings of this research, and from observations made by Errinton (1986), Sneddon (2003a, 2003b, 2006), Djenar (2006) and Djenar & Ewing (2015). The
following conceptual models are theory-building exercises that need to be further tested. They set the background for further research on Indonesian diglossia.

In the Indonesian sociolinguistic domain (σδ) spectrum, CIAL act as independent variables that determine the degree of formality within the σδ spectrum by virtue of the speakers’ relative freedom and choice to employ FI or CI AL. Thus, in Indonesian diglossia, AL function as the independent variable, while the σδ is the dependent variable, its degree of formality being determined by the speakers employing AL to suit pragmatic intentions (Figure 7).

![Diagram](image.png)

*Figure 7. Indonesian diglossia: sociolinguistic domain (σδ) as a function of Aspect of Language (Aλ)*

This differs from the traditional Fergussonian (1959) view of diglossia, according to which the sociolinguistic domain (σδ) determines the linguistic range within which speakers can operate. Thus, AL (Aλ) is the dependent variable that operates as a function of the independent sociolinguistic (σδ) variable (Figure 8):
Figure 8. Traditional Fergussonian diglossia: Aspect of Language (Aλ) as a function of sociolinguistic domain (σδ)

How do these abstractions translate into real-life situations? Table 1 depicted Fergusson’s concept of how Sociolinguistic domains are supposed to be translated into sociolinguistic situations in real life. According to Fergusson’s theory of diglossic languages, the sociolinguistic situation strictly dictates the speaker’s choice of language. As mentioned previously, in the case of the Indonesian language situation there is a switching of the independent variable between the speaker (the agent of AL) and the sociolinguistic situation. In real-life Indonesian speech, the use of FI is often not conducive to these social ends and sounds unnaturally stilted, perhaps due to familiarity, informality and the importance of relationship building (Sneddon 2003b). Thus, speakers do not adhere to the sociolinguistic boundaries that are supposed to contain the language variants within their respective. What traditionally is considered a formal social context might be imbued with AL to reduce formality. Fergusson’s diglossia (see Table 1) would translate differently into Indonesian diglossia (Table 11) due to the malleability of language use. Factors of culture as well as sociopolitics (e.g., democratisation), ideology (e.g., egalitarianism, communalism) and social media, to name just a few, all have a bearing upon the ‘relaxation’ of language use. The distribution of FI and SI sociolinguistic usage in Table 11 is extrapolated from the corpus, as well as from observations made by Errington (1986), Sneddon (2003a, 2003b, 2006), Djenar (2006) and Djenar & Ewing (2015). Since data from this study indicated that, in practical terms, pure CI no longer really exists as a language in Indonesian σδ (Figure 7), a CI column is omitted and the L diglossic variant is represented by the M column.
Table 11

**Indonesian Diglossia**

<table>
<thead>
<tr>
<th>Type of communication</th>
<th>FI (H)</th>
<th>SI (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sermon in church or mosque</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Instruction to servant, workmen, subordinates</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Personal letter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political speech</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Business communication (written)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business communication (spoken)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>University lecture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University tutorial (discussion groups)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Teacher – student discussion in class</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Teacher – student conversation outside class</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Conversation with friends, family, colleagues</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>News broadcast</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Films, soap operas, comedy shows</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Newspaper editorial, news story, caption on picture, political cartoon</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Contemporary novels; men’s, women’s, youth magazine</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Advertisement</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

*Note. H = ‘high’ diglossic variant; M = ‘middle’ diglossic variant.*

4.3 **Conclusion**

The premise of this thesis is the assertion that Indonesian is a diglossic language (Errington, 1986) (Sneddon, 2003b, 2006) consisting of Formal Indonesian (FI) – the H diglossic variant – and Colloquial Indonesian (CI), the L diglossic variant. Moreover, Sneddon (2003b) postulated that H and L might be merging into an M diglossic variant. Thus, this research frames the M hypothesis – testing whether an M diglossic variant exists – using data in the form of a corpus of authentic sources (not generated for pedagogy) that represents the language spoken in everyday, informal situations (an SI corpus). To test the M hypothesis, an alternative hypothesis (CIAL:SI) was framed to measure the prevalence of CI language features in the SI corpus.

Four interrelated research aims thus formed the basis of the research outlined in this chapter:

1. To document that a spoken informal Indonesian language exists that contains linguistic features outside the paradigm of the standard grammar of FI. These linguistic features are subsequently identified as the linguistic features of CI (CIAL).
2. To identify and describe the linguistic features of Colloquial Indonesian using the International Phonetic Alphabet (IPA).

3. To test the Spoken Indonesian (SI) M hypothesis.

4. To determine the prevalence and sociolinguistic scope of CIAL in SI as a spoken language.

The data consists of a corpus of over 14,000 words from 48 data sets across four media categories (contemporary literature, billboard advertisements, transcripts of interviews and audiovisual media). The patterns of CIAL from these data were identified and depicted using the IPA. Corpus data using authentic sources of everyday spoken and written Indonesian support Sneddon’s observations that standard Indonesian FI has merged with CI to form an informal spoken M diglossic variant, SI. A CI:FI ratio was applied to measure the CIAL component in the data sets. The majority of data sets (94%; 45 of 48) fall within a CIAL:SI ratio range of <.2 to >.7. The results of this research indicate an overall CIAL:SI mean ratio of .39, with a range of 58 (.17 to .75).

The attributes of CI as a natural spoken language give SI its quality as an informal spoken language. Findings of CI in news reports (Appendix 6) also show that the sociolinguistic scope of SI has reached beyond the traditional range that a non-H variant is supposed to occupy. CIAL has made it ‘upward’ into the domain of FI, into sociolinguistic domains that were traditionally, or supposedly, reserved for FI. A comparison of classical Fergussonian diglossic language H-L use (see Table 1) with that of Indonesian diglossia (Table 11) shows that SI has a wider sociolinguistic range than that of the classical H-L juxtaposed model. This introduces the notion that SI functions on a continuous spectrum (Figure 6).

This chapter presented the findings of the corpus. The next chapter presents and discusses the findings of the data generated from the questionnaire-survey research instrument. The questionnaires were distributed among native Indonesian speakers, and this data provides information about frequency of usage and the demographic scope of CI across different parts of Indonesia.
Chapter 5

The Frequency and Demographic Range of Spoken Indonesian: Findings and Discussion
5. Introduction

This chapter presents the findings of a questionnaire (Appendix 7) that was distributed to 98 native Indonesian speakers \((n = 98)\) from 35 different regions/cities. The purpose was to examine the extent of CI usage and its spread across different parts of Indonesia. The questionnaire asked participants about their language choices (FI–CI) in informal everyday language usage and collected demographic and geographic information. Since CI has its origins in Jakarta, it was expected that its usage would score highly in Jakarta and surroundings. Of particular interest was the extent of CI usage outside of its native origin, Jakarta, in regional areas and cities. The questions, designed to simulate everyday speech acts in everyday situations, were composed by the author based on the findings of this research and utilising the author’s native-speaker competency.

5.1 The demographic reach of CI: Questionnaire

A component of this research (Method 3b) involved gathering information using a questionnaire as the research instrument. Questionnaires were distributed to 98 native Indonesian speakers in Australia between 2015 and 2016. These individuals were Indonesian students at Australian universities and Indonesians at social and community events (festivals and social gatherings). The research instrument was given ethics clearance by the University of the Sunshine Coast (July 2015: S/15/760).

The participants were de-identified, although the questionnaire gathered information about their age, occupation, education and residence. The participants came from 35 different regions/cities across Indonesia. Finding out where participants live is of special interest, as it provides an indication of the geographical extent of CI usage in regional Indonesia, beyond its linguistic origins in Jakarta. In view of this question, participants from Jakarta and the surrounding satellite cities or regions (JABOTABEK–Jakarta–Bogor–Tanggerang–Bekasi) are collated separately. These places, including Bogor, whose indigenous population is the Sundanese ethnic group, come under the linguistic influence of the capital in the use of informal spoken Indonesian because of their proximity to the city. The data were organised into two tables for analysis: one for Jakarta and surroundings, and another for regional Indonesia. The rationale of this classification is that Indonesian, and, for that matter, CI–SI, is a mother tongue and sole everyday language of the Jakarta region, whereas in regional Indonesia bahasa Daerah (BD) (regional languages) are the mother tongue and often also the social language outside the working context. By separating these two categories, the expected high-CI content of the Jakartan population would not raise the overall average of CI statistics across all data sets, and it allowed the use of CI outside Jakarta to be analysed separately.
The main part of the questionnaire posed 18 questions to native Indonesian participants asking them if they would use FI or CI in informal everyday situations. The 18 questions gave participants choices of FI phrases and CI phrases for a particular informal scenario. An example of a question is:

‘If you are asking a friend/colleague what they are doing:
   a. (FI phrase)       b. (CI phrase)’.

From these choices the participants indicated their usage in everyday non-formal situations. After the 18 FI–CI multiple-choice questions, participants were asked whether on a daily basis they use Indonesian or a regional language. These questions asked the participants to indicate their usage of Indonesian (BI) vis-à-vis regional languages (BD). In this way, information was gathered on how much CI is used by people whose mother tongue is not Indonesian. In today’s Indonesia, Indonesian is the mother tongue only to people born in certain urban areas, while a regional language (bahasa Daerah) would be the mother tongue of those who are born and raised in regional areas.

Another section of the questionnaire required the participants to indicate if they would find themselves more often in formal or informal language situations. An intervallic scale of 1 to 7 was used, with 4 being the midpoint. The 3.5–4.5 range was used to indicate interaction in a balanced formal-informal social situation (Figure 9).

```
<table>
<thead>
<tr>
<th>Formal language</th>
<th>Balanced usage</th>
<th>Informal language</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

*Figure 9. The intervallic scale of the research instrument*

Participants were also given writing space to comment on the use of SI and state their opinion about the questionnaire. The last question asked participants to indicate whether their answers reflected ‘the reality of everyday spoken Indonesian’ or ‘how Indonesian ought to be spoken’. This question determined the participants’ notion of what informal spoken language is.

### 5.2 Jakarta data: analysis and discussion

The participants from Jakarta scored highly in the use of CI (CI .91) in their daily language usage (Table 12). Shaded cells indicate higher FI/CI usage.
Table 12  
*FI and CI Usage by the Jakarta Participants*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Gender</th>
<th>Edu.</th>
<th>Occupation/Area of work</th>
<th>Context of language use</th>
<th>FI</th>
<th>CI</th>
<th>FI–CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>JKT 1</td>
<td>65</td>
<td>F</td>
<td>Sec.</td>
<td>Cook</td>
<td>BI&amp;BD</td>
<td>.11</td>
<td>.89</td>
<td>6.5</td>
</tr>
<tr>
<td>JKT 2</td>
<td>66</td>
<td>M</td>
<td>Sec.</td>
<td>Lab.</td>
<td>BI&amp;BD</td>
<td>.17</td>
<td>.83</td>
<td>7</td>
</tr>
<tr>
<td>JKT 3</td>
<td>37</td>
<td>F</td>
<td>Tert.</td>
<td>Admin.</td>
<td>BI</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>JKT 4</td>
<td>27</td>
<td>F</td>
<td>Tert.</td>
<td>Admin.</td>
<td>BI</td>
<td>.33</td>
<td>.67</td>
<td>5</td>
</tr>
<tr>
<td>JKT 5</td>
<td>27</td>
<td>F</td>
<td>Tert.</td>
<td>Admin.</td>
<td>BI</td>
<td>.11</td>
<td>.89</td>
<td>6</td>
</tr>
<tr>
<td>JKT 6</td>
<td>32</td>
<td>F</td>
<td>Tert.</td>
<td>Admin.</td>
<td>BI&amp;BD</td>
<td>.06</td>
<td>.94</td>
<td>5</td>
</tr>
<tr>
<td>JKT 7</td>
<td>29</td>
<td>F</td>
<td>Tert.</td>
<td>NGO staff</td>
<td>BI</td>
<td>.11</td>
<td>.89</td>
<td>7</td>
</tr>
<tr>
<td>JKT 8</td>
<td>23</td>
<td>F</td>
<td>Tert.</td>
<td>NGO staff</td>
<td>BI</td>
<td>.06</td>
<td>.94</td>
<td>7</td>
</tr>
<tr>
<td>JKT 9</td>
<td>20</td>
<td>F</td>
<td>Tert.</td>
<td>Student</td>
<td>BI</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>JKT 10</td>
<td>45</td>
<td>F</td>
<td>Tert.</td>
<td>Teacher</td>
<td>BI</td>
<td>1</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>JKT 11</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>BI</td>
<td>.06</td>
<td>.94</td>
<td>6</td>
</tr>
<tr>
<td>JKT 12</td>
<td>77</td>
<td>M</td>
<td>Tert.</td>
<td>Retired</td>
<td>BI</td>
<td>.61</td>
<td>.39</td>
<td>4.5</td>
</tr>
<tr>
<td>JKT 13</td>
<td>44</td>
<td>F</td>
<td>Tert.</td>
<td>Home</td>
<td>BI</td>
<td>.11</td>
<td>.89</td>
<td>4</td>
</tr>
<tr>
<td>JKT 14</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>BI</td>
<td>.35</td>
<td>.65</td>
<td>2</td>
</tr>
<tr>
<td>JKT 15</td>
<td>28</td>
<td>M</td>
<td>Tert.</td>
<td>Superv.</td>
<td>BD</td>
<td>.11</td>
<td>.89</td>
<td>7</td>
</tr>
<tr>
<td>JKT 16</td>
<td>35</td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>.06</td>
<td>.94</td>
<td>3</td>
</tr>
<tr>
<td>JKT 17</td>
<td>45</td>
<td>M</td>
<td>Tert.</td>
<td>Public serv.</td>
<td>BI&amp;BD</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>JKT 18</td>
<td>42</td>
<td>F</td>
<td>Tert.</td>
<td>Employee</td>
<td>BI</td>
<td>.11</td>
<td>.89</td>
<td>6</td>
</tr>
<tr>
<td>JKT 19</td>
<td>45</td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>.06</td>
<td>.94</td>
<td>6</td>
</tr>
<tr>
<td>JKT 20</td>
<td>26</td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>JKT 21</td>
<td>47</td>
<td>F</td>
<td>Masters</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>.28</td>
<td>.72</td>
<td>4</td>
</tr>
<tr>
<td>JKT 22</td>
<td>30</td>
<td>F</td>
<td>Masters</td>
<td>Public serv.</td>
<td>BI</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>JKT 23</td>
<td>25</td>
<td>F</td>
<td>Masters</td>
<td>Business</td>
<td>BI</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 13 below is a summary of the data from table 12.

Table 13

**Table 13 Summary Statistics on language use in Jakarta area**

<table>
<thead>
<tr>
<th>Location</th>
<th>Gender</th>
<th>Occupation</th>
<th>Language Use</th>
<th>Home</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>JKT 24</td>
<td>F</td>
<td>Tert. Business</td>
<td>BI BI BI</td>
<td>0 1 6</td>
<td></td>
</tr>
<tr>
<td>JKT 25</td>
<td>F</td>
<td>Tert. Public serv.</td>
<td>BI BI BI</td>
<td>.17 .83 4</td>
<td></td>
</tr>
<tr>
<td>TGRNG1</td>
<td>F</td>
<td>Tert. Admin.</td>
<td>BI&amp;BD BD BI&amp;BD</td>
<td>.06 .94 6</td>
<td></td>
</tr>
<tr>
<td>TGRNG2</td>
<td>F</td>
<td>Tert. Vet.</td>
<td>BI BI BI</td>
<td>1 7</td>
<td></td>
</tr>
<tr>
<td>TGRNG3</td>
<td>F</td>
<td>Tert. Admin.</td>
<td>BI&amp;BD BI BI BI</td>
<td>.6 .4 4.5</td>
<td></td>
</tr>
<tr>
<td>BEKASI</td>
<td>F</td>
<td>Tert. Industry</td>
<td>BI&amp;BD BI&amp;BD BI&amp;BD</td>
<td>1 4</td>
<td></td>
</tr>
<tr>
<td>Bogor 2</td>
<td>M</td>
<td>Sec. Industry</td>
<td>BD BI&amp;BD BI</td>
<td>.4 .6 6</td>
<td></td>
</tr>
<tr>
<td>Bogor 3</td>
<td>M</td>
<td>Tert. Publisher</td>
<td>BD BI BI</td>
<td>.89 .11 1</td>
<td></td>
</tr>
<tr>
<td>Bogor 4</td>
<td>F</td>
<td>Postgr. Publisher</td>
<td>BI BI BI</td>
<td>.22 .78 6</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** FI (Formal Indonesian), CI (Colloquial Indonesian), BI (Indonesian), BD (regional language), shaded data = higher usage.
The high prevalence of CI in Jakarta was to be expected, because CI, which has its roots in the Jakarta Betawi Malay dialect, is the mother tongue of the population born in the capital. Despite this, only 58% of the Jakarta participants indicated that they used only Indonesian (BI) at home, while 30% indicated that they used both BI&BD at home, and 12% indicated they used only BD. This reflects the high number of regional migrants who have resettled in the capital and continue to use their regional mother tongue at home.

Eleven participants (33%) from Jakarta scored 1 (100%) in their choice of CI usage in daily informal interaction, while 26 participants (78%) scored a high usage of CI above .8. Thirty per cent of participants indicated that they interact more in formal or balanced formal–informal environments every day, while 70% indicated that they interact more in informal environments. A 30% daily formal language use suggests that factors such as professional settings require people to communicate in FI in their daily interaction.

Three participants (9%) scored a higher usage of FI. Two of these were above middle age (53 & 77 years old), and one was 30 years old (Tangerang 3). From a sociolinguistic perspective it is understandable for post-middle age participants to use more FI where conservative values associated with generational factors and seniority incline speakers towards ‘proper speech’. A speaker’s use of FI in daily informal interaction could be a display of civility and refinement. Using FI with strangers establishes distance and is combined with appropriate intonation and politeness. The familiarity of CI is often reserved for family, friends and colleagues, or when a person wants to demonstrate solidarity within a particular class.

5.2.1 Comments by participants from Jakarta
The reactions and attitudes towards FI and CI reflected in the participants’ written comments are revealing. For many participants this exercise was a personal revelation about their own use of language. They were amused and surprised to discover that their use of CI is actually not standard or ‘proper’ Indonesian.

Participant Jakarta 5 remarked, as if surprised, ‘In actual fact the daily language we use is an informal language, and we are so used to it’ (‘ternyata bahasa yang kita gunakan sehari-hari adalah informal dan kita sudah terbiasa dengan itu’). Participant Jakarta 7 said, ‘Knowingly or not, we indeed do not use good and proper Indonesian’ (Sadar atau tidak, kita memang tdk menggunakan bahasa Indonesia yg baik & benar). The participants’ lack of awareness that their language use is not ‘standard’ was a recurring theme. As Participant Jakarta 19 observed, ‘We feel normal when using everyday informal
language, but when it is explained with FI comparisons, it feels funny’ (Sewaktu memakai bahasa sehari-hari informal kita merasa biasa saja tetapi sewaktu diperjelas dengan dibandingkan bahasa Indonesia yang baku, menjadi terasa lucu).

All Jakartan participants indicated that they had answered the questionnaire according to ‘the reality of how Indonesians speak in daily interaction’ (kenyataannya bagaimana orang Indonesia berbicara sehari-hari). Participant Jakarta 9 commented that the questionnaire was amusing because ‘it is rare to see where FI is still used in our surroundings’ (sangat jarang melihat sekitar masih menggunakan bahasa baku).

The comparison between FI and CI does not seem to have generated discussion among the participants. Participant Jakarta 22 remarked that the survey was ‘interesting, considering that it is rare for anyone to be researching or studying the use of formal Indonesian language and everyday informal language’ (Menarik, mengingat jarang yang meneliti atau mempelajari pemakaian bahasa Indonesia yang baku dan sehari-hari [non baku]).

To some participants, such as Jakarta 20, the survey was amusing ‘because the everyday language that we use is seldom discussed as the subject of a question’ (Karena bahasa sehari’yg kami gunakan jarang digunakan sebagai topik pertanyaan). And to participant Jakarta 18, writing spoken language felt ‘strange’: ‘Spoken language is different from the written language. When the spoken language is written, it feels rather strange’ (Bahasa percakapan berbeda dengan bahasa tulisan. Ketika percakapan sehari-hari ditulis, jadi agak aneh rasanya).

Participant Jakarta 12, the only participant who scored higher in FI use (.61) than in CI use (.39), believed that his answers were explained by the fact that ‘Indonesian is spoken every day’. The participant, who is 77 years old, also perceived that Indonesian is constantly changing and that youth culture is a contributing factor: ‘The use of everyday Indonesian evolves through the ages. Every new generation, especially the youth, creates new words. For example, the word galau [headache, entangled thoughts, stress] has only been used in the last two years’ (Pemakaian bahasa Indonesia sehari2 berkembang setiap zaman. Setiap generasi baru terutama anak2 remaja menciptakan kata2 baru. Umpamanya kata ‘galau’ [pusing, pikiran rumit, stress], baru dua tahun terakhir).

The survey was revealing for some participants. Participant Jakarta 15 remarked that ‘as an Indonesian I just realised that the language that is used every day is not standard formal Indonesian’. (sebagai orang Indonesia saya baru sadar kalau bahasa yang digunakan sehari-hari bukan E.Y.D.). Similarly, Participant Jakarta 23 remarked how ‘without realising it, we use a spoken language daily
that is far different from standard formal Indonesian. For this reason, this research is interesting’ (Kita menggunakan bahasa sehari yang tanpa disadari sgt jauh berbeda dgn bhs EYD. Jd sgt menarik yg diamati). Participant Tanggerang 3, who scored higher in FI use (.6) than in CI use (.4), remarked that the survey was ‘providing an awareness and understanding of the Indonesian language that we use in everyday life and whether it is in accordance with the (proper) grammar or is more like everyday spoken Indonesian’ (lebih menyadarkan quizioner lebih memahami penggunaan bahasa Indonesia yang kita pakai dalam kehidupan sehari-hari apakah sesuai dengan tata bahasa atau bahasa Indonesia percakapan sehari-hari).

5.3 Regional Indonesia data: Analysis and discussion

While regional languages are still spoken, the position of Indonesian in regional Indonesia as a daily language is well established. Movement of people from one part of Indonesia to another is a significant contributing factor in the greater use of Indonesian in regional Indonesia in recent decades (Steinhauer, 1994). The term ‘regional Indonesia’ does not necessarily refer to remote areas, though it may. A city like Yogyakarta can hardly be considered ‘regional’ in the broader meaning of the term. Here, ‘regional’ refers to the regions outside Jakarta’s immediate cultural sphere (Table 13). This might be difficult to define, since Jakarta’s cultural sphere extends to all parts of Indonesia through the power of mass media and the internet. However, for the purpose of this research, satellite areas that geographically surround Jakarta (JABOTABEK–Jakarta–Bogor–Tangerang–Bekasi) were considered to be under the immediate geographic cultural influence of Jakarta (see Table 13). The research pre-empted a lower result in the use of CI in this regional category. This is because many people in the region, especially in non-urban areas, speak Indonesian as a second language. It used to be that school was their first encounter with Indonesian, which meant that FI was the first language they encountered. This is no longer true due to mass media and the internet. Most regional young people today are much more comfortable using both Indonesian and CI than were their parents when they were the same age.

Table 14
FI and CI usage among participants from regional Indonesia

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Gender</th>
<th>Edu.</th>
<th>Occupation/Area of work</th>
<th>Context of language use</th>
<th>FI</th>
<th>CI</th>
<th>FI–CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDAN 1</td>
<td>64</td>
<td>M</td>
<td>Tert.</td>
<td>Business</td>
<td>At home, Socially, At work</td>
<td>.5</td>
<td>.5</td>
<td>6</td>
</tr>
<tr>
<td>MEDAN 2</td>
<td>30</td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>.3</td>
<td>.7</td>
<td>5</td>
</tr>
<tr>
<td>Area</td>
<td>Type</td>
<td>Gender</td>
<td>Level</td>
<td>Field</td>
<td>Employ.</td>
<td>Ex.</td>
<td>Exp.</td>
<td>Num.</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>----------------</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>PALEMB 1</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Home</td>
<td>BI</td>
<td>BI&amp;BD</td>
<td>N/A</td>
<td>.33</td>
</tr>
<tr>
<td>BANTEN 1</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>BANTEN 2</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Publisher</td>
<td>BI</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>.6</td>
</tr>
<tr>
<td>BIMA 1</td>
<td></td>
<td>M</td>
<td>Dipl.</td>
<td>Business</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>.5</td>
</tr>
<tr>
<td>BANDUNG 1</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Public serv.</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>BI</td>
<td>.17</td>
</tr>
<tr>
<td>BANDUNG 2</td>
<td></td>
<td>F</td>
<td>Postgr.</td>
<td>Teacher</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>.1</td>
</tr>
<tr>
<td>BALI 1</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>.17</td>
</tr>
<tr>
<td>BALI 2</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>0</td>
</tr>
<tr>
<td>Cianjur</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Vet.</td>
<td>BD</td>
<td>BI</td>
<td>BI</td>
<td>.17</td>
</tr>
<tr>
<td>Cirebon</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Industry</td>
<td>BD</td>
<td>BD</td>
<td>BD</td>
<td>0</td>
</tr>
<tr>
<td>Deli Serdan</td>
<td></td>
<td>M</td>
<td>Postgr.</td>
<td>Teacher</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>.5</td>
</tr>
<tr>
<td>Demak</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Teacher</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.3</td>
</tr>
<tr>
<td>Gorontalo</td>
<td></td>
<td>M</td>
<td>Postgr.</td>
<td>Teacher</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>.9</td>
</tr>
<tr>
<td>YOGYA 1</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Graphic design</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>YOGYA 2</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Editor</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.33</td>
</tr>
<tr>
<td>YOGYA 3</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Business</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.11</td>
</tr>
<tr>
<td>YOGYA 4</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Teacher</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>1</td>
</tr>
<tr>
<td>YOGYA 5</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.28</td>
</tr>
<tr>
<td>YOGYA 6</td>
<td></td>
<td>F</td>
<td>PhD</td>
<td>Lecturer</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>.3</td>
</tr>
<tr>
<td>YOGYA 7</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Factory</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>YOGYA 8</td>
<td></td>
<td>F</td>
<td>Masters</td>
<td>Business</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.06</td>
</tr>
<tr>
<td>YOGYA 9</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>YOGYA 10</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Business</td>
<td>BI</td>
<td>BI</td>
<td>BI</td>
<td>.12</td>
</tr>
<tr>
<td>Jember</td>
<td></td>
<td>F</td>
<td>Tert.</td>
<td>Home</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>BI</td>
<td>.06</td>
</tr>
<tr>
<td>Jombang</td>
<td></td>
<td>M</td>
<td>Postgr.</td>
<td>Principal</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.5</td>
</tr>
<tr>
<td>Klaten</td>
<td></td>
<td>F</td>
<td>Postgr.</td>
<td>Teacher</td>
<td>BD</td>
<td>BI</td>
<td>BI&amp;BD</td>
<td>.47</td>
</tr>
<tr>
<td>KALTENG</td>
<td></td>
<td>M</td>
<td>Postgr.</td>
<td>Publisher</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>.2</td>
</tr>
<tr>
<td>Mataram 1</td>
<td></td>
<td>M</td>
<td>Masters</td>
<td>Teacher</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>BI</td>
<td>.11</td>
</tr>
<tr>
<td>Mataram 2</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Teacher</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>Mataram 3</td>
<td></td>
<td>M</td>
<td>Postgr.</td>
<td>Lecturer</td>
<td>BD</td>
<td>BI</td>
<td>BI</td>
<td>.27</td>
</tr>
<tr>
<td>Mataram 5</td>
<td></td>
<td>M</td>
<td>Tert.</td>
<td>Teacher</td>
<td>BI</td>
<td>BD</td>
<td>BI&amp;BD</td>
<td>0</td>
</tr>
<tr>
<td>Area</td>
<td>No.</td>
<td>Gender</td>
<td>Level</td>
<td>Qualification</td>
<td>BI &amp; BD</td>
<td>BD</td>
<td>BI</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>--------</td>
<td>------------</td>
<td>-------------------------</td>
<td>---------</td>
<td>------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Mataram</td>
<td>30</td>
<td>F</td>
<td>Tert. Teacher</td>
<td>BI &amp; BD</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mataram</td>
<td>30</td>
<td>F</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Mataram</td>
<td>33</td>
<td>M</td>
<td>Postgr. Lecturer</td>
<td>BI &amp; BD</td>
<td>.44</td>
<td>.56</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mataram</td>
<td>29</td>
<td>F</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>0</td>
<td>1</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Makasar</td>
<td>48</td>
<td>M</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>.6</td>
<td>.4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Makasar</td>
<td>40</td>
<td>F</td>
<td>Tert. Business</td>
<td>BI &amp; BD</td>
<td>.28</td>
<td>.72</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Mojokerto</td>
<td>42</td>
<td>M</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>.7</td>
<td>.3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Narmada</td>
<td>27</td>
<td>M</td>
<td>Tert. Teacher</td>
<td>BI &amp; BD</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>NTB</td>
<td>50</td>
<td>M</td>
<td>Postgr. Publisher</td>
<td>BD &amp; BI</td>
<td>.3</td>
<td>.7</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Sumatra</td>
<td>30</td>
<td>M</td>
<td>TAFE House</td>
<td>BI &amp; BD</td>
<td>.6</td>
<td>.4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Padang</td>
<td>41</td>
<td>M</td>
<td>Postgr. Publisher</td>
<td>BI &amp; BD</td>
<td>.7</td>
<td>.3</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>47</td>
<td>M</td>
<td>Postgr. Business</td>
<td>BI &amp; BD</td>
<td>.44</td>
<td>.56</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Papu</td>
<td>52</td>
<td>M</td>
<td>Sec. Educ.</td>
<td>BI &amp; BD</td>
<td>.94</td>
<td>.06</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>31</td>
<td>M</td>
<td>Bachelor. Educ.</td>
<td>BI &amp; BD</td>
<td>.56</td>
<td>.44</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>55</td>
<td>M</td>
<td>Postgr. Educ.</td>
<td>BI &amp; BD</td>
<td>.94</td>
<td>.06</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>47</td>
<td>M</td>
<td>Dipl. Educ.</td>
<td>BI &amp; BD</td>
<td>.89</td>
<td>.11</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>50</td>
<td>M</td>
<td>Bachelor. Educ.</td>
<td>BI &amp; BD</td>
<td>.94</td>
<td>.06</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>40</td>
<td>M</td>
<td>Bachelor. Educ.</td>
<td>BI &amp; BD</td>
<td>0</td>
<td>1</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>50</td>
<td>M</td>
<td>Postgr. Educ.</td>
<td>BI &amp; BD</td>
<td>.89</td>
<td>.11</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>45</td>
<td>F</td>
<td>Bachelor. Teacher</td>
<td>BI &amp; BD</td>
<td>.72</td>
<td>.28</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>54</td>
<td>F</td>
<td>Postgr. Educ.</td>
<td>BI &amp; BD</td>
<td>.78</td>
<td>.22</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>41</td>
<td>M</td>
<td>Postgr. Educ.</td>
<td>BI &amp; BD</td>
<td>.61</td>
<td>.39</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>37</td>
<td>M</td>
<td>Postgr. Educ.</td>
<td>BI &amp; BD</td>
<td>.78</td>
<td>.22</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Papua</td>
<td>57</td>
<td>M</td>
<td>Postgr. Admin.</td>
<td>BI &amp; BD</td>
<td>.44</td>
<td>.56</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Pontianak</td>
<td>45</td>
<td>M</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>.84</td>
<td>.16</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Purwokerto</td>
<td>44</td>
<td>F</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>.4</td>
<td>.6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Riau</td>
<td>45</td>
<td>M</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>1</td>
<td>0</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Samarinda</td>
<td>57</td>
<td>M</td>
<td>PhD Publisher</td>
<td>BD &amp; BI &amp; BD</td>
<td>.23</td>
<td>.77</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Semarang</td>
<td>43</td>
<td>M</td>
<td>Postgr. Teacher</td>
<td>BI &amp; BD</td>
<td>.33</td>
<td>.67</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Semarang</td>
<td>53</td>
<td>M</td>
<td>Tert. Teacher</td>
<td>BI &amp; BD</td>
<td>.7</td>
<td>.3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>SUMBAR</td>
<td>32</td>
<td>F</td>
<td>Postgr. Publisher</td>
<td>BI &amp; BD</td>
<td>.4</td>
<td>.6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Surabaya</td>
<td>25</td>
<td>F</td>
<td>Masters Business</td>
<td>BD &amp; BD</td>
<td>.11</td>
<td>.89</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
Table 15 below is a summary of the data from Table 14.

Table 15

<table>
<thead>
<tr>
<th>Summary Statistics on Language Use in Regional Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speakers using more FI than CI</strong></td>
</tr>
<tr>
<td>32%</td>
</tr>
<tr>
<td><strong>Speakers using more CI than FI</strong></td>
</tr>
<tr>
<td>61%</td>
</tr>
<tr>
<td><strong>Speakers using balanced FI-CI</strong></td>
</tr>
<tr>
<td>7%</td>
</tr>
<tr>
<td><strong>Speakers interacting more in formal environments</strong></td>
</tr>
<tr>
<td>24%</td>
</tr>
<tr>
<td><strong>Speakers interacting in balanced formal/informal environments</strong></td>
</tr>
<tr>
<td>20%</td>
</tr>
<tr>
<td><strong>Speakers interacting more in informal environments</strong></td>
</tr>
<tr>
<td>56%</td>
</tr>
<tr>
<td><strong>Speakers who use only BI at home</strong></td>
</tr>
<tr>
<td>28%</td>
</tr>
<tr>
<td><strong>Speakers who use only BD at home</strong></td>
</tr>
<tr>
<td>18%</td>
</tr>
<tr>
<td><strong>Speakers who use BI&amp;BD at home</strong></td>
</tr>
<tr>
<td>54%</td>
</tr>
<tr>
<td><strong>Speakers who use only BI socially</strong></td>
</tr>
<tr>
<td>39%</td>
</tr>
<tr>
<td><strong>Speakers who use only BD socially</strong></td>
</tr>
<tr>
<td>6%</td>
</tr>
<tr>
<td><strong>Speakers who use BI&amp;BD socially</strong></td>
</tr>
<tr>
<td>55%</td>
</tr>
</tbody>
</table>

Note: FI (Formal Indonesian), CI (Colloquial Indonesian), BI (Indonesian), BD (regional language).

The use of CI still outweighs that of FI in informal settings in regional Indonesia (.61 to .32, with .07 participants indicating 50–50 CI–FI usage) (Table 13). As in Jakarta and surrounds, age seemed to be a determining factor, with 85% of participants younger than 40 indicating they use more CI, while among participants aged over 40 years, only 40% indicated that they used more CI than FI as an informal spoken language. This suggests that the younger generation in regional Indonesia is more inclined to speak in CI.

The use of BI at home in regional Indonesia (82%) – in a traditionally BD domain – and in social situations (94%) indicates that there has been an increase in the use of Indonesian in the region. This is probably the result of regional Indonesian cities becoming more cosmopolitan, where there is movement of people with multi-ethnic backgrounds. Families of inter-ethnic marriages in regional cities may be more inclined to use Indonesian as the home language. The impact of popular youth culture on Generation Y from traditional BD language environments may also make them more
inclined to speak Indonesian at home. All participants who stated they speak a regional language at home and with their friends indicated that Indonesian was the main language at work. Again, this may reflect the multi-ethnic composition of the modern Indonesian work environment, even in remote regions such as Papua.

Some of the participants (32%) indicated that they use more FI as the informal language, with 80% working as educators or in education. The inclination to use FI among this professional group is not surprising, given that they represent institutions committed to upholding FI as the ‘good and proper language’ (bahasa baku yang baik dan benar).

Another factor affecting CI seems to be provinciality. These data suggest that the island of Lombok marks the eastern linguistic boundary of the area in which CI is prevalent. One hundred per cent of participants from Lombok indicated that they use more CI than FI, whereas only 26% of participants from islands east of Lombok indicated that they use more CI than FI. Some of these regions are known to have their own dialect of Malay. Kupang in West Timor is one example. While there are no data from this research from Kupang, Errington (2014) noted how Kupang Malay is commonly used in informal settings, as opposed to other forms of Indonesian. These provincial regions would be equally exposed to CI through media and internet, but perhaps the strong sense of regional identity that is represented by their local Malay dialect prevents CI from gaining adherence among locals. Why would it be that a region like Lombok readily absorbs CI, while another like Kupang does not? A possible reason might be that the local language of Lombok is the non-Malay Sasak. While still belonging to the Austronesian language tree, Sasak is only distantly related to Malay, and for most people from Lombok and Kupang, the language of the other place is unintelligible. Thus CI is an addition to their linguistic repertoire that poses no existential threat to the local language, Sasak, because it does not alter or affect it. The two languages remain separate, independent linguistic entities. However, in the case of Kupang Malay, which is still a Malay-based language, incorporating the Jakartan CI into their local Malay dialect will alter and undermine (locals may say ‘erode’) the character and linguistic independence of their own Malay variant, posing an existential threat to the uniqueness and integrity of their local language and, hence, their identity.

5.3.1 Comments by participants from regional Indonesia

The contention that Indonesians do not categorically distinguish between FI and CI is in line with some of the responses. Many participants seemed surprised to be ‘caught out’ in their choices of language use. Participant Yogya 5 wrote that the questionnaire was ‘interesting’ and that it was ‘the first time’ she had ‘filled in a survey about everyday language use’ (Cukup menarik pertanyaannya
Participant Yogya 1, who scored 100% in CI use, commented that ‘in daily reality people are more comfortable with the non-formal language when communicating in their environment’ (secara realitas sehari-hari orang lebih nyaman menggunakan bahasa non-formal dalam berkomunikasi di lingkungannya). Participant Yogya 2 commented that ‘there are many words in usage that, although I did not realise it are not standard formal Indonesian’ (Ada banyak penggunaan kata yang tanpa sadar bukan bahasa Baku). Participant Makassar 2 also expressed surprise: ‘The variety of answers in this survey made me choose dialogues that I haven’t been conscious of, and that makes it interesting’ (Karena pilihan jawaban pada angket ini bervariasi dan menempatkan saya akan pilihan percakapan yg tidak saya sadari dan itu menarik). Likewise, Participant Surabaya admitted that she realised her Indonesian language was ‘no longer in accordance with the proper grammar system’ (Ternyata bahasa Indonesia saya sdh tdk sesuai dengan kaedah tata bahasa... Hahaha... Baru deh sadar!!!). Participant Banten 1 noted that the survey made him ‘aware of the reality of the language’ that he was ‘using every day’ (saya menjadi sadar tentang kenyataan penggunaan bahasa saya sehari-hari). Participant Mataram 5 commented that the questionnaire was both entertaining and interesting. It was entertaining because ‘the [CI] choices were very “young”’ and ‘interesting’ because he didn’t realise that his everyday language was ‘like that’ (lucu karna pilihannya bahasa muda dan menarik karna saya tidak sadar kalau bahasa sehari-hari ternyata seperti itu).

Other participants also supported the claim that CI is more common in natural speech and remarked that FI is not really used as a spoken language. Participant Papua 10 commented that CI is widely used every day across all ages: ‘Colloquial Indonesian is used by older people, young people and children’ (Karna penggunaan Bahasa Indonesia dan gaul diucapkan baik org tua, muda, anak-anak). Participant Yogya 9 supported the notion that CI is the more common informal language, ‘because in everyday life the formal language is very seldom used in speaking’ (karena dalam kehidupan sehari-hari, bahasa formal sangat jarang digunakan secara lisan). Similarly, Participant Yogya 3 found the survey interesting because ‘Indonesians have forgotten standard formal Indonesian language’ (Orang Indonesia telah lupa menggunakan bahasa Baku). Participant Bali 1 simply stated that she was ‘not used to using standard formal Indonesian’ (Tidak terbiasa menggunakan EYD). Similarly, participant Papua 13 commented that the survey was ‘interesting because the (Indonesian) language use does not necessarily follow proper spelling’ (Menarik karena belum tentu bahasa yang kita gunakan belum tentu sesuai ejaan yang disempurnakan / kaidah bahasa Indonesia yang baik dan benar). Participant Mataram 7 commented that the questionnaire was interesting because he ‘realised that formal Indonesian as a spoken language’ is ‘not pleasant to listen to, and understanding the vocabulary of informal Indonesian makes it easier to use it in context’ (Menarik karena tanpa disadari pemakaian
Conversely, Participant Papua 14 noted that ‘CI is not commonly used’ in her home town of Jayapura: ‘Instead people have their own version of an informal language’ (Di Jayapura-Papua jarang menggunakan Bahasa sehari-hari di atas, jadi rasa lucu. Contoh: minta dong jadi minta kah, dst). Some participants commented that the relative use of FI and CI depended on the social situation. Participant Yogya 7 wrote: ‘Without knowing it, we speak Indonesian in accordance with whoever we speak to’ (Secara tidak sadar kita berbahasa Indonesia sesuai dgn siapa kita bicara). Participant Mataram 3 also noted that the use of some of the expressions in the questionnaire depended ‘on context and situation – who speaks to whom, about what, in what situation’ (ada pengungkapan yang tergantung pada situasi dan kondisi – kepada siapa berbicara dan apa kondisinya). Participant Papua 6 stated that he used more FI in everyday interaction and acknowledged that CI and regional languages were ‘used by people as an everyday language, but in particular settings’ (Bahasa yang digunakan adalah bahasa baku sebagai Bahasa nasional dan ada tempat tertentu untuk menggunakan Bahasa sehari-hari sebagai Bahasa ibu/Bahasa daerah). Participant Mataram 6 commented that she found the questionnaire interesting because she ‘almost never’ used ‘formal Indonesian’ when she speaks. She had only just realised this fact. The participant said she used ‘Formal Indonesian only in formal situations and when talking to non-native Indonesian speakers’ ([Menarik] karena selama ini saya hampir tidak pernah memakai bahasa Indonesia yang baku saat berbicara. Baru nyadar [saya hanya menggunakan bahasa Indonesia baku dalam situasi formal dan saat berbicara dengan non-bahasa Indonesian speakers]).

Another participant (Papua 7), who indicated that he used more FI than CI, commented that ‘good and proper Indonesian will become extinct because it is not seriously guarded and taught at school and due to the prevalence of the use of dialects in the cities as well as in the regions’ (Bahasa Indonesia yang sempurna (EYD) di Indonesia sudah akan punah karna tidak serius dijaga dijarkan di sekolah2 secara baik karna dialek baik di kota maupun di daerah).

Participant Papua 8 indicated 100% usage of CI in everyday interaction. The participant perceived that CI represents the everyday language (Question 22). This participant also expressed that the survey was interesting (Question 23) because it demonstrated how ‘the citizens of Indonesia do not know how to use good and proper language’ (Sebagai warga Negara Indonesia tidak bisa mempergunakan Bahasa Indonesia yang benar/baik). Participant Mataram 2, who also scored 100% in usage of CI in everyday interaction, commented that the questionnaire was interesting because of
his background as an Indonesian teacher and because he saw the difference between what he was teaching students (Formal Indonesian) and ‘the spoken language that is actually in use’ (Karena latar belakang saya sebagai guru. Cukup menarik melihat perbedaan apa yang saya ajarkan [bahasa Indonesia Formal] dengan apa yang beredar di masyarakat sekarang). Similarly, Participant Mataram 9 commented that the questionnaire ‘showed real language usage by Indonesians, which is far from the proper language rules and regulation (grammar) as directed by the KBBI (Official Indonesian Language Dictionary)’ (Pertanyaan menunjukan kenyataan penggunaan bahasa yang dipakai orang Indonesia. Bahasa yang jauh dari ketentuan atau aturan berbahasa yang baik menurut KBBI).

The discrepancy between how Indonesian is taught and how it is spoken was also noted by Participant Mataram 4, who remarked that ‘in the teaching of BIPA (Indonesian for non-L1 learners) the teachers stress bahasa Baku, yet in everyday speech people emphasise the use of informal Indonesian (bahasa sehari-hari). Nonetheless, to be able to use bahasa sehari-hari, students must also know standard formal Indonesian’ ([Menarik] karena dalam pembelajaran BIPA, guru lebih menekankan pada bahasa Indonesia baku, tetapi dalam percakapan sehari-hari orang Indonesia lebih menekankan penggunaan bahasa sehari-hari. Akan tetapi untuk mampu menggunakan bahasa sehari-hari siswa harus tahu bahasa baku). Participant Mataram 10 commented that ‘the reality of teaching (Indonesian) in class differs from what is actually used in daily life and conversation (standard versus non-standard)’ (Kenyataan mengajar di kelas dan yang digunakan dalam percakapan kehidupan sehari-hari berbeda [standard vs non standard]). Putting it even more bluntly, Participant Narmada 1 remarked that it was ‘hilarious to read the formal spoken Indonesian options in the questionnaire, yet this is what we teach in BIPA (Indonesian language for non-native speakers)’ (Lucu sendiri ketika membaca bentuk formalnya padahal itu yang kita ajarkan di BIPA).

5.4 Conclusion

The questionnaire results show that CI is consistently and extensively used as an informal spoken language. The research instrument generated data from native Indonesian speakers about their daily language use (FI–CI) and has given some indication of the demographic and geographic spread of CI. The data show that CI is the dominant language in and around Jakarta, and that it features strongly in regional Indonesia. The data indicate that age and occupation are significant factors in determining FI–CI usage. Most teachers or education-related professionals stated that they preferred to use FI, while those aged under 40 leaned heavily towards CI. Commentaries from participants indicated that many Indonesians are surprised and bemused at the suggestion that they speak in FI. Some
participants who teach Indonesian to foreigners even conceded that, although they teach FI to foreigners, they actually would not use it themselves in everyday interactions.

This chapter concludes the presentation of the main findings of this research with regard to establishing the linguistic and sociolinguistic paradigms of SI in the context of diglossia. The next chapter will propose a pedagogic design that incorporates the teaching of CI. The teaching of FI to foreign learners of Indonesian is well established, but students are very rarely exposed to CI. Both FI and CI serve their respective sociolinguistic roles, and an understanding of both is necessary for students to acquire well-rounded language skills. The next chapter applies the knowledge of CI gained from this research to the development of a pedagogic model to complement the teaching of FI, as a contribution to a more holistic approach to Indonesian language teaching.
Chapter 6

Teaching Spoken Indonesian – the Practical Application of this Research
6. **Introduction**

The chapter explores the practical application of the findings of this research – which indicate that Spoken Indonesian (SI) is the contemporary spoken language - and how it can be incorporated into a pedagogy of the Indonesian language.

6.1 **Towards the teaching of SI**

This research is motivated by the misrepresentation of the Indonesian language in textbooks and pedagogy. FI is often depicted as the exclusive language, including the language of everyday informal interaction. This depiction results in a spoken language that is stilted and unnatural. This adversely affects language acquisition by foreign learners of Indonesian. The practical purpose of this thesis is to investigate and define the informal spoken language that has produced a description of SI. The last section of this chapter is dedicated to this purpose and presents the practical dimension of this research: to incorporate SI into the pedagogy of the Indonesian language. The following section analyses the current state of Indonesian textbooks, compares the CI content of the informal spoken language with the findings of this research, and provides a pedagogical model that incorporates SI.

The point of an inclusive pedagogy of SI is that it increases the foreign student’s ability to engage in meaningful social interactions imbued by culture. This ability to interact across cultural boundaries is also referred to as intercultural competence (Byram, 2003; Deardorff, 2011). It is the assertion of this thesis that FI is limited in facilitating the ability to interact across the culture of the Indonesian Target Language (TL). Competency in CI increases the language and cultural competency of foreign students learning Indonesian.

A dynamic view of culture prompts a reconceptualisation of the pedagogical approach to L2 acquisition (Liddicoat & Crozet, 2000). Language learning in itself is meaningless in isolation; that is, if it does not include the study of TL culture to equip the students to interact across the cultural boundaries of the TL. Knowledge in FI alone will not equip the foreign learner to the TL culture. This is because FI is a constructed administrative and technical language. It does not contain the organic expressions and speech acts that CI possesses.

CI, on the other hand, evolved as a spoken language and contains the language features that reflect the pragmatics, semantics and the psychology of the speakers. Thus, a foreign learner who becomes competent in CI will need to venture into the TL schema. Learning about the TL schema goes deeper than just learning about the appearances of a culture, i.e. the arts and social sciences, which have been the focus of learning Indonesian through FI. The premise of an SI pedagogy that incorporates CI
alongside FI is to broaden and deepen the understanding of Indonesian culture beyond its socio-artistic dimension. According to this premise the foreign student learns to ‘think like a native speaker’ in order to speak like a native speaker. This is a deeper process of traversing cultural boundaries and involves understanding the TL from the perspective of the Target Culture (TC).

Naturally, intercultural competence is not automatically acquired merely during the language learning process. Learning CI outside the TC does not result in native speaker competency. A pedagogy that includes CI only provides exposure and the linguistic foundation, which can then be developed when the student is immersed in the TC environment. In-country experience is necessary to develop fluency and the acquisition of language skills.

### 6.2 The state of Indonesian language teaching in Australia

In the Australian context, the motivation of this research was to develop an Indonesian language pedagogic model which aligns with the Australian Government 2012 White Paper: *Australia in the Asian Century*. The White Paper categorises Indonesian as one of Australia’s national priority languages. It emphasises the need for greater accessibility to language learning for Australian students to develop Asian relevant capabilities, as the need to engage with Asia is paramount to securing Australia’s future economic and strategic interests. The capacity for Australia to build deeper ties with Asia will be hampered if there is not an increase in proficiency in languages other than English. Proficiency in more than one language is a basic skill of the 21\(^{st}\) century (Australian Government, 2012, p. 170).

The National Objectives 10 and 11 of the White Paper state that:

10. Every Australian student will have significant exposure to studies of Asia across the curriculum to increase their cultural knowledge and skills and enable them to be active in the region.

11. All Australian students will have the opportunity, and be encouraged, to undertake a continuous study in Asian language throughout their years of schooling. (Australian Government, 2012, p. 170)

The Australian Government White Paper states that the National Objectives 10 and 11 are to be implemented through the Australian Curriculum\(^4\).

---

\(^4\) At the time of writing, the author of this thesis is employed by the Department of Education and Training Queensland as the C2C assistant manager responsible for the development of Indonesian language resources for the Australian Curriculum.
6.3 Prescriptive grammar versus descriptive grammar

Understanding the two opposing views on the term *grammar* – traditional prescriptive grammar and modern descriptive grammar – is fundamental to understanding the grammatical perspective of this thesis. From the traditional point of view, grammatical ‘correctness’ is that which adheres to the ‘proper’ rules, while modern descriptive grammar eschews the concept of correctness in favour of the view that language is an ever-changing rule-based system in which the rules are violated and reinvented by users of the language.

Prescriptive grammar has been the dominant system used in language teaching and learning. Most written world languages throughout history have been subject to the regime of prescriptive grammar, while descriptive grammar is a modern phenomenon that arose through academic – quantitative and qualitative – study of language(s). Modern linguists challenge the ‘absoluteness’ of prescriptive grammar. The following quote from Fromkin’s account of the beginnings of English prescriptive grammar reflects this perspective:

In 1762 Bishop Robert Lowth wrote *A short introduction to English grammar with critical notes*. Lowth prescribed a number of new rules for English, many of which were influenced by his personal taste. Before the publication of his grammar, practically everyone – upper class, middle class and lower class – said *I don’t have none, You was wrong about that* and *Matilda is fatter than me*. Lowth, however, decided that ‘two negatives make a positive’ and therefore one should say *I don’t have any*; that even when *you* is singular it should be followed by the plural *were*; and that *I not me, he not him, they not them* and so forth should follow *than* in comparative constructions. Because Lowth was influential... many of these new rules were legislated into English grammar, at least for the prestige dialect. (Fromkin, 2005, pp. 11–12)

The descriptive systems of Transformational-generative grammar (Noam Chomsky) and Systemic Functional grammar (Michael Halliday) have become the standard systems of linguistic analysis. With the democratisation of Indonesia, there appears to be an emerging academic desire for more ‘linguistic democracy’, an outlook that is shared in this thesis. Sneddon recounts that even linguists in Indonesia are voicing their disaffection with the status quo of prescriptive grammar:

Indonesian language will continue to suffer as a subject so long as spoken Indonesian is not taught. Unfortunately the education system is highly dominated by conservative linguists who are steadfast in maintaining the status quo that only Formal Indonesian should be taught. One prominent Indonesian linguistic scholar, Harimurti Kridalaksana, has placed the blame for the
situation on the attitude in the education system that formal language is the only variety deserving positive evaluation. He has stated that there should be a more positive view of non-standard forms of the language, calling for the claim that formal language is applicable in all situations to be done away with. Instead it should be acknowledged that it has limited functions, being complementary to colloquial speech (Sneddon, 2003b, p. 141)

As the only variety of the language associated with education, formal Indonesian is the only variant usually considered worthy of being taught to foreigners. The teaching of the spoken language to foreigners would be next to unimaginable to most native-speaking teachers of the language. This is due to three factors. The first factor is that the attitude among educators is for the most part a reflection of the ‘traditional prescriptive grammarians’ who refuse to acknowledge the prevalence and relevance of spoken Indonesian. The common perception among educators of Indonesian is that it is better for foreign students to err on the correct or even formal side than to employ careless, incorrect or substandard usage (Kohler & Mahnken, 2010). This perception is held notwithstanding the fact that informal speech is used by native Indonesian teachers themselves when it is appropriate to the social situation (Sneddon, 2003b). This prescriptive attitude has directly generated the other two factors of non-representation of spoken Indonesian. The second factor is the lack of teaching and learning resources for spoken Indonesian, which in turn results in the third factor: the lack of fluency in spoken Indonesian among foreign teachers. This deficiency can also be partly blamed on Indonesians when hosting foreigners who are learning the language in Indonesia. Sneddon has made a cultural observation:

Indonesians have few expectations that foreigners visiting the country will learn their language. When they meet a visitor making an effort, they are usually most pleased and encouraging. With the addition of a cultural avoidance of criticism, Indonesians tend to be effusive in praise of the struggling foreigner’s attempt at their language. *Sudah Lancar* (you’re fluent) is heard by a great many foreign visitors who can produce a few hesitant phrases. In comparison with native speakers of English and French, who expect proficiency in their language from visitors and who require high standards before they consider using words like ‘fluent’, Indonesians are extremely tolerant. This leads many learners to believe that they actually are good at the language (Sneddon, 2003b, p. 17).

Steinhauer (1994) also warned that flattery by Indonesians to foreigners that their Indonesian is fluent is to be viewed with some scepticism. This tolerance of Indonesians, plus a culture of avoidance of criticism, has also adversely affected the quality of Indonesian language-learning textbooks. Often it is foreign initiative and capital that produces Indonesian language-learning textbooks in Australia. These foreign authors may have consulted with native-speaking Indonesian parties, who, more often
than not, give sub-standard professional advice or fail to rectify fundamental shortcomings in language content that the foreign authors may propose.

The negative effect of non-fluency amongst foreign teachers of Indonesian, combined with a pedagogy that does not cater for the teaching of spoken Indonesian, trickles down to the students. The effect of an FI-centric pedagogy is that Indonesian language students in Australia are only aware of formal Indonesian and are for the most part poorly informed about the spoken Indonesian language. When visiting Indonesia they encounter a language that is substantially different to what they have learnt at school or university. This is of course disappointing, as one would expect to be able to communicate with ease after many years of studying the language. ‘Students then express dissatisfaction with the way they have has been taught, claiming that real Indonesian is being ignored’ (Sneddon, 2003b, p. 19).

The imposition of FI in L2 Indonesian teaching has partly been responsible for the decline of Indonesian teaching in Australia. An inorganic curriculum cannot be dynamic, nor will it be interesting, because it does not reflect the organism of a living language. Due to this lack of dynamism, learning Indonesian can be mundane (Kohler & Mahnken, 2010). The FI-centric pedagogy contrives conversation and frames a youth culture that is unreal or half-true at best. This inadequate modelling of the richness of Indonesian through a pedagogy skewed towards FI has resulted in negatives for both contemporary Indonesian culture and foreign learners of the target language. It impacts negatively on contemporary Indonesian culture, which is denied a transparent window through which it can represent itself, and on the foreign learner, as it draws the curtain on the vibrant linguistic and cultural reality of contemporary Indonesia. Thus, CI mostly remains an unknown quantity for the foreign student, and this helps to maintain a cultural gap between Indonesia and those in Australia who endeavour to learn about Indonesia. Therefore, the current pedagogic status quo of FI has created a misconception that only ‘very formal Indonesian is actually Indonesian. The colloquial language used by millions of Indonesians every day is either not recognised as existing or is regarded as something other than a form of Indonesian’ (Sneddon, 2003b, p. 18).

The FI taught to foreign students is primarily a formal language used mainly for administrative purposes, with little capacity to convey the emotions of everyday experiences. It has created a misconception amongst some that Bahasa Indonesia is a language bereft of emotions. ‘Anyone who has spent time with Indonesians can hardly be anything but puzzled at claims that their language is turgid, humourless and bereft of emotions’ (Sneddon, 2003b, p. 18).
At its very core, the motivation of this thesis is the compelling need to have the living, spoken Indonesian language represented in education, because its exclusion ultimately misrepresents Indonesian linguistics and thus also contemporary Indonesian culture. Sneddon (2003b, p. 20) claims that ‘if Indonesian language educators are to effectively address the needs of learners of the language, informal Indonesian must be taught as well as formal Indonesia.’

The next section compares the language content of informal spoken Indonesian as represented in textbooks with the informal language contained in the corpora of the research underpinning this thesis.

6.4 A review of language content in Indonesian L2 textbooks

This section is a critical review of the informal Indonesian language that is represented in Indonesian language textbooks. The conversations in the excerpts below are meant to reflect everyday informal interactions in Indonesian. These textbooks are part of Indonesian classroom collections and are used by high school and university teachers as resources for students.

The first review is of the textbook Suara Siswa – Stage 1 (1993), produced by the National Indonesian Language Curriculum Project sponsored by the then Curriculum Centre, which was based in Carlton, Victoria. In Suara Siswa there is an attempt to represent spoken Indonesian in the use of interjections and some CI vocabulary. However, dialogues that supposedly represent daily conversations still use FI’s predicate and pronominal morphology.

The title of chapter four in Suara Siswa is ‘Suka nggak’. Here, nggak is a lexical item from CI, yet sub-headings in the same chapter include: Kenalkan Aku... / Suka berlibur, kan?, where base-words are inflected in FI morphology. On page 5, the phrase ‘Maaf, Ida, akan terlambat ke kelas, nih. Sampai nanti’ (Sorry Ida, (I) will be late for class, (interjection)) uses the tense akan, contrary to CI usage, in which the modal auxiliary mau and bakal are more common. The CI word telat (to be late) is more common in informal speech. Although terlambat is rare, it is not totally out of use.

A dialogue on page six makes some attempt at natural dialogue:

Anwar: Selamat Sore. Mau ke mana? (Good afternoon. Where are you going?)

---

5 At the time of writing, the author has taught Indonesian language in eight Queensland primary schools, high schools and universities for 19 years (since 1998). The books described in this thesis were actively used in the classroom and were part of the classroom and library collection of the school / university / institution. The author also canvassed the Indonesian language resources at the Queensland LOTE (Languages Other Than English) Centre in 2011. The LOTE Centre was a Queensland state-run library and resource development institute that provided language learning materials to schools.
Carol: *Jalan-jalan saja.*  
(Just going for walk)

Anwar: *Bisa bahasa Indonesia?*  
(You can speak Indonesian?)

Carol: *Ya, bisa sedikit.*  
(Yes, a little bit)

Anwar: *Sudah lancar. Di mana belajar?*  
(You’re fluent. Where did you learn?)

Carol: *Di sekolah.*  
(At school)

Anwar: *Kenalkan, saya Anwar* (they shake hands) ...’  
(Let me introduce myself, I am Anwar)

The above is an example of CI linguistics features being inaccurately applied, as the corrections below illustrate (in bold):

*Selamat Sore. Mau ke mana?*
*Jalan-jalan saja.*
*Bisa bahasa Indonesia?*
*Ya, bisa sedikit.*
*Sudah lancar. Di mana belajar?*
*Di sekolah.*
*Kenalin, saya Anwar* (they shake hands) ...

On page 10 there is a letter supposedly written by an Indonesian youth, inviting an Australian youth to become pen pals. The heading of the section containing the letter is *Salam Perkenalan*. The reads as follows:

*Banjarmasin, 4 April,*
*Salam Teman,*

*Saya membaca nama dan alamat Anda di majalah ‘Bobo’. Mau nggak kita kenalan lewat surat? Apakah Darren bisa berbahasa Indonesia? Saya belajar bahasa Inggris tetapi belum bisa... Saya mau menjadi insinyur... Hobi saya banyak. Saya bermain bulu tangkis, basket, membaca buku, bermain piano dan surat-menyurat.*

(Friendly greetings. Hello, how are you? You’re just fine right? I hope you are well. I read your name and address in *Bobo* magazine. Would you like to get to know each other by mail? Can you speak Indonesian, Darren? I want to learn English but I am not good at it yet. I want to become an engineer. I have many hobbies. I play badminton, basketball, read books, play the piano and write letters.)

The use of the personal pronoun *Anda* and the question word *apakah* is very formal and impersonal between teenagers. The phrase *Baik-baik saja bukan?* (You’re just fine right?) is not a natural
expression. Additionally, the overuse of the FI *me*- and *ber*- prefixes does not sound natural for a casual correspondence between teenagers.

The following dialogues are from the textbooks *Bahasa Tetanggaku*’ (White, 1994) and *Speak Standard Indonesian – A Beginner’s Guide* (Fan, 1990):

a) Sri: *Kenalkan, nama saya Sri*  
   (Let me introduce myself. My name is Sri.)

   (Good afternoon, Sri. My name is Yanti. Very happy to meet you. How are you?)

   Sri: *Baik-baik saja, dan bagaimana kabarmu?*  
   (Well thank you, and how about you?)

   Yanti: *Saya juga sehat-sehat saja.*  
   (I’m also well). (White, 1994, p. 7)

b) Arifin: *Selamat Sore Nyonya*  
   (Good afternoon Mrs.)

   Nyonya: *Selamat Sore. Siapa nama Anda?*  
   (Good afternoon. Who is your name?)

   Arifin: *Nama saya Arifin*  
   (My name is Arifin)

   Nyonya: *Anda punya saudara laki-laki?*  
   (Do you have a male sibling?)

   Arifin: *Ya, saya punya dua orang saudara laki-laki*  
   (Yes, I have two male siblings)

   Nyonya: *Berapa umur mereka?*  
   (How old are they?)

   Arifin: *Seorang berumur 20 tahun, dan seorang lagi berumur 15 tahun.*  
   (One person is 20 years old and the the other is 15.) (Fan 1990, p. 28)

The formal title *Anda* is not used in everyday language. *Anda* is used in indirect speech for addressing an audience; it is mostly found in advertisements and public announcements. The use of the classifiers *orang* and *seorang* is not common in everyday language.

c) Budi: *Hei, Saudara Leo, sedang apa?*  
   (Hey, brother Loe, what are you doing?)

   Leo: *Saya sedang membaca koran*  
   (I am reading the newspaper.)

   Budi: *Apa yang sedang saudara baca?*  
   (What are you reading?) (Fan 1990, p. 155)

The FI portrayed in the dialogue above is far removed from spoken Indonesian. The title *Saudara* is not used in daily speech, especially between two people who know each other. The phrase *sedang apa?* is also an FI idealisation. SI corpora based on live recordings by Sneddon (2006) demonstrate
that the progressive predicate *sedang* is not in use in daily speech. The question phrases *lagi apa?* and *lagi ngapain?* are the equivalent CI phrases. The FI *me-* prefix and the syntax of *Apa yang sedang saudara baca* are also FI idealisations. The simpler variant *Lagi baca apa?* is more natural. The same analysis applies to the following in Fan (1990, p. 214):

\[
d) \quad \text{Djoko: Apa yang sedang Anda kerjakan Tini?} \quad \text{(What are you doing, Tini?)} \\
\quad \text{Tini: Saya sedang mengetik surat ini Tuan} \quad \text{(I am typing this letter, Mr.)} \\
\quad \text{Djoko: Tono sedang mengerjakan apa?} \quad \text{(What is Tono working on?)} \\
\quad \text{Tini: Dia sedang memeriksa surat-surat yang keluar} \quad \text{(He is checking the letters that are going out.)}
\]

This trend toward idealising FI conversation as an everyday language is found in many Indonesian language learning textbooks. A review of textbooks at the LOTE centre in Brisbane, Queensland, in May 2010 included other common titles: Ian J. White’s (2002) *Keren!*, the *Bagus Sekali* (Newnham, Julie & Soehodo, 2002) textbook series by CIS Heinemann and the *SuaraSiswa* series released by the (National Indonesian Language Curriculum Project, 1993).

However, there are signs that CI linguistics are being acknowledged. In the introductory chapter, one beginner’s book by Dr Liam Yock Fan (1990), *Speak Standard Indonesian*, makes mention of the ‘informal or Jakartan dialect’, as the author describes it, and offers some language features of CI. Nevertheless, the language content in the lessons themselves reverts to an idealised FI variant.

Another recent textbook that addresses the socio-linguistic issue of H and L variants is *Bersama-sama Selalu* (Taylor & Day, 2006).

### 6.5 Comparing the CI content in Indonesian language textbooks with corpus data

This section will compare the results of some authentic sources of SI with the ‘informal spoken’ language that is represented in Indonesian language textbooks. Samples of the depiction of informal dialogues in Indonesian language textbooks have been taken from seven textbooks. The seven textbooks have been selected on the basis that these materials claim to contain representation of informal language or colloquial dialogues. To gauge how SI is represented in these textbooks, the CI in textbooks is compared with the CI content of the data samples from the corpus of this study. What is referred to as an ‘authentic source’ consists of language materials taken from real life samples (written and/or audiovisual) that have not been changed, moderated, added to, or simplified in any way.
These seven selected textbooks are in circulation in Australian secondary and tertiary Indonesian language courses. Table 16 provides information about the samples sizes, the CI content and the CI/SI ratio of the seven textbooks and the three authentic source data samples (Numbers 8, 9 and 10).

Table 16

A Comparison Between CI Content in Textbooks and the Corpus

<table>
<thead>
<tr>
<th>Sample no.</th>
<th>Name of resource</th>
<th>Resource type</th>
<th>Sample size</th>
<th>CI A</th>
<th>CI L content</th>
<th>CI:SI Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Let’s Speak Indonesian (Rafferty, 2014)</td>
<td>Textbook</td>
<td>276</td>
<td>18</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Bahasa Tetanggaku (White, 1994)</td>
<td>Textbook</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Indonesian – An Introductory Coursebook (Robson &amp; Kurniasih, 2010)</td>
<td>Textbook</td>
<td>188</td>
<td>14</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Bagus Sekali! 2 (Newham &amp; Soehodo, 2002)</td>
<td>Textbook</td>
<td>255</td>
<td>20</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Beginning Indonesian through self-instruction (Wolff, 1992)</td>
<td>Textbook</td>
<td>326</td>
<td>10</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Speak Standard Indonesian (Fan, 1990)</td>
<td>Textbook</td>
<td>70</td>
<td>1</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Bahasa Indonesia – Langkah Baru 3 (Johns, 1996)</td>
<td>Textbook</td>
<td>215</td>
<td>32</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Sule 100% Lucu (Youtube)</td>
<td>Data Sample – TV</td>
<td>320</td>
<td>140</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Coblos Cinta (Youtube)</td>
<td>Data Sample – Novel</td>
<td>122</td>
<td>50</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Wayang Cepot (Youtube)</td>
<td>Data Sample – TV</td>
<td>484</td>
<td>142</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>

Note. Textbooks CI range: 0–0.15, data samples CI range: 0.29–0.44; Textbooks CI mean: 0.06, data samples CI mean: 0.38.

This comparative study shows that the everyday informal dialogues of the Indonesian language textbook fail to represent SI as the diglossic spoken variant both qualitatively (phonologically, morphologically and lexically) and quantitatively. In both range and mean the CI content of the textbooks is substantially below the CI content of the authentic source data samples. Moreover, most
of the CI in these textbooks are interjections, emotive particles and discourse particles. There are no actual CI phonological realisations, morphological markings or lexis (apart from some isolated exceptions of the use of enggak [no, not]). Some of these CI lexical items also include contractions such as tapi (‘but’; short form of tetapi) and the affirmative lexis ya, a word of Dutch origins which exists as a spoken word only (not used in the FI domain).

Textbook Sample No. 7 has the highest CI content compared to other textbooks. This is to be expected, since textbook sample No. 7 differs from the other selected textbooks in that it specifically claims to be an introduction to spoken informal Indonesian. Despite that, the CI content of sample No. 7 at 0.15 is statistically still lower than the lowest SI data sample corpus (data sample No. 10 with a CIAL content of 0.29).

6.6 The descriptive-pragmatic process

This study proposes a process of identifying the sociolinguistic paradigms of SI as the spoken variant and incorporating CI into a pedagogical framework. It is proposed that a ‘descriptive-pragmatic’ (D-P) process (Figure 12) should be applied as a process of appropriate language selection towards the design of an SI pedagogical model. This is the task of the pedagogy designer who has in mind the need to create Indonesian language materials in the context of diglossia and its application to different social situations.

Figure 10. The Descriptive-Pragmatic Process
The (D-P) process will consist of identifying SI as the spoken variant and then incorporating it into a pedagogical framework. This means there are two stages in the process: first, a descriptive qualification, and second, a pragmatic selection process.

The process is first descriptive because it will apply a sociolinguistic approach in which language is identified and analysed as it is, without any value judgment and without a prescriptive agenda. Grammar can be taught as a relative constant – as an identification of common patterns – but not as absolute rules.

The descriptive approach will then be filtered through a pragmatic ‘sieve’, meaning that for the purposes of teaching the most used and useful language, boundaries have to be drawn around CI to retain only a mainstream variety, and non-mainstream colloquialism will be excluded. Examples of mainstream usage are:

1. Always address a stranger with a title (Pak [Bapak], Bu [Ibu], Mas, Mba, Bang [Abang], De [Adik]). The function of each personal pronoun title will also have to be spelled out, with the use of Pak or Bu for addressing older people, Mas or Mba for one’s peer, Bang for tradespeople or workers, and De for younger strangers.

2. The personal pronouns kamu and lu (elu) should only be used with young people you know well, or if they initiate it. It is important to point out that the use of kamu can also be impolite, and it is better to use proper names or terms of address in situations requiring some formality, respect or politeness.

One of the objectives of this research was to establish pragmatic uses such as the examples above, using data from corpora gathering. Corpora obtained from native speakers will help establish parameters of pragmatic competence and conventions. The ability to use terms of address is an essential skill in Indonesian, because terms of address follow pragmatic functions of ‘solidarity’ and ‘politeness’. During the independence era titles such as ‘Bung’ and ‘Saudara’ were common to establish a sense of solidarity and ‘brotherhood’, in the collective struggle of the time. As titles of address, the use family words (Bapak [father], Ibu [mother], Mas [older brother-Javanese], Mba [older sister-Javanese], Abang [older brother-Sumatran], Adik [younger sibling], testifies to the strong family values of Indonesian culture. The socio-cultural significance of using such titles is the sense of ‘togetherness’ created between speaker and addressee.

6.6.1 The descriptive-pragmatic process: Defining mainstream SI

Sneddon (2003b, p. 11) stated that since ‘a prestige variety of informal Indonesian has emerged, the terms “standard” and “formal” must be clearly differentiated…””. In the context of Indonesian
diglossia, Sneddon’s ‘standard’ spoken language is what is referred to here as mainstream SI and FI is the formal language. The social domains of FI and SI are presented in Table 11 (Chapter 4). These categorisations of FI–SI social domains serve as a guide in allocating SI to its appropriate social domains, which in the context of the Notional-Functional pedagogic model are the thematic categorisation or themes of the syllabus.

Before SI is assigned to its appropriate sociolinguistic themes, the CI language content needs to be selected for appropriateness and filtered through a pragmatic sieve. For the purpose of education, boundaries have to be drawn around SI to retain only an appropriate language variety that excludes inappropriate colloquialisms. Non-mainstream lexicon, such as fad lexicon, sub-culture colloquialism and slang will be filtered through the prescriptive sieve. Extreme slang, gangster talk and obscure teen speak, such as bahasa prokrem, bahasa gaul and bahasa abege, will not qualify.

Table 17

The Prescriptive Sieve: Filtering Non-Mainstream Colloquialism

<table>
<thead>
<tr>
<th>Prescriptive sieve for SI</th>
<th>Pass</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used at home (talking to family)</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Used when talking to strangers (shopping, travelling)</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Used at work (office) in informal setting</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Used between colleagues in informal setting (outside work)</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Used by student-teacher, student lecturer in informal setting (outside classroom)</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Used between relatives and friends</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Polite to use with strangers</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Impolite to use with strangers</td>
<td>✘</td>
<td></td>
</tr>
<tr>
<td>Bahasa Prokem/Gaul/ABG</td>
<td>✘</td>
<td></td>
</tr>
</tbody>
</table>

6.7 Contextualising diglossia in the Notional-Functional Syllabus Model

As part of teaching SI to foreign learners of Indonesian, it is important that students acquire sensitivity to the sociolinguistic situation of language use and appreciate how to make appropriate language choices for each social occasion and context (Sneddon, 2003b). To achieve this end, it is proposed in this study that the Notional-Functional Syllabus Model (Johnson, 2003b) is a suitable pedagogic design that can cater for the different sociolinguistic language situations of diglossia.
In the Notional-Functional model, the pedagogic unit structure is organised into ‘themes’ of situational settings. These situational settings can be arranged to correspond with the ‘formal’ and ‘informal’ settings of H and L/M diglossic-sociolinguistic domains. Thus, for example, FI can be taught in units where the situational setting is formal with a topic such as ‘Doing Business’ and contain language functions such as ‘making appointments and arranging business meetings’, while an informal social setting using SI can cover a theme like ‘At the Traditional Market’, containing language functions such as ‘shopping, asking for prices and bargaining’. Table 18 is an example of how a language unit outline that is based on a Notional-Functional syllabus model may look to cater for the teaching of both FI and SI in their respective sociolinguistic domain.
### Table 18

**Diglossic Themes in a Notional-Functional Syllabus**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Topics</th>
<th>Roles</th>
<th>Language Activities</th>
<th>Notions</th>
<th>Functions</th>
<th>Exponents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theme: Doing business in Indonesia (FI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Country</td>
<td>Country</td>
<td>Colleague to colleague</td>
<td>Making phone calls to FL country in relation to business arrangements / business trip</td>
<td>Greetings, Introductions</td>
<td>Framing questions / requesting information</td>
<td>Selamat pagi…permisi</td>
</tr>
<tr>
<td>Office Conferences Meetings</td>
<td>Arranging transport / making bookings</td>
<td>Business partners</td>
<td></td>
<td>Dates, Times</td>
<td>Giving information</td>
<td>Saya mohon ingin memperkenalkan diri…</td>
</tr>
<tr>
<td></td>
<td>Writing business inquiries and proposals</td>
<td>Secretary to boss</td>
<td></td>
<td>Numbers, quantities and amounts</td>
<td>Agreeing and disagreeing</td>
<td>Saya sangat berminat untuk…</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Modal verbs, Time markers</td>
<td>Making arrangements</td>
<td>Apakah menurut Bapak/Ibu ada peluang untuk…?</td>
</tr>
<tr>
<td><strong>Theme: At the Traditional Market (SI)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Country</td>
<td>Buying</td>
<td>Customer with seller</td>
<td>Going to the market / shops looking for souvenirs</td>
<td>Greetings</td>
<td>Asking about items</td>
<td>(Met) Pagi…</td>
</tr>
<tr>
<td>Market Shop</td>
<td>Shopping</td>
<td>With friends / host families</td>
<td></td>
<td>Numbers / prices / currency</td>
<td>Asking about prices</td>
<td>Saya cari / nyari/ perlu / butuh / kepingin mau beli</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sizes, colours, quality (adjectives)</td>
<td>Asking about items/details</td>
<td>Berapa harganya yang ini? Berapa ini?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Modal verbs</td>
<td>Expressing wants/needs</td>
<td>Boleh kurang?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pronoun: yang</td>
<td>Bargaining</td>
<td>Ada yang…?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tag question: …nggak?</td>
<td></td>
<td>Bikinans mana ini?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bagus nggak mutunya?</td>
</tr>
</tbody>
</table>
6.8 Conclusion

The theme-based organisation of the Notional-Functional syllabus is the model that is proposed in this chapter as a design that can be associated with the different sociolinguistic settings of diglossia. What is of essence is that formal and informal diglossic social settings are categorised separately, such that H and M/L variants are assigned to their respective social settings. This section of the chapter completes the narrative of this thesis: to provide practical application of the findings of this study, which has explored Indonesian diglossia. The findings of this research, as presented in Chapters 4 and 5, indicate that an FI and a CI have become components of an M diglossic variant of Spoken Indonesian that is spoken as the everyday informal language. Indonesian language textbooks, however, have represented the everyday spoken informal language as FI. The discrepancy between prescriptive pedagogy and reality has been the motivation for this thesis. Incorporating SI in Indonesian language teaching will greatly improve the comprehension and communicative competence of students of the Indonesian language. Appendix 2 of this thesis is a compilation of Indonesian lexis and collocations towards the construction of a reference for foreign students learning CI. The intention is to present CI collocations and phrases in context. Research has shown that L2 learning is more effective when lexis is arranged in collocations and formulaic sequences as it provides contextual and semantic markers (Boellstorff, 2004; Oxford, 1996).
Chapter 7

Conclusion
7 Overview

This thesis set out to explore the linguistic and sociolinguistic nature of the Indonesian language. The investigation has been framed within the conceptualisation of the sociolinguistic theory of diglossia. Errington (1986) and Sneddon (2003a) stated that the Indonesian language is in a state of diglossia.

The research had three objectives:
1. To study the Indonesian L diglossic variant CI and, as such, to provide a description of the linguistic features of CI.
2. To use a mixed methods approach to analyse a corpus of authentic, informal spoken language to test the M hypothesis, in order to establish if an M diglossic variant (SI) exists.
3. To provide a pedagogic model for the teaching of CI linguistic features in the context of the SI diglossia.

The research investigated whether Indonesian is a diglossic language – consisting of H and L variants – and tested the M hypothesis that the H and L variants are merging into a M standard spoken variant. To develop this research within sociolinguistic diglossia theory, the Indonesian language had to be categorised into H (High) and L (Low) diglossic variants. Subsequently, this research used the terms Formal Indonesian (FI), denoting the H diglossic variant, and Colloquial Indonesian (CI) as the L diglossic variant. The sociolinguistic functions of FI and CI were elaborated in Chapters 2 and 4.

Sneddon (2003b) proposed that the interaction of the FI and CI variants has resulted in a merging of the two diglossic variants into a Middle (M) informal/semiformal spoken variant. This M Indonesian diglossic variant is referred to in this thesis as Spoken Indonesian (SI). To test this hypothesis of a merger, a corpus was collected from data samples of the everyday informal spoken language. The linguistic features in this corpus that did not abide by the rules of grammar of Formal Indonesian (or known in Indonesian as bahasa Baku) are identified as CI.

This research set out to investigate this question using qualitative linguistic description (IPA) and quantitative descriptions (CIAL:SI). The findings indicated that a substantial proportion of language in the corpus is CI. Every data set contained CIAL. The majority of data sets (94%) had a CIAL:SI ratio range of 58 (0.17–0.75). The overall CIAL:SI mean was 0.39. These results confirm that an L diglossic variant exists in the form of CI and that the (positive) M hypothesis can be accepted; there is a merging of the H and L variants into an M standard spoken variant.
As discussed in Chapter 2, CI evolved in Jakarta from the Betawi language. The dissemination, and hence the ‘homogenisation’, of SI by the national media (television, film and literature) has rendered it the new national spoken language. The second set of data generated from the questionnaire research instrument (Chapter 5) showed that CI strongly features in the language spoken in Jakarta and surroundings. This indicates that SI is the dominant spoken language. The data also indicated that SI is commonly spoken in other regions in Indonesia. Demographic factors, particularly age and occupation, significantly affect the use of CI, with the younger generation using more CI and people in the education profession using more FI.

The practical application of this thesis is to incorporate SI into Indonesian language pedagogy (Chapter 6). The thesis proposes the Notional-Functional pedagogic model, since it allows H and L diglossic situations to be categorised into the thematic organisation of the work units. The CI elements of SI have largely been neglected in Indonesian language pedagogy, and incorporating SI in its full form will greatly improve the way foreign students acquire the Indonesian language.

7.1 Problems encountered

An interesting challenge was co-indexing and glossing between Indonesian and English phrases and vice versa. For example, how should the infinitive particle ‘to’ be co-indexed vis-à-vis the Indonesian translation? Should it be assumed that this particle is inherently imbedded in the Indonesian verb thus co-indexed (makan, → to eat), or should it be regarded as a null element (makan, → toØ eat)?

Another example is a phrase like jangan, begitu. Should jangan be co-indexed [do not be], or [do not be]. Does Jangan as a negative-imperative-predicate lexical entity already contain within it the verb ‘to do’ and the copula ‘to be’? The same question applies to the English infinitive particle ‘to’.

Should tidak, mau, makan, sekaran, be glossed as [do not, want, to eat, now] or [do not, want, to eat, now]? What about begitu in jangan begitu? Is begitu a demonstrative pronoun, to be translated simply as that? At the semantic level it can behave like a pronoun-predicate compound, translating as do like that, as in the phrase begitu, saja, which can be glossed as [just, do it, like, that]. Or should the predicate [to do] and the expletive–anaphora [it] be co-indexed with begitu [just, do, it, like, that], since they are insinuated and semantically contained at the Deep Structure level? A fascinating challenge was translating ngapain?, which semantically contains [to do], [what] and can be grammatically described as [+active verb +interrogative +progressive gerund (-ing)], while at the same time as a question phrase it can have a null personal pronoun, NP. A full translation of ngapain,? can be [what are you doing,?]. So should it be glossed as [what, are, you, doing,?] or [what, are, you, doing,?]? What about the predicate-adjectival compound berdua? Would berdua, [the two],
or berdua [theØØ two] be more accurate? This thesis has taken a pragmatic approach and opted to treat unrepresented lexis as null elements. This is an interesting technical issue that warrants further discussion within the greater community of Indonesian linguists.

The research did not delve into the function of intonation as a phonological marker of semantics and pragmatics. Consider the following quote from Crystal (2008, p. 252):

Intonation performs several functions in language. The most important function is as a signal of grammatical structure, where it performs a role similar to punctuation in writing. The marking of sentence, clause and other boundaries, and the marking of questions and statements, may be made using intonation. For example, a contrast between rising and falling pitch in the phrase: _He’s going isn’t he?_ (asking) and _He’s going isn’t he!_ (telling). As such intonation also communicates personal attitude: sarcasm, puzzlement, anger, etc. In sociolinguistics, intonation can also signal social background.

Intonation in Indonesian, as in English, plays a very important linguistic role, which is not captured in the written orthographic form of Indonesian. The Indonesian speech acts discussed in this thesis interact dynamically with intonation. This discussion and analysis would be more meaningful to readers who are familiar with Indonesian and can imagine the intonation of a given phrase.

Another matter that posed some challenges was finding the right methods of quantitative analysis. No conclusion was derived from the discourse as to whether the linguistic variants FI and CI are inherently dependent or independent variables. There are two possible perspectives on this. From an ‘outward-macro’ perspective, which views FI and CI as subsets of the SI corpora or subsets within the universe of bahasa Indonesia, they appear to be dependent variables that are activated by the language appropriateness of the sociolinguistic environment – akin to Ferguson’s view on diglossia (see Chapter 2). On the other hand, from a ‘from-within-micro’ view, the language variants (i.e. FI and CI) can be considered as independent variables. This perspective is based on the independence of the individual speaker as the generator of FI or CI language who chooses language use with intent to create certain social, pragmatic and semantic outcomes. No decision was made on the questions of the independent/dependent nature of FI and CI, and this thesis has opted to measure the quantitative relationship in terms of ratio. Establishing the quantitative nature of the data will determine statistical methods for future research.
7.2 Future research

This thesis provides the platform for further research in the following areas:

- The influence of regional languages on CI. A preliminary study on this topic was conducted during this research (see Appendix 5) and needs further investigation.
- The role of prosody in constructing pragmatics and semantics in CI. A phrasal prosodic example is the collocation *boleh juga*. When the last syllable of *juga* is vocalised as a long vowel [aː], as in the phrase *boleh juga: begitu*, it bears the semantic value ‘it may also be so’. A prosodic change to the pronunciation of *juga*, vocalising the first syllable as a long vowel [uː], gives the phrase *boleh juga* a different meaning: ‘that’s not bad!’, as in the phrase *boleh juːɡa itu*.
- Further study on phonemic variations and semantic shifts in homographs. For example, in the second syllable of the word *sedang*, the unrounded [a] is the modal auxiliary progressive tense marker. When the unrounded [a] is vocalised as a schwa [ə], the meaning of the word becomes ‘medium’.
- The rapid evolution of Anglicism in CI (also known as ‘Indolish’). These include ‘CI-nised’ English words and abbreviations, the use of English words in a phrase, and the outright use of English phrases. Some examples include: *ngecek* (to check), *ngefek* (to have effect), *BT* (bad tempered), *lagi mood* (to be in a mood for something), *lagi nggak mood* (to be in a bad mood), *kepo* (an acronym for ‘knowing every particular object’, ‘to be curious’).
- A philological study constructing a timeline of the linguistic (lexical and morphological) evolution of Malay from Ancient Malay to classical Malay to modern Indonesian, and a timeline of the evolution of spoken Indonesian during the 20th century.
- A philological study on how different orthographies (Pallava-based Hanacaraka, Arabic and Latin) affect syllabic constructions of lexis.
- Further philological studies on the affinities between Malay and other Austronesian languages in Southeast Asia and the Pacific.
- The incursion of CI into FI sociolinguistic domains, such as news articles and politicians using CI in public speaking.

7.3 Final remarks

My observational impression is that there appears to a breaking down of sociolinguistic barriers in Indonesian. Many language varieties – diglossic, dialectal or regional languages – seem to be freely and dynamically merging and blending in everyday language usage. Indonesians would liberally use words, terms and expressions from any available Indonesian language variant mentioned above – with the addition of English - to express themselves. Whichever language gets the point across best gets
used. With this merger of CI and FI into SI, peppered with words from the regional languages, the latest Gaul or English, Indonesian is rapidly transitioning out of the category of diglossia and the traditional Fergussonian divide might be rendering itself irrelevant. My forecast is that CI and FI are well imbedded as part of everyday Indonesian in the form of SI, which itself will continuously and dynamically evolve, spurred by the trends, the demands and the culture of each era.
References


Errington, J. (2014). In search of Middle Indonesian: Linguistic dynamics in a provincial town. In G. V. Klinken & W. Berenschot (Eds.), In search of middle Indonesia – Middle classes in provincial towns (pp. 199. Leiden, Netherlands: Brill.


Appendix 1

SI Corpus
A. **Corpus: Contemporary Literature**

**Word Counts**

**Corpus from Text 1**  
*Kartun Riwayat Peradaban Jilid I*  
*Bab 1-7: Dari Ledakan Besar Hingga Alexander Agung*  
*Total Word Count: 244*  
*Words with CI Linguistics: 90 (CIλ x 100 / n = 36.9%)*

Page 28: 6  
*Nak!* Jangan rakus melahap kadal itu! Nanti demam *lho*!  
Memang itu tujuannya, *ma*!  
Apa *sih* “rakus” itu?

Page 29: 5  
*Kok* semuanya mati *sih*?

Page 38: 2  
*Hei! Terus* kami mesti makan apa?  
Page 40: 4 : 5  
Yang lain lagi jadi *doyan* makan ikan dan mencoba-coba berenang  
*Sana kelelep aja*, jangan *nakut-nakutin* ikan.  
*Eh, apaan tuh* ribut-ribut?

Page 45: 1  
*Colek* - Jangan lelet dong!!

Page 56: 3 : 4  
*Ngejar* antelop saja *deh*, lebih enak!  
*Nggak* ada gunanya *nih*!

Page 66: 2 : 3  
Kamu *ngapain sama* putriku sabtu malam kemarin?  
*Nggak macem-macem*, sungguh!!  
*Hayo*, kenapa ayam *nyebang* jalan?? *(kikikik)*  
*Huu*. Itu *udah* basi!

Page 80: 2 : 4  
*Cewek bahanol oy! Hehehee *ngiler*  
Bukan tipe *cewek* seleraku, kurasa…  
Kamu yakin ini bukan *cuma* dongeng nenek tua?

Page 81: 1  
*Kamu suka tidak?*  
Sama sekali *nggak* mirip aku *deh*!

Page 82: 4  
*Rasanya* kita harus *bagi secuil buat* mereka!  
Page 87: 4
**Heh! Jangan macam-macam** denganku Bung!

Page 88: 3

**Mendingan sih,** tapi masih kurang enak

Page 93: 3 : 5

Ini perlu latihan *dalu – serr – swiing – ceprot*

Kapak emas?! Bagai *sih,* tapi lembek!

Page 94: 2

*Nah,* coba kita bahas… kamu *tuh* adik perempuanku, tapi kamu ber-*ehm-ehm* dengan pamanku, dan pamanku melakukan dengan saudara perempuannya sendiri, dan aku juga begitu. Jadi kamu apaku?

**Emang gue pikirin.**

Page 139: 1

**Hai! Sini deh! Nggak** apa-apa kok!

133: 3

Orang kulit putih memerintah orang kulit hitam? Belum pernah ada kejadian begini!

**Nggak bakal awet!**

Page 170: 8

**Eh – itu kan** untuk Tuhan?

Tidak apa-apa – aku *kan* perantaraNya!

Page 171: 6

Hakim keliling pertama!

**Nggak** jauh *sih*

Page 208: 1

Ayo ke selatan, untuk merayakan milenium baru!

**Boleh juga nih,** liburan!

Page 228: 3 : 4

Besi *nggak* bagus!

**Rasain nih… waduh**

Page 240: 2 : 4

**Lo,** ini *kan gampang*? Anak kecil juga bisa!

Bagus. Ada tidak anak kecil di sini yang bisa mengajari kita?

The range of \((CI^6 \times 100 / n^6) – (CI^1 \times 100 / n^1)\); (the percentage range of CI linguistic in Spoken Text 1) at random intervals = 39.7% – 35% = 4.7%

\[
CI^\lambda : n
\]

1. 30 : 84 = 35%
2. 43 : 114 = 37.7%
3. 46 : 118 = 39%
4. 57 : 148 = 38.5%
5. 63 : 165 = 38.2%
6. 79 : 199 = 39.7%
7. 84 : 220 = 38.2%
8. 90 : 244 = 36.9%

Text 2
Kartun Riwayat Peradaban Jilid II
Bab 8-13: Dari Berseminya China Hingga Runtuhnya Romawi
Total Word Count: 86
Words with CI Linguistics: 27 (CIλ x 100 / n = 31.4%)

Page 47: 7
*Benar nih* ingin tahu?

Page 106: 7
Ada buku yang *nggak* penting *banget buat* dibaca

Page 146: 1
Yang penting *duitnya dulu*!

Page 163: 2
Apa di bawah sini?
Fiuh…*udahan aja yuk*…

Page 178: 1
*Nah*, mungkin kita *nggak jago* berfilsafat, tapi sudah pasti kita hebat dalam perang dan pembangunan! Apa artinya?

Page 182: 3
‘Yo-de-lay-he-hoo’ itu maksudnya apa *sih*?

Page 190: 2
*Nggak dong*, kalau kamu punya pasukan sendiri!!

Page 192: 3
Mungkin kita harus mempertimbangkan lagi soal Gladiator itu…
Apa? *Ngaku* kalah?

Page 234: 1
*Wah!* Ada yang *ngomong* bahasa Patagonia!
Atau bahasa Pamfligionia *ya*?
Atau mungkin juga bahasa Petalum
*Nganggur niye*

Page 279: 3
*Ngerti* kamu? *Jangan macam-macam sama* Diocletianus!!
The range of (CIλ \(^\frac{1}{4}\) x 100 / \(n^4\)) – (CIλ \(^\frac{4}{4}\) x 100 / \(n^4\)); (the percentage range of CI linguistic in Spoken Text 2) at random intervals = 37.5% – 24.6% = 12.9%

*CIλ : n*
1. $3 : 8 = 37.5\%$
2. $10 : 29 = 34.5\%$
3. $12 : 44 = 27.3\%$
4. $15 : 61 = 24.6\%$
5. $18 : 68 = 26.5\%$
6. $27 : 86 = 31.4\%$

Text 3
Kartun Riwayat Peradaban Jilid III
Bab 14-19: Dari Bangkitnya Arab Hingga Renaissance
Total Word Count: 126
Words with CI Linguistics: 42 ($\frac{CI\lambda \times 100}{n} = 33.3\%$)

Page 70: 3 : 5 : 6
Lihat? *Ente mau coba juga kan?*
Menontonnya *sih* suka mungkin…
Negus, mereka telah menaklukan seluruh Arab! Mereka telah mengalahkan bangsa Romawi!
Bangsa Mesir! Bangsa Persia! Bangsa Suriah!
Ngg… Unta bisa naik gunung *nggak*?
*Nggak* bisa *lah*!

Page 71: 3
Dan kalian berjanji akan mengirimkan gandum setiap tahun, jangan lupa *ya*!
Ya ya… *nggak usah diingetin*…

Page 76: 5
Apa, apa?
Oh, ehem… *nggak ada apa-apa*!

Page 107: 7
Kalian sadar *nggak*? Kalian bisa membunuh beberapa orang saja, mengambil alih pemerintahan, memajaki satu bangsa dan menjadi benar-benar kaya?
Wah… *nggak tuh… beneran nih*?

141: 1
Ini *nggak* adil!

Page 157: 3 : 6
Peduli *amat*! Aku *nggak* bisa baca!
*Kampungan* kamu Cloitaire! Kamu *nggak ada apa-apanya*!
*Lagi-lagi* begini…

Page 159: 7
*Dasar* orang-orang Bahlul!

Page 175
Bantu aku *ngurus* rumah *dong*!!
Page 272: 2
Aku **duluan**!

Page 278: 8: 9
Lihat? *Gampang kan*!
Sketsanya sudah selesai **belum**?
Aku masih mengerjakan hitungan matematikanya!

Page 285: 4 : 7
*Nah*, ini garis yang bagus!
Orang ini **cuma omong doang**!

293: 2
**Ngomel**

The range of \((Cl\lambda^2 \times 100 / n^2) - (Cl\lambda^3 \times 100 / n^3)\); (the percentage range of CI linguistic in Spoken Text 3) at random intervals = 36.4% – 32.7% = 3.7%

\[Cl\lambda : n\]
1. 15 : 43 = 34.9%
2. 24 : 66 = 36.4%
3. 33 : 101 = 32.7%
4. 42 : 126 = 33.3%

---

**Text 4**

**Party Girls – Pesta Semalam Suntuk**

**Total Word Count**: 647

**Words with CI Linguistics**: 183 \((Cl\lambda \times 100 / n = 28.3\%)

Page 94 -
Dria – rumah
“kak, kakak **ngapain**?” Tegur Dria melihat Videl keluar dari kamar mama yang **nggak** pernah **ditempatin** sambil membawa kotak berlian mama.
“Bukan urusan kamu! Kamu **udah** bilang **kan** kita **urusan** urusan kita masing-masing!” bentak Videl.
“Ini bukan **cuma** menyangkut urusan kakak. Ini urusan mama juga!” bentak Dria.
*“*Udah deh! **Nggak usah kebanyakan bacot.** Kamu **nggak tau kan** masalahnya apa? **Nggak usah sok ikut campur deh jadi** anak kecil!” Videl makin jengkel.

“Dria **udah gedeh**! Dria bisa **bedain** mana yang maling mana yang **enggak**!” balas Dria. Videl melayangkan sebelah tangannya yang tidak memegang kotak untuk menampar pipi Dria, tapi Dria berhasil mengelak.
“**Memangnya** kakak siapa berani sentuh-sentuh Dria? Papa-mama **aja nggak** pernah **maen** tangan sama Dria!” bentak Dria.


"Tapi bukan begini kak caranya!" bantah Dria.

"Sok pintar banget sih kamu!"

"Emang gw pintar. Emang gw itu lo yang saking begonya kuliah enam tahun gak lulus!" semprot Dria.


Anggri - Mobil Anggri


Yaya – Pintu Keluar Bioskop 21

"Nah, lo dah gw temenin nonton Doraemon kan…, ujar Nizar sambil menghabiskan sisa popcorn di tangannya. "Sekarang, lo tolong gw ya… please…"

"Sip! Apaa?" tanya Yaya. Jantung Yaya berdebar keras. Pasti Nizar pengen ngajak makan atau cieeehh…


"Banyak nggak?” tanya Yaya.

Tenang aja. Nggak banyak-banyak amat kok, "jawab Nizar sambil ketawa ringan.

"Ya udah, kerjainnya di mana nih?” tanya Yaya.

"Di rumah gw aja deh...” jawab Nizar.


"Lo ke sini naek apa?” Nizar balas tanya.


“Nih, salinin ya buat gw! Lo kan udah janji mau nurutin permintaan gw…,” jelas Nizar sam bil melemparkan novel tebal itu ke pangkuan Yaya.

The range of \((\lambda_1 \times 100 / n^1) – (\lambda_2 \times 100 / n^2)\); (the percentage range of CI linguistic in Spoken Text 4) at random intervals = 30.3%–25.5% = 4.8%

Text 5
Entertaiment Salah Gaul - Cerita Horor
Total Word Count: 492
Words with CI Linguistics: 103 (CI \(\lambda \times 100 / n = 20.1\%\))

Page 109 – 112

Lanjut, mereka berangkat jam 3 sore dan berangkat sama manajer mereka. Perjalanan dari Seoul ke Bojong Jengkol itu kira-kira butuh waktu 8 jam, satu-satunya akses menuju Bojong Jengkol adalah dengan cara melewati kuburan!! KUBURAN!! *sengaja diulang biar serem*.

Mau nggak mau mereka harus lewatin kuburan itu. Kuburan itu namanya “Tanah Gaul”. Gue nggak tau deh kenapa itu kuburan dinamain “Tanah Gaul”. Mungkin penghuninya semasa hidup sering nongkrong di Seven Eleven (Sevel) atau PIM #entahlah.

Sebenarnya ada alternatif lain buat nggak lewatin kuburan, yaitu naik elang. Masa iya mereka mau naik elang? Mereka F(x) apa tutur tinular?

Tepat jam 10 malam, mereka sampai di depan kuburan Tanah Gaul. Victoria mulai ngomong, “kalian semua harus hati-hati ya, tempat ini gelap, perhatikan langkah kalian”. Member F(x)
yang lain ketakutan dan cuma bisa bilang ‘Iya’, karena tempat itu bener-bener gelap dan serem.


Luna keringetan, Sulli ketakutan. Krystal kesakitan, Victoria terus berdoa, tapi Amber _malah nyeletuk_, “_Please dehhhhh, nggak_ akan mau lagi _gue_ promo ke Bojong Jengkol”. Mereka semua masih berpelukan sambil duduk, posisinya _tuh kaya_ sinetron ratapan anak tiri pas bagian anak tirinya diinterogasi _sama_ ibu tiri.

Sial, mahluk itu semakin mendekat. Selangkah demi selangkah dan mahluk itu _udah_ semeter di depan member F(x). Semua member F(x) menggigil. Makhluq itu _ngeluarin_ sesuatu dari kantongnya!! *pasang muka cemas*

_Ternyata_ makhluq itu _ngeluarin_ senter dan diarahkan ke member F(x). Makhluq itu _ngebantu_ semua member F(x) _buat_ berdiri dan mahluk itu _bilang_ “Aduh Victoria!! _Kan_ aku _udah_ bilang CD album Hot Summer nya jangan _sampe_ ketinggalan di mobil. Ini _malah_ ketinggalan, jadi _aja_ aku _balik dulu_ ke mobil tadi”.

_Ternyata_ makhluq itu adalah manajer F(x) yang sejak tadi _nggak_ diketahui keberadaannya. Mereka pun lega dan mengelus dada. Eits tapi ternyata, celana Amber basah!! Mungkin dia _ngompol kali ya_, hihhi…

The range of \( (CI^3_{\lambda} \times 100 / n^3) - (CI^8_{\lambda} \times 100 / n^8) \); (the percentage range of CI linguistic in Spoken Text 5) at random intervals = 25.5%–20.1% = 5.4%

\[
CI^\lambda : n
\]
1. \( 25 : 103 = 24.3\% \)
2. \( 37 : 165 = 22.4\% \)
3. \( 59 : 231 = 25.5\% \)
4. \( 68 : 289 = 23.5\% \)
5. \( 78 : 345 = 22.6\% \)
Text 6
Entertainment Salah Gaul – Maknae Sayang Nenek
Total Word Count: 106
Words with CI Linguistics: 36 (CIλ x 100 / n = 33.9%)


Nenek: Kamu cantik ya
Krystal: Makasih
Nenek: Mirip nenek, tapi waktu nenek masih muda
Krystal: Masa sih? Yah, kalo gitu gue nggak mau tua ah!

Dana: Nenek kalo udah dewasa mau jadi apa?
Nenek: Gue udah dewasa cu, malah udah kelewat dewasa
Dana: Nenek bohong! Nenek masih balita kan?
Nenek: Maksud lo?
Dana: Iya, tuh gigi nenek ajaa belum tumbuh?
Nenek: Ini ompong cu namanya

Nenek: Aduh cu! Pasang musiknya jangan kenceng-kenceng dong
Seohyun: Kenapa nek? Kepala nenek pusing ya?
Nenek: Bukan cu! Nenek takut goyang. Tuh kan cu, kalo nenek goyang bebek gimana?

The range of (CIλ² x 100 / n²) – (CIλ¹ x 100 / n¹); (the percentage range of CI linguistic in Spoken Text 6) at random intervals = 36.7%–27.6% = 9.1%

\[CIλ : n\]
1. 8 : 29 = 27.6%
2. 18 : 49 = 36.7%
3. 29 : 83 = 34.9%
4. 36 : 106 = 33.9%

Text 7
Shinjuku Premium Salon Kuningan City Jakarta - Review and Experience
www.cominica.net/2012/07/shinjuku-premium-salon-kuningan-city.html
Total Word Count: 991
Words with CI Linguistics: 238 (CIλ x 100 / n = 24%)

Oke! Aku rasa beberapa orang mungkin sudah lihat post ku kemarin tentang rambut baru aku. Hehe, sorry yah kalau kesannya norak banget, soalnya aku emang udah pengen banget rambut ku diginni dari dulu. lol jadi pas udah tercapai yah rasanya tuh seneng banget! Terus jadinya lebay gitu yah kesannya, wakakakak XD

Nah, banyak banget yah yang nanyain soal rambut baru ku dan soal salonnya. Mungkin masih belum banyak yang tau yah soal Shinjuku salon di Jakarta, karena masih tergolong baru. Mungkin untuk bagi blogger atau yang sering baca blognya Stella, dll udah tau duluan soal Shinjuku salon ini. Tapli gak papa lah yah aku perkenalin salon ini dengan versi ku, lol *banyak gaya

Awalnya itu aku agak malas ke salon karena MAHAL, ya mahal. Aku bukan anak orang kaya yang bisa nyalon tiap bulan dan belum punya penghasilan tetap (sedih), hahahaha jadi yah kalo mau ke salon harus mikir-mikir dulu. Tapi, temanku Miharu sering banget ngajakin aku untuk ke Shinjuku ini. Dia sering cari di google dan makin pengen abis baca blog nya si tella xD. Dia pengen banget rambutnya blonde kaya Mugi (K-On), maklum cosplayer.

Hahahaha lalu Stella bilang katanya di Shinjuku ada promotion gitu, aku pikir yah mungkin boleh ajaih aku coba kalo harganya cocok, ya gak? Aku coba email Ko Audid dan nanya-nanya, tapi dibalesnya agak lama sepertinya dia gak sempat buka karena sangat sibuk. Lalu, Miharu akhirnya telp dan tanya langsung ke Shinjuku soal promo nya dan kita akhirnya deal untuk datang kesana Rabu kemarin untuk Konsul langsung sama Ko Audid nya karena rambut tiap orang kan berbeda yah.

Shinjuku Premium salon ini letaknya ada di Kuningan City lantai 1, aku sebel banget ya karena ada 3 in 1!! Dan jauuhh banget dari rumah ku, maklum lah tinggal di pelosok desa LOL. Nah aku janjian untuk jemput Miharu di pluit junction, aku sampe bawa boneka tiger yang didandanin kaya orang in case sesuatu terjadi, ahahhahahahha.

Dan akhirnya kita sampai, tapi telat karena MACET banget T____T , bener-bener cape di jalan deh, sampai juga muka uda kaya orang stress. lol. Waktu sampai, mal nya sih masih sepi banget yah tapi ada beberapa salon yang udah buka disana, salah satunya Shinjuku premium salon ini. Waktu liat dari jauh, interior salonnya didominasi warna gelap orang in case sesuatu terjadi, ahhahahahahha.


atau seperti rambut Kyary. Nah itu bisa aku dapetin tanpa di bleach lagi karena tingkat terangnya udah cukup buat dapetin rambut itu. Hebat banget yah, mereka bisa langsung tau
harus campur warna rambut apa. Lalu, aku bilang lagi kalo aku pengen di highlight tapi bawahnya aja di ombre gitu. Terus mau di highlight pink juga, hahaha aduh banyak mau nya banget yah aku?

Kokonya bilang, bisa aja kalo mau digituin, dia bilang coba di blok aja biar lebih beda kesannya. Biasa kan banyak yang di gradasiin yah, dia bilang karena aku berani lebih baik coba yang beda aja apalagi kan aku suka sama Jepang, jadi gak masalah kalo aku pede2 aja. Lalu untuk soal Pink nya, dia bilang sayang karena akan cepat hilang dan kalo nambah uang lagi budget ku gak ada jadi, Kokonya malah suruh aku pake cat rambut Sasha itu loh di eksperimen sendiri, karena dia lihat aku bisa bleach sendiri kan, highlight kan lebih gampang karena cuman sedikit. Kalo ngga berani, kokonya mau bantuin, ya ampun baik banget kan?? Aku belum pernah seumur-umur ke salon ketemu stylist sebaik ini loh, bener-bener mau bantuin kita.

Nah mulai make over nya, kokonya gunting rambut ku. Dia bilang, teknik gunting dia berbeda dengan yang lain. Kalo untuk warna semua salon mungkin bisa, tapi untuk gunting dia punya teknik dan konsep sendiri. Awalnya aku gak mau digunting sih, tapi gak ada salahnya kan di coba apa lagi aku dapat kesempatan di make-over sama kokonya, hehe.


Nah, pendapatku mengenai Shinjuku Premium salon ini:

♥ Staffnya baik dan ramah! Ini bikin aku ngerasa nyaman aja di salon ini, semuanya baik-baik. Bukan cuman aku aja yang merasa begini, temenku Miharu juga. Oia mampir ke blognya kalo mo liat rambut dia yah.

♥ Pelayanannya 100% professional! Biarpun budgetdutu sedikit tapi service dan pelayanannya benar-benar memuaskan, beneran loh aku gak lebai kali ini. XD
♥ Tempatnya bersih, wangi dan nyaman. Sayang dingiininbanget, hahah tapi staffnya baik sekali mengijinkan aku untuk tetap pake salon cape nya terus, lol.


♥ Tempatnya jauh, buatku sih. Hahahah tapi aku pasti akan kesana lagi karena aku udah puas banget sama pelayanan dan hasilnya.

Nah buat kalian yang penasaran sama harganya, nih aku lampirin disini. Menurutku sih harganya itu udah affordable banget yah dengan banyak diskon pula. Dan produk yang mereka pake itu Loreal, loreal itu termasuk yang bagus loh untuk produk di Salon.

The range of (C1λ2 x 100 / n2) – (C1λ8 x 100 / n8); (the percentage range of CI linguistic in Spoken Text 7) at random intervals = 31.1% – 21% = 10.1%
Text 8
Coblos Cinta
Total Word Count: 122
Words with CI Linguistics: 46 (CIλ x 100 / n = 37.7%)

Page 3
“Sori kita lagi ngumpulin tanda tangan buat Bella, buat calon kandidat ketua BEM. Boleh minta tanda tangannya nggaak?” kata Sasha dengan nada menggoda dan manja.
“Gue masang seratus di MU, Put!” kata seorang cowok dengan suara mengambang.
“Eiit, syarat dari bandar, tiap masang harus tanda tangan ini dulul!” kata putri cepat.

Page 4
“Menurut gue sih siapa aja boleh nyalonin diri – “
“Emangnya lo yakin, bisa jadi pemimpin yang baik kalo kepilih?”
“Itu kan kalo kepilih”.
“Berarti lo engga yakin menang?”
“Bukan gitu!, makshu gue..”
“Oke makasih”.

Page 8
“Duuh, besok pengunguman, kok gue yang deg-degan banget ya?”
“Ya iyalaah, emangnya pemilihan cover girl pake penjurian segala?”
“eh, siapa lagi sih yang masukin selain gue?”
“Kayaknya Aldi juga ikutan deh. Tadi gue liat dia juga minta tanda tangan sana-sini”.

The range of (CIλ³ x 100 / n³) – (CIλ¹ x 100 / n¹); (the percentage range of CI linguistic in Spoken Text 8) at random intervals = 37.7% – 21% = 16.7%

\[ CI\lambda : n \]
1. \( 11 : 52 = 21\% \)
2. \( 27 : 84 = 32.1\% \)
3. \( 46 : 122 = 37.7\% \)
Menjadi Pacar Setia

“Bang, enakan mana kuliah di luar ama di sini?”
“Ya enakan di sini lah… lebih deket. Jadi bisa makan masakan Nyokap kita yang enak itu”.

“Pasti si Cumi sinting itu yang cerita ya?”
“Maya itu enang suka cerita. Tapi belum tentu dia lah. Serumah juga kan udah tau. Emang udah berapa lama loe pacaran sama Truly?”
“Gak lama sih. Cuma kalo ukuran jadian yang sebelumnya, ini yang paling lama… Yah gitulah”.
“Kenapa loe? Lagi ada masalah ya?”
“Itu, kalau manusia tuh enang udah dari sononya, harus berpasangan ama satu orang ama? Jadi, satu cowok ama satu cewek, gitu”.

“Ooh… orang ketiga nih. Pertanyaan loe aneh banget. Menurut loe sendiri kenapa?”
“Ya itu ada bagusnya ada nggaknya kok.”
“Wah belum kepikiran deh. Tapi kayanya Truly yang bikin gue setia kayak gitu.”
“Bukannya itu justru bagus? Ato loe ngerasa keberatan, sampe-sampe ada orang ketiga dari sisi loe?”
“He. Iya kali yah. Kenapa ya ada orang yang selingkuh dan ada yang setia?”

“Kalo dari ilmu yang gue pelajari sih, tingkah laku monogamus itu dipengaruhi oleh hormon vasopresin di otak. Orang yang setia bisa jadi karena hormon itu. Kenapa ada yang selingkuh? Karena cuma 3% dari bangsa mamalia yang setia ama satu pasangan alias monogamus. Dan manusia gau termasuk di dalamnya lho.”

“Ah, yang benar loe’??!! Jadi emang benar ya kalo kita ini pada dasarnya mahluk yang suka selingkuh. Heuheuheu… terus kenapa jadi ada kebiasaan harus punya pasangan cuma satu?”
The range of \((CI \lambda^1 \times 100 / n^1) – (CI \lambda^6 \times 100 / n^6)\); (the percentage range of CI linguistic in Spoken Text 9) at random intervals = 53.8% – 42.5% = 11.3%

\[
CI \lambda : n
\]
1. \(35 : 65 = 53.8\%
2. \(67 : 140 = 47.8\%
3. \(96 : 214 = 44.8\%
4. \(123 : 270 = 45.5\%
5. \(161 : 350 = 46\%
6. \(178 : 419 = 42.5\%

Text 10
Dari Presiden ke Presiden
Total Word Count: 172
Words with CI Linguistics: 62 \((CI \lambda \times 100 / n = 36\%\)

Page 1-4
“Cihuuuuuy kita udah dapet presiden baru Mak!”
“Halah!! Kirain ada apaan… kalo itu mah gue juga udah tau!”
“ooo… Udah tau yaaa?!?”
“Gado-gadonya boleh juga nih Mak! Mau dong!”
He… he… he… Tumben lu beli gado-gado gue, biasanya lu kemari cuman nongkrong doang!”
“Abis… girang ada presiden baruuuu”
“Girang mah, bole aja… Menaruh harapan juga boleh… yang penting presiden kita kudu bisa memenuhinya…”
“Memenuhi apa, mak? Bak mandi? Kok ngomongnya jadi serius gitu siiih…”
“Ya memenuhi harapan rakyat laaah… terutama soal ekonomi gituuuu…”
“Kalo sekaran kan yang terjepe masih banyaaak!”
“Yang masih bisa gaya juga banyak, Mak! Contohnya kayak saya gini…”

Page 12
“Sebentar lagi Ramadhan… sembako pada mahal, Pak…”
“Biarrlah, Bu… Ramadhan atau bukan, kita udah terbiasa puasa…”

Page 13
“Nah, yaaaa… masih punya TV!!! Berarti bukan termasuk penduduk miskin dong?”

Page 18
“Sembako naik lagi… kompor udah keren pake gas, akhirnya cuman dipake buat masak air panas!!”

Page 24
“Hnmrmmmph… Nasiib… dari minyak tanah, gas, sembako, sampe berobat ke dukun juga mesti antri!!!”
Maap, Mak… berhubung tarip angkot turunnya cuma dikit, kembalian nya permen aja, yah…”

Kenapa, Bu?! Kok kayak orang lagi sakaw gitu…”
“Gas emang kayak NARKOBA!! Pertamanya GRATIS, kalo udah ketagihan belinya MAHAL!!! Langka lagi…”

The range of (CI $\lambda$ x 100 / $n$) – (CI $\lambda$ x 100 / $n$); (the percentage range of CI linguistic in Spoken Text 10) at random intervals = 42.4% – 33.3% = 9.1%

$CI \lambda : n$
1. 36 : 85 = 42.4%
2. 38 : 101 = 37.6%
3. 40 : 113 = 35.4%
4. 46 : 128 = 36%
5. 47 : 141 = 33.3%
6. 52 : 151 = 34.4%
7. 62 : 172 = 36%

Text 11
Tiga Manula Jalan-jalan ke Pantura
Total Word Count: 188
Words with CI Linguistics: 50 (CI $\lambda$ x 100 / $n$ = 26.6%)

Page 1-5
Kisah tiga Manula kali ini kita awali dari sebuah sumur. Sumur? Ya… sumur!
“Waaal!! Ngapain lu? Kecemplung yaaa?!!!”
“Ndak apa-apa, kok! Aku cuma lagi rendeman saja…”
“Aku lagi kangen sama desaku, Tingal… sudah 40 tahun aku ndak pulang…”
“Lalu, apa hubungannya sama sumur Wal?”
“Dulu di desa Tingal ada sebuah sumur. Setiap hari aku selalu membantu ibuku nyuci di sana…”
“Oalaaah… Waluyoooo!! Kok airnya diobok-obok!”
“Belajar nyelam, Mbok!”

“Aku memang anak yang baik. Ibuku sayang sekali sama aku”
“Bocah edan!!”
“Weee… ndak kena!”
“Aku keepingin mudik…”
“Tunggu dulu… Emangnye waktu muda lu mirip Michael Franks gitu, Wal?”
“The Lady wants to know…”
“Tenang, Wal… Nggak usah sedih… Gue anterin mudik mau, nggak?”
“Yang bener, Liem? Diantar ke Tingal?”
“Tingal itu ada di mana, Wal?”
“Nal, itu dia… Nip! Aku lupa! Kayaknya di Jawa Tengah… tapi mungkin juga Jawa Timur…
“Aneh… Perasaan barusan lu belum pake baju?”
“Ya udah… kita telusuri aja jalur mudik Pantura!!” Siapa tau, lu inget di mana Tingal”
“Kalo gitu, tunggu apa lagi? Ayo berangkat!!”
“Lho? Udah siap?”

Tips: Bawa air minum agak lebih! Selain untuk minum, juga buat jaga-jaga kalau radiator mobil Anda membutuhkannya.
“Kenapa, Liem? Ada yang salah? Bawa Aqua yang banyak, kan?!”
“Sekalian aja, bawa dispersernya…. Bodoh!!”
Tips: Jangan lupa bawa jaket, sweater, atau pakaian penghangat lainnya. Perjalanan panjang siang malam di dalam mobil berAC bisa menyebabkan masuk angin!”
“Lebih pantes lagi kalo lu bawa senter sama Kentongan Wal!!”
The range of (CIλ3 x 100 / n3) – (CIλ4 x 100 / n4); (the percentage range of CI linguistic in
Spoken Text 11) at random intervals = 32% – 26.6% = 5.4%

\[ CIλ : n \]
1. 16 : 55 = 29%
2. 23 : 81 = 28.4%
3. 42 : 131 = 32%
4. 50 : 188 = 26.6%

Text 12
5cm
Total Word Count: 182
Words with CI Linguistics: 54 (CIλ x 100 / n = 29.7%)

Page 80-87
“Gimana kalo kita nggak ketemu dulu sementara?”
“Ta, lu bercanda kan?”
“Serius ni!”
“Aku setuju!”
“Aku juga ngerasain hal yang sama, Ta.”
“Kita terlalu sering bersama, sampe udah lupa gimana rasanya sendirian”
“Sesekali kita perlu keluar dari ruang nyaman kita!”
“Gimana yang lain?”

“Gue juga mau! Gue harus selesain skripsi, enak aja semua udah lulus, gue belom!”
“Arial?”
“Boleh”
“Kalo Riani?”
“Nggak… gue nggak mau kehilangan kalian…”
“Nggak lah ni… nggak akan sampai segitunya…”
“Percaya deh, sama kita!”
“Sudah, jangan cengeng! Kita udah sering bertindak spontan gini kan?”
“Terus… untuk berapa lama?”
“Enam bulan gimana?”
“Nggak mau! Kelamaan ah! Nggak mau gue!”
“Gimana kalo tiga bulan aja”
“Setuju!”
“Ok juga”
“Boleh”
“Gimana Riani?”
“Iya deh, gue ngikut”
“Oke kita sepakat ya, jadi selama tiga bulan ke depan kita nggak ketemu”
“Nggak nelpon, sms atau komunikasi apapun…”

“Tiga bulan dari sekarang… kita ketemu lagi tanggal 14 Agustus”
“Hehehe aku nggak kebayang apa jadinya hidup kalian tanpa diriku…”
“Lebih damai pastinya!! Sini lo, Juple, gue peluk sampe remuk”
“Teman-teman, malam ini kita berpisah, untuk bertemu lagi!”
“Selalu bersama dalam hati…”
“Walau jarak memisahkan”
“Sampai kita tua nanti”
“Ingatlah akan hari ini!”
“Ya! Ingatlah hari ini!!!”

The range of \((CI_{\lambda}^2 \times 100 / n^2) – (CI_{\lambda}^4 \times 100 / n^4)\); (the percentage range of CI linguistic in Spoken Text 12) at random intervals = 39.2% – 29.7% = 10.5%

\[
CI_{\lambda} : n \\
1. \ 13 : 41 = 31.7% \\
2. \ 29 : 74 = 39.2% \\
3. \ 49 : 130 = 37.7% \\
4. \ 54 : 182 = 29.7%
\]

B. **Corpus: Television Shows and Film**

**Word Counts**

**Dialogue 1**

Total Word Count: 154
Words with CI Linguistics: 75

*trima* kasih, *trima* kasih, biasanya saya *dilemparin* bata, sekarang bunga
Yaaa….Pa Edi *molor* nih,….Pa Edi, udah imsyak…udah imsyak, *wah bedung* ada yang *maling*
weh, udah mulai *nih*
udah mulai….tidur aja, *kebanyak* *njabalen* begini *nih*
udah mulai juga
saya kemana *nih*? Ke Jerman?
kesana
ya udah…*bantuin* dong, saya *nga* *ngelihat* *nih*….orang *dongo*, yang *nga* lihat gue *kenapa* *lu* *yang* *gue*
tuntun? *Udah* *sono* *sono*
tanya *dulu* *dong* kenapa
kenapa *telat*?
*ya* begitu-ten…. *nga* boleh *nyanyi* tadi *katanya*, *terus* kita *dikonci* berdua-duaan di kamar mandi
Oh... olahnya Sule ini?
Iya
entar saya bales..... akhirnya Amipur membawa lari Isabella ke negaranya, bagaimana kelanjutan ceritanya awal kisah kita mulai dari TKP
Kenape sih lu sedih banget sih?
mau diputusin
aduh lu cengeng banget, cowo mau diputusin ama cewe nangis!? Aduh
udah udah, ceritanya sambil duduk, sambil duduk
gue tuh ganteng....(got pushed by Ami and chair collapsed when sat upon)...kamu tega (Andre cries)...(Ami tries to console)...nga usah pegang-pegang

Dialogue 2
Total Word Count: 408
Words with CI Linguistics: 219

Setelah merubah penampilan...( ada cewe nih...ada cewe...) munculah seorang wanita...jadi korban pertamanya Thomas
kasih, kasih rayuan ...pantun
pantun pantun.... sipirling....hai
hai
elu dong...itu cewek agresip
kok tawanya gitu banget?
itu agresip memberikan senyal-senyal cinta
masa gitu...ahaha
nga papa
Hai
Iya mas
siapa nih....ah ini cewe sarap apa yang ini nih?
bukan lu, ini anak orang kaya, kelihatan dari penampilannya
elu nga ngelihat sih, elu dari depan, gue mah dari belakang kelihatannya laen, kaya pohon beringin ketup angin
halo
halo
siapa namanya?
susah ngomong
namanya siapa?
nama saya Ajeng
Ajeng?
Sonya
Sonya
anak orang kaya ya?..... gue rasa orang gila ini...
situ kan pasti dukun, kok tau kalo saya orang kaya, situ dukun ya?
dikatain dukun...orang gue playboy
ini temen gue, ini anak orang kaya
kok dia tau kalo saya orang kaya
kelihatang...lebar kaya rumah.....rumah orang kaya kan lebar-lebar....gede, kaya rumah lu kan gede
Iya emang anak orang kaya
ngerayuin gue supaya...
ok, lu ikinin gue.....hai cewe...
hai cewe
Iya apa toh
kodok burik
apa urusannya ama kodok burik?
ini pantun mah, udah
elu nyinggung gue Dre... jangan pake burik-burik
ok...kodok sawan les balet
Dre, yang bener-bener bantuin nih
ini kan udah bener gue
.... selama gue lahir nyanpe sekarang gede, baru denger ada kodok les balet
itu kan pantun bro
ngaco nih
...elu jangan terganggu bro, langsung fokus aja
setelah melihat kamu, rasanya aku...jiji.....maksudnya jatuh cinta
jatuh cinta dong
ngerayu tuh yang bener le....ngerayu kaga bisa
udah lama Pak nyuci Pak
udah lama
oh
gini loh men, aku tuh maunya kalo kenalan sama cowo itu aku pingin diajak kemana gitu loh...
ajak jalan bro, ajak jalan bro
jalan-jalan ke mol ok?
ha kalo ke mol mabok, aku harus minum obat anti mabok dulu
emang naek kapal ke mol?
elu jangan ajak ke mol kalo cewe-cewe begini...
loh emang napa sih mas?
mol itu jadi tepung
makin jauh nih...diadon..diadon..set..masukin...tuh...trus dicolokin...set...keluar jadi kueh putu..tuh mau?
mau...walaaahhh
belanja...tempat belanja
shopping-shopping-shoping
tapi duit saya kan kebetulan diluar negri semua, dolar semua, jadi saya nga sempet nukerin ke mesenjer
kok mesenjer?
tempat nukerin duit apa sih?
money changer....elu kalo kelihatan udik jangan kelihatan banget dong...
mae..kalo masalah duit mah gampang...papiku kan juga da uang, entar aku tinggal telpon papi harus kirim...
sekarang dong duitnya...sekarang..nanti kita jalan-jalan
tapi gue duitnya nga bawa
ya udah, aku ikut sama kamu ya....gue jalan dulu ya
ok bro
ayo sayang
yang ono....gue ngapain?
heh, sudah dapet duitnya
sip
itu dia
itu cewe lu kemanain?
gue cumplungin ke sumpur

Dialogue 3
Total Word Count: 160
Words with CI Linguistics: 105
aduh...udah jam sekin, bro gue kemana belum nyampe-nyampe juga nih gimane, kelihatan kaya nga gue?...emas semua!
elu mau ngapain pake baju beginian?
mau ketemu ama cewe
orang mah kalo orang kaya pake jas, elu pake baju dramben, nora lu ah biar nga malu, cewe gue kan anak unindra bro...biar keren gini
elu sih nga malu, gue yang malu
dia mah nga papa yang penting gue pd aja deh
tapi duit ada kan?
ali itu dia abis
elu keren banget bro, cuma masalahnya kalo nga punya duit gimana lu mau ajak jalan?
dah lah...sekarang kita malak aja lah
nga boleh malak...itu perbuatan dosa.....nyopet aja
itu sama aja
kalo nga kita bagi tugas
gimana?
gue yang mancing perhatian dia...elu yang dari belakang
itu mantap bro
iya nga?
gue dari belakang biar nga kelihatan...elu ajak ngobrol tuh orangnya
gue ajak ngobrol terus yah, abis itu ambi set
bukan begitu, elu ajak ngobrol, elu bikin kopi
nyopetnya kapan?
inget, supaya lu dapet duit lagi

Dialogue 4
Total Word Count: 238
Words with CI Linguistics: 74

In this dialogue the actors Andre and Sule were performing authority roles and it is note worthy that BB linguistics come into greater interplay in accordance with BI socio-linguistics were people in authority are more likely to employ BB to assert their status.

He, ngapain kamu disitu?
Kelilipan Pak
Sini saya tiupin
Pantesan sakit Pak
Lu kelilipan beginian matalu segede ape?
Mana saya tau
Heh, disini kita waktunya latihan ... kita ini adalah life guard, kamu tau artinya life guard?
ah.. tau Pak
apa?
lap ged
life guard....life itu artinya hidup, guard itu artinya penjaga, life guard?
Penjaga hidup
iya, betul.... saya akan mengajarkan kamu bagaimana cara menyelamatkan orang, seandainya kamu
ada di tepi pantai, di lautan kamu lihat orang begini-begini
oh...minta tolong Pak
kamu dadain... itu artinya dia mau pergi
kalau saya lagi ada di pantai, ada orang dada
kamu dadain lagi...gitu ya
langsung dia pergi? Iya Pak?
Mati! Orang tenggelam
catatnya suruh dadain
kamu itu junior, saya senior disini
ngapain kamu pake ngajar-ngajar junior...hah? yang lebih senior disini saya! Kamu liat di lautan,
ikan apa aja, siapa yang nga kenal ama gue? Semuanya kenal! Orang cumi-cumi aja cium tangan
sama gue...enak aja pe
kaka keluarga muaer ya?
enak aja...betok...saya keluarga betok
lu jangan gitu...masa keluarga muaer...lu liat mukanya kaya yang julung-julung
ini anak baru kak, baru masuk sini
ktp dong ktp?
ien...ktp kasih
kak pake ktp pak?
ini aturan men!
emang kita bayar pajak?
ini aturan...ini galaknya ini aturan tau nga
lu tau ini bawa apa? Heh...ini bawa apa?
Wajan Pak, penggorengan Pak
pinter, berarti lu nga ketara...ini wajan buat menolong seseorang
gimana caranya?
seasat ada yang kelelep, kita naik ini...kita masuk diatasnya
langsung tuh kita Pak, kita nyamperin

Dialogue 5
Total Word Count: 247
Words with CI Linguistics: 107

idih...udah kelewat sepuluh hari...
dalang nga usah pake lilin kali...cari obor...obor...obor
nih kadonya nih, suruh buka satu satu dulu nih...awas
ngucapin selamat dulu, ama pengantin sunat...selamat...selamat yah
Bapak...saya nga mau sunat lagi Pak
kami dari semua tim opera van Java mau mengucapkan selamat ulang tahan ya
entar entar kadonya buka.....bukan itu bukan buat telpon
ini berat ini, pasti mahal ini
kaya mau ujan bunyinya.....
Dalang biasanya nga pake sendal
entar entar biasanya berapa kali bayar ini? Biasanya lu ngasih ngasih hadiah?
tiga kali bayar
tu kan....nga nga jadi saya
bayarnya pagi siang ama malam
abis sama aja
udah itu tes...siapa ingin mengucapkan sesuatu mungkin
ayo makan dulu, biarin yang ulang tahan ma makan aja dulu....dari bu Sri ni...bu Sri bu rt bu rt
dah, Dalang suruh ngomong, kasihan
yang ulang tahun siapa?
ulang tahun kok nga ngaku loh?
udah kelewat tuh
ini...nah...nah
ini yang bagus ini......ini pres on pres
hah?
pres on pres...timen ini
assalamualaikum waruhmatulah wibarakatu
yang bunyi nih yang ini
pada kesempatan ini saya mau mengucapkan...
ayo kacang rebusnya kacang rebusnya
dah..ayo silakan..dah
langsung kasih pidato
masih kurang le...sekalian lah
entar dulu entar dulu entar dulu...biar kelihatan keren, biar kelihatan keren
kaga salah nih...baju udah dicopot paling disiram terigu nih
kaga
kacamata lu puyeng banget kok?
biar keren kita ada baju...ambil baju Dre
ini minnya kelewatan ini mintar sini nih
ngapa sih anak-anak
biasanya Aril ama Luna, ini Luna
ngelunjak lu
biar lebih wibawa
ni ada apa sih
biar keren...kalo dia ada ulang tahun....udah gagah

Dialogue 6
Total Word Count: 79
Words with CI Linguistics: 42

Penonton, perkenalkan saya ini ada...orang gendangnya nga bener nih...orang lagi ngomong jangan
main gendang dulu
ini latihan le musiknya...masing masing aja...latihan
apa lu kata...saya disini berperan...orang lagi ngomo...noh...orang gendang ganggu tuh...orang lagi
ngomong
ey...tukang gendang keluar aja dulu...ini mau ngomong nih
nga, kalo nga bisa tangannya potong aja dulu biar nga mukul mukul gendang...enak aja... saya ini
Aceng...punya temen namanya Sekak
kan saya udah tau .. saya bacain tadi
saya lagi nungguin si Sekak nih...biasa dia lagi di got....Sekak...Sekak

Dialogue 7
Total Word Count: 73
Words with CI Linguistics: 35

Assalamualikum
Bagi dong Pak ama orang yang nga berdaya Pak, keluarin duinnya yang di laci Pak...
he he mas...nga ada orangnya kok
oh begitu...kenalu lu nga bilang dari tadi kadal
Ri, ada tukang mina-minta, badanya gemuk, sehat lagi
masa
kita kerjain yu, tunggu sinia ya
cepetan Pak...cape nih
salamualikum
pindah apa ya?
bagi Pak
lama-lama gue copotin nih pintu...ada protetmu...sarap lagi
lari
aduh...cape (sambil berlari)
rasain lu pemelas, nyari duit kaga buta, dasar

Dialogue 8
Total Word Count: 236
Words with CI Linguistics: 175
Ben, pelan-pelan dong you jalanye, aye sedang meriang nih, Ben... brengsek lu ah ah... diem-diem aja lu di situ, moral aja terus, lu taunya sanpe, ah... pake meriang segala, udah tau orang mau ngungsi, mau ngikut, jage diri lu baek-baek, gue nga buang aja udah bagus lu... ah... let's go aduh Ben... Ben... tobat ah... aye bisa mati di jalanan nih... aduh... slowly... slowly tiger... slowly Ben... Ben... plosotan Ben... pelan pelan dong you jalanye, aye sedang meriang nih, Ben... brengsek lu ah... diem-diem aja lu di situ, molor aja terus, lu taunya sampe, ah... pake meriang segala, udah tau orang mau ngungsi, mau ngikut, jage diri lu baek-baek, gue nga buang aja udah bagus lu... ah... let's go aduh Ben... Ben... tobat ah... aye bisa mati di jalanan nih... aduh... slowly... slowly tiger... slowly

Ben... Ben... plosotan Ben... pelan pelan Ben... aduh aduh aduh... bisa nyangkut aye nih... Ben... mau dibawa kemana sih Ben... Ben... mau ke mana?

Sorry dongo... memang nasib lu... c'mon tiger... mudah-mudahan nga dicaplok buaya lu pake lewat sungai lagi... aduh... dingin... aye sedang meriang nih... Ben... kira-kira dong... kau kira aku ini ikan kapus... Ben... apaan tuh?... pada ngambang nih gituan... Ben... lekasana dongo shut up! Merendem aja situ terus... ama gituan aja takut... bencet aja... ah... masa nga ancur... eh norok aja situ terus

Ben, pelan-pelan dong you jalanye, aye sedang meriang nih, Ben... brengsek lu ah ah... diem-diem aja lu di situ, moral aja terus, lu taunya sanpe, ah... pake meriang segala, udah tau orang mau ngungsi, mau ngikut, jage diri lu baek-baek, gue nga buang aja udah bagus lu... ah... let's go aduh Ben... Ben... tobat ah... aye bisa mati di jalanan nih... aduh... slowly... slowly tiger... slowly Ben... Ben... plosotan Ben... pelan pelan Ben... aduh aduh aduh... bisa nyangkut aye nih... Ben... mau dibawa kemana sih Ben... Ben... mau ke mana?

Sorry dongo... memang nasib lu... c'mon tiger... mudah-mudahan nga dicaplok buaya lu pake lewat sungai lagi... aduh... dingin... aye sedang meriang nih... Ben... kira-kira dong... kau kira aku ini ikan kapus... Ben... apaan tuh?... pada ngambang nih gituan... Ben... lekasana dongo shut up! Merendem aja situ terus... ama gituan aja takut... bencet aja... ah... masa nga ancur... eh norok aja situ terus

Dialogue 9
Total Word Count: 456
Words with CL Linguistics: 272

Nek...
apa...
kopi segelas
heh... saya duluan... enak aja lu... cepet nek
brengsek lu
woh... apaan nih... nenek kok gimana sih... mentang-mentang kopii ngutang dicampur sisik... (minum air panas)... aduh diseduh lagi gue
aduh kok gue yang lu sembur... jiggong lu nih
hai nek, mau ngracunin apa mau gimana... kalo nga senang jangan begitu dong... mulut saya diseduh maap deh Min, nga sengaja nga sengaja sih ampe dua kali
maklum nenek udah piken
... bagus tuh... lu bakal ketiban rejeki
lah ketiban rejeki... yang jelas gue ketiban sisik... ampe keseduh mulut gue
air panas lu pake kumur
sh sh... cewe tuh kemari In... tuh... tuh... asyik
cakep juga ya cewenya
ini cewe baru
baru sih baru... tapi ati-ati tu... orang tuanya galak kaya macan
ah... baru kate kaya macan... kepuh gue ketikin... heh... sekaran begini... ayo kita taroan ya... seorang gocep... limapuluh... limapuluh perak kita gontian tegor tuh cewe... siapa yang tegorannya disautin... dia yang menang... ok?
nenek ikut nek? (nenek meludah)... ini duit gue nih... ini duit gue... ni doku gue nih
lu ikut nga... ya udah
ini baru cewe nih...gue dulu yang negor ah...pasti gue yang menang taroan nih...gue dong...cakep ya!...hebat juga nih...pantes buat bini gue nih...ah...bini gue dong...iya siapa kek duluan deh
dah Ida
amit-amit
ya hahaha...galak o
ilang gocapan gue
tengil pada lu ah
alo Ido, mo kebanyo, ngasih kiy, randang duo
hahaha apa loe kate, muka lu kaya kuto
ah...tua-tua nga tau diri
orang bigul...liat gue...alo wido...lupa ama gue ya, sombong banget
Cuma segitu aja nih? Nga ada lagi nih? Lu lihat kalo gue yang panggil pasti nyahut...kalo kaga lu
boleh iris kuper gue...tolong! tolong! tolong! (siul)
kirain siapa
aha....menang gue
hebat lu Min, pake jimat apaan sih lu?
Jimat palang becat
Emang jago lu Min, tapi lu bayarina kita semua dong
jangan takut, gue yang traktir
cepe dong
ah..baru kate cepe...tuh bakal lu...nenek mau minjem lagi
dah pulang yu...
ok jangan kuatir...gue yang traktir...gue yang bayar....jadi berapa duit semua nek...
empat ratus perak
ah nga salah itung?
kalo...bohong itu pantangan tuh...masa bohong
aduh...cilaka...menang tiga ratus...si Aneng pinjem seratus...bayar minuman empatratus...aduh kalo
gitu gue yang rugi dong....ah...dongo juga gue dah
Ida...Ida..
Iya Pak
sini cepet
ada apa Pak
kemana aja sih
ke pasar
masa ke pasar begitu lama...orang yang belanja aja satu truk nga begitu lama
ya ileh...bapak segitu curiganya...emangnya orang belanja seperti orang makan...begitu ada makanan
langsung diganyang...orang kan mesti tawar-tawar dulu Pak...supaya jangan kemahalan
pinter aje lu nyari alasan
ada apa sih ribut-ribut melulu...Pak...alasan apa sih Kleopatra sampe bunuh diri
masa bodo...mati ke...mampus ke...sodara gue juga bukan...motong aja lu kalo orang tua lagi
ngomong...ei...dengerin ya...ini soal penting...jangan coba-coba bergaul ama anak kampung
sini...Bapak denger dia pada brengsek...gelandangan...bikin ruih kota
itu kan baru kedengerannya...cuma info...kenyataannya kan belum tentu begitu Pak
Diem..kalo dibilangin lantas lu ngejawab...ade ama kakak sama gilenye...sana dapur
Alah...udeh..ini kan hari minggu...tentsnya banyak orang belanja...lagian anak baru pulang sedikit
telat aja udah begitu ributnya...belum juga sampe pulang pagi
coba-coba pulang pagi

Jakarta undercover #1
Dialogue 10
Total Word Count: 298
Words with CI Linguistics: 113
Gimana sih pada belum siap..ayo... lima menit lagi mulai... kalo sampe semua kacau, balik semua ke jalanan ya!...
He sayang...ada lapor?
Kasian banget deh mbo...kenapa sih mbo kok dia seneng banget di tempat gelap?
You know what...lu sadar nga?... kayanya lu dikutuk deh Vi, semua orang didekat lu nga ada yang normal...gue banci ade loits
Gue udah nga tahan lagi man...lu tau kan kenapa gue nari kayanya gini...cuma buat Ara...tambahin lagi sekarang yang ngejagain dia udah nga mau jaga lagi...tetangga gue juga udah mulai nanya-nanya
Viki, Viki liat gue, lu cukup nari sekali lagi aja mbo, itu udah cukup untuk biayain Ara, Viki, lu mesti kuat untuk Ara, biarin aja mereka semua nanya-nanya, nga ada yang penting mbo...
Yang nanya itu ada yang polisi man, kalau mereka tau gue dicari-cari polisi gimana? Yang ngejagain Ara yang ngerawat Ara itu siapa?
Vik, gue minta lu dengerin satu omongan gue kali ini...nda ada yang gue biarin satu pun orang nyaktin lu...itu janji gue sama lu...gue janji siapapun itu nga akan mbo...heh, apa ini...my baby don’t cry...never...Viki, Viki...no...c’mon...cheer up
Makasih ya
Anything for you baby
Viki, kita harus dance, cheer up, dance baby, yo, let’s go...masih keburu
Bye ya Ra..aman nga ya?
Apa disini? Mbo...gue udah lama disini...ini aman, ini cuma tempat buka siang untuk meeting Jefrey...
He how are you? Gimana? Having a good time?
Jeff, itu emangnya siapa? Itu yang ditengah namanya Viki, dia baru sebulan disini
Gue mau bawa dia malem ini ya
Gue cuma bisa tawarin lu private dance, selebihnya susah Min, sorry
Kenape?
Ya banyak yang mau sama dia...dia tuh nga bisa
Enjoy ya...thank you
Gue bakal dapetin dia malem ini
Taroan ya
Bot, terus terang ya, gue paling nga enak ambil duit lu
Seratus juta...
I’ll take it
Siapapun yang menang lu kasih gue cash

Buruun cium gue
Dialogue 11
Total Word Count: 139
Words with CI Linguistics: 56

udah deh cepetan cium nga usah pake drama-drama lagi
cepat cium
ciumnya pake lidah dong biar seru
ini peringatan buat kalian juga!
Mereka ketawan lagi ya
Mereka tuh kayanya nga pernah kapok deh
Tapi kan Pak, ciumannya diluar jam pelajaran Pak
Diem...tetep saja...kalian ini masih pake seragam...di lingkungan sekolah lagi.... gimana kalo ade-ade
kelas kalian niru-niru
Wo...apes ye
Ini sudah ketiga kalinya Reni melaporkan perbuatan kalian
Mampus lo!
Kenapa kalian tidak mencantoh teman kalian sendiri? Tuh liat...Ardi sama Desi, walaupun mereka sudah lama akrab, tapi mereka belum pernah cium-ciuman...
Hahaha...Bapak nga tau aja mereka pernah atau nga
Weh...fitnah lu Jo..nga pernah..nga pernah nga pernah
Nga pernah? Ngapain aja lu berdua?
Wan, kaya nga pernah aja lu sama Jeru...
Pacarnya masih perawan nih...
Gila...cuku abis lu berdua
Apa-apaan kalian..bubar sana!
Kalian terus cium
Cabut yu Di...
Kamu sih genit...gosong nih...

Perawan #1
Dialogue 12
Total Word Count: 334
Words with CI Linguistics: 197

Mel...jorok lah...
Nga usah diberesin...
Jorok banget sih lo?!
Nga berani kan?
Berani gue...
Cupul...cupul..cupul..nga berani
Inceng banget..bandel sih lu...culut
Bian..ini kenapa sih disini
Cing lu..apa-apaan sih lu
Gue da dada Bian
...biar satu sama
Coba liat... oh lala..gedean punya gue...
Minumannya Keting empuk...aduk..aduk
Sayang...liat atas dong
Apaan sih?
Stell, Stell, Stell....gue...gue mau ngelepas keperawanan gue
Gatel juga kan lu...gue bilang juga apa...elo nga bakal bertahan lama
Elo bercanda kan Ket?
Elo mau ngelepas sama siapa?
Siapa aja yang berani bayar mahal
Sekarang?
Disini aja nih...
Eh Lala...yang itu tuh La...
Kayanya nga deh...ini nih tipe tipe om yang takut sama bininya...sakin takut sama bininya...itunya sampe nga bisa berdiri
Nah..itu tuh seniman tuh...
Bau...
Kalo lo mau belajar ngerasa...sama dia tuh makanya bau...karena bau itu adalah bagian dari rasa
Gue jadi bingung deh...lu itu sebenernya mau seks, atau duit sih?
Ya duit dong...
Kalo duit kita salah...dia miskin...
Guys...
Dia berani lima...
Lima juta?
Minta limabelas!
Apa nga kemahalan?
Ya mentok-mentoknya sepuluh lah...tapi cash...
Ok deh...gue coba...
Gini aja...sepuluh juta...ya? ok?
Entar ya...
Sepuluh...
Cash?
Transfer lewat atm... gue nga punya rekening nih...
Pake rekening gue... kapan?
Sekarang...
Sekarang?
Dimana?
Toilet...
Gimana? Elo tuh ngasih nomernya bener nga sih?
Bener kok... emang kenapa?
Nga masuk lho ke nomer gue...
Nga mungkin elu udah check lagi belum?
Elu check deh
Udah checkin tanya Bia
Elu serious Bi?
Ngapain sih kita bercanda?
Elu berdua bercanda kali...
Eh... ini kan duit lu... gue kan cuma ngechekin...
Coba deh lu check sekali lagi
Gini aja... tadi kita sepakat ya... mendingan kita ngomong lagi ke omnya terus transfer ke rekening gue yang lain
Serious lo?
Omnya mana?
Omnya udah pergi... gimana dong...
Elu gila ya... gue kan udah bilang jangan biarin omnya pergi sebelum kita berdua balik...
Terus kita harus gimana? Gimana dong... jadi gue sia-sia dong... duit segitu kan banyak buat gue...
gue bisa beli handphone... gue bisa beli tas...
Tunggu tunggu dulu... elu bilang apa?
Gue mau beli handphone kaya lu... yang ada kameranya
Ya udah gini aja deh... ini gue pinjemin duit gue nih...
Ah... lu berdua bercandain gue...

Lagu (Aku Perawan):
kawan kita bergandengan... namun kita tak sejalan... kawan kita berlarian... namun tak satu
tujuan... wo... o... o... aku perawan... wo... o... o... ku tak sejaman... wo... o... o... aku perawan ... wo... o... o... Biar
aja orang bilang aku ketinggalan...
Kesucian bagiku penting untuk di pertahankan
Dan aku yakin tuhan pun berikan aku tempat yang
Terindah...
C. **Corpus: Billboard Advertisement**

Figure 11 Advertisement 1

**CI Linguistics:**

Kopi Nikmat... TIDAK bikin kembung

1. bikin – CI omission of *mem-* prefix
2. kembung – CI lexical item

Total word count : CI linguistics = 5 : 2
CI Linguistics:
Lu Musti Coba
1. Lu – CI lexical item
2. Musti – CI alternative spelling
Total word count : CI linguistics = 3 : 2
CI Linguistics:
NONTON LANGSUNG MOTOGP 2012

1. NONTON – CI omission of me- prefix
Total word count : CI linguistics = 3 : 1
CI Linguistics:
GAK LUPA DIRI, WALAU SUDAH MASUK TIVI
INI BARU COWO
1. GAK – CI lexical item
2. BARU – CI alternative semantic usage
3. COWO – CI lexical item
Total word count : CI linguistics = 10 : 3
CI Linguistics:
TRI memang Jagoan Intenet!
   1. Jagoan – CI lexical item
Total word count : CI linguistics = 4 : 2

Figure 16: Advertisement 6
CI Linguistics:
REBUT GRATISNYA, SEBARIN SEMANGATNYA
GRATIS GAK ABIS-ABIS
KIRIM 2 SMS GRATIS RIBUAN SMS
CUMA Rp1.000 GRATIS RATUSAN MENIT NELPON
  1. SEBARIN – usage of CI suffix –in
  2. GAK – CI lexical item
  3. ABIS-ABIS – CI elision of consonant ‘h’
  4. CUMA – CI lexical item
  5. NELPON – CI omission of me- morphology
Total word count : CI linguistics = 17 : 5

Figure 17: Advertisement 7
CI Linguistics:
Mau Cepat Silahkan Lewat, Mau Hemat Nelpon & SMS Pake XL
  1. Silahkan – non BB / alternative spelling 
  2. Nelpon – CI omission of me- prefix 
  3. Pake – CI alternative spelling 
Total word count : CI linguistics = 9 : 3

Figure 18: Advertisement 8
CI Linguistics:
Jangan Sumpah kalo cuma jadi Sampah!
1. kalo – CI lexical item
2. cuma – CI lexical item
Total word count : CI linguistics = 6 : 2

Figure 19: Advertisement 9
CI Linguistics:
Ice Cream tentrem
1. Ice Cream – English lexicon
2. tentrem – CI alternative spelling
Total word count : CI linguistics = 3 : 1

Figure 20: Advertisement 10
CI Linguistics:
Streaming Gak Pake Lelet!
1. Streaming – English lexical item
2. Gak – CI lexical item
3. Pake – CI alternative spelling
4. Lelet – CI lexical item
Total word count : CI linguistics = 4 : 3

Figure 21: Advertisement 11
CI Linguistics:
Depatin SuksesMu Di Sini
   4. Dapetin – CI morphology of -in suffix
Total word count : CI linguistics = 5 : 1

Figure 22: Advertisement 12
CI Linguistics:

Ngapain mahal kalo bisa murah!

1. Ngapain – CI morphology of nga-___ -in affix
2. Kalo - kalo – CI lexical item

Total word count : CI linguistics = 5 : 2

D. Corpus: Interviews and Recordings

The corpus of interviews and recordings have been obtained from Sneddon, James Neil (2006) *Colloquial Jakartan Indonesian*. Pacific Linguistics, ACT.
Appendix 2

Semantics and Pragmatics of CI Lexis
and Collocations
This appendix is a document-in-progress that provides CI lexicon and collocation with the purpose of compiling a CI reference. These lexical items will be analysed in terms of the meaning they provide. As single lexical items they contain certain semantic properties. In collocation with other lexical items, some create specific speech acts that perform particular language functions.

The dissection and analysing of lexicon is also known as *componential analysis*: “Componential analysis breaks a word down into a list of the components present in its meaning” (Malmkjaer 2004: 340).

The componential analysis in this section will utilise the theoretical framework of Generative Grammar, Aspect Theory and Generative Semantics. The premise of these theories is that there are different linguistic levels to syntax. For our purpose here, only two of these levels are relevant: Deep Structure (DS) and Surface Structure (SS). At the DS level the semantic properties are manifested while at the SS level the lexical components are represented:

![Diagram of Lexicon, Deep Structure, and Surface Structure](image)

“The syntactical component of a grammar must specify, for each sentence, a deep structure that determines its semantic representation” (Chomsky in Malmkjaer 2004: 200).

Widdowson (in Lewis 2006: 14) classified the nature of lexicon into *signification*; the codified meaning of word (dictionary definition) and *value*; the meaning that may be similar or different to *signification*, when words are used with extended or metaphorical meanings. The *signification* here correlates to the SS of Aspect theory while the *value* is the semantic representation imbedded in DS.

Lexicon at the SS may or may not represent the semantics of the DS. This is true for spoken Indonesian as well as in English. We can take an example of the colloquial English expression “hanging out”. At DS, semantically it represents the value [+to spend time with people/at a place] and not as the SS signification suggests: [+ hanging/suspended off the ground +outside]. While another example, “making tracks”, is a figurative expression where a link can still be made between DS value and SS signification: [+to leave +to go], [+leaving footprints on the ground]. At SS, to ‘make tracks’ also represent a DS meaning of leaving footprints as a mark. In modern colloquial speech the DS value has also extended to ‘leave’ or ‘go’ in general, and it is inconsequential if the Agent performing the action would leave visible footprints or not. An example of an expression in bahasa Indonesia is the phrase “bukan main”. At DS level its meaning is [+incredible +amazing] while at the SS level its signification is [+not play]. A link can perhaps be extrapolated between the DS and SS representations in so far that the object of description is of no ‘play matter’, it is serious and no joke. The learner of Indonesian may or may not infer the semantic components of speech acts and expressions. Of course, the more remote the obvious link between DS and SS values, the more difficult it is to make sense.
Each lexical item and its collocation - if there are any – will be given a language function (LF) - which describes the DS value, the SS signification and the speech act, a syntactical structure (Syntax) - if relevant, an Indonesian speech act example (SA) and an English-Speech-Act-Correlation (ESAC), which is also the translation of the Indonesian example.

The LF definition in this section may be describing either semantic, language function or grammatical function. In the following example +question word and +conjunction describe a grammatical function, [+yes or no] and [+constituent] describe language functions while [+what] describes signification:

**apa**

i. LF: [+what] [+question word: +yes or no; +constituent]
    Syntax: a. Apa_{NP}?  
    SA: a. Apa_{NP} ini tempatnya?
    ESAC: a. Is_{NP} this the place?

ii. LF: [+conjunction: or]
    Syntax: NP apa NP?
    SA: N_{NP} Ini, hpnnya kakak; apa_{NP} hpnnya; mamah_{m}?
    ESAC: Is_{NP} this, your or mum’s mobile phone?

What often is difficult to grasp for students of Indonesian is that the syntactical properties of Indonesian do not correlate with the syntax of L1. In the [+yes or no] example above the syntax of *apa_{NP} ini tempatnya?* Do not correlate with the English translation *is this the place?* In spoken Indonesian often the question word *apa* is dropped all together while in BB *apakah* is used instead. Either way the English syntax contains the copula *to be* and the WH-question word is not operational. In the [+constituent] example *apa ini?* the Indonesian copula is also a null element. The NP *kakak* has an SS signification ‘older sister’ in CI, and in this context, an LF value as second person pronoun.

A typical syntactical feature of CI is its Null Elements. The most common Null Elements in CI is the pronoun NP and the copula verb ‘to be’. When it is necessary, to help the reader with deciphering of the semantic translation, linguistic categorisations of Null elements are indicated as Ø and o. Co-indexing using the letters, will also be used to help connect Indonesian to English translation. The FI correlants are also included if relevant or otherwise marked as Ø. When a lexical item has a grammatical or language function within syntax, then the syntactical cohorts are marked as [____].

With lexical items that do not generate set-speech-act-phrases, then only signification and CI sample is given with the ESAC below it. The ESAC will be focussed on DS semantic value. An SS-signification translation is deemed futile in itself as a tool for semantic description, although it may be at times presented solely for its amusement/interest value. For example, the speech act ‘*ada-ada, aja, kamu,*’is correlated with the ESAC ‘Ø,Ø,You_{m} are_{Ø} such_{Ø} a joker_{Ø},’ whereas its SS signification ‘there-is-there-is just_{Ø},’ gives very little clue to its inherent meaning.

ESAC words that best translate the LF value are co-indexed corresponding to the CI sample. The focus of this section is not comparative linguistic features but DS semantic representation. So in the case of *Andi, orangnya, bener, baik, banget_{s},* (*Andi, is_{Ø} a really_{Ø} nice_{Ø} person_{Ø},*) there might not be a discussion on the ‘why’ of the –*nya* suffix and the DS value of *baik* as ‘nice’ as opposed the SS signification of ‘good’. There is also the feature of double adverbial ‘*bener, ADJ banget_{t},*’ both co-indexed_{t} to ‘really_{Ø}.’ The intention here is to expose the reader to syntactical features of spoken Indonesian and to the ways in which lexical item manifest DS value in context, and provide the closest equivalent English speech act with the same – where possible – DS value.

Much of these CI speech acts presented here have evolved from Betawi Malay. Amongst the Betawi people, Malay (later renamed into ‘Bahasa Indonesia’) was the mother tongue where it naturally and organically evolved into a dynamic spoken language. This contrasts with Bahasa Baku (Formal Indonesian), the derivative of Classical Malay - the royal idiom of the Riau Court, which was mainly
the recording tool of religious scriptures and later on adopted by the Dutch as an administrative language. The Dutch education planners contributed to the structurisation and organisation of Formal Indonesian grammar. Herein a problem has developed. Foreign learners of Indonesian are for most part taught Formal Indonesian. Formal Indonesian has strong ‘European flavours’ in its syntax and grammar, which would have been the product of the Dutch linguists’ compiling and arranging of Indonesian grammar, hence the sense of ‘artificial-ness’ about Formal Indonesian as a spoken language. Foreign learners of Indonesian have difficulty in understanding spoken Indonesian because the syntax, speech act and language functions of spoken Indonesian are not taught in conventional language textbooks. The syntax and speech acts of spoken Indonesian are very unique and more often than not, incomparable to the speech acts of the learners’ L1, which for majority is English. This section will provide comparative spoken Indonesian to English lexicon and speech acts that perform correlant language functions. It is important to note that many CI linguistic items have been absorbed into Indonesian via MSI, but not necessarily into FI at its highest sociolinguistic domain.

**CI Lexicon, Speech Acts and Language Function**

*abis*
Elision of FI *habis*, signification: finish.

i. **LF**: [+conjunction: because]
   [+explaining the course of action as a consequence of circumstance]

   Syntax: [___] *abis* [___]

   **SA**: [Saya, beginiin], *abis*,[terpaksa]

   **ESAC**: [I did like this] because [I had to]

   The predication of the anaphoric demonstrative *begini* → *beginiin* is explained under the entry *beginiin*.

ii. **Collocated with** *gitu* → *abis gitu?*, it performs as an interrogative speech act: ‘and then?’

   **LF**: [+interrogative +constituent: and then?]

   **Syntax**: *abis gitu* [___]?

   **SA**: *Abis gitu* [gimana jadinya,]?

   **ESAC**: What happened then?

   Students of Indonesian may try to insert the FI SS equivalent of the LF *and then?* → *dan lalu?* which has nothing wrong with it. But to translate the CI SS signification of *abis gitu?* Will produce finish like that (*gitu* being an elision of *begitu*). DS meaning can almost be inferred from the SS of the phrase *abis gitu gimana jadinya: finish like that how become?*

iii. *abis*, similar to the collocation *abis gitu*, usually collocated with *dari* → *abis dari* is a CI alternative to the FI *sesudah dari*.

   **LF**: [+preposition: after/from] [just been]

   **Syntax**: a. *abis dari* [___]
   b. *abis* VP

   **SA**: a. *Abis dari* [situ, kita, kei, ...]
   aii. *Abis dari* mana, barusan kita?
   b. *Abis*VP ngebantuin, ade kita pulang

   **ESAC**: a. From, [there, we, went, to, ...]
   aii. Where, did, NP, just come, from?
b. After helping our younger brother, we went home.

Abisan
Variation of abis definition i. (LF: [+conjunction: because] [+explaining the course of action as a consequence of circumstance]), almost functions as an interjection of emphasis.

SA:
- A. Abisan, terpaksa begini
- B. Abisan, mau gimana lagi?  
- C. Abisan, dibilangin, nga percaya

ESAC:
- a. Because, NP were forced, like this
- b. What else can be done? (SS: because, want, how, else)
- c. NP were told, but wouldn’t believe, NP

Ada
Besides the standard locative and possessive verb, there are many speech acts involving the lexical item ada. Many are in collocation with tag question marker kan or ngga(κ) and interjections. Explanation on tag question markers and interjections are under their respective entries. Below are some SA of ada with LF and ESAC description:

<table>
<thead>
<tr>
<th>SA</th>
<th>LF</th>
<th>ESAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ada, kan?</td>
<td>[+interrogative +yes/no]</td>
<td>There are, some aren’t there?</td>
</tr>
<tr>
<td>Ada, nga, sih?</td>
<td>[+interrogative +yes/no]</td>
<td>Is there, any or not, Ø?</td>
</tr>
<tr>
<td>Ada, nga, yah?</td>
<td>[+interrogative +yes/no]</td>
<td>I wonder if, there are, some, Ø?</td>
</tr>
<tr>
<td>Ada, dong!</td>
<td>[+declarative]</td>
<td>Of course, there are some, Ø!</td>
</tr>
<tr>
<td>Ada, apa, sih?</td>
<td>[+interrogative +constituent]</td>
<td>Ø, What’s the matter, Ø?</td>
</tr>
<tr>
<td>Ada, sih</td>
<td>[+declarative]</td>
<td>Ø, Ø, Not telling, you, Ø</td>
</tr>
<tr>
<td>Ada, deh</td>
<td>[+declarative]</td>
<td>There is, some, actually, Ø</td>
</tr>
<tr>
<td>Ada, aja</td>
<td>[+declarative]</td>
<td>Ø, Ø, You are, such, a joker, Ø</td>
</tr>
</tbody>
</table>

(Me)ngada-nda
[a intransitive verb] [to fib]

Seada-adanya
[+adverb] [to make do with what there is]

Ngadain (mengadakan) [+transitive verb] [to put on an event]

Aduh
Signification: ‘auch’. It is also an interjection that gives emphasis when complaining or being upset.

LF: [+interjection +complain +upset]
SA: Aduh, kok begini sih?
ESAC: Oh no, why is it like this, Ø?

Agak

LF: [+adverb]
Syntax: agak AP
SA:
- a. Agak, spmahal, harganya
- b. Agak, kurang, ini
- c. Yang ini, agak, gede

ESAC:
- a. Rather, expensive, Ø
b. This isn’t quite enough.
c. This one is rather big.

**aja**

Elision of *saja*. Signification: only, just. In spoken Indonesian it maintains its original semantics but is common as as an interjection. A number of collocations of *(s)aja* are spoken speech acts:

<table>
<thead>
<tr>
<th>SA</th>
<th>LF</th>
<th>ESAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Disini, aja</em></td>
<td>[+declarative +suggestive]</td>
<td>Just VP₀ here,</td>
</tr>
<tr>
<td><em>Begini, aja</em></td>
<td>[+declarative +suggestive]</td>
<td>Let’s₀ just VP₁ Ø like this₀</td>
</tr>
<tr>
<td><em>Tetep aja</em></td>
<td>[+declarative +affirmative]</td>
<td>Still VP₀ allowed₀ NP self₀ be</td>
</tr>
<tr>
<td><em>Mau, aja</em> + VP₀!</td>
<td>[+reprove]</td>
<td>Ø NP₀ allowed₀ NP be</td>
</tr>
<tr>
<td><em>Mau, aja</em> + kalo* gratis</td>
<td>[+declarative +agreement] Ø, Ø, NP‘llØ have₀ some₀ ifᵣ it’s₀ freeᵣ</td>
<td></td>
</tr>
</tbody>
</table>

**alah/ya alah/ya ileh**

Derivative of Allah (Arabic for *God*). Functionally identical with *aduh*.

LF: [+interjection + complain + upset + dissapointment]

**ama**

Elision of *sama*. FI signification: same. CI signification: with.

CI sample: *Ama, siapa, jalannya_k?*
Who₀, did₀ NP₀ goᵣ withᵣ?

**amat/amet**

Signification: very.

<table>
<thead>
<tr>
<th>SA</th>
<th>LF</th>
<th>ESAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Susah, amat!</em>?</td>
<td>[+interrogative +wondering]</td>
<td>How₀ difficult, can₀ it₀ be₀ Ø!</td>
</tr>
<tr>
<td><em>Gede, amet!</em>?</td>
<td>[+interjection +impressed]</td>
<td>Howᵣ big, is₀ that₀ Ø!</td>
</tr>
<tr>
<td><em>Masa bodo, amat</em>ᵣ</td>
<td>[+declarative +indifference]</td>
<td>Iᵣ really, don’t care</td>
</tr>
</tbody>
</table>

**ambruk**

Signification: collapse.

CI sample: *Begitu, kena tabrak, langsung, ambruk, itu_k grubuk_*
Theᵣ shackᵣ collapsed, the moment, the carᵣ hit, itᵣ

**ampe**

Allomorphic variant of *sampai*. Signification: until.
In spoken Indonesian it maintains its signification meaning but also function as the phrase ‘to the point’:

*Sakin, kecewanya, ampe, pasrah_*
NP₀ were₀ soᵣ disappointed, to the pointᵣ that NP₀ give upᵣ
**anak-anak**
Besides the standard meaning ‘children’, in colloquial Indonesian it can mean ‘friends’, ‘everybody’:

*Anak-anak, lagi dimana?*
Where is everybody?

**anteng**
Signification: calm.

CI sample:  
*Andi orangnya, bener, anteng, banget j*
Andi is Ø a Ø really, calm person,

**apa/apaan**
Besides the standard WH- question word signification: what, in colloquial Indonesian it can manifest a DS value: or.

i. LF:  
[+what] [+question word: +yes or no; +constituent]
Syntax:  
a. ApaNP?
 b. apa NP?
SA:  
a. Apa NP ini tempatnya?
 b. Apa NP ini?
ESAC:  
a. I is this the place?
 b. What is this?

ii. LF:  
[+conjunction: or]
Syntax:  
NP apa NP?
SA:  
NP ini, hpnya kakak, apa_k NP hpnya, mamah_m?
ESAC:  
Is this your, or_k mum’s mobile phone?

There are a number of interesting variations of DS values as speech acts in collocation with other lexical items:

<table>
<thead>
<tr>
<th>SA</th>
<th>LF</th>
<th>ESAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apa, bagus, nyai_k?</td>
<td>[+interrogative +constituent]</td>
<td>What is so good about that?</td>
</tr>
<tr>
<td>Apaan, sih?</td>
<td>[+interrogative +reprove]</td>
<td>What is all about that?</td>
</tr>
</tbody>
</table>
| Ikan, apaan_k, yang_k, bisa, terbang_m? | [+interrogative +doubt] | What kind of fish can fly?
| Nga, tau, tuh_k, apaan_t | [+declarative +wondering] | No idea what is all about.
| Apapun, yang, terjadi_k... | [+declarative] | Whatever happens...
| Ta, apa_lah_k | [+declarative] | Ø, Ø, Ø, Never mind |
| Apa, lu_k? | [+interrogative +challenging] | What do you want?
| Apa, tuh_k? | [+interrogative +constituent] | What is that?
| Yang merah, apa, yang jingga_k? | [+interrogative +conjunction: or] | The red one, or the orange one? |
apes
Signification: \textit{bad luck}

CI sample: 
\begin{quote}
\textit{Apes, banget hari ini.}
\end{quote}
It’s been really bad today.

asal
FI signification: \textit{origin}. In CI, it is an elision of \textit{asalkan} (\textit{so/as long as}), and it has another DS value of: \textit{sembarangan} (\textit{any (senseless) thing}).

i. LF: \ [+conjunction phrase: \textit{so/as long as} \ ]

Syntax:  
a. \textit{asar [__], [__]}
b. [___], asal [___]

SA:  
a. \textit{Asalkan kita ulet}, [pasti berhasil]m
b. [Kita mau aja main], asal m [dibayar]

ESAC:  
a. \textit{As long as we work hard, we’ll surely succeed}m
b. We will Ø play, so long m we get paid

ii. LF: \ [+adverb: \textit{any(senseless)thing}]  

CI samples: 
\begin{quote}
\textit{Asoi banget!}
\end{quote}
NP+copula Ø just saying, anything!  

\begin{quote}
\textit{Jangan asal-asal, begitu dong!}
\end{quote}
Don’t, just do such nonsense, Ø!

asoi
Variation of \textit{asyik}, signification: \textit{great; fun}.

CI sample: \textit{Asoi, banget!} - This is so great!

atuh
Betawi/Sundanese origin.

LF: \ [+emphatic exclamation +concern]  

Syntax: \ [+conjunction phrase: \textit{so/as long as} \ ]

SA:  
a. \textit{Hati-hati, atuh!}
b. \textit{Jangan, gitu, atuh!}

ESAC:  
a. Careful, Ø!
b. Not, like Ø that, Ø!

babe
Signification: \textit{father}.

CI sample:  
\begin{quote}
\textit{Babe, lu, kerja dimana?}
\end{quote}
Where, does<sub>3</sub> your, dad, work?
**bacot**
Signification: mouth.

CI sample: 
*Jangan, sembarangan, bacotinya!*  
Ø,ØWatch your mouth!  
(SS: Not anything, your mouth!)

Note: In FI the suffix –iya is a third person agreement marking, possessive marker or quasi definite article / determiner. In CI it is used more freely as any personal pronoun (1st, 2nd) marker. Eventhough in this case it is not necessarily clear if –iya functions as a second person possessive marker or determiner.

**bagi**
Fl signification: for. CI signification: asking for something (identical to mina).

CI sample: 
*Bagi, dong,*!  
- Can, I have, some,Ø!  

**bagi-bagi**
Contraction of FI membagi-bagikan, signification: to share.

CI sample: 
*Bagi-bagi, rejeki,*  
Sharing, good fortune.

**baju**
Signification: clothing

CI sample: 
*Gue, perlu baju, baru nih,*  
I, need, new, clothes, Ø

**bakal(an)**
Alternative to FI akan. Signification: will [+future AUX marker].

CI sample: 
*Gue, bakal, dapetk nih kali ini,*  
I'll get it this time

*Himbauin, tuh, nasehatinya, kalo nga bakalan celaka, nanti,*  
Take heed, of the, advice, or else, there'll be an accident, Ø

**balik**
The most common FI forms are kebalikan(iya) (in contrary) and dibalik (turned around). In CI it’s an alternative to: 1. Kembali (return) 2. pulang (go home).

CI sample: 
*Balik, yo,*  
Let’s go, back

*Balik, duluy aya,*  
I’m going home, now, Ø, Ø?

*Kapan, baliky?*  
- When, are you coming back?

*Balikin, duit, nya*  
Give, the, money, back
**bandel**  
Alternative to FI nakal. Signification: *naughty.*  
CI sample:  
\[ \text{Ade}_i \text{g} \text{ue}_j \text{bandel}_k \text{banget}_l \text{waktu}_m \text{kecil}_n \]  
My brother, was really naughty when he was a child.

**banget**  
Alternative to FI sekali (very).  
CI sample:  
\[ \text{Enak}_i \text{banget}_j \text{pangsit}_k \text{nya}_l \]  
The wontons are really nice.

**bangsa**  
FI signification: *people* (of a nation). In spoken Indonesian it also has an interesting alternative usage meaning *type*, often with suffix –nya attached (*bangsanya*)  
CI sample:  
\[ \text{Yang}_i \text{bangsa}_k \text{nya}_l \text{kayak}_m \text{ini}_n \text{kita}_o \text{perlu}_p \]  
It’s this type that we need.

**bareng(an)**  
Signification: *together.* (FI: bersama-sama)  
CI sample:  
\[ \text{Jalan}_i \text{barengan}_j \text{yo}_k \]  
Let’s go together!

**baru**  
FI signification: *new.* In CI it is a contraction of the phrase baru saja: *just, recently.* In this context it often also takes the suffix –an: *barusan.* In another CI context baru is also an interjection of emphasis.  
CI samples:  
\[ \text{Barusan}_j \text{nyampe}_k \]  
Just arrived.

\[ \text{Ini}_j \text{dia}_j \text{yang}_k \text{baru}_l \text{dicari}_m \]  
This is the one that I’m looking for!

**bawel**  
Signification: *talkative; irksome complaining.* Inherited from Betawi (possibly from Dutch: babble?)  
CI samples:  
\[ \text{Bawel}_i \text{lu}_j \]  
You talk/wine too much.

**becus**  
It differs from the standard benar in so far that it only has a specific signification: *proper.*  
CI samples:  
\[ \text{Yang}_i \text{becus}_j \text{dong}_k \text{ngomong}_l \text{nya}_m \]  
Speak properly!

\[ \text{Engga}_j \text{becus}_j \text{ini}_k \]  
This isn’t right.
**begini**
A variation from the standard phrase *seperti ini* (like this). Some common everyday speech acts using *begini* include:

\[ \text{Begini, lho...} \quad \text{This, it, how, it, is, O...} \]

\[ \text{Begini, aja,... (Of course alternatively also: gimana kalo begini...)} \quad \text{How, about, this, O...} \]

**beginian**
An extended variation of *begini* with a more specific signification: *this kind* (standard Indonesian: *seperti yang ini* or *yang seperti ini*)

CI samples:
\[ \text{Yang, beginian, ngn, mau, ah} \]
\[ \text{O, don't, want, this, kind, O} \]

**beginin**
Another extended variation of *begini*, this time it has gone through predication. Signification: *do it like this* (Standard Indonesian: *lakukan seperti ini* (*beginikan*? – if this is a legitimate FI version, then it is argued here that it is due to the influence of the CI *beginin*, since the active FI *mengbeginikan* is not a lexical item)).

CI sample:
\[ \text{Dibeginin, aja...} \quad \text{Just, do, it, like, this,...} \]

**begitu**
In standard Indonesian it is a demonstrative pronoun, signification: *like that, as such*. In speech, *begitu* has a number of different DS and like the previous form *beginin*, a predicate form *begituin*:

\[
\begin{align*}
\text{i. LF:} & \quad [+ \text{noun: the instant}] \\
\text{SA:} & \quad \text{Begitu, kita, tiba, langsung, ujian}_m \\
\text{ESAC:} & \quad \text{It, rained, as soon as, we, got, there} \\
\text{ii. LF:} & \quad [+\text{adverb +augment}] \\
\text{SA:} & \quad \text{Begitu, macet, nya, sampe, berjam-jam, di, jalan}_n \\
\text{ESAC:} & \quad \text{The, traffic, was, so, bad, that, it, took, hours, on, the, road} \\
\text{iii. LF:} & \quad [+\text{transitive verb}] \\
\text{SA:} & \quad \text{Jangan, dibegituin, dong!} \\
\text{ESAC:} & \quad \text{Don't, do, it, like, that, O!} \\
\end{align*}
\]

**bego**
Signification: *stupid*.

CI sample:  
\[ \text{Bego, banget,} \quad \text{So, stupid,} \]

**bejibun**
Signification: *many*. 

---
CI sample:  
*Orang itu bejibun-jibun duitnya.*  
That man (woman) has lots of money.

**bekas**
In standard Indonesian signification: *old item.* In CI it also functions like the English the prefix *ex-* in the context of *previous/former.*

CI sample:  
*Dia, tuh bekas cowo gue.*  
He is my ex-boyfriend.

**bejen**
Signification: *famous.*

CI sample:  
*Itu orang lagi bejen.*  
That person is famous at the moment.

**belaga**
Signification: *pretending* (FI: *berpura-pura*). A typical context for *belaga* is to pretend ignorance:

*Belaga, nga, tau, aja.*  
NP + copula + just + pretending + not + to know.

Note: *Belaga pilon* is also a set-phrase DS: *pretending ignorance.*

**belagu**
Signification: *conceited attitude.*

CI sample:  
*Jangan belagu, mentang-mentang baru menang.*  
Don’t be cocky because you’ve just won.

**bengong**
Signification: *blank look/empty gaze.*

CI sample:  
*Ngelamunin, apa bengong aja?*  
What are you daydreaming about staring at the clouds?

**berantem**
Signification: *fighting/quarrelling.*

CI sample:  
*Ade-kakak itu berantem terus.*  
Those two siblings are always quarrelling.

**bercanda**
CI alternative to FI *bergurau,* signification: *to joke.*

CI sample:  
*Cuma bercanda doang, jangan marah dong.*  
I was just joking, don’t be angry.
**bete**

Started as a colloquial fad word, a contraction from English *bad temper*, but it’s meaning is more like *bad mood*. Fad words are generally not given treatment here as they come and go, but this one became very common and stood the test of time. It is worth noting since the student of Indonesian will most certainly encounter it in daily speech. Other alternatives are *suntuk* and *jenuh*.

CI sample:  
\[\text{lagi, bete, nih} \quad \text{I’m in a bad mood,}\]  

**betulan**

Even though it has a pronominal –an marking, it functions more like an adjective, usually interchangeable with standard meaning *asli*. Signification: *authentic, real*.

CI sample:  
\[\text{Ini, jem, rolex betulan, bukan?} \quad \text{Is this a real rolex watch?}\]  

**biang**

Not specifically a CI item. It originated from an old Malay expression *biang keladi*. In spoken Indonesian as a single lexical item it means the origin of something – usually troublesome; *the culprit*.

CI sample:  
\[\text{Ini, dia, biang, nya} \quad \text{This is the culprit!}\]  

**biar**

Standard Indonesian signification: *to leave alone, let be*. In CI it also has signification: *so that; in order to*. Alternative to FI: *supaya*.

CI sample:  
\[\text{Biar, cepet, selesai, lebih, baik, mulai, sekarang} \quad \text{It is better to start now, so that finish quicker!}\]  

**biarin**

Common in speech with its predicate –in suffix. Preserves the standard signification (*dibiarkan: let it be*) and as a speech act it can also transform to DS meaning: *don’t care*.

CI sample:  
\[\text{A: Nga, kasi, aja, dia?} \quad \text{B: Biarin, salah, sendiri,}! \quad \text{A: You don’t pity, NP, B: I don’t care, it’s own fault!}\]  

**bisa**

There are no semantic changes in CI for the modal verb *bisa*. It is in high usage and there some common set phrases derivative of *bisa*:

**SA**  
**ESAC**  
**Bisa, aja, ngomong!**  
**NP, can, talk!**
Bisa, aja, tuh orang!  How_{NP_1} manages_{NP_1}! [+exclamation +wonderment]
Bisa, jadi?  It_{NP_1} is_{NP_1} possible_{NP_1}?
Apa, bisa, ya?  Ø, Is_{NP_1} it_{NP_1} possible_{NP_1}?
Bisa, nga?  Can{NP_1} it{NP_1} be_{NP_1} done_{NP_1}?

A predicate form of bisa: bisain, is a speech act that means: to make happen:

Bisain, dong?!  Can, we_{NP_1} make, it, happen, Ø, ?!

blepotan
Signification: messy.

CI sample:  Makan, bisa, ampe, blepotan, gitu_{m}!  How_{NP_1} can{NP_1} eat, so_{NP_1} so_{NP_1} messy_{NP_1}?
            Ngomongnya, blepotan, lu_k  You_{NP_1} are_{NP_1} speaking, nonsense_{NP_1}

bodo
An elided form of bodoh (stupid) but also a short form of the speech act masa bodoh (don’t care).

CI sample:  Bodo, emang, gue_k pikirin_{i}  I_{NP_1} don’t_{NP_1} care, as if_{NP_1} I_{NP_1} would lose_{NP_1} sleep_{NP_1} over_{NP_1} it_{NP_1}

boke
Signification: penniless.

CI sample:  Traktirin, dong, gue_k lagi boke_{m} nih_{m}  Can_{NP_1} you_{NP_1} shout, Ø, I_{NP_1} ’m_{NP_1} broke_{NP_1}, Ø_{NP_1}

bonyok
Signification: crushed.

CI sample:  Mobil, kecelakaan, sampe_k bonyok_{k} begitu_{m}?!  That_{NP_1} car_{NP_1} got_{NP_1} really_{NP_1} crushed_{NP_1}, in_{NP_1} the_{NP_1} accident_{NP_1}!

brengsek
Brengsek is a description for any troublesome, unwanted behaviour. Anything from naughty, nuisance, belligerence, to pugnaciousness can be referred to as brengsek.

CI sample:  Dasar, brengsek, tuh orang_{i}  Ø, That_{NP_1} NP_{1} is_{NP_1} a_{NP_1} real_{NP_1} trouble maker_{NP_1},

buat
The standard Indonesian signification: to make. In CI it is often a substitute for the preposition untuk (for). In the past it was common only in speech, but now it also appears in written FI domain.

CI sample: Buat, apa? - What for?

bukan main
A standard and common Indonesian phrase. Origins very likely from the spoken language. FI proponents would argue otherwise.

DS: [+incredible] [+amazement]
SS: [+not play]
Syntax: bukan main AP-nya
E.G.: bukan main, indahnya AP, pemandangan ini
ESAC: this scenery is amazingly beautiful AP

bule
Signification: caucasian foreigner. A curious word worth mentioning as Westerners are sure to encounter it. Nowadays non-caucasian Westerners are also referred as bule.

CI sample: Orang, bule, tinggi-tinggi. Europeans are all tall

bunting
Signification: pregnant. From a FI perspective it should only be used in reference to animals but among the Betawi speech community it is used for both human and animals. FI purists insist should only be used in reference to animals. It is used more freely in spoken Indonesian: to marry.

CI sample: Bini, gue, lagi, bunting. My wife is pregnant

buntut
CI alternative for FI ekor (tail). In figure of speech buntut is used to express that a matter has not ended: ‘there is more to it’.

CI samples: Ini, pasti, ada, buntutnya. This, surely, has n’t ended here.

Buntut-buntutnya, nanti, juga, minta, uang, lagi. In the end, NP will be asking, for money again.

buru(an)
A short form of terburu-buru (to be in a hurry). It is commonly used to urge someone to hurry.

CI sample: Buruan, telat, nanti. Hurry up, we’ll be late

busyet
Exclamation of bewilderment, surprise, disapproval.

CI sample: Busyet, nga, kebayang itu!
Would never have imagined it.

**cabut**

Standard signification: *to pull out*. It is common in Jakartan CI meaning *to get going*, *to leave*. It is interchangeable with *jalan*.

CI sample: *Cabut, yo’j* . Let’s go.

**cakep**

A curious word that seems to be an allomorphy of *cakap*; in standard Indonesian more commonly known in the form *bercakap-cakap* (*to chat*). In CI it has signification *good looking*, but unlike *cantik or ganteng* it is not gender specific. It is also often used to show affection to children.

CI sample: *Aduh ini anak cakepnya m!*
You are so gorgeous (child!)

Note: Potentially adding to further confusion is that in Malaysian Malay *cakap* means *to talk*.

**canggih**

A CI word that has found itself into mainstream BI usage, but not into upper diglossic domain of FI. Signification: *advanced*, often used in reference to technology.

CI sample: *Ini mobil, Jepang baru yang ter, canggih o*
This is the latest and most, advanced, Japanese car.

*Canggih juga pikiran nya*
His/her thinking is ahead of time.

**cemplung**

One of many onomatopoeiac CI lexical items with origins from the regional languages. Signification: *to fall into water*.

CI sample: *
Hati-hati, jangan sampe kecemplung*
Careful, don’t fall into the water.

**cerewet**

Signification is identical to *bawel: talkative; irksome complaining*

CI sample: *Ibu, tiri, ku cerewet, banget*
My step, mum, is such nag

**cewe(k)**

Originally a CI item confined to the speech of Jakartan youth, it has become common in everyday Indonesian throughout Indonesia across social domains. The last consonant ‘k’ only manifests as a glottal stop [ʔ]. Signification: *a girl, girlfriend*.

CI samples: *Anaknya, Pak, Rus cewe apa cowo?*
Ø Has Mr Rus got a girl or boy?
Cewe, gue, kuliah\textsubscript{t} di UI\textsubscript{m}  
My\textsubscript{t} girlfriend\textsubscript{t}, studies\textsubscript{t} at UI\textsubscript{m}  

ciprat  
Onomatopoeic lexical item of Betawi origins. Signification: splashed.  
CI sample:  
\textit{Masa, gue, diciprat\textsubscript{t} ade, gue\textsubscript{m} ama\textsubscript{m} pistol\textsubscript{t}, air\textsubscript{p}}  
I can’t believe, my\textsubscript{m} little brother\textsubscript{t} sprayed me\textsubscript{t} with\textsubscript{t} a\textsubscript{t} water\textsubscript{p}, pistol\textsubscript{t}.  

colong  
Also of Betawi origins, an alternative to FI mencuri. Signification: to steal.  
CI samples:  
\textit{Motor\textsubscript{n}ya\textsubscript{t}, Mas\textsubscript{t}, Tri\textsubscript{t}, dicolong\textsubscript{m}, semalem\textsubscript{n}}  
\textit{O\textsubscript{t}, Tri\textsubscript{t}’s\textsubscript{j} motorbike\textsubscript{t}, was\textsubscript{t} stolen\textsubscript{n}, last night\textsubscript{n}.}  
\textit{Kasihan\textsubscript{n}, kecolongan\textsubscript{t}, dia\textsubscript{k}}  
Feel sorry, for\textsubscript{t}, him\textsubscript{k}, someone\textsubscript{t} stole\textsubscript{t} from\textsubscript{t} him\textsubscript{k}.  

congek  
Not as absolute as the FI tuli, it is more like to have very bad hearing. Often also figuratively used when someone is not listening.  
CI sample:  
\textit{Congek, ya\textsubscript{j}? Udah\textsubscript{t}, berapa\textsubscript{t}, kali\textsubscript{m}, dibilang\textsubscript{i}\textsubscript{n}?}  
Aren’t you listening\textsubscript{j}? How many\textsubscript{t} times\textsubscript{m} have\textsubscript{t} you\textsubscript{t}, been\textsubscript{t} told\textsubscript{t},?  

copot  
Can be, but not always a synonym to the FI lepas. Signification: to come off.  
CI sample:  
\textit{Tutup\textsubscript{n}ya\textsubscript{t}, copot\textsubscript{t}, terus\textsubscript{t}, nih\textsubscript{m}}  
This\textsubscript{t}, lid\textsubscript{t}, keeps\textsubscript{t}, falling\textsubscript{t}, off\textsubscript{t}, \text{Ø}\textsubscript{m}.  

cowo(k)  
The opposite-pair of cewek. Signification: male, boyfriend.  
CI sample:  
\textit{Cowo, itu\textsubscript{t}, anak\textsubscript{k}, Pasar Baru}  
That\textsubscript{t} guy\textsubscript{t}, is\textsubscript{t} from\textsubscript{t}, Pasar Baru  
\textit{Cowo\textsubscript{t}, si\textsubscript{k}, Ani\textsubscript{k}, siapa\textsubscript{m}, sih\textsubscript{h}?}  
Who\textsubscript{m} is\textsubscript{t}, Ø\textsubscript{k}, Ani\textsubscript{k}’s\textsubscript{j} boyfriend\textsubscript{k}?  

cuek  
Initially a Jakartan CI word that appeared in the 70s. It has persisted and become part of everyday spoken vocabulary. Signification varies to context: to be indifferent, not to worry, to be care free.
CI samples:

- *Cuek, banget, itu, orangnya*
  That is so happy go lucky.

- *Ayo, kita, samperin, cuek, aja*
  Let's approach, don't worry.

**Cuma**

Alternative to *hanya*. Signification: *only, just.*

CI samples:

- *Masacuma, segitu?*
  Really? Only that much?

- *Gue, Cumaada, limapuluhribu*
  I've only got fifty thousand.

**Cuman**

A CI alternative to Indonesian conjunction *tapi*. Signification: *but[+conjunction]*.

CI sample:

- *Tadinya, sih, janjian, jam, empat, cuman, karena, macet, jadi, nya, ditunda,*
  At first we were supposed to meet at four, but because of the traffic jam, it got postponed.

**Dablek**

CI alternative for *kepala keras*. Signification: *obstinate*.

CI sample:

- *Dasardablek, nga, bisa, dibilangin*
  Typical, stubborn, can't tell anything.

**Dah**

Elision of *sudah*, can be used as an interjection of resignation.

CI sample:

- *Udah, dah, jangan, dipikirin*
  Let go of it, don't think about it.

**Dandan**

This CI item is now a standard word. FI equivalents are *mengenakan pakaian untuk mencantikan diri*. Signification: to dress up.

CI sample:

- *Si, Yani pinter, dandan, dia, selalu, kelihatan, cantik*
  Yani is good, she always looks pretty.

**Dasar**

The standard signification is: *base, foundation*. In CI it also has signification: *typical*. Usually used as an interjection when complaining about the nature or character of someone or something.

CI sample:

- *Dasar, brengsek*
  Typical, nuisance.
degeg
Onomatopoeia, signification: the heart throb. FI equivalent: jantung berdebar.

CI sample:
Begitu degeg, nonton film serem, ampe ngga bisa tidur.
My heart was throbing so hard watching that spooky movie, that I couldn’t sleep.

deh
An interjection particle of agreement or suggestion, typically said at the end of a phrase.

LF: [+interjection +agreement +suggestion]
Syntax: [____] deh.
SA:
  a. Iya, deh, setuju.
  b. Begini, aja, deh!
ESAC:
  a. Alright, then, I agree.
  b. How about this, okay!

demen
A CI alternative to standard suka or FI menyukai. The DS emphasis of demen is on enjoyment. So, in the context of liking a person or a car model, then suka would be more correct. Signification: to like, to enjoy.

CI sample:
Demen, nga, ama, jajanan yang disini?
Do you like the food here?

digituin
Not a word in itself but a short form of dibegituin (look under begitu). Worth noting as a common speech act and interesting as the demonstrative itu is predicated in its form (be)gitu. An analogous comparison would be turning the demonstrative pronoun that, somehow, into a verb: “tolike-thating”.
In FI the ‘to do’ verb is needed: dilakukan begitu. Signification: to be done like that.

CI sample:
Nga usah digituin ah!
Don’t do it like that, okay!

doang
Signification: just, only. Another typical popular Betawi lexical item. Alternative to standard saja. Often works in double adjectival phrase: CI: Cuma ini doang; FI: Hanya ini saja; ‘only just this’.

CI samples:
Jangan putus asa, doang, gitu, doang.
Don’t lose heart, just, that! 
Cuma, ada, segini, doang.
There is only this much.

dong
Another Betawi origin word that has become imbedded in everyday speech. In itself it has no semantic meaning and functions as an interjection particle when asking or requesting for something. Like most interjection particles, it precedes a phrasal clause.
LF: [+interjection +asking +requesting]
Syntax: [____] dong.

CI samples: Mau_i dong_j! - ØiCan_oI_o have_o some_o Øj!
Coba_i dulu_j dong_k! - Try_o some_o first_O_k!

dongo
A CI alternative to Indonesian bodoh. Significaiton: stupid.
CI sample: Aduh_i dongo_jbanget_k gue_l! - Øi How ok stupid_l am_oI_o!

doyan
A synonym to suka but specifically in reference to taste, mostly heard in the context of food.
Signification: to like (taste preference).
CI sample: Gue_i kurang_j doyan_m ama_o duren_m
I_l don’t_l like_l duriann_oØj so_l much_j

duit
CI alternative to uang. Signification: money.
CI samples: Berapa_i duit_j ini_k? - How muchØj is_o it_k?
Bisa_i pinjem_i duit_k nga_l?
Can_l borrow_l some_o money_k Øj?
Dasar_i mata_k duitan_m! - Typical_k money_l hungry_l!

dulu(an)
The standard dulu is a past tense marker. In CI it is also a short form of lebih dahulu: to perform something first in ordinal sequence. It is a fairly common speech act. Often the use of dulu in the context of taking leave does not easily translate into English, as the use of dulu is a speech habit that does not necessarily translate as ‘going first’. Duluan as an isolated phrase signifies ‘to be going now’. Examples a and b below are such cases, while in cand dit suggests both ‘to be going now, first, before someone else’, and in e it simply means ‘to do something first’.

CI samples:
a. Kita_i mau_pulang_k dulun_i
Wej_Ohave_o to_o go_o home_k now_o Øi
b. Duluan_i ya_l . I’mgoing now_Oj

c. Kita_i duluan_l ya_k nanti_k ketemu_m disana_n
We’ll_l go_o first_k Øk, meet_m you_o there_n_later_l

d. Kamu_i duluan_l aja_k . Why_o don’t_o you_l, go_o first_l

e. Yang_l ini_k dulun_k . This_l one_l first_k
elu/lu
Jakartan colloquial of kamu (you). It became standard for the New Order generation and has continued to this day. It is on the extrema of informal social linguistics, only appropriate if used with close associations. One reproaching expression that involves lu is:

Lu lagi, lu lagi!

The closest literal translation is: (not) you again! But here the use of lagi (again) does not necessarily imply the same mistake or wrongdoing has been repeated, or that the offender has consecutively done another blunder. It is just used as a speech habit of the set expression.

emang
An elided for of memang, literally translated as indeed. In a phrase like memang, benar, begitu, it keeps its original semantic property as indeed: that is correct, indeed. In some other spoken usage, often also taking on the suffix -nya, it is more difficult to translate emang. It functions more like an interjection for interrogatives verifying a fact. Similar to English use of really?!

LF: [+interjection +verifying]
Synt: Emang(nya) [___]?
     Or
     [___] Emang(nya)?

CI samples: Emang, lu, tau,?

Do you really know?

Emang, kamu, udah, pernah, ke, mana,?
Have you really been there?

Emangnya, kenapa,?
Why? (What’s wrong with that?)

Bagus, emangnya,?
Is it nice, really?

Emang, bagus,?
It is nice, really

(In this last example emang bagus, while SS signification of emang is indeed, at the linguistic DS value it functions more like the anaphoric referential pronoun ‘it’. In the above example, Emang is co-indexed as a null element Ø. From a semantic point of view it has a language function [+agreeing +confirming] and can be co-indexed as It is nice, (indeed))

embat
Colloquially use is similar to sikat. Literal SS of sikat is brush but DS semantics is something like ‘go for it’. Most common predicates associated with the speech act are take and hit.

CI sample: Embat, aja
Just, go VP fot it (VP: take, hit)
**enak**
Standard DS value: nice/delicious. Associated with taste and sensation. Some common expressions involve the use of *enak.* All of these examples below have the connotation of objection (to unfairness, etc), cynicism or sarcasm:

\[
\text{Enak, aja!} \quad - \quad \text{No way!}
\]

\[
\text{Enak, kamu, ya_k?}! \\
\text{You’re enjoying, yourself, aren’t you?} \quad \text{(Note that } ya_k,?! \text{ functions as a tag question)}
\]

\[
\text{Enak, banget!} \quad - \quad \text{NP_0 is_0having is good_0!}
\]

**engga(k)**
CI variant to *tidak.* *Engga(k)* is the most common negator in spoken Indonesian. There is no agreed spelling of *engga(k),* as unlike with other CI lexical items, it doesn’t function at all in the written FI domain. Often in speech there is only a faint trace of the phoneme /ə/ and it manifests as [ŋa]. Orthographic form varies from *enggak, engga, nggak, ngga, gak* to *ga.* Other less common alternatives are *kaga(k)* and *nda(k).* Corpora from recordings by Sneddon (2006: 273) has relative frequency of *engga(k)⁻ at 95.6%, 1.6% for *kaga(k), 0.8% for *nda(k) and 2% for *tidak.*

Some common speech acts involving *engga(k)* include:

\[
\text{Mau nga mau harus, dijalankan_k} \quad (\text{mau, nga, mau, SS: want, not, want}) \\
\text{Like it or not}, \text{NP_0 must, keep going_k}
\]

\[
\text{Engga, salah_j tuh_k?} \quad (\text{SS: not, wrong, that_k}) \\
\text{Ø,Ø,Sure_0 about_0 that_0?}
\]

\[
\text{Engga, anak_p, engga, bapak_k, sama_l, aja_m} \\
\text{Ø, Father_k and,Ø, Son, are_0 just_m the_0 same_l}
\]

Another important CI usage of *engga* is that it functions as a question particle:

\[
\text{Bisa, nga_j?} \quad - \quad \text{Can, it_0 be_0 done_0?}
\]

**entar**
CI alternative to *nanti* (later). Also common is its spoken form in aphaeresis [ntar].

Some speech act samples of *entar*:

\[
\text{Entar, dulu, ya_k} \quad (\text{SS: later, first, ok_k}) \\
\text{Won’t be long_0k}
\]

\[
\text{Entar, aja_j} \quad (\text{SS: later, just_j}) \\
\text{Why_0 not_0 later,}
\]

**gampang**
CI alternative to *Flmudah.* Signification: easy.

CI samples: \[
\text{Ini, nga_j, gampang_k, lho_l}
\]
This isn’t easy (you know).

In some speech acts it is used in assuring or abating one’s worry:

\[
\text{Gampang}_k, \text{ jangan}_k \text{ kuatir}_k
\]

Easy\(_k\), not\(_k\) to\(_k\) worry\(_k\)

\[
\text{Gampang}_k, \text{ bisa}_k, \text{ diatur}_k
\]

Easy\(_k\), it\(_k\) can\(_k\), be\(_k\) arranged\(_k\)

By itself, semantically it can be interpreted as ‘no worries’:

\[
\text{Gampang}! \quad - \quad \text{No worries}!
\]

**ganti**

CI alternative for Frubah. Signification: change, replace.

CI samples:

\[
\text{Ganti}_b, \text{ batre}_n, \text{ nya}_t
\]

Change, the\(_t\) batteries\(_t\)

\[
\text{Minta}_y, \text{ ya}_m? \text{ Nanti}_t, \text{ saya}_s, \text{ ganti}_m
\]

Can\(_t\) I\(_t\) have some\(_s\)? I\(_t\)’ll\(_m\) replace\(_m\) it\(_t\) later\(_t\)

**gantian**

Extension of ganti and CI alternative to F\(l\)bergiliran. Signification: to take turns.

CI sample:

\[
\text{Gantian}, \text{ dong}! \quad - \quad \text{Take turns}!\]

**gara-gara**

CI alternative to F\(l\) oleh karena. Signification: because of.

CI sample:

\[
\text{Gara-gara}, \text{ sibotak}_s, \text{ dompet}_t, \text{ gue}_u, \text{ jadi}_r, \text{ hilang}_n
\]

Because of\(_r\), baldy\(_t\), I\(_t\) lost\(_u\) my\(_u\) wallet\(_t\), I\(_r\)’ll\(_r\) get\(_r\) it\(_t\) later\(_t\)

**gebuk**

A typical Betawi word that has become common in Jakartan CI. Signification: bash.

CI sample:

\[
\text{Maling}_n, \text{ nya}_t, \text{ digebuk}_s, \text{ orang}_r, \text{ kampung}_m
\]

The\(_r\) thief\(_t\), got\(_t\) bashed\(_s\) by\(_t\) the\(_t\) kampong\(_m\) folks\(_t\)

**gede**

CI alternative to besar. Signification: big, large. Gede is an acquisition from Betawi, Sundanese, Javanese and Balinese. All these regional languages have gede as a lexical item. It has become more common in speech according to corpora from recordings by Sneddon (2006: 282) where gede has a relative frequency of 88.6% while the BB variant besar is at 11.4%. The samples \(b\) and \(c\) below include the enhancing morphological suffix –an (FI adverb lebih) and excessive affixes ke- -an (FI adverb terlalu).

CI sample:

\[
\text{a. Gede}_t, \text{ amat}_i, \text{ rumah}_n, \text{ nya}_t
\]

How\(_t\) big\(_i\), is\(_i\) that\(_i\) house\(_n\), ?!
b. *Gedean* yang mana?
   *Which one is bigger?*

c. *Baju* ke *(gede)*
   That, dress/shirt, is too big.

**gr**[ge-er]
A Jakartan slang lexical item from the 1980s that has persisted as common use. It is an *abbreviation* of *gede rasa* (to be proud). Worth mentioning as learners of Indonesian are sure to encounter it.

CI sample:  
*Jangan, gr* j*ya!* -  
Don’t, get, big headed!

**gembel**
CI item for Fl: *orang tunagana dan tunawisma* (homeless person).

CI sample:  
*Ada* Banyak *gembel* di *daerah* kota,  
There are, lots, of, homeless, people, in, the, city, area.

**gencet**
Signification: *to squash, to oppress.*

CI sample:  
*Kasihan* tuh *gencet* terus *ama* preman,  
That, poor, business, is always, extorted, by, the, gangs.

**gepeng**
Signification: *crushed flat."

CI sample:  
*Kaleng* nya *digencet* ampe *gepeng*  
Squash, the, can, until, they are, flat.

**gue**
Signification: first person personal pronoun.

CI sample:  
*Gue, kan temen* lu,  
I, your, friend.

**gimana**
Contraction of *bagaimana* (how). A very common lexical item used in many speech acts. Below are a number of examples:

*Gimana, dong, ini? / Gimana, ya?*  
Ø, Ø, Ø, So, what, to, do?

*Gimana lgi, mestinya?*  
What else should, be, done (have, been, done)?
Gimana, pun, kita harus coba
No matter what we must try

Gimana, sih,?!
What's this,?! [+questioning +reproaching]

Gimana, jadinya,? - So what's the outcome,?

**gini**
Elision of begini (like this). There are some common expressions involving gini in compound with interjection particles:

Gini, aja,... / Gini, deh
How about this,... [+suggesting]

Gini, lho,...
It's like this,... [+explaining]

Gini, dong,...
This is how to do it,... [+corrective suggesting]

Gini-gini, aja
Not much has happened,... / just the usual...

**gitu**
Elision of begitu (like that). Similar to gini, expressions and speech act are identical with a change in emphasis from the demonstrative pronoun ini to ini. Some peculiar DS values include:

Gitu, lho,...
That's what I mean meant...

Gitu, dong,...
That's the way,...

**hajar**
Signification: to strike.

CI sample: Gara-gara gitu, kena, hajar, gue
I bore the brunt, because of it

**idih**
CI interjection particle. Signification: [+interjection +questioning +disapproval +challenging]

CI samples: Idih, gitu, aja, marah
Ø, You are upset just because of that?

Idih, kok, gitu, sih?
Ø, How come it's like that?
incer
Signification: to chase, to target.

CI sample: Itu buronan diincer jangan sampel, lepas, Stay on that, fugitive, don’t let him go.

ini
A standard lexical item (this [+demonstrative pronoun]). It is worth noting how it manifests in speech some acts:

Ini, dia. Here, it is that (Both pronouns ini and dia represent the anaphoric pronoun it, while concurrently ini also represents the existential role of the pronoun here, functioning as a contracted form of di sini)

Yang, ini, kan? - It’s this one, isn’t it?
(The tag question marker isn’t it is co-referential with the expletive pronoun)

It’s okay
Yang, ini, dong? - Can it be this one?

jadi
In standard Indonesian, jadi is the base word of the predicate menjadi: to become. At SS in Indonesian, jadi can also translate into as the English copula to be. Where an action is in question, in this context Jadi signifies that the predicate will occur after all. In spoken Indonesian, semantically jadi also functions as the conjunction particle ‘so’. Example a below is of the latter meaning and examples b and c are of the former:

a. Jadi, sekarang, gimana? - So, what now?

b. Jadi, nga, enak? - I feel bad

c. Jadi, nga, kita? - Are we still, VP?

jago
It is unclear whether jago is the base word of the nominalised form jagoan (a person who is the master of his art; the ‘tough guy’) or if it came about as a contraction of jagoan. Signification: to be good/skilled at something. In CI it is a substitute for the FI pandai in reference to someone’s abilities.

CI samples: Si, Aro jago, maen, bola; Ø, Aro is, good, Ø, at football;

Jago, juga, tuh, orang; That guy, is, pretty, good;

jalan
The standard jalan in its predicate DS value means to walk and in its nominal form means street. What is seldom taught is its CI DS value to go, often used as a substitute for pergi or berangkat:
Jalan, yo(k)  
Let’s go.

Jalan, dulu, ya(k)  
I’m going now.

**Jambak**

A CI lexical item of Jacanese/Balinese origin which in FI would be phrased as *menarik rambut*. Signification: to pull hair.

CI sample:  
Silat, apa-an ini pake jambak, segala?  
What hair-pulling marital art is this?

**Jangkan**

In standard Indonesian *jangkan* is the imperative forbidding *do not*. With the suffix –kan attachededit changes language function and becomes the compound conjunction phrase: *let alone:*

Jangkan, hapsip, lawan, polisi aja berani.  
NP dare to challenge police let alone security guards.

**Jiper**

CI alternative to standard Indonesian *takut*. Signification: afraid.

CI sample:  
Lihat, macan di hutan, jiper juga, gue.  
I got scared, alright, seeing, the tiger in the jungle.

(In the above example, *juga* is co-indexed with *alright* as they both function as interjections performing the same exclamatory role eventhough they have different SS significations)

**Jiplekan**

The standard signification of the active FI form *menjiplak* is ‘to trace’. The nominalised form *jiplekan* is mainly a CI item. Signification: fake.

CI sample:  
Yang ini jiplekan.  
This one is fake.

**Joget**

CI alternative for Flberdansa. Signification: to dance.

CI sample:  
Begitu Pallapa maen, semua padajangsung, joget.  
As soon as Pallapa played, everyone got dancing.

(Again, despite different SS signification, *langsung* is co-indexed with *got* as they performing the same semanticDS significations)

**Jorok**

Signification: gross.

CI sample:  
Cuci tangan dulu, jangan jorok!  
Go and wash your hands, don’t be gross.
**juga**
Standard Indonesian signification is *also*. In spoken Indonesian also function as an interjection of emphasis:

*Nekad, juga, tuh, orang—That guy, really, got guts*

**kabur**
CI alternative item for standard *lari* (*to run*) but also has an additional semantic value *menghilang* (*to disappear*):

CI sample:  
*Waktu polisi dateng, udah kabur maling nya.*  
*When, the police came, the thief already escaped*

**kaga(k)**
A typical Betawi lexical item, alternative to *tidak*. Signification: *no* [negator]). Some common speech acts involving *kaga(k)*:

*Kaga, mau—Don’t, want to*  
*Kaga, karuan—Chaotic, not right*  
*Kaga, kepikir—Did not, think of it*

**kalem**
CI alternative to the standard *tenang*. Signification: calm.

CI sample:  
*Si, Abdi orang nya, emang kalem.*  
*Abdi is really calm person*

**kali**
Not to be confused with standard signification *creek*. In CI it is a contraction of *barangkali: maybe*. Some common speech acts involving *kali* pertains to the notion possibility:

*Iya, kali—yes, maybe / Possible*  
*Kurang, tau, kali, aja—Not too sure, could be*

**kalo**
The allomorphic spoken form of FI *kalau*. Signification: *if* [+conditional tense marker].

CI samples:  
*Kalo, gitu—* If, that’s the case

**kan**
A contraction of *bukan*, it is a very common tag question particle that is hardly taught to learners of Indonesian.

LF:  
*+particle +tag question*
Syntax:  [___] kan!?
Kan [___]!?

CI samples:

a.  Bener, kan!?
  True,isn’t it,?!

b.  Bagus, kan!?
  It’s good,isn’t it,?!

c.  Suka, kan!?
  You liked,it didn’t you,?

d.  Tapi, kan, lucu,...
  But, it’s funny,....

kangen
Cl alternative of standard rindu from Javanese and Sundanese origin. Signification: to miss.

CI sample:
Kangen, banget nih ama, anak
I’m really missing my child.

kapok
A CI lexical item of Betawi/Javanese/Sundanese origin. Signification: to have enough,to have learnt one’s lesson.

CI sample:
Nga, mau lagi ah, udah kapok gue,
Not having anymore of that, I’ve had enough.

karuan
Signification: proper, coherent. Used mainly in collocation with engga as a negating reference.

CI sample:
Engga, karuan, ngomong, nyai
He is talking jibberish.

kasih
A CI alternative to FI memberi. Signification: to give. A common phrase of kasih in collocation is kasih tau (to let know):

Kasih, tau ya – let me know ok

kata
In standard Indonesian signification is word. In spoken Indonesian it is more commonly used in a predicate as an alternative to bilang: to say. With the –nya suffix it forms the expression ‘it is said’.

CI samples:
Iya, deh, apa lu, kata – Alright, whatever you say
Gue, kata, juga, apa! – What did I say?
Apa, kata, nya,? – What did you say,?
Katanya, sih, – It is said, that...

kawin
CI alternative to FI menikah: to marry.

CI sample:
Kapan, kawin, nya,?
When, are you getting married,?
**kaya**
CI alternative to standardseperti:like; similar.

CI sample: Kaya, anak kecil aja. Just like; a0 kid.k
Kaya, apa sih orang,nyaa_m? - What is0NPm like,?

**keburu**
The standard expression terburu-buru means to be in a hurry. The CI form is keburu-buru and its short form keburu has different semantic values. The first meaning is: to make it in time (alternative to FIsempat), and the second meaning: it’s too late, it’s already done (alternative to FItelanjur)

CI samples: Keburu, nga? Udah, telat, nih_m. Will0NPm make it on time, Ø? It’s already, late, Ø_m
Udah, keburu, masukin, uang_l. Too late, already, put, money, in_k

**kek**
Signification: Interjection particle exclaiming indifference:

Mau, ke, Jerman, kek, Jepang, kek, nga, pengaruh_o. Ø, Whether, to, Germany, or, Japan, it’s doesn’t matter,!

**kelar**
Signification: finish.

CI sample: Udah, kelar, belum_k. Ø, Is0 it0 finished, Ø_k

**kelelep**
CI variation to standard tenggelam (to drown) but only for the context of animated objects. So to describe the sinking of a ship, tenggelam will still be used.

CI sample: Hati-hati, jangan, kelelep_k. Careful, don’t drown_k

**kelewatan**
Signification: to have gone too far. It is used both literally and figuratively (as an alternative toketerlaluan).

CI sample: Bener-bener, kelewatan, tuh, orang_l. Ø, He, ’s0 really gone too far.

**kelilipan**
A unique CI lexical item. Signification: to have something stuck in the eye.

CI sample: Mata, gue, kelilipan, nih_l. Something is0 stuck, in, my, eye,
**kenceng**
CI item of Javanese origin. It has various significations in context: **tight**, **fast**, **loud**.

CI samples:  
*Kencengnya;i ngebut;i tuh; mobil;am*!  
How;i fast;i that;i car;in driving;i!?

*Ikati; pinggang;nya;i kekenceng;man;i nih;am*  
This;i belt;i is;i too;i tight;i Ø

*Jangan; kenceng-kenceng;main;i lagunya;am!*  
Don’t;i play;i the;in song;i so;i loudly;i!

**kepengen**
Inherited from Betawi Malay, an allomorphical expansion of *ingin*. Even though signification of *ingin* is *to wish*, CI signification of *kepengen* is *to want*.

CI sample:  
*Kepengen;i juga;i sih;i ke; bioskop;am*  
I when;in want;in to;go;in to;the;in cinema;in

(Note that SS signification of * juga* is actually a null element Ø, but its DS value correlates with the auxiliary *do*, as an interjectionary exclamation emphasising ‘willingness’)

**kepleset**
Signification: **to slip**.

CI sample:  
*Gara-gara; nginjek;i kupas;i pisangampe;i kepleset;am*  
I slipped; in because; of; stepping; on; banana; peel; Ø

**keren**
A Jakartan item that has become standard in spoken Indonesian. Signification: **cool** (appearance).

CI sample:  
*Itu; mobil;i keren;banget;i*  
That;i car; really;i cool;i

**kerjain**
This expression item, with *kerja* as a base word, has a unique DS value in CI. In FI it would take the phrase *mempermainkanseorang;to do; a prank*.

CI samples:  
*Kita; kerjain;i yo;i*  
Let;i's;i do;i a; prank;in on;me

*Masa; gue; dikerjain;i ama;i si;i Adi*  
Would; you; believe; Ø; Ø; Adi; did; a; prank; in me;i

**ketara**
Alternative to the standard *jelas*. Signification: **clear**.

CI sample:  
*Nga; ketara;i ini;i maksudnya;i apa;am*  
It's;i not;i clear; what;i is;i meant;i by;i this;
ketawan
An allomorph of ketahuan. CI signification: to be found out.

CI sample: Ketawan, ya, ternyata! - You got busted, after all.

ketemu
The more common CI alternative to FI bertemu. Signification: to meet.

CI samples: Jam, berapa, ketemu, nya?; What time, are we meeting?

Ketemu, dimana, kita? - Where are we meeting?

ketimbang
The base word timbang denotes lexical items that pertain to measuring weight. In its form ketimbang, it is a CI alternative to the standard dari pada. Signification: compared to, rather than.

CI sample: Ketimbang, beli, yang, ini, sekarang, kita, lihat-lihat, aja, dulu. Let's just have a look around, first, rather than buying this one now.

kirain
The base word kira is a standard lexical item. In its CI inflected form kirain, it is a common alternative to FI menduga. Signification: to assume.

CI sample: Kirain, lu, tau. I thought, you knew.

kocak
Signification: funny.

CI sample: Badut, nja, kocak, banget. That clown is really funny.

kok
A common exclamation particle [+questioning +interjection +surprise].

CI samples: Kok, baru, tau?; How come, you only just found out?

Kok, mau, sih?; Why would NP want to do O?

kuping
CI alternative to FI telinga. Signification: ear.

CI sample: Kuping, gue, conge, nih. My ears, are deaf.
**kudu**
CI item from Betawi/Javanese/Sundanese. Alternative to *harus/mesti*. Signification: must/should.

CI sample: $Kudu, dibenerin\_ini$  
... This needs to be fixed.

**lagi**
The standard signification is *again*. In CI it has a different function as an alternative to FI continuous marker *sedang*.

**LF:** [+auxiliary +progressive tense]  
**Syntax:** [NP lagi VP]

CI samples:  
$Lagi\_apa\_an$? – What are you doing?  
*Sama, lagi, ada, di, Ancol* – Everyone, it’s in Ancol

There are some speeches act involving *lagi* where it is used as an interjection particle functioning as an affirming exclamation:

$Iya, lagi\_!$ – yes of course!  
*Engga, lagi\_!* – It’s not like that!

**lagian**
A contraction of *lagipula* [+conjunction +nonetheless +despite +besides] that is used as an interjection [+remark +pointing out] rather than a conjunction. It is used to point out an ‘lesser’ or ‘undesireable’ avoidable outcome despite some other facts that should possibly nullify it.

CI sample:

*Lagian, udah, dibilangin, tetep, ngotot, jadinya, celaka, sekarang*  
Besides, NP has been told, but was still stubborn, now it’s gone wrong.

**lah**
*Lah* is another interjection particle. It is used as an exclamation [+agreement] or [+resignation].

CI samples:  
*Udah, lah, biarin, aja* – Don’t worry about it, let it go.  
*Terserah, lah* – Up to you.  
*Iya, lah* – Of course.

**lambat**
CI alternative to the standard *pelan*. Signification: slow.

CI sample:  
*Lambat, nya*,... - How, slow...

**ladenin**
CI alternative to FI *melayani* (to serve) or to FI *menghimbau* (to take heed, to entertain somebody’s idiosyncracy).
CI samples:  

\text{Udah, diladenin, belum?}  

\text{Nga, usah, diladenin, yang itu} \text{m}

\text{Don't, Ø, take heed, of that, one} \text{m}

\text{lekas}

CI alternative to FI \text{cepat} but only used as a command. Signification: \text{quickly, hurry}.

CI sample:  

\text{Lekas, udah, ditungguin, tuh!}

\text{Hurry, they, already, waiting, Ø} \text{m}

\text{leker}

Dutch origin, alternative to \text{sedap (delicious)}.

CI sample:  

\text{Leker, banget, ini, kueh, -}

\text{This, cake, really, delicious} \text{i}

\text{loh}

An interjection particle that expresses [+surprise] and/or [+confusion]

CI sample:  

\text{Loh, kenapa, ini,?}  

\text{Ø, What happened, here} \text{m}

\text{lolos}

Signification: \text{escape}.

CI sample:  

\text{maling, nya, lolos, -}

\text{the, thief, got, away} \text{k}

\text{loncat}

CI alternative to the standard \text{lompat}. Signification: \text{jump}.

CI sample:  

\text{Iku, kanggur, pada, loncat, kemana, mana,}

\text{Those, kangaroo, are, jumping, everywhere} \text{m}

\text{macem}

An allomorphical variation of \text{macam (type, kind)}. There is one common expression involving the repeating of \text{macem} \rightarrow \text{macem-macom}. Signification: \text{wasting time, playing around, causing trouble}.

CI sample:  

\text{Jangan, macem-macom, -}

\text{Don't, mack, around} \text{k}

\text{mah}

An interjection particle used in statements asserting a fact:

\text{Ini, mah, kurang, -}

\text{This, Ø, isn't, enough} \text{k}!

\text{makanya}

A short form of \text{maka oleh karena itu}. Signification: \text{therefore, that is why}. 
makasih
A commonly used short form of terima kasih. Signification: thank you.

malah(an)
CI alternative of FI sebaliknya. Signification: instead.

males
An allomorphy of malas (lazy) and in spoken usage it is also used to mean ‘don’t feel like it’.

maling
CI alternative for FI pencuri. Signification: thief.

mampet
CI alternative to the standard tersumbat. Signification: clogged.

mampir
Of Betawi and Javanese origin; alternative to singgah (drop by).

mana
A standard word with signification where. There are several expressions and speech acts in collocation with mana where it functions as a negator:

Manabisa,! – No way, ! That won’t work
Mana, mau,! – NP will, want, to
Mana, ada,! – There is, no, such thing
Mana, tau,? – NP would, n’t, know
Mana, inget, – NP would, n’t, remember
Mana, tahan,! – Irresistible!
**mantep**
Signification: good, great, solid. Alternative to hebat.

CI sample:  
Mantep, juga, ini, mobil,–This, car, quite, good,

**maruk**
Betawi alternative to rakus (greedy)

CI sample:  
Jangan, maruk,–Don’t, be, greedy,

**masa**
A short form of Masya Allah (oh my God!—surprise/astoundment interjection), not to be confused with the standard signification time/period/era. In CI it translates as an exclamation really?! And there are also various speech acts in collocation with masa:

Masa, sih,?–Really,?  
Masa, begitu, aja, marah,?–You’re, really, angry, just, because, of, that,?  
Masa, apa, tau,?–You’re, really, don’t, know,?  
Masa, bodo,–I, don’t, care,

**mau**
Standard signification is to want. In CI it is also an alternative to akan. Signification will [+tense +future]. Mau differs to bakal(an) – also an alternative to akan – in that it functions like a future progressive tense while bakal(an) indicates that a future event will be taking place ([+eventive]).

CI samples:  
Mau, jadi, apa, kalau, begini, terus,?  
What, will, become, if, you, keep, going, like, this,?

Abis, dari, sini, kita, mau, ke, mall,  
After, this, we, ‘ll, go, to, the, mall,

**melek**
Signification: to have the eyes open / to be awake.

CI sample:  
Eh, udah, melek, enak, nga, tidurnya,?  
Hey, you’re, awake, did, you, sleep, well,?

(Note that nga, here functions as a tag question and is therefore coindexed with the auxiliary did, that generates the question phrase)

**meleng**
Signification: not looking / not paying attention.

CI sample:  
Jangan, meleng, kalo, nyopir,  
Stay, focused, when, you’re, driving,

(Rather than translating Jangan, meleng, at its SS signification ‘do not not paying attention’, it is semantically translated as its DS value equivalent Stay focused,)
**melulu**
CI alternative to the standard selalu (always). It differs in syntactical position to selalu. It follows the verb or adjective [VP/AP melulu] instead of [selalu VP/AP].

CI sample:  
Bisanya, ngeluh\textsubscript{j} melulu\textsubscript{k}  
All\textsubscript{l} NP\textsubscript{o} can, do\textsubscript{o} is\textsubscript{o} complain\textsubscript{j}  

(All\textsubscript{k} in this context is a contraction of ‘all of the time’)

**mending**
Signification: as if.

CI sample:  
Mending, bagus\textsubscript{j}, lumayan\textsubscript{k} aja\textsubscript{l} engga\textsubscript{m}  
As if, it’s\textsubscript{o} good\textsubscript{i}, it’s\textsubscript{o} not\textsubscript{m} even\textsubscript{e} average\textsubscript{k}

**mendingan**
CI alternative to the standard collocation lebih baik (it is better that…).

CI sample:  
Ketimbang mahal-mahal beli, mendingan nyewa aja  
Rather than buying so expensively, it’ll be better if we just rent

**mentang-mentang**
A typical CI item that functions as a conjunction phrase. Signification: only because.

CI sample:  
Sok, tau\textsubscript{j}, mentang-mentang\textsubscript{k}, lulusan\textsubscript{l}, luar\textsubscript{m} negri\textsubscript{n}  
Only because\textsubscript{o}, NP\textsubscript{p} graduated\textsubscript{q}, overseas\textsubscript{r}, NP\textsubscript{s} thinks NP\textsubscript{t} knows it all\textsubscript{u}

**mentok**
Signification: dead end.

CI sample:  
Nga, bisa\textsubscript{j}, jalan\textsubscript{k}, lagi\textsubscript{l}, udah\textsubscript{m} mentok\textsubscript{n}, nih\textsubscript{o}  
Can’t\textsubscript{p} go\textsubscript{q} any further\textsubscript{r}, \textsubscript{s} it’s\textsubscript{t} a\textsubscript{u} dead end\textsubscript{v}, Ø

**molor**
Alternative to tidur (sleep).

CI sample:  
Molor\textsubscript{j}, aja\textsubscript{k}, bangun\textsubscript{l}\textsubscript{!}  
All\textsubscript{n} you\textsubscript{o} do\textsubscript{o} is\textsubscript{o} sleep\textsubscript{p}, wake up\textsubscript{p}\textsubscript{!}

**muat**
CI item with signification to fit or adverbial phrase large enough.

CI sample:  
Muat, nga\textsubscript{j}, mobil\textsubscript{k}, nya\textsubscript{l}\textsubscript{?}  
Is\textsubscript{m} the\textsubscript{c} car\textsubscript{r} big\textsubscript{e} enough\textsubscript{g}?
**mudah-mudahan**
Standard signification of the base word *mudah* is *easy*. In its expanded form *mudah-mudahan* it is an alternative to *semoga* (*hopefully*).

CI sample:  
*Mudah-mudahan, nga, ada, hambatan*  
*Hopefully, there are no obstacles*.

**mumpung**
A CI conjunction phrase with signification: *since already, while at it*.

CI sample:  
*Mumpung, udah, disini, kita, mampir, di, warung, yo*  
*Since, we are already, here, let’s drop by the warung*.

**nah**
LF: interjection particle. Signification: [+confirming] [+realisation].

CI sample:  
*Nah, kan, dibilangin, nga, percaya*  
*I Ø, Ø, told you, you would, not believe it*.

**nebeng**
Signification: *getting a ride*.

CI sample:  
*Kapan, jalanya, lu? Boleh, nebeng, nga?*  
*When, are you going? Can, I get a ride?*

**nekad**
Signification: *blind and reckless bravery*.

CI sample:  
*Perhitungan, dong, jangan, nekad, gitu*  
*Consider, carefully, don’t, just, blindly brave, it*.

**ngaco**
Allomorphy of *kacau* (*chaos*). Signification: *not coherent; reckless*.

CI sample:  
*Tu, orang, ngomong, nya, udah, ngaco*  
*That, man, talk, is, not coherent, anymore*.

*Jangan, ngaco, gitu, dong*  
*Don’t, be so, reckless*.

**ngambek**
Signification: *to be upset, sulking*.

CI sample:  
*Jangan, ngambek, dong, kan, cuma, bercanda*  
*Don’t, get upset, just, joking*.
**ngapain**
Inflected predicated form of the question word *apa* (*what*). Signification: It is the common CI alternative to the question phrases *sedang apa?* (*what are you doing?*) and it can also be the question phrase *untuk apa?* (*what for?).

**CI sample:**
Lagi, ngapain in? – What are you doing?
Ngapain in digituin? – What are you doing that for?

**ngapel**
Signification: visiting girlfriend.

**CI sample:**
Malem minggu, ya ngapel lah
It’s Saturday night, of course I’m visiting my girlfriend.

**ngasil → asal**

**ngawur**
Of Javanese origin, it is an alternative to *sembarangan* (*any senseless thing; nonsense*)

**CI sample:**
Ngawur, aja, yang bener, tong
Nonsense, Ø, do properly

**ngebub**
Signification: speeding.

**CI sample:**
Nga, usah, ngebut juga, nyampe
No need to speed, we’ll get there all the same.

**ngibul**
CI alternative to FI *membohong* (*to lie*).

**CI sample:**
Ah, dasar, tukang, ngibul
Typical liar

**ngomong**
CI alternative to FI *berbicara* (*to talk*).

**CI sample:**
Ngomong, apa, sih?
What NP talking about?

Nga, mau, ngomongin itu, sekarang
NP don’t want to talk about it now

**ngomong-ngomong**
An interesting extension of *ngomong* with a change of DS value: *by the way*.

**CI sample:**
Ngomong-ngomong, apa kabar? – How are you by the way?
ngobrol
CI alternative to Flbercakap-cakap. Signification: chatting.
CI sample:  Mau, ngobrolin$_i$ apa$_o$? — What$_k$ do$_o$ you$_o$ want$_o$, to$_o$ chat$_i$, about$_i$?

ngotot
CI alternative to Flbersikeras. Signification: to insist. Ngotot has extra semantic value than bersikeras in that it is an emotional insistence.
CI sample:  Ya$_i$ udah$_i$ nga$_o$ usah$_i$ ngotot$_m$  
Alright$_i$, no$_i$ need$_o$ to get worked up$_m$.

ngumpet
CI alternative to Flbersembunyi. Signification: hiding.
CI sample:  Ngumpet$_i$, dimana$_i$, sih$_i$ tuh$_i$ anak$_m$?  
Where$_i$, is$_o$$_i$ that kid$_m$ hiding$_i$?

nih
An allomorphy of ini that like its antecedent still functions as a demonstrative pronoun and has also evolved into an interjection particle with LF [+emphasis]. Sample a below is an example of nih as a demonstrative pronoun and the examples b, c, and d are examples of nih as interjection particle.

a. Nih$_i$, ambil$_i$  — Here$_i$, take$_i$ it$_i$

b. Yang$_i$, ini$_i$, nih$_k$  — It$_i$’s$_o$ this$_i$ one$_i$, Ø$_k$

c. Mau$_i$, nga$_o$, nih$_k$?  — Do$_o$ you$_o$ want$_o$, to$_o$ Ø$_o$ Ø$_k$?

d. Apa$_i$, nih$_k$?  — What$_i$, part$Ø$_$k$?

nongkrong
Signification: hanging out.
CI sample:  Kita$_i$, nongkrong$_i$, di$_k$ tempatnya$_m$ Rusli$_i$ yo$_m$  
Let$_o$’s$_i$ hangout$_i$ at$_k$ Rusli$’s$$_m$ place$_i$.

nongol
CI sample:  Kok$_i$, tumben$_i$, tiba-tiba$_k$ nongol$_i$  
How come$_i$, NP$_o$ suddenly$_i$ turned up$_i$?

numpang
Signification: getting a ride. In collocation with other words, this CI item has semantic values that is asking for a favour. So, numpang tanyameans asking for direction, numpang duduk means asking permission to sit down for a short while.

CI samples:  Boleh$_i$, numpang$_i$, ke$_k$ Bogor$_i$ nga$_o$?  
Can$_i$, I$_o$ get a ride$_i$, to$_o$ Bogor$Ø$$_i$?

Permisi, Pak$_p$, numpang$_k$, tanya$_i$...  
Excuse, me$_o$ Sir$_i$, Ø$_i$ may$_o$ I$_o$ ask$_i$ you$_o$...
**nyaris**
CI alternative to the standard hampir. Signification: almost.

CI sample: 
\[
\text{Gara-gara, ngebut, jadi nyaris kecelakaan, tadi.}
\]
Because, NP was speeding before, NP almost had an accident

**nyasar**
CI alternative to FI phrase kehilangan jalan. Signification: to have lost one’s way.

CI sample: 
\[
\text{Wah, nyasar, nih, adakah dimana, sih kita?}
\]
Where are lost, we are where are we?

**nyata**
CI alternative for the standard jelas. Signification: clear.

CI sample: 
\[
\text{Nga, nyata, ini, petunjuk, nya.}
\]
This instruction is, very clear

**nyelonong**
Signification: to sneak.

CI sample: 
\[
\text{Kapan, nyelonong, masuk, k?}
\]
When did you sneak in?

**nyeletuk**
Signification: to interrupt conversation.

CI sample: 
\[
\text{Nyeletuk, aja, }
\]
No one asked you to butt in

**ogah**
A typical Betawi item, alternative to the standard phrase tidak mau. Signification: don’t want.

CI sample: 
\[
\text{Ogah, ah, ngapain, ke, sana?}
\]
Don’t want to, why should I go there for?

**omel**
CI alternative to FI memarahi. Signification: to scold.

CI samples: 
\[
\text{Bisa, nya, ngomel, melulu}
\]
All that he/she does is get angry

\[
\text{Diomelin, lagi, aku, gara-gara, telat}
\]
I got in trouble, again, because, I was late
orang
There is a curious shift in language function from the standard signification person. In CI it is also used as an interjeciton particle; LF: [+complaint].

CI sample: Orang, yang, sepele k aja; nga m, ngerti, apa lagi, yang, ruwet, Ø, can’t, even, understand, the, simple, stuff, let alone, the, hard, ones.

pada
In FI, pada is the preposition used in reference to time and period. In CI, it is a plural particle marker:

Pada, ke mana, semuanya,? - Where, did, everyone, go, to,?

Kursi, nya, pada, penuh, ini, duduk, di, mana, ya,? 
All, the, chairs, are, full, Ø, where, can, we, sit, Ø,?

padahal
This is a CI preposition non existent in FI. Signification: but actually, where as a matter of fact.

CI sample: Ngaku, nya, sih, udah, kuliah, ini, padahal, masih, SMA, 
NP, says, that, NP, goes, to, university, but actually, NP, is, still, in, highschool.

pake…segala
Pake is an allormoph of the standard pakai(to use). In collocation with segala it forms a seperable questionphrasethat is an alternative to question marker kenapa. Signification: why?, what for?.

CI sample: Pake, ngomong, ke dia, segala,?!
Why, did, NP, talk, to, NP, for,?

paling
In standard FI paling is the superlative marker. In CI it is also an adverbial phrase marker. Signification: most likely.

CI samples: Paling, iya, lah, 
Most likely, so, Ø, Ø,

Paling-paling, ketemu, di, moh, lagi, 
Will, probably, meet, at, the, mall, again.

pas
A common CI speech act, alternative to FI seketika. Signification: the moment.

CI sample: Pas, lu, telpon, si, Ana, dateng, 
Ø, Ana, came, when, you, rang.

payah
Signification: useless, hopeless.

CI samples: Payah, ini, keadaan, nya, this, situation, is, hopeless.
Payah, lu_3 gitu aja, nga_m bisa_n
you, are asshole, can’t even do that.

**pinter**
CI alternative to FI panda that has become standard. Signification: *smart, good at.*

CI samples: 
- Kamu tuh memang yang, paling_m pinter_m di, kelas_p
  You, are indeed the smartest in class
- Kakak, gue, bener-bener, pinter, ngelukis_m
  My sister, is really good at painting

**pokoknya**
A common spoken CI pronominal phrase marker. Signification: *the main thing is…, no matter what….*

CI samples: 
- Pokoknya, saya, ngeka mau tau_m
  I don’t care, no matter what
- Pokoknya, harus, bisa_k
  No matter what, it has to be done_k

**puyeng**
CI alternative to the standard pusing. Signification: *headache, dizzy.*

CI sample: 
- Pala, ku, puyeng_k nih_l
  I’m feeling dizzy

**rada**
CI alternative to the standard agak. Signification: *rather.*

CI sample: 
- Yang, ini_rada-rada_susah_l
  This one is rather difficult

**reseh**
An allomorphy of risih. Signification: *a hassle.*

CI sample: 
- Males, ah, reseh_k
  Don’t feel like it, it’s a hassle

**ribet**
Signification: *complicated.* This is a recent CI item that has become common replacing the older ruwet.

CI sample: 
- Yang, ini, ribet_k
  This one is complicated

**sama**
In standard Indonesian it means *the same.* In CI it is an alternative to Fldengan (with) and oleh(by).
CI samples:  
*Sama* siapa jalaninya? – Who are going?

*Tadi* kita dibantu sama ibumu – Mum helped us earlier.

**sebel**
CI alternative to FI *kesal*. Signification: irritated.

CI sample:  
*Aduh* ngeselin ini! – This is annoying!

**segen**
CI item of Javanese origin that is an alternative to the FI phrase *tidak berminat* (*don’t feel like*) and also alternative to FI * hormat* (*respect*).

CI samples:  
*Lagi* segen ngapa-ngapain ini
I don’t feel like doing anything.

Megawati, segen ama Gusdur
Megawati respects Gusdur

**sepele**
Signification: simple matter.

CI sample:  
*Jangan* khawatir, sepele itu
Don’t worry, that’s a simple matter.

**sewot**
Signification: bad mood, narky.

CI sample:  
*Lagi* sewot dia
NP is in a bad mood.

**sih**
*Si* is an interjection particle that puts emphasis a question or statement. In samples *a*, *b*, *c* and *d*, *sih* is co-indexed with the English lexical item that conveys the same DS emphasis. Samples *e* and *f* are examples where *sih* functions as speech acts in collocation with other lexical items.

*a*. Iya* sih*
Yes, actually

*b*. Mau juga* sih*
NP want, some

*c*. Dimana* sih*
Where, actually is it?

*d*. Ngeliatin apa* sih*
What is it that you’re looking at?

*e*. Kamu* sih*
It’s your fault!

*f*. Masa* sih*
Really?

**sok**
Signification: self-complacent.

CI samples:  
*Sok* lu... (you are) cocky...
Sok, tau! – Ø, You think you know it all!

suka
Standard signification: to like[+mdal auxiliary]. In CI it also has DS value that indicate habitual action:

Saya, kalo, pergi, kuliah, suka, naek, bis.
I, Ø, take, the, Ø, to, uni, sometimes.

Dulu, kita, suka, ketemu, setiap, minggu.
We, used, to, meet, every, week.

susah
CI alternative FI sulit and sukar. Signification: difficult. The predicate inflected form is also common in speech.

CI samples: Susah, amat, sih, ini? – Why is this so difficult, Ø?
Nyusahin, diri, – Making it difficult, for oneself.

Nyusahin, orang, – Making it difficult, for others.

tanggung
Signification: to might as well finish/completesomething.

CI sample: Tanggung, ini, – Might as well finish it.

tau
This item is an allomorphy of the standard tahu (to know). There are some common speech acts involving tau.

a. Tau, nga,?! – Do you know?
b. Tau, nga, sih,? – Do you know it or not?
c. Tau,?! – I don’t know?!
d. Tau, ah,! – I don’t care!

telat
CI contraction of the standard terlambat (to be late).

CI sample: Maaf, telat, – Sorry, I’m late.

terus
Standard signification is straight ahead. In CI it is also substitute of lalu (and then…) and a substitute for selalu (always).

CI samples: Terus, gimana,? – And then, what happened?
Kok, salah, terus, sih,? – Why, is it always going (going) wrong, Ø?
**terus-terusan**
Signification: *to keep doing.*

CI sample:  
\[Tadi,\text{ sayajer} \text{terus-terusan} \text{aduk}_{\text{i}} \text{ ampe}_{\text{m}} \text{ kentel}_{\text{n}} \]
\[I_{\text{j}} \text{ kept}_{\text{t}} \text{ stirring}_{\text{i}} \text{ it}_{\text{o}} \text{ until}_{\text{m}} \text{ it}_{\text{o}} \text{ got}_{\text{o}} \text{ thick}_{\text{n}}\]

**timpuk**
CI alternative to *lempar* (*to throw*).

CI sample:  
\[Nimpukin, \text{ batu}_{j} \quad \text{To}_{\text{o}} \text{ throw}_{i} \text{ rocks}_{j}\]

**tuh**
Variant and elision of *itu* [+demonstrative pronoun].

CI samples:  
\[Iya, \text{ tuh}_{i} \quad \text{That}_{j} \text{ right}_{i}\]
\[Nga, \text{ salah}_{i} \text{ tuh}_{k} \quad \text{Is}_{o} \text{ that}_{k} \text{ right}_{j} \quad \text{(Sure about that?)}\]

**tukang**
Standard signification is *trade person*, in CI it is often used to refer to a habit.

CI sample:  
\[Tukang, \text{ ngibal}, \text{ tuh}_{k} \text{ orang}_{i} \quad \text{That}_{k} \text{ guy}_{i} \text{ is}_{o} \text{ a liar}_{j}\]

**tulen**
CI alternative to the standard *asli*. Signification: *authentic*.

CI sample:  
\[Yang, \text{ ini}_{j} \text{ tulen}_{k} \quad \text{This}_{j} \text{ one}_{i} \text{ is}_{o} \text{ authentic}_{k}\]

**uber**
CI alternative to the standard *kejar*. Signification: *chase*.

CI sample:  
\[Tadi, \text{ diuber}_{i} \text{ anjing}_{k} \quad \text{NP}_{\text{o}} \text{ was}_{\text{o}} \text{ chased}_{i} \text{ by}_{\text{o}} \text{ a}_{\text{o}} \text{ dog}_{k}\]

**udah**
Elision of *sudah* (*already, have done*). There is no fundamental semantic shift, but there are some common collocation speech acts of *udah*.

CI sample:  
\[Udah, \text{ lum}_{j} \quad \text{Is}_{o} \text{ it}_{o} \text{ done}_{i} \text{, } \text{Ø}_{j}\]
\[Udah, \text{ ah}_{i} \quad \text{That’s enough}_{j}\]
\[Udah, \text{ lah}_{j} \quad \text{Let it go}_{j} \quad \text{Don’t worry about it}_{j}\]
\[Udah, \text{ dong}_{j} \quad \text{That’s enough}_{j} \]
\[Udahen, \text{ yuk}_{j} \quad \text{Let’s}_{j} \text{ finish up}_{j}\]
**usah**
Semantic property is akin to *perlu* but only works in the negative anteceded by *engga*. When preceded by *ya!* It is an exclamation of objection.

CI samples:  

- *Nga, usah* — No need.
- *Nga, usah, ya*! — No way! / You wish!*

**waduh**
A variation of *aduh(auch)* but in CI it is used more so as an interjection particle [+concern +dismay].

CI sample:  

- *Waduh, ini, bakalan jadi masalah nih*  
  *Oh no*, this is going to be trouble.

**wah**
An interjection particle expressing surprise or concern.

CI samples:  

- *Wah, apa nih?* — Wow, what's this?
- *Wah, kalo gini terus repot nanti*  
  *Oh no, if it continues like this it'll be trouble.*

**waktu**
CI alternative to *Fiketika*. Signification: *when* [+conjunction +past tense marker].

CI sample:  

- *Waktu, nyampe kemarin kamu belum ada*  
  When I arrived yesterday you weren’t here yet.

**yang**
There is no DS or SS shift of *yang* in CI. Signification: +pronoun + anaphoric reference. The samples below are some examples of *yang* in collocation with other lexical items creating common speech acts:

- *Yang, benar,?!* — Really?!
- *Yang, benar, dong nih,?!* — Be serious!
- *Yang, mana,?!* — Which one,?
- *Yang, apaan,?!* — What type,?
- *Yang, kaya, apa nih,?!* — What does it look like?
- *Kerjain, yang, benar,!* — Do it properly!

**yuk/yu/yo**
Elision of *ayo(let’s...)*. It is also used as a particle to indicate leave taking.

CI sample:  

- *Jalan, yok* — Let’s go,
- *Yo* — See you, good bye,
Appendix 3

Extended Discussions on
CI Linguistic Features
**Extended discussions on CIAL: comparative FI – CI morphology and syntax**

Bahasa Indonesia (BI) is an agglutinative language where affixes attach to the root words, which are more commonly referred to as base words in BI. It can also have an inflectional character when morphological affixes trigger substitutions of consonants when attached to base words. X⁰ will represent BI base word in the phrase structure syntax. All BI base words are freemorphs and in the unmarked form, they are inherently word classes on their own with one or more functions; these are represented as N, V, A, P. In the phrase structure syntax the word classes are considered as phrases, thus NP denotes to Noun Phrase, VP as Verb Phrase, etc. The phrase definition is inherited from generative syntax where Nouns can have other elements such as articles, and in the case of BI, Nouns can have classifiers and quantifiers. The indices ₀, ₁, ₂ mark the θ-roles of the AGENT (subject), PATIENT (direct object) and BENEFICIARY (indirect object) respectively. The phrase structure syntax generated by the morphology is not a rule of ‘grammar’; it is a descriptive mechanism of the most common patterns. If it were attempted to be instituted into a rule, then it will become evident that various linguistic variables will block its functioning. The indexations H & L mark lexicon or sentences as H diglossic code and L diglossic code. Every main verb in a phrase structure can be preceded by an auxiliary (Aux) in the form of mood or aspect. ([___+Want/+Intention/+Future/+Progressive]).

BI affixes are bound morphs. The prefix ke- and di- besides being a morphological prefix, also have a word class function as preposition. Allomorphy and allphony take place in many BI base words between the FI and CI variant in the form of elisions and reduction of diphthongs into single vowels and the conversions of vowels into schwa. The following description of Indonesian morphology uses the model from Lieber’s Organisation of the lexicon.

**The active verb**

1a. FI: mə - ; [____ [X⁰]] [V, +active]

The FI morphological variations of the active me- prefix include: mə-, məm-, məɲ-, məŋ-.

1b. CI: n- / ŋ- / ŋ- ; [_____ [X⁰]] [V, +active]

Comparative examples (FI – CI):

1c. Hmengaku – ¹ngaku

1d. Hmembantu – ¹bantu
1e. \textsuperscript{1}mencari – \textsuperscript{1}cari / nyari
1f. \textsuperscript{1}mengasuh – \textsuperscript{1}ngasuh
1g. \textsuperscript{1}menonton – \textsuperscript{1}nonton

The \textsuperscript{1} morphological derivations simply demonstrate simplifications, namely in the omission of the me- prefix. Not every morpheme of the me- prefix is abandoned and a trace of it remains in CI as outlined in rule 1b and examples 1c, 1e, 1f. We can co-index the trace as ; : me,ngaku – ,ngaku;
me,ngasuh – ,ngasuh; me,nonton - ,nonton. Interestingly in example 1e an allomorphy of n- has emerged. This is a feature of CI that applies the same inflection to the first letters ‘c’ and ‘s’ consonants of a base word. The CI n- derivation of the ‘s’ consonant is otherwise just a shortening of the me- prefix.

The [V, +active] morphology generates SVO syntax: \( S \Rightarrow \NP_0 V \NP_1 \)

E.g.:
\[
\begin{array}{c}
\text{\textsuperscript{1}Pak jaga menangkap pencuri} \\
\NP_0 \\ V \\ \NP_1
\end{array}
\]

\[
\begin{array}{c}
\text{\textsuperscript{1}Pak jaga nangk\textsuperscript{ə}p maling} \\
\NP_0 \\ V \\ \NP_1
\end{array}
\]

\textit{The intransitive verb}

FI: bər- ; \[ ___ \[X^v]\] \[V, +intransitive\]
CI: \Ø- ; \[ ___ \[X^v]\] \[V, +intransitive\]

The [V, +intransitive] morphological marking generates the phrase structure syntax:
\( S \Rightarrow \NP_0 V (PP) (\NP_1) \)

\textbf{anak-anak bersepeda ke sekolah}
\[
\begin{array}{c}
\NP_0 \\ V \\ PP
\end{array}
\]

\textbf{anak-anak na\textsuperscript{ə}p sepeda ke sekolah}
\[
\begin{array}{c}
\NP_0 \\ V \\ PP
\end{array}
\]
Some base words are inherently already intransitive and can therefore function independently without morphological derivation in CI/MSI (e.g., $^H$menonton – $^L$nonton, $^H$bermain – $^L$main), while other base words will need auxiliary predicates to function such as naik or pakai (e.g., $^H$bersepeda – $^L$naik sepeda, $^H$bersepatu – $^L$pake sepatu).

The elision of /me-/ in CI/MSI also occurs on base words with [V, + active] markings. Thus the consonant ‘p’ just undergoes an allophony from [p] → [m]. (e.g., $^X$n$^p$akai - $^H$memakai - $^L$make [mə]) (the schwa allophony will be discussed later in this section), $^X$n$^p$ilih - $^H$memilih - $^L$milih).

The active-dative verb

FI: mə- –kan; [ ____ [X$^a$] ____ ] [V, +active, +dative]

CI: n-/ŋ-/ŋə- -in; [ ____ [X$^a$] ____ ] [V, +active, +dative]

Comparative examples (FI – CI):
$^H$mengambilkan – $^L$ngambilin
$^H$membelikan – $^L$ngebeliin
$^H$mencarikan – $^L$nyariin
$^H$menyalahkan – $^L$nyalahirin
$^H$menakutkan – $^L$nakutin

The [V, +active, +dative] morphological marking generates the benefactive phrase structure syntax:

$$ S \Rightarrow NP_0 V NP_2 NP_I $$

E.g.: $^H$Si Anton sedang membelikan Dewi bunga

NP₀ Aux V NP₂ NP_I

$^L$Si Anton lagi ngebeliin Dewi bunga

NP₀ Aux V NP₂ NP_I

In CI this syntax is often used in the progressive tense context and hence the modal progressive AUX lagi precedes the main verb. Certain base words will ‘block’ the $ S \Rightarrow NP_0 V NP_2 NP_I$ syntax from generating. This is in accordance with the premise of the *Lexical Hypothesis*, that ultimately it is the lexicon that triggers grammatical structures. Consider the predicate base words sebar, siar, and the
compound bagi-bagi. These are not benefactive lexical items that share the semantic element of 'sharing and/or spreading'. If we inflect these with the \([V, +active, +dative]\) morphological marking, then the \(S \Rightarrow NP_0 V NP_2 NP_1\) syntax becomes ungrammatical:

\[
*TVRI \text{ menyiarkan seluruh negara berita} \\
NP_0 \quad V \quad NP_2 \quad NP_1
\]

\[
*Bung \text{ Karno menyebarkan rakyat pemikirannya} \\
NP_0 \quad V \quad NP_2 \quad NP_1
\]

\[
*Kepala \text{ sekolah membagi-bagikan semua anak-anak hadiah} \\
NP_0 \quad V \quad NP_2 \quad NP_1
\]

To make these sentences grammatical the NP_1 will need to be promoted, the NP_2 demoted and a preposition linking the two patient NPs:

\[
S \Rightarrow NP_0 V NP_1 PP NP_2
\]

\[
^{^H}TVRI \text{ menyiarkan berita ke seluruh negara} \\
NP_0 \quad V \quad NP_1 \quad PP \quad NP_2
\]

\[
^{^H}Bung \text{ Karno menyebarkan pemikirannya kepada rakyat} \\
NP_0 \quad V \quad NP_1 \quad PP \quad NP_2
\]

\[
^{^H}Kepala \text{ sekolah membagi-bagikan hadiah kepada semua anak-anak} \\
NP_0 \quad V \quad NP_1 \quad PP \quad NP_2
\]

\[The \text{ nominative copula verb}\]

\[
\text{FI: \quad adalah; \quad[V, +nominative]} \\
\text{CI: \quad } \emptyset \quad [V, +nominative]
\]

The nominative verb generates the phrase structure syntax:
Pak Syamshudin adalah Rektor Universitas Satma Jaya

\[ \text{Pro}_0 \ \text{V} \ \text{NP}_0 \]

It must be noted that the intonation in CI compensates for the ellipsis by a prolongation of the last syllable of the PRO as if to fill in the Ø gap.

**The passive-dative verb**

FI: \( \text{di–} -\text{kan}; [____ \text{[X]} _____] [\text{V, +passive, +dative}] \)

CI: \( \text{di–} -\text{in}; [____ \text{[X]} _____] [\text{V, +passive, +dative}] \)

Comparative examples (FI – CI):

\( ^\text{H}\)diambilkan – \( ^\text{L}\)diambilin
\( ^\text{H}\)dibelikan – \( ^\text{L}\)dibieliin
\( ^\text{H}\)dicarikan – \( ^\text{L}\)dicariin
\( ^\text{H}\)disalahkan – \( ^\text{L}\)disalahin
\( ^\text{H}\)ditakutkan – \( ^\text{L}\)ditakutin

The [V, +passive, +dative] morphological marking in traditional FI grammar generates the phrase structure syntax:

\[ S \Rightarrow \text{NP}_2 \ \text{V} \ \text{NP}_0 \ \text{NP}_1 \]

Hence the sentence;
*? Dewi sedang dibelikan Anton bunga
NP₂ Aux V NP₀ NP₁

The marking *? questions the grammaticality the sentence. While it may conform to traditional FI rule of syntax, if we adhere to a modern linguist’s definition of grammaticality the sentence would appear ungrammatical. Beyond the classroom grammar lesson, it is an extremely stillted phrase not in daily usage.

The CI version has quite a different syntax:

₁Dewi lagi dibeliin bunga sama Anton
NP₂ Aux V NP₁ PP NP₀

4.6.1 The null-element parameter
A typical differentiation between FI and CI is the null-element parameter which include null-subject, null-object and predicate ellipsis. In CI the null-elements are understood from the general context of the conversation. Often the null-element has already been mentioned and subsequently need not be repeated. Unlike in other pro-drop languages such as the conjugations in Romance languages, where the subject-verb-agreement morphology contains the information about PRO, in CI there are no inflection or morphological markings that compensate for the null-elements.

The following are some comparisons of FI sentences with null-elements in CI. The examples 1h – 1l are answers to the question “kamu mau coba yang itu?” Examples 1m – 1r are answers to the question “kamu sudah pernah pergi ke Banten?” All CI null-elements are grammatical:

1h. ₁Saya mau coba yang itu
NP₀ Aux V NP₁
1i. ₁ _____ mau coba yang itu
Ø₀ Aux V NP₁
1j. ₁Saya mau coba _____
NP₀ Aux V Ø₁
1k. ₁ _____ mau coba _____
Ø₀ Aux V Ø₁
1l. ₁ _____ mau _____
Ø₀ Aux Øᵥ Ø₁
4.6.2 Interrogatives and Wh-words

As much as prescriptive textbooks teach that the wh-word ‘apakah’ marks an open question, it is virtually non-existent in the spoken informal language and therefore ungrammatical. The following FI question is an example of an ungrammatical question in CI: “*apakah yang sedang Udin laukan?”. Natural and grammatical in spoken CI is “lagi apa si Udin?”, or “Si Udin lagi ngapain?”. Note that there is flexibility in BI wh- phrase structure where a wh- word can begin and end a sentence:

\[ \text{Apa itu yang di atas meja?} \]
\[ Wh_0 \text{ Det Aux PP NP}_1 \]

\[ \text{Yang di atas meja itu apa?} \]
\[ \text{Aux PP NP}_1 \text{ Det } Wh_0 \]

Note too that the locative predicate ada [V, +locative] is optional and a predicate ellipsis can take place:

\[ \text{Apa itu yang (ada) di atas meja?} \]
\[ Wh_0 \text{ Det Aux V PP NP}_1 \]
In CI a declarative sentence can become a question without a change of phrase structure but with the use of prosody and intonation. Thus the sentence *Si Ade lagi tidur* can be both a statement and a question depending on the prosody and intonation employed by the speaker.

### 4.6.3 Comparative, superlative and excessive parameters

Variations in constructing the comparative, superlative and excessive parameters are characteristic differences between FI and CI. The examples below use the FI adjective *besar* and the more common CI Equivalent *gede*.

#### 4.6.3.1 Comparative

In FI the comparative is constructed by preceding an adjective with the adverb *lebih*, whereas in CI the suffix `–an [[Adj] ___ ] [Adv, +comparative]` marks the comparative.

```
^1_Lebih besar
adv  adj
```

```
^1_Gede an
adj  suffix
```

#### 4.6.3.2 Superlative

The superlative in FI can be formed by the *ter-* prefix `[[Adj]] [Adv, +comparative]` and the adverb *paling*. In CI only the latter is common.

```
^1_Terbesar
prefix adj
```

```
^1_Paling gede
adv  adj
```

#### 4.6.3.3 Excessive

In FI the excessive parameter is constructed by adding the adverb *terlalu* to an adjective. In CI the affixes `ke- –an [ ___ [Adj] ___ ] [Adv, +excessive]` typically mark the excessive parameter.
4.6.4 Elisions and allophony

Many but not all BI lexis that are common in everyday use undergo elision and/or allomorphy. Some words simply cannot undergo elision as it changes the meaning, such as sikat (brush) → ikat (to tie), and becak (rickshaw) → becek (wet ground). Elision and allomorphy in MSI lexicon can be traced to the Javanese and Sundanese influence (either directly or indirectly via Betawi Malay). Javanese and Sundanese lexica that have an identical counterpart in FI (from Riau Malay) are often elided and have an allomorph variation.

4.6.4.1 /a/ → /æ/

| antar | antər |
| asam  | asəm |
| balas | baləs |
| benar | benər |
| bosan | bosən |
| cepat | cepət |
| dalam | daləm |
| dapat | dapət |
| dekat | dekət |
| dengar | dengər |
| desak | desək |
| diam | diəm |
| gemas | geməs |
| garam | garam |
| gatal | gatal |
| hangat | angət |
| hantam | antəm |
| hisap | isəp |
| hitam | itəm |
| ikat | ikat |
| ingat | ingət |
injak
kencang
kental
kesal
lancar
lapar
lipat
macam
malam
malas
matang
 pagar
 panjang
 pantas
 pinjam
 pintar
 pesan
 sambal
 sambar
 santan
 santap
 sebal
 segan
 selamat
 sempat
 senang
 sesak
 simpan
 tajam
 tampang
 tanam
 tangkap
 teman
 tenggelam
 tukar
4.6.4.2 /aː/ → /e/: 
- bangkai bangke
- cabai cabe
- capai cape
- cerai cere
- gulai gule
- lantai lante
- pakai pake
- pandai pande
- pantai pante
- ramai rame
- rantai rante
- satay sate
- sampai sampe
- santai sante
- tokai toke

4.6.4.3 /aː/ → /aɔ/: 
- kain kaɔ
- lain laɔn
- main maɔn

4.6.4.4 /aɪk/ → /aː/: 
- baik baɔ
- naik naɔ

4.6.4.5 /ɪk/ → /a/: 
- tarik tarɔ
- naik naɔ

4.6.4.6 /ia/ → /a/: 
- durian durɔn
A grammatical body of knowledge pertaining to the pattern of the a → /a/ & /e/ allomorphy does not yet exist. A very brief observation suggests that words with the last-syllables /ah/ (cah/, kah/, mah/, nah/, rah/, sah/), do not undergo the schwa conversion; /bocah/, /lincah/, /pecah/, /langkah/, /tingkah/, /rumah/, /ramah/, /musnah/, /punah/, /fitnah/, /darah/, /marah/, /parah/, /asah/, /basah/, /susah/. The last-syllables /tak/, /lak/ and /af/, also do not appear to undergo the schwa conversion; /kotak/, /botak/, /basah/, /asah/, /tolak/, /mutlak/, /maaf/. There is yet no evidence that there are lexica with these last-syllables that undergo the schwa conversion. With other last-syllables, it seems that the conversion is an ‘irregular’ grammatical process. The last-syllable /sak/ for example may or may not undergo conversion; /rusak/ → /rusak/, /masak/ → /masak/, /sesak/ → /sesak/, /desak/ → /desak/.

4.6.4.7 /u/ → /o/  
belum belom  
mabuk mabok  
jatuh jatoh  
keruk kerok  
ketuk ketok  
separuh separo  
supir sopir  
surga sorga  
taruh tareo  
tegur teger  
telur telor  

4.6.4.8 /au/ → /o/  
hijau ijo  
kacau ngaco  
kalau kalo  
kerbau kebo  
pulau pulo  
saudara sodara  
saus sos  

4.6.4.9 Elision of first letter phoneme /h/  
habis abis  
hafal afal  
halus alus
hampir    ampir
hancur    ancur
handuk    anduk
hangat    angat
hantam    antam
hantu      antu
hati      ati
hidup      idup
hijau      ijo
hilang    ilang
hisap      isap
hitam      itam
hitung    itung
hujan      ujan
hutan      utan
hutang    utang

4.6.4.10 Elision of middle/last letter phoneme /h/
bohong    bo’ong
jahl      jail
jahit      jait
jodoh    jodo
lihat      liat
pahit      pait
sahut      saut
tahi      tai
tahun      taun
taruhan    taroan

4.6.4.11 Elision of /s/
satu      atu
saja      aja
sama      ama
sampai    ampe
sudah      udah
Appendix 4

Some issues of English L1 Interference in Learning Indonesian
Some practical considerations on English-Indonesian parametric comparisons

This section discusses the key parametric non-transfers between English and Indonesian. Pointing out similarities between Indonesian and English is a useful tool in teaching Indonesian to students of English speaking backgrounds. More difficult is to try to explain grammar and linguistic elements of Indonesian that do not occur in English. This requires employing the tools of universal linguistic analysis. Using a universal linguistic framework transcends the limitation of a grammar that is specific to a particular language. The following discussions addresses issues that students of Indonesian often grapple with as their natural linguistic expressions – that of English – are often blocked when they try to translate their L1 (English) parameters into L2 (Indonesian).

The copula null-element ‘adalah’ vis-a-vis the existential-locative ‘ada’

The English copula to be has two forms in Indonesian; adalah and ada. Adalah is used for the nominative clause, linking two Noun Phrases and is an FI item. The equivalent of the FI clause ‘Pak Tri adalah, seorang petani’ (Pak Tri is a farmer) is ‘Pak tri Ø, Ø petani’ in Cl. In this paper the copula adalah is considered a null-element in CI. Unlike in English, the Indonesian continuous tense structure is not constructed with the copula. Thus the clause ‘they; are, swimming’ translates as ‘mereka, Ø, lagi, berenang’. The progressive tense marker lagi (sedang) is also constructed with passively inflected predicates to fulfil the tense function indicating ‘a current phenomenon’. The clause ‘lagi, terkenal, dia’ translates as ‘he is, famous, at the moment’. It can be argued here that there is also an ellipsis of the English copula-gerund compound being, which would have represented the adalah null-element.

The verb ada, which can perform the role of the English copula ‘to be’, is a proper verb in Indonesian that is used to indicate location and presence. Like the copula adalah, there are instances where ada is also omitted. In the example ‘lagi, dia, manak?’ (where, are, you?), lagi infers both the progressive tense and the locative ada. The same dual function of lagi applies to the example ‘kita, lagi, di, pasar’ (we, are, at, the market).

In phrases where the copula to be is phrased with the demonstrative pronoun it (e.g., it is, it was, was it?), the null-copula-element in a phrase like ‘it was raining’, is inextricably collocated with the expletive it, both of which are null-elements in Indonesian. This also applies in cases where the copula is collocated with it as an anaphoric reference. Discussions on these null elements are covered in Part C of this section.

The subject and personal pronoun null-element
When context is established between speakers, the subject of a clause, most of the time being the personal pronoun, is often omitted in spoken Indonesian. When not contextualised, it is not possible to establish the subject in an isolated written clause such as ‘nga, suka NP Ø’ (‘NP does not like, it Ø’ - a response to the question ‘suka, ngaj NP Ø?’ (‘do you like, it Ø?’)). Unlike pro-drop languages where the conjugation of the predicate provides the subject-NP information, in Indonesian it is solely the known context of the speakers that establishes the subject of the pronoun null-element. Exception can be made in common direct questions such as ‘lagi, di mana?’ (‘I where you?’) and ‘mau, kej mana?’ (‘J where you going?’). These are common speech acts where it is a given that the null pronoun is the second person personal pronoun. If other than a second person personal pronoun subject is intended then it is not an ellipsis in the clause: ‘nga, mau ahk’ (‘I don’t want, any he/she?’); ‘mau, juga sih’ (‘I actually want, some’).

The ‘it’ pronoun

At 2012 ASILE presentation in 2012, Dr George Quinn of the Australian National University gave a presentation on translations of the demonstrative pronoun ‘it’ into Indonesian. Dr Quinn pointed out that students of Indonesian have the tendency to translate the demonstrative pronoun ‘it’ as its Indonesian counterpart ‘itu’. Dr Quinn demonstrated that a direct translation of the demonstrative pronoun ‘it’ as ‘itu’ will often not result in meaningful translations. It might seem that a phrase such as ‘that’s not it’, can be translated easily enough as “bukan itu”. But careful consideration reveals that ‘itu’ in this example stands for both personal pronouns ‘that’ and ‘it’. Whereas in a phrase such as ‘is it time?’, ‘itu’ cannot be applied as a translation of ‘it’.

The following discussions extends Dr Quinn’s demonstration of non-parametric transfers between ‘it’ and ‘itu’ by considering the fundamental English linguistic functions of the pronoun ‘it’ and providing samples how these functions correlate in Indonesian. Three grammatical functions of the demonstrative pronoun ‘it’ will be discussed: ‘it’ the expletive, ‘it’ the anaphoric reference and ‘it’ the co-referential pronoun.

‘It’ the expletive

The expletive or dummy element feature occurs when demonstrative pronouns are used in the subject positions in a clause but they do not function as a demonstrative pronoun and they have no inherent semantic meaning. As such, they also have no thematic role assigned by predicate. Consider the following examples:

a. It is good
b. It is raining; it was raining

In example a, ‘it’ represents an NP that is known from context or already mentioned in a previous sentence (e.g., ‘the food is good’). In example b however it is not clear what ‘it’ stands for. Is the ‘day’ raining? The ‘sky’ raining? Or the ‘now’ raining? It is a requirement of English syntax that the subject position of a clause must not be empty and ‘it’ in example b fills in the subject position to complete the syntactical structure without representing any NP whatsoever. This expletive function does not exist in Indonesian. The subject position will simply be null (akin to the pro-drop) and the specific time reference fulfil the role that the present/past tense conjugation of the copula performs: “it’s raining” (Ø, Ø, lagi, hujan) and “it was raining” (Ø, Ø, tadi, hujan). In both of these examples the expletive ‘it’ (and copula to be: ‘is’, and ‘was’) are null-elements in Indonesian. The past tense conjugation ‘was’ is filled with time-reference marker tadi and the progressive tense gerund ing has correlation in the auxiliary tense marker lagi.

The same principle applies with other expletive cases that refer to time. The phrases “it is late” or “it’s already late” are translated as “sudah telat”. In cases where the expletive is in subject position of an adjectival clause but no explicit NP is specified, such as “it is cold” (the ‘weather’ is cold?, ‘the ‘day’ is cold?, the ‘place’ is cold?), then the adjective itself suffices in Indonesian: ‘dingin’. Otherwise if the speaker has a particular Noun-subject in mind then it is explicitly mentioned: “dinding cuacanya”, “tempatnya dingin”, etc (“the weather is cold”, “the place is cold”).

Another common expletive form is the passive phrase. These are translated in Indonesian as the passively inflected base word: “it is said” – “dikatakan / katanya”, “it is known” – “diketahui”, “it is believed” – “dipercayai”.

The following are examples of other common expletive speech acts of phases. The Indonesian forms translate the main theme of the phrase: “it is so” – “iya begitu / memang begitu”, “it seems like…” – “kayaknya…”, “it is possible” – “mungkin”, “it sounds like…” – “kedengerannya”

‘It’ the anaphoric reference

‘It’ as an anaphoric reference derives meaning from an antecedent NP. In a phrase such as “it was good”, ‘it’ refers to a known antecedent subject NP that might be ‘the food’ or the ‘the film’. In Indonesian the known subject NP is simply restated: “bagus, filmnya”, (it, was good). The same statement can also be translated without an NP: “memang bagus”. The speaker here adds emphasis by using ‘memang’ and the NP may be omitted. The intention of the speaker here is [+affirmative], in agreement or response to another speaker. If the same statement is rephrased as a question (“is it good?”), then the yes-no question marker ‘nga’ follows the adjective: “bagus nga?”. The subject NP may be excluded if already known from context. The equivalence in FI would need to have the
subject NP included: “apakah bagus filmnya?”. If the intention of the question is asking for confirmation, rather than a simple open question, then ‘memang’ - or its elided form ‘emang’ - is used again: “bagus emangnya?”.

The omission of subject NP in Indonesian also applies to open anaphoric questions containing ‘it’ that are phrased with modal and tense auxiliaries: “do you like it?” → “Ø, Ø, sukaø, Ø, ngaø?”; “have, you, tried, it?” → “pernah, Ø, cobåø, belumø, Ø?”.

Where the ‘it’ anaphoric word refers to a known NP ‘matter’ and is placed in subject position in passive predicate structure, then the Indonesian structure states the tense that the phrase hints and the third person passively inflected base word: “it is done” → “sudah dilakukan”; “it will be” → “akan (ter)jadi”. In adverbial phrases such as “it could be” and “it is certain”, then the adverb alone expresses such speech acts in Indonesian (in this case being [+possibility; +certainty]): “mungkin”; “pasti”. The demonstrative pronouns ‘ini’ or ‘itu’ can be part of this structure functioning as anaphoric words, thus: “ini mungkin” or “pasti itu”. But this is counter intuitive to the native speaker since ‘ini’ and ‘itu’ translate more accurately to the pronouns ‘this; and ‘that’. Instances where ‘itu’ is treated as the pronoun ‘it’ sounds foreign and occurs possibly due the influence of English syntax conventions on modern Indonesian. ‘Ini’ can function as the demonstrative pronoun anaphoric ‘it’ in phrases like “Ini diaq,” (“here; it; is it”). Both pronouns ‘ini’ and ‘dia’ represent the anaphoric pronoun ‘it’, while concurrently ‘ini’ also represents the existential role of the pronoun ‘here’.

In cases where the anaphoric ‘it’ phrase accompanies relative pronouns, such as “it, is, that, one,” the relative pronoun ‘yang’ and the demonstrative pronoun ‘itu’ is used in the Indonesian: “Ø, Ø, yang, ituq”. In its question form “is it this one?”, intonation suffices in Indonesian for the conversion without any change in syntax. Alternatively the question markers ‘bukan’ and FI ‘apakah’ can be used: “yang ini bukan?”; “apakah yang ini?”. The tag question “it’s, this, one, isn’t, it?” is translated with the tag question marker ‘kan’ (curiously a contraction of ‘bukan’ but has acquired independence functioning as a tag question marker): “Ø, yangq, ini, kanq?”. It is worth noting that the tag question marker isn’t it is co-referential with the pronoun-copula compound It’s. Such anaphoric reference is not represented in the Indonesian translation. In the example above the pronoun iniq corresponds to demonstrative pronoun thisq.

Another anaphoric structure of ‘it’ is where it is positioned in object NP position: “I, amq, eatq, ingq, itq.”. Again, the Indonesian equivalent may omit the object NP or include it: “saya, Ø, lagiq, makanq, Øq?”; “saya, Ø, lagiq, makanq, pisangnya,” (assuming that ‘it’ here stands for banana). In similar phrases where the anaphoric ‘it’ is part of a preposition phrase such as: “did, notq, thinkq, ofq, itq”, then the predicate is expressed in its passive form: “Ø, ngaq, kepikirq, Øq, Øq.”.

Confusion about the infinitive ‘to’ particle

It is a feature of the English language that the infinitive verb is accompanied by the particle ‘to’. The infinitive English verb is syntactically positioned after other verbs either as the main verb following
The predication of question words and demonstrative pronouns

The predication of question words is a feature in Indonesian that does not occur in English. Every student of Indonesian learns that ‘mengapa’ means ‘why’. What often is not taught is that ‘mengapa’ is the question word ‘apa’ that has technically undergone predication through the attachment of the active prefix ‘meng-‘. Curiously ‘mengapa’ does not function as a predicate word class and it does not assign thematic roles. On the other hand the passive predicate form of ‘apa’, ‘diapakan/diapain’ assigns theta roles. The same applies to the predicate question phase ‘ngapain’. Other predicated question words that take theta roles is ‘mana’: ‘dikemanain/digimanain’.

The demonstrative pronouns ‘ini’ and ‘itu’ can also be predicated: begituin, beginiin, dikesiniin, dikesituin. The movement or action involved over time in these predicated forms convert the demonstrative pronouns ‘ini’ and ‘itu’ into dynamic verbs.

Some question framing without Wh- question words

An aspect of question framing without the use of question words is rarely taught in conventional Indonesian language textbooks. It is common in very language that intonation itself without the aid of question words can form questions. But there is no universally used orthographic description for depicting intonation and prosody, hence textbooks resort to overusing question words as question markers. One notable Indonesian question marker that is overused in textbook but in actually spoken Indonesian rarely occur is ‘apakah’. The standard explanation is that ‘apakah’ forms yes-no question. Yet, intonation itself can create yes-no question without the question marker ‘apakah’: “apakah ini punyamu?“ → “ini punyamu?”. In CI there are three other common tag questions for yes-no questions. The question markers ‘nga’, ‘kan’ and the adjectival phrase ‘emangnya’.
The question marker ‘nga’ is what the listener will hear in everyday spoken Indonesian instead of ‘apakah’. “Nga” follows the adjective while ‘apakah’ precedes the adjective. The essential intention does not change between ‘apakah’ and ‘nga’: “apakah baik orangnya?” → “baik nga orangnya?”

The tag question ‘kan’ – an elision of ‘bukan’ - behaves rather like the English tag question when the copula phrase is repeated in its negative form: “it is good isn’t it?” → “bagus kan?”. As discussed in sections A and B, the Indonesian subject and copula (and the copula not being an antecedent to adjectives) are often null elements in Indonesian. ‘Kan’ in the above example, stands as the tag-question marker for the anaphoric pronominal and copula null elements. In the case of the question marker ‘emangnya’, while the function is still forming yes-no question types, the intention shifts from asking for confirmation (‘kan’) to inquiring for open information where the adjective is used as a possible axiom: “bagus kan?” [+asking +confirmation]; “bagus emangnya” [+asking +inquire].

A peculiar CI question phrase uses the markers “Pake [   ] segala”. These separable question phrase markers are an alternative to the question marker “kenapa [   ]”: “kenapa [harus pergi ke sana]?” → “Pake [harus pergi ke sana] segala?”.

Translating expressions and speech acts
A common tendency of English speaking students that are learning Indonesian is to translate expressions from English to Indonesian at the SS signification, not realising that the DS value is not transferred. For example, the expression ‘to have a go’ will have an SS translation as ‘ke punya suatu pergi’, which has no inherent DS value whatsoever. What students often do not realise is that expressions have a predicate equivalent, which in this case is to try. Once this transfer from an idiomatic expression to its predicate equivalent has been made, then a DS value transfer from L1 to L2 can be made, translating the English predicate to try into its Indonesian equivalent → coba. So now the DS value of ‘to have a go’ can be translated as ‘mencoba’. The same goes for translating from Indonesian into English. Idiomatic expressions and speech acts such as ‘ngapain aja?’ will translate as ‘what are you doing?’ at the SS signification, which has no inherent DS value in English. The DS value in English of ‘ngapain aja?’ is ‘what have you been doing?’ There was no means by which the students would have deciphered its DS value. There are many idiosyncratic Indonesian parametric features that the student would need to appreciate when translating Indonesian expressions. In the above example of ‘ngapain aja?’ the student would need to be familiar with the concept of the personal pronoun null-element and the predication of question words – features which do not occur in English. Secondly, the students need to familiarise themselves with set expressions in Indonesian that are created through the collocation of lexical items.
Appendix 5

The Influences of Regional Languages on CI
**Comparative lexicon: CI – regional languages**

The diglossic language demarcation of Indonesian is a default product of the wider Indonesian linguistic culture that is rooted in the regional cultures (*Budaya Daerah*). Built into the quasi caste hierarchies of the Sundanese-Javanese-Balinese cultures are ‘high’ (*halus*) and ‘low’ (*kasar*) levels of lexis. This ‘high-low’ linguistics psychology has possibly been transferred into the Indonesian language, whose majority of speakers is from the Sundanese-Javanese-Balinese cultures. Thus, the complexities of Indonesian diglossic sociolinguistics have linear correlation to the sociolinguistics of the major regional languages.

The Indonesian non-usage of the second person pronoun ‘you’ is a prime example. Instead names or terms of address (Bapak, Ibu, Mba, Mas, Bang, Kang, Neng, etc – the last four being direct adaptation from Javanese and Sundanese) are used. In spoken phrases, a second person pronoun can be omitted all together: ‘*dari mana?*’ (‘where are you from?’ or ‘where have you just come from?’) and ‘*sudah lama di sini?*’ (‘have you been here long?’).

In very informal situations, or when talking to a child, the word *kamu* (or *lu*; between friends - originally in Jakarta only) is used. Conversely, very formal situations might use the second person pronoun *anda*, which in daily speech is rather awkward, and could even be regarded as artificial hence aloof and offensive.

“Which language register a speaker uses depends on the relative status of the two speakers to each other, as well as the formality of the situation and what they are talking about. A whole discourse can be in only one of the levels, but speakers can also move between them within sentences, sometimes picking a higher or lower register word for special emphasis.” (Johnson 2011: *The Economist* Nov 16)

MSI (spoken Indonesian), the language that people use on everyday-basis, is a mixture of FI (*Bahasa Baku*), CI and even English. As is the case with the use of language levels in regional languages, “variables such as degree of intimacy, age, sex, marital status, physical and event setting, speech act features, presence and identity of audience and even types of clothes participants are wearing – all differentially interact to control register choice” and the possible mix of lexis from different language levels (Zurbuchen 1981: 106-107).

Indonesians can direct their speech ‘up’ and ‘down’ during a conversation. There exists an innate mastery of language usage among Indonesians that traverses the delineation of Indonesian social space. Honorific and depreciating vocabularies, even though theoretically diglossically subdivided,
can co-occur when the context is appropriate. “The language levels are like a ladder: the important thing is not who is on each step, but on being able to go up and down” (Hunter 1988: 11).

Many modern day Indonesian speakers take the input from the regional languages for granted. Standard Indonesian from the 1950s is markedly different from today’s standard-formal language. Many of the lexis from CI and the regional languages have penetrated into the sociolinguistic realm of FI and accepted as standard Indonesian by virtue of common usage. Words that originally were only used in informal CI speech are now regular features in both spoken and written journalism (FI domain). The following is an example of some CI words (of regional language origin) that are now commonly used in FI sociolinguistic domain. The subscript index \text{SUN,JAV,BAL,BET} denote to - the regional origin in – Sundanese, Javanese, Balinese and Betawi respectively: \text{boro-boro}_{\text{SUN/BET}} (jangankan\text{FI} – ‘not even’), \text{molor}_{\text{SUN/BET}} (tidur\text{FI} – ‘sleep’), \text{ngoto}_{\text{JAV/BET}} (bersikeras\text{FI} – ‘adamant’), \text{ketimbang}_{\text{BAL/BET}} (dari pada\text{FI} – ‘rather than’), \text{pada}_{\text{BET}} (\emptyset \text{FI} – [+pronominal particle +plural marker]}, \text{mending}_{\text{JAV}} (lebih baik\text{FI} – ‘it is better that’), \text{padahal}_{\text{SUN}} (sebenarnya\text{FI} – ‘actually’), \text{kerasan}_{\text{SUN/JAV}} (beta\text{FI} – ‘to feel at ease/at home’), \text{kapok}_{\text{SUN/JAV,BAL/BET}} (\emptyset \text{FI} – ‘to have had enough’), \text{gede}_{\text{SUN/JAV,BAL/BET}} (besar\text{FI} – ‘big’), \text{lumrah}_{\text{SUN/JAV}} (biasa\text{FI} – ‘usual’).

The linguistic features of CI can be described as a syncretism of Malay with the regional languages and dialects of Indonesia. This section endeavours to provide some further understanding of the influences that the regional languages have on Indonesian. Surprisingly, there has not been much research undertaken in this area. Grijns (1991: 106) stated that ‘there is still limited data that provides information about speech habits in the Indonesian language that have been acquired from modernism and regional languages’.

There is a substantial lexical input from Sundanese, Betawi and Javanese in CI-SI. Where there is a constant at the lexical level with its Malay origins (common base word), many linguistic features that operate on the words are non-Malay in origin. The table 12 compares the phonological, morphological and syntactical isomorphism in CI/SI that is derived from Sundanese, Javanese and Balinese. As discussed in Chapter 2, many of these linguistic inputs were absorbed into CI-SI indirectly via Betawi-Malay: ‘many borrowings from Betawi-Malay come from Javanese’ (Grijns 1991: 107). The Betawi people themselves trace their ancient origins to the ancient Sundanese culture. Jakarta, prior to conversion into Islam, was known as Sunda Kelapa, a port city of the Taruma Negara and Padjajaran Sundanese Kingdoms. Even during this early period, mixed ethnic marriages with Malays, Javanese, Indian and Chinese had taken place. Later on, under Islamic rule it became known as Jayakarta, where there was a greater influx of Muslim Javanese from the Central Javanese Sultanate of Demak. Under the Dutch, Jakarta was renamed Batavia, (from which the Betawi people received their name sake). It
is not clear how dominant Malay was in Jakarta prior to the Dutch era. Being a lingua franca across the archipelago, Malay would have been used especially in the trading/port suburbs of the old Jakarta. Undoubtedly Sundanese was common, as it still is today, around the periphery of Jakarta, while areas with a concentration of a particular ethnic group (i.e., Chinese, Javanese) would have continued speaking in their dialects/languages. Yet it was the Dutch who instituted Malay as an administrative language and possibly reinforced it and elevated Malay as the hometown or mother tongue of Batavia. The multi-ethnic presence continued and flourished during this time. The Dutch encouraged the entrepreneurship of the Chinese and Indians, thus China towns and Indian sections were part of the characteristics of Batavia. These cultures left a lasting influence on the Betawi culture (even today, some traditional Betawi customs and dances bear unmistakably Chinese traits, and the Betawi people today still count in Hokien-Chinese). The Dutch also brought in coolies from Bali, hence the presence of ‘Kampong Bali’ in several Jakartan areas. It is in this melting pot environment of Batavia that the Betawi Malay dialect evolved. There is no written record of the development of Betawi Malay, but at the dawn of the modern era of the twentieth century, Betawi Malay was already long established as the native Malay dialect of the Betawi people. As discussed in Chapter 2, it was this Betawi Malay that provided the linguistic platform for modern CI to sprout, and its flowing and ‘natural’ sounding character in turn provided SI its spoken dynamics.

Once phonological and morphological constancy had taken place in SI, lexical injections from the regional languages continued. Javanese, Sundanese and Balinese are hierarchical languages - there are levels of language variants that pertain to specific sociolinguistic domains (diglossia within diglossia). The basic subdivision is halus and kasar, semantically translating into ‘refined’ and ‘common’ (Kromo-Ngoko in Javanese). There are also further subdivisions. Where a lexical item has been absorbed into CI/SI, it is most likely to be the ‘common’ variant of the regional language in question. Sociolinguistically speaking, the ‘spoken’ (CI) element of SI often correlates with the spoken ‘common’ regional languages. Comparative linguistic features are expounded in section
Table 19. Comparative linguistic features

<table>
<thead>
<tr>
<th>Base Word [X&quot;]</th>
<th>CI-SI</th>
<th>Sundanese</th>
<th>Balinese</th>
<th>Javanese</th>
<th>FI</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ajar [V, + active]</td>
<td>ngajar</td>
<td>Ngajar</td>
<td>Ngajah</td>
<td>ajar</td>
<td>mengajar</td>
<td>To teach</td>
</tr>
<tr>
<td>Akan [Aux, + Tense]</td>
<td>mau/bakal/ bakalan</td>
<td>bade, arek</td>
<td>larak, pacing</td>
<td>arep, bade</td>
<td>akan</td>
<td>Will (auxiliary: future marker)</td>
</tr>
<tr>
<td>Aku [Pronoun]</td>
<td>aku, saya, gue</td>
<td>Abdi</td>
<td>Titian</td>
<td>aku, dalem</td>
<td>aku</td>
<td>I (1st person personal pronoun)</td>
</tr>
<tr>
<td>Alir [N]</td>
<td>aliran</td>
<td>Aliran</td>
<td>Embahan</td>
<td>Iline</td>
<td>aliran</td>
<td>Stream</td>
</tr>
<tr>
<td>Antar [X”]</td>
<td>anter</td>
<td>Anteur</td>
<td>Ateh</td>
<td>terke</td>
<td>antar</td>
<td>To take somebody somewhere</td>
</tr>
<tr>
<td>Antar [V, + active]</td>
<td>nganter</td>
<td>ngajapkeun</td>
<td>Ngatehin</td>
<td>ngeterake</td>
<td>mengantar</td>
<td>To take somebody somewhere</td>
</tr>
<tr>
<td>Asah [V, + active]</td>
<td>ngasah</td>
<td>Ngasah</td>
<td>Nyangihin</td>
<td>ngasah</td>
<td>mengasah</td>
<td>To sharpen (knife, etc)</td>
</tr>
<tr>
<td>Bantu [V, + active]</td>
<td>ngebantu, nolongin</td>
<td>Ngabantu</td>
<td>Nulungin</td>
<td>nulung</td>
<td>membantu</td>
<td>To help</td>
</tr>
<tr>
<td>Baru [Adj]</td>
<td>baru [adj; interjection]</td>
<td>Anyar</td>
<td>Mara</td>
<td>anyar</td>
<td>baru</td>
<td>New</td>
</tr>
<tr>
<td>Beli [V, + active]</td>
<td>beli, ngebeli</td>
<td>Meuli</td>
<td>Meli</td>
<td>tuku</td>
<td>membeli</td>
<td>To buy</td>
</tr>
<tr>
<td>Beli [V, + active, + benefactive]</td>
<td>ngebeliin</td>
<td>mangmeulikeun</td>
<td>Meanin</td>
<td>nukoake</td>
<td>membelikan</td>
<td>To buy somebody something</td>
</tr>
<tr>
<td>Bikin [V, + active]</td>
<td>bikin, ngebikin</td>
<td>ngadamel</td>
<td>Ngae</td>
<td>ndamel</td>
<td>membikin</td>
<td>To make</td>
</tr>
<tr>
<td>Bingung [V, + active, + causative]</td>
<td>ngebingu- ngin</td>
<td>ngalieurkeun</td>
<td>mingungin</td>
<td>bingungake</td>
<td>membingung- kan</td>
<td>To confuse, confusing</td>
</tr>
<tr>
<td>Cebur [V, + active, + dative]</td>
<td>nyeburin</td>
<td>nyeburkeun</td>
<td>nyeburin</td>
<td>njegurake</td>
<td>menceburkan</td>
<td>To plunge into water</td>
</tr>
<tr>
<td>Cebur</td>
<td>kecebur, kacemplung</td>
<td>macemplung</td>
<td>kejegur</td>
<td>tercebur,</td>
<td>Plunged</td>
<td></td>
</tr>
<tr>
<td>Verb Type</td>
<td>Action</td>
<td>Meaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[V, + passive, + accidental]</td>
<td>kecemplung</td>
<td>tercemplung</td>
<td>into water</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curi [V, + active]</td>
<td>nyuri, nyolong</td>
<td>mamaling</td>
<td>nyolong</td>
<td>mencuri</td>
<td>To steal</td>
<td></td>
</tr>
<tr>
<td>Curiga [V, + active]</td>
<td>nyurigain, nyangka</td>
<td>nyenidrine</td>
<td>nyujanani</td>
<td>mencurigai</td>
<td>To suspect</td>
<td></td>
</tr>
<tr>
<td>Coba [V, + active]</td>
<td>nyobaan</td>
<td>mintonin</td>
<td>nyoba</td>
<td>mencoba</td>
<td>To try</td>
<td></td>
</tr>
<tr>
<td>Daki [V, + active]</td>
<td>naik, naek, manjat</td>
<td>manekin</td>
<td>manjat</td>
<td>mendaki</td>
<td>To climb</td>
<td></td>
</tr>
<tr>
<td>Damping [V, + active, + transitive]</td>
<td>ngedampi-ningin</td>
<td>ngarendengan</td>
<td>nampin</td>
<td>njejeri</td>
<td>mendampingi</td>
<td>To accompany</td>
</tr>
<tr>
<td>Harus [V, + modal]</td>
<td>harus, kudu</td>
<td>kudu</td>
<td>apang, mesti</td>
<td>kudu</td>
<td>harus</td>
<td>Must</td>
</tr>
<tr>
<td>Hormat [V, + active, + transitive]</td>
<td>ngormatin</td>
<td>ngajenan</td>
<td>ngormatin</td>
<td>ngormati</td>
<td>menghormati</td>
<td>To respect</td>
</tr>
<tr>
<td>Jual [V, + active]</td>
<td>ngejual</td>
<td>ngajual</td>
<td>ngadep</td>
<td>ngadol</td>
<td>menjual</td>
<td>To sell</td>
</tr>
<tr>
<td>Jewer [V, + active]</td>
<td>ngejewer</td>
<td>ngajewer</td>
<td>ngobes</td>
<td>njewer</td>
<td>menjewer</td>
<td>To twist someone’s ear</td>
</tr>
<tr>
<td>Kenal [V, + active, + transitive]</td>
<td>kenalin</td>
<td>kenalkeun</td>
<td>kenalin</td>
<td>kenalke</td>
<td>kenalkan</td>
<td>To introduce</td>
</tr>
<tr>
<td>Kirim [V, + active]</td>
<td>ngirim</td>
<td>ngirim</td>
<td>ngirim</td>
<td>ngirim</td>
<td>mengirim</td>
<td>To send</td>
</tr>
<tr>
<td>Marah [V, + active]</td>
<td>marah, ngambek</td>
<td>ambek</td>
<td>pedih</td>
<td>duka</td>
<td>marah</td>
<td>angry</td>
</tr>
<tr>
<td>Mau [V, + modal]</td>
<td>mau, pengen</td>
<td>hayang</td>
<td>nyak, demen</td>
<td>gelem, arep</td>
<td>mau</td>
<td>Want</td>
</tr>
<tr>
<td>Langkah [V, + active, + transitive]</td>
<td>ngelangkah-in</td>
<td>ngalengkahin</td>
<td>nglangkahin</td>
<td>jumangkah</td>
<td>melangkahin</td>
<td>To step over</td>
</tr>
</tbody>
</table>
It is extrapolated in this section, on the basis of linguistic, demographic, historic and geographic considerations, that these three regional languages have had the biggest influence on CI/SI, many through the intermediary of Betawi Malay. The linguistic features of concern here are phonological, morphological, lexical and syntactical isomorphism.


The most substantial linguistic contribution towards Betawi Malay comes from the Sundanese language (Grijns 1991: 89). There are obvious reasons for this: geographically, the Betawi culture has been surrounded by the Sundanese culture. As mentioned earlier, Jakarta itself in ancient times was a Sundanese port of the Tarumanagara and Padjajaran Kingdoms. There had always been a

<table>
<thead>
<tr>
<th>+ transitive</th>
<th>Pakai</th>
<th>pake, make</th>
<th>make</th>
<th>nganggon</th>
<th>nganggo</th>
<th>memakai</th>
<th>To wear, to use</th>
</tr>
</thead>
<tbody>
<tr>
<td>[V, + active]</td>
<td>Paksa</td>
<td>maksa</td>
<td>maksa</td>
<td>maksa</td>
<td>meksa</td>
<td>memaksa</td>
<td>To force</td>
</tr>
<tr>
<td>[V, + active]</td>
<td>Pilih</td>
<td>milih</td>
<td>milih</td>
<td>memilihin</td>
<td>milih</td>
<td>memilih</td>
<td>To choose</td>
</tr>
<tr>
<td>[V, + active]</td>
<td>Pulang</td>
<td>pulang, balik</td>
<td>balik, mulang</td>
<td>mulih, mantuk</td>
<td>bali, mulih</td>
<td>pulang</td>
<td>To go home</td>
</tr>
<tr>
<td>[V, + intransitive]</td>
<td>Saja</td>
<td>saja, aja, melulu</td>
<td>wae, melulu</td>
<td>dogen, kewanten</td>
<td>wae, thok</td>
<td>saja</td>
<td>Only, just</td>
</tr>
<tr>
<td>[ADV, ADJ]</td>
<td>Sedia</td>
<td>nyediain</td>
<td>nyadiakeun</td>
<td>nyediang</td>
<td>nyawisake</td>
<td>menyediakan</td>
<td>To provide</td>
</tr>
<tr>
<td>[V, + active, + dative]</td>
<td>Susun</td>
<td>nyusun</td>
<td>nyusun</td>
<td>nyusun</td>
<td>nata</td>
<td>menyusun</td>
<td>To stack, to layout, to organise</td>
</tr>
<tr>
<td>[V, + active]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Discussion on lexical and linguistic derivatives of CI/SI*

It is extrapolated in this section, on the basis of linguistic, demographic, historic and geographic considerations, that these three regional languages have had the biggest influence on CI/SI, many through the intermediary of Betawi Malay. The linguistic features of concern here are phonological, morphological, lexical and syntactical isomorphism.


The most substantial linguistic contribution towards Betawi Malay comes from the Sundanese language (Grijns 1991: 89). There are obvious reasons for this: geographically, the Betawi culture has been surrounded by the Sundanese culture. As mentioned earlier, Jakarta itself in ancient times was a Sundanese port of the Tarumanagara and Padjajaran Kingdoms. There had always been a
demographic continuum in this area, and for that matter, the Betawi people are related to the Sundanese people. As Batavia’s role as a trading port expanded, it developed its own unique culture through cross marriages with the many different peoples that converged there.

The linguistic affinity between Sundanese and Malay-Indonesian might also be attributed to the Sriwijaya period. Most of the Sundanese territory was part of the Sriwijaya Empire while Central and East Java were not. Sundanese is closer to Indonesian than Javanese is. While the Malay-Indonesian language is a direct descendant of Sriwijayan Malay, Sundanese derived from an older Austronesian-Malay stem. The many common lexical and linguistic features between Sundanese and Indonesian could suggest that there was substantial linguistic input into Sundanese by Sriwijayan Malay.

At the lexical level, Javanese is the regional language that has had the biggest linguistic input on contemporary Indonesian language. This is primarily due to the ‘Javanisation’ of the modern Indonesian administration. This has resulted in a certain bias in the development and construction of standard Indonesian (Bahasa Indonesia yang baik dan benar – ‘good and proper Indonesian’).

**Lexis and linguistic features in review**

The [AUX, + TENSE] marker akan (akan\_FI, - mau\_CI\_MSI – bade\_SUND\_JAV) is often not used in that function outside the domain of Bahasa Baku. What is interesting here is that in CI/SI akan is substituted with mau, which in standard FI grammar is the auxiliary [V, + Modal] ‘to want’. The semantic property of mau in CI/SI is to express both the notion of ‘want’, but also as a [AUX, + TENSE] future marker. In this second role it is simply a substitute of akan and might or might not imply a subjective ‘wanting’. For example, a phrase such as ‘kita lagi mau jalan’ (we want to go), could indicate that collectively the agentive first personal pronoun might subjectively ‘want to go’, or simply ‘are about to go’. Thus semantically the phrase could mean ‘we are about to go’. Akan seldom occurs in CI/SI and the function of mau as [AUX, + TENSE] occurs just as often as [V, + Modal] that semantically conveys a subjective ‘wanting’. ‘Mau hujan’ (it will rain) is another such example of mau as future tense auxiliary. The CI/SI [AUX, + TENSE] function of mau correlates with the Javanese and Sundanese lexical item bade, which performs the identical function as the CI/SI mau, both as [V, + Modal] and [AUX, + (future)TENSE] (Mangunsuwito 2002: 21) (Tamsyah, Purmasih, Purmawati & Supratman 1998: 32). The Betawi-Malay pengen and kepengen (Gil 2001: 8), which are allomorphological derivatives from the base word ingin, also function as [V, + Modal] and [AUX, + TENSE]. Both of these Betawi-Malay lexical items are also operational in CI/SI as an alternative to mau. While there is no readily available philological evidence that the Betawi-Malay pengen is a semantic derivative from the Javanese and Sundanese bade, linguistic (and biological) genetic link between Javanese, Sundanese and Betawi-Malay would strongly suggest semantic antecedence of pengen in the regional languages of Java. The modern usage of mau would then just be a natural switching between the
isomorphic alternative lexical items (pengen ↔ mau_{MODAL\_VERB}) by contemporary speakers of Indonesian.

For the majority of Indonesian speakers who live on Java Indonesian is a second language. The semantic and functional isomorphism between bade, mau (and pengen) would also be just another natural ‘switch’ when they speak Indonesian and hence they have possibly reinforced the [AUX, + TENSE] function of the lexical item mau. The Betawi-Malay lexical items bakal/bakalan that function as future [AUX, + TENSE] has also been absorbed into CI/SI and is in common usage.

One of the synonyms of mau [V, + modal] in Balinese is (demen_{JAV/BAL}). In CI/SI and Betawi-Malay, demen is a synonym of the [V, + modal] suka. It is very likely that this is another lexical transfer that has gone through a semantic shift from mau (to want) to suka (to like). Demen is not part of the corpus of FI.

The CI/SI [V, + active, + causative] inflection of bingung possesses the prefix characteristic of Sundanese and the suffix characteristic of Balinese: (membungin{FI} – ngebungin{CI,MSI} – ngalieurkeun{SUN} – mingungin{BAL} – bingungake{JAV}). This is Sundanese-CI-SI-Balinese [V, + active] and [V, + active, +benefactive/ + causative/ + dative] relationship is a recurring pattern and is discussed in the following paragraphs. On a side note, in CI/SI the collocation ‘bikin bingung’ is also frequently used instead of ngebungin.

The [V, + active] form of the base word bantu (membantu – ngebantu/ nolongin_{CI,MSI} – ngabantu_{SUN} – nulungin_{BAL}) and its alternative lexical item tolong, show morphological influence from Sundanese and Balinese. The CI/SI [ŋə-] and [___-in] markings resemble the Sundanese and Balinese equivalents respectively. Isolated incidences of linguistic isomorphism can be treated as circumstantial, but persistent likeness and consistency in pattern transfer would then suggest linguistic linkages. In the case of CI/MSI [ŋə-], it shows consistent pattern distribution and correspondence in Javanese, Sundanese and Balinese. In Balinese, Sundanese and Javanese, base word lexical items that begin with vowels and the consonant ‘k’ (such as ajar, antar, asah, kirim) have the [V, +active] morphological derivative /ŋ-/ (ng-). The CI/SI phonological prefix [ŋə-] (nge-) that precedes [V, +active] lexical items that start with consonants (except ‘k’), the comparative table suggests strong correlation with the equivalent Sundanese [V, +active] [ŋa-] (nga-). In Prawirasumantri (1979: 23-24) it is listed as a Sundanese [V, +active] morphological prefix. The correlation in this morphological pattern between CI/SI and Sundanese is consistent, acknowledging the [ŋə-] to [ŋa-] allomorphy between the CI/SI and the Sundanese counterpart. The CI/SI [ŋə-] [ŋə-] phenomena may appear as an aphaeresis in the Jakartan speech habit of the FI form [məŋ-], but the strong presence of [ŋə-] in the surrounding regional languages as a [V, +active] marking in its own right suggests that the
SI phenomenon has been acquired through the Sundanese-Javanese-Balinese via the intermediary of Betawi-Malay. Some other lexical items from the comparative table that show pattern consistency include: damping [V, + active, + transitive] (mendampingi\textsubscript{FI} – ngedampingin\textsubscript{CL/MSI} – ngarendengan\textsubscript{SUN} – nampingin\textsubscript{BAL} – njejeri\textsubscript{JAV}); langkah [V, + active, + transitive] (melangkah\textsubscript{FI} – ngelangkahin\textsubscript{CL/MSI} – ngalengkahan\textsubscript{SUN} – nglangkahin\textsubscript{BAL} – jumangkah\textsubscript{JAV}) and hormat [V, + active, + transitive] (menghormati\textsubscript{FI} – ngormatin\textsubscript{CL/MSI} – ngajenan\textsubscript{SUN} – ngormatin\textsubscript{BAL} – ngormati\textsubscript{JAV}).

The pattern of the CI allomorphy /m-/ in base words starting with the consonant ‘p’ also shares with Sundanese, Balinese and Javanese: i.e.; pakai\textsubscript{Xn} [V, + active] – memakai\textsubscript{FI, make} – maksa\textsubscript{CL/MSI} – maksa\textsubscript{SUN} – maksa\textsubscript{BAL} – meksa\textsubscript{JAV}. Also: pilih\textsubscript{Xn} [V, + active] – memilih\textsubscript{FI} – milih\textsubscript{CL/MSI} – milih\textsubscript{SUN} – milih\textsubscript{BAL} – milih\textsubscript{JAV}. The phonological form of pakai is identical between CI/SI and the Sundanese: [mə].

The A Comprehensive Indonesian-English Dictionary\textsuperscript{iii} gives the definition of the base word cebur as: (onom) o of s.t. being plunged or thrown into water. Its SI [V, + active] inflection nyeburin is identical to the Balinese form. The elision and allomorphy of the phoneme /c/ into: [n-]; [____ [X\textsuperscript{n}]] [V, +active]\textsuperscript{iv} is common in Sundanese and Balinese. In Prawirasumantri (1979: 27) on Sundanese Predicate Morphology, the phoneme /p/ (in BI written as ‘ny-’) and its allophone [n] are active inflections on predicate base words that start with the phonemes /c/ and /s/. In MSI this corresponds with the allophone [n] as allomorphy on lexical items [V, + active] that start with the consonants ‘c’ and ‘s’\textsuperscript{v}. The same allomorphy is consistent in other lexical items that come under this parametric description (i.e., coba, curiga, susun, sedia).

It is possible that the strong presence of Balinese towns (kampong Bali)\textsuperscript{vi} have left their linguistic marking on Betawi Malay in the form of the suffix –in in the structures: [n-____-in], [ŋ-____-in] and [ŋə-____-in]; [____ [X\textsuperscript{n}]] [V, +transitive]\textsuperscript{vi} (Sneddon 2003: 154). The suffix –in appears in Balinese as a [V, +transitive] inflection, and more common in morphological markings that corresponds with the FI structure: [mə-____-i]; [____ [X\textsuperscript{n}]] [V, +transitive] than the FI structure of: [mə-____-kan]; [ ____ [X\textsuperscript{n}]] [V, +active, +dative]. In SI the –in suffix operates on both the [V, +transitive] and [V, +active, +dative] structures mentioned above. Some comparative examples from the table: membelikan\textsubscript{FI} – ngebeliin\textsubscript{CL/MSI} – meanin\textsubscript{BAL}; mencoba\textsubscript{FI} – nyobain\textsubscript{CL/MSI} – mintonin\textsubscript{BAL}. 

---

\textsuperscript{iii} The A Comprehensive Indonesian-English Dictionary (Denford et al. 1996) is a valuable resource for understanding the vocabulary of Indonesian.

\textsuperscript{iv} This example highlights the use of the phoneme /c/ in Indonesian and its elision into /n/ in Balinese.

\textsuperscript{v} The phoneme /c/ and /s/ in Indonesian often undergo allomorphic changes, and the Balinese form shows a characteristic /-in/ suffix.

\textsuperscript{vi} Balinese towns have left significant linguistic influence on Betawi Malay, particularly through the use of the –in suffix in Balinese and its adoption in Betawi Malay.
In CI/SI the –in suffix also operates on passive di- inflections as well as on base words: kenal [V, + active, + transitive] (kenalkanFI – kenalin_CI/SI – kenalkeun_SUN – kenalin_BAL – kenalkeu_JAV). The –in suffix is not known in Javanese as a predicate marking, and in Prawirasumantri (1979) it is also not listed as a Sundanese predicate morphology. Again, extrapolating on the process of elimination, taking into consideration historical demographics and identical functional linguistic feature, it leaves compelling evidence of the Balinese origins of the –in suffix in Betawi Malay and hence SI.

It is interesting to note that phonologically, the –keun suffix of nyeburkeun, the Sundanese form of the [V, +active, +dative] of the base word cebur is closer to the FI form menceburkan. The common origins of the –kan / -keun suffix can be traced to Classical Malay or perhaps even to an older genetic stem.

The SI forms kecebur / kecemplung [V, + passive, +accidental] of the base word cebur show morphological affinity with the Sundanese and Javanese equivalent kacemplung and kejegur respectively, while at the lexical level the alternative base word cemplung is the same as with the Sundanese kacemplung and the Balinese macemplung. The variation to the FI prefix ter-: [_____ X^a] [V, +passive] in MSI is ke- ([kə]), which corresponds to the Javanese prefix ke- and Sundanese ka-.

The ke- prefix [kə], is characteristic of Betawi-Malay (Grijns 1991: 127) and in Prawirasumantri (1979: 135) it is listed as the Sundanese [V, +passive] morphology. The oldest known form of this morpheme is ka- in Sriwijayan Malay serving the same inflectional function. It is therefore extrapolated here from identical linguistic form and function, and geographic and historic proximity, that the Javanese prefix ke- and Sundanese ka- [V, +passive] markings are precursors to the Betawi-Malay and hence the SI phone [kə].

Onomatopoeia is commonly an injection from the regional languages too. The A Comprehensive Indonesian-English Dictionary denotes the word cemplung as being of Javanese origins and gives the definition of: (Jv onom) 1. Plop! the sound of s.t. thrown or falling into water. The Javanese entry in Kamus Indonesia-Daerah is a different base word: jegur. The entry cemplung is absent in The Learner’s Dictionary of Today’s Indonesian, and Tuttle’s Concise Indonesian Dictionary gives the definition: to plunge into water. A reverse search, looking up ‘plunge (into water)’ in Gramedia’s Kamus Inggris-Indonesia (The English-Indonesian Dictionary) gives a totally different base word celup. A reverse search, looking up ‘kejegur’ under the Javanese base word jegur in Mangunsuwito’s (2002: 84) Kamus Bahasa Jawa-Indonesia (The Javanese-Indonesian Dictionary) is consistent with Sugiato’s Kamus Indonesia-Daerah.
The base word *daki* and its [V, + active] marking has interesting alternative lexical forms in CI/SI, showing influences from both Javanese and Sundanese, contrasting with the FI form *mendaki*. A FI phrase ‘*saya sedang mendaki gunung*’ would not adhere with sociolinguistic competence in CI/SI domains. ‘*Saya lagi naik gunung* ’ are more appropriate in CI/SI social domains. The choice between *manjat* and *naik* would depend on the steepness of the climb. *Manjat* would emphasise a steep climb while *naik* would indicate a progressive-tempered (walking, or driving) climb, or the general idea of going up a mountain. In speech, *[naik]* also often takes a schwa allomorphy in the form of *[naə]*, at times with a silent ‘k’ or a glottal stop [ʔ]. The comparative table points to *manjat* being of Javanese origins and *naek* coming from Sundanese. The comparative table also correlation in the schwa allomorphy that is common in SI with the equivalent lexical item in the regional languages (i.e.; antar_{FI} - anter_{CI/MSI} [antar] – anteur_{SUN} [antar]).

A synonym of *harus* [V, + modal] or *mesti* [V, + modal] in CI/SI and Betawi-Malay is *kudu*. This is a direct injection from Sundanese and Javanese: harus_{FI} – harus/kudu_{CI/MSI} – kudu_{SUN} – mesti_{BAL} – kudu_{JAV}, while the Balinese *mesti* corresponds with the Indonesian base word. Other showings of Sundanese injections are in the lexical items *pulang* [V, +intransitive]: pulang_{FI} – pulang/bali_{CI/MSI} – balik_{SUN} – mulih_{BAL} – mulih_{JAV} and *selalu* [ADV, ADJ]: selalu_{FI} – melulu_{CI/MSI} – melulu_{SUN} – dogen_{BAL} - wae_{JAV}. In CI/SI *balik* is a familiar alternative lexical item for *pulang*, commonly used in phases like ’*mau balik dulu*’ or ‘*kapan baliknya?*’. The lexical item *saja* can semantically be applied in the context of ‘just/only’ and also in the context of ‘always’. In the latter context the alternative *melulu* is common in CI/SI such as in a phrase ‘*nyanyi melulu bisanya*’. The table shows it to be a direct loan word from Sundanese.

The typical first person personal pronoun *gue* of Betawi-Malay and Jakartan-CI has no traces in the regional languages. At this point it is deemed as an original Betawi-Malay composition. The corresponding FI *aku* has equivalence in Javanese as *aku* and *dalem*. The Javanese *dalem* emphasises the ‘inner-ness’ of the self while the Sundanese *abdi* in Indonesian has semantically evolved to mean ‘to devote’. Thus, *mengabdikan* is to make something “one’s own”; making it a part of oneself → to devote.

Lexical item transfer and semantic evolution are aptly demonstrated by the base word *marah*: marah_{FI} – marah/ngambek_{CI/MSI} – ambek_{SUN} – pedih_{BAL} – duka/lesu_{JAV}. The Betawi-Malay and CI-SI *ngambek* is a direct input from Sundanese. In Balinese the ‘stinging sensation’ of being upset is expressed in *pedih*. The notion of being upset is expressed as *duka* in Javanese which in Indonesian is understood more in the context of ‘grief’. Perhaps this reflects the Javanese cultural value that anger is not an emotion that one should entertain, thus the only ‘upset’ that is to be exercised is to grieve.
Many phrase constructions and speech habits in SI/CI have isomorphic correlations in the regional languages. For instance, the use of *apa* as the conjunction *atau* (‘or’) which resembles the function of *napi* in Sundanese and Balinese. Another example is the common usage of *dulu* (‘first, now’) in everyday speech acts phrases such as *pulang dulu* (‘I am going now’) or *coba dulu* (‘try it first’) has counterparts in the regional languages (*mulih dumu*<sub>BAL</sub>, *mintonin dumu*<sub>BAL</sub>). The common usage of *saja* (‘just, only’) in everyday phrases also has equivalents in the regional languages: *biarkan saja*<sub>FI</sub> – *biarin aja*<sub>CI/MSI</sub> – *atepkeun bae*<sub>SUN</sub> – *depin dogen*<sub>BAL</sub> – *jarna wae*<sub>JAV</sub> (‘let it be’); *di mana saja*<sub>FI</sub> – *di mana aja*<sub>CI/MSI</sub> – *di mana bae*<sub>SUN</sub> – *dija dogen*<sub>BAL</sub> – *ing ngendi wae*<sub>JAV</sub>. Another unique phrase construction in SI/CI is the use of *orang* (person/people) as an interjection of emphasis: ∅<sub>FI</sub> - *orang* *udah saya bilangin*<sub>MSICT</sub> - *tiang anak suba bilang*<sub>BAL</sub> (‘(person) I have told him’). Also interesting is the isomorphic usage of *baru* (‘new’) as an interjection in speech acts and expressions between CI/SI and Balinese. The standard word class of *baru* in BI is an adjective, but in CI/SI it is also used as an approval-interjection particle in a phrases such as: *itu baru, bagus*<sub>CI/MSI</sub> – *keto, mara, luwong*<sub>BAL</sub> (‘that is new, good’).

Summary

The discussions above and the analysis of the comparative table between CI-SI and the languages of Javanese, Sundanese and Balinese have demonstrated consistent lexical, phonological and morphological isomorphism between CI-SI and Javanese-Sundanese-Balinese. It is purported by this research that the following linguistic features have been transferred/acquired by CI-SI, mostly via Betawi-Malay, from Javanese, Sundanese and Balinese:

1. The acquisition of the unstressed schwa phone [ə] on the second syllable.
2. The [V, +active] morphological derivative /ŋ- (⟨[ŋ]⟩).
3. The [V, +active] morphological derivative /ŋ- (⟨[ŋ]⟩) and /ŋə- (⟨[ŋə]⟩).
4. The suffix –in.
5. Javanese prefix *ke-* and Sundanese *ka-* [V, +passive] [kə].
6. The aphaeresis and allomorphy /m-/ in base words starting with the consonant ‘p’.
7. Alternative lexical items and synonyms that are not part of the FI corpus, such as: *pulang*<sub>FI</sub> – *pulang/balik*<sub>CI/MSI</sub> – *balik*<sub>SUN</sub>; *saja* [ADV, ADJ]: *saja*<sub>FI</sub> – *aja/melulu*<sub>CI/MSI</sub> – *melulu*<sub>SUN</sub>.
8. Parametric extension of the function of lexical items such as with the case of *mau* from [V, +Modal] to → [AUX, + TENSE].
9. Usage of certain lexis as interjection particles outside the standard word class in the transfers of expressions and speech acts from the regional languages into CI/SI: *itu baru bagus* /CI/MSI – *keto mara luwong* /BAL.
Appendix 6

The Use of CI in News Reporting
**CI in news reporting**

There are signs that some Indonesians can no longer distinguish between standard FI lexicon and CI lexicon. Traditionally, bahasa Baku, or FI, the standard Indonesian language, was the language exclusively used in News presentation, be it in the print or Audio-Visual media. Today, some CI lexicon can be found even in news reporting. There is increase blurring of what is standard FI and CI by native speakers, many CI items have been ‘promoted’ and deemed acceptable to operate in the H sociolinguistic domain. Many of the younger generation - who have grown up in post New Order and hence have not had FI relentlessly inculcated in them through education and the all-controlling state media - might not be aware anymore of what is traditionally standard FI and that CI items that are in common everyday usage are actually not standard FI. Many CI lexical items – and linguistic features for that matter - are such an integral part of everyday Indonesian, that they are simply considered as ‘Indonesian’. While historically CI exclusively pertained to Betawi Malay - and subsequently Jakartan Indonesian, which included trendy speech and subculture idioms such as bahasa Gaul - an array of CI items and linguistic features have ‘transcended’ the ‘lower’ L sociolinguistic domains and are today in standard every day usage. By virtue of common usage, it seems that many CI items may be becoming standardised.

The following news excerpts from the internet show some examples of usage of what are traditionally CI lexical items. The lexical items below are not listed in older Indonesian dictionaries as they were not considered as standard and ‘proper’ Indonesian. Kramer, A., Koen, W. (1993) Tuttle’s Concise Indonesian Dictionary, which was first published 1966, follows a traditional perspective of standard Indonesian and does not list these CI items. On the other hand, Stevens and Schmidgall-Tellings (2004) A Comprehensive Indonesian-English Dictionary, includes these CI items. The more modern approach of the latter dictionary, which is written primarily for foreign students of Indonesian language, has taken a practical stance and included all lexical items that are part of common speech. Nonetheless, these CI lexical items and linguistic features are still not listed in the Indonesian national official dictionary (Kamus Besar Bahasa Indonesia).

Source: www.tribunnews.com

- Air Asia Ditolak Naik Ketinggian Gara-gara Ada Garuda (28/12/2014)
- Menhub Bakal Kembali Mengkaji Operasi dan Bisnis Ais Asia (28/12/2014)
- Bingung Tahun Baruan Ke Mana? Nonton Nanial Idol Aja, Yuk! (28/12/2014)
- Ringkasan Pertandingan Liga Inggris, Minggu: Trio Teratas Cuma Imbang (28/12/2014)
- “Biasanya masang foto dirinya atau sama pacarnya” (29 Dec 2014)
Menteri Perhubungan *Nyatakan* Posisi Pesawat Air Asia Tak Jauh Dari Garis Pantai (29 Dec 2014)

Source: www.nasional.sindonews.com

“*Ketimbang* Datangi ISIS, *Mending* ke Daerah Bencana” (27 Dec 2014)

Source: www.liputan6.com/

- *3 Cara Keren* Masak Telur di Microwave (2 Jan 2015)
- *Top 5 Lifestyle: Perayaan Tahun Baru Teraneh Bikin Penasaran* (2 Jan 2015)

Source: www.antaranews.com

- Jokowi: “*Ngapain* kita berutang” (9 Jan 2013)
- Jokowi: “*Saya capek ngurusin* prosedur dan administrasi MRT” (12 Apr 2013)
- Ramadhan Pohan: Front Pembela Islam “*ngawur*” (23 Jul 2013)
- Jokowi: “*baru lima menit blusukan, Zuckerberg capek*” (13 Oct 2014)
- Dari pada *buat* korupsi, *mending* untuk pesta rakyat (19 Oct 2014)
- Seluruh Jakarta bakal diguyur hujan hari ini (5 Jan 2015)
- Sebuah Airbus A320 *nyaris tabrakan* dengan “drone” (4 Maret 2016)
- Sabine Atlaoui: “Saya *capek, enggak bisa ngomong* banyak...” (12 Mar 2016)
- Sultan imbau masyarakat Yogyakarta *ngalah* tidak masuk Malioboro (9 Jul 2015)
- Wakil Ketua MPR nilai LGBT contoh kebebasan yang *kebablasan* (21 Feb 2016)
- Diskon *gede-gede-an* kaca film mobil di IMMS (12 Apr 2016)
- *Cuma* tersisa 100 harimau sumatera di TN Gunung Leuser (25 May 2016)
- Pimpinan MPR minta Presiden tak *cuma* blusukan ke pasar (8 Jun 2016)
- Polisi Los Angeles pilih BMW i3 *ketimbang* Tesla Model S (9 Jun 2016)
- *Gara-gara* selfie, siswi SMK jatuh dari jembatan (10 Jun 2016)
- KPU *nyatakan* hasil verifikasi dukungan KTP sebaiknya terbuka (11 Jun 2016)
- Yusril *ogah* komentar Ahok: *Bakal* calon Gubernur... (22 Jun 2016)
- Polisi Kediri tangani peluru *nyasar* (11 Aug 2016)
- Mensos luncurkan mobil *antigalau* di Surabaya (14 Aug 2016)
- Ayo ikut *nonton bareng* pertandingan Tontowi/Liliyana (16 Aug 2016)
- Menpora *galau* jika Malaysia sponsori Rio Haryanto (24 Feb 2016)
The presence of CI lexical items in H diglossia domain raises many pressing questions. Does it mean that CI lexical items and CIAλ by virtue of common use are becoming accepted into the standard FI vocabulary list? These findings are significant more so because the news site that many of these examples were obtained from is the Indonesian government official news website Antara News and government institutions are traditionally the bastion of FI language use and are agents for promoting ‘good and proper’ Indonesian language.

Since FI is not a natural spoken language, it takes a conscious effort for an Indonesian to speak in pure ΔFI and it is very possible that to unwittingly ‘slip’ into using CI is mere honest human mistake. One the other hand perhaps these findings may suggest that the post New Order generation of Indonesians are losing their ability to adequately distinguish the differences between ΔFI and ΔCI.

Lastly, there is the Javanese factor. The Javanese are the dominant ethnic group in the institutions of the Indonesian state and make up the majority of speakers of the Indonesian language. Many of these CI lexical items found in the news reports are originally Javanese words (absorbed into CI directly from Javanese or indirectly via Betawi Malay). It is very likely that CI lexical items that have origins in Javanese are not ‘filtered out’ by speakers whose first language is Javanese. This is especially evident in some of the quotes some uttered by the high ranking officials and politicians including the president himself.
Appendix 7

Research Instrument: Survey
Angket Penggunaan Bahasa Indonesia Sehari-hari

Angket ini adalah pengumpulan data untuk riset S3 saya (Halim Nataprawira) tentang penggunaan bahasa Indonesia sehari-hari. Tujuan Angket ini adalah untuk mengukur/menilai penggunaan bahasa Indonesia sehari-hari di kota-kota besar Indonesia. Penggunaan bahasa Indonesia sehari-hari yang dimaksudkan adalah kosa-kata, pengistilahan dan pembicaraan informal jika Anda:

- Berbicara dengan keluarga (adik-kakak, orang tua) di rumah
- Berbelanja di toko/pasar atau membeli makanan di warung/restoran
- Bercengkrama dengan teman-teman sekolah/kuliah atau rekan-rekan kerja
- Naik kendaraan umum
- Berbicara dengan sanak-saudara yang sepehatan

Kontak
Tim riset ini terdiri dari peneliti utama Halim Nataprawira (mahasiswa PhD), Dr Richard Curtis (pengamat utama) and Dr Michael Carey (pengamat kedua):

Halim Nataprawira  Email: hnatapra@usc.edu.au  Phone: 07 3870 3843
Dr Richard Curtis  Email: rcurtis@usc.edu.au  Phone: 07 5430 1254

Pengalaman Pengikut Serta

Dampak dan keuntungan
Tidak ada risiko atau dampak negatif dari mengisi angket ini. Tidak ada pula upah ataupun keuntungan pribadi dari keikutsertaan di angket ini. Kami sangat berterima kasih jika Anda mengisi angket ini karena dengan demikian Anda membantu kami dalam penelitian kami.

Keluhan dan kekhawatiran

Jika ada keluhan tentang pelaksanaan angket ini, Anda dapat menghubungi peneliti utama atau jika Anda ingin menghubungi pihak di luar regu peneliti. Anda dapat menghubungi ketua dari Human Research Ethics Committee at the University: (c/- the Research Ethics Officer, Office of Research, University of the Sunshine Coast, Maroochydore DC 4558; telephone (07) 5459 4574; email humanethics@usc.edu.au).

Jenis kelamin: ________________ umur Anda: ________________
Tingkat pendidikan: ________________ Pekerjaan: ________________
Tempat tinggal Anda (nama kota saja): ________________

Tanggal: ________________

Harap dilingkari penggunaan bahasa yang paling lumrah Anda pakai dalam situasi-situasi berikut ini
1. Jika Anda meminta sesuatu dari teman atau keluarga, Anda bilang:
a. minta dong   b. Berikanlah kepada saya   c. kasih saya aja

2. Jika Anda bertanya kepada teman/rekan apa yang sedang mereka lakukan:
a. Sedang apa? b. lagi ngapain?
c. lagi apa?   d. Apakah yang sedang kamu lakukan?

3. Jika Anda ingin mengundang teman/rekan berjalan-jalan keluar:
a. jalan-jalan yuk   b. Marilah kita berjalan-jalan
c. Apakah kamu ingin pergi keluar?   d. mau jalan-jalan ngak?

4. Jika Anda menanyakan suatu harga kepada penjual:
a. Berapakah harganya barang ini?   b. yang ini berapa Mas harganya?

5. Jika Anda tawar-menawar dengan pedagang:
a. Bisakah Anda mengurangi harganya?   b. boleh kurang nga Mas?
c. kasih diskon dong Mba...   d. Kurangkanlah harganya

6. Jika Anda ingin menghentikan suatu kegiatan:
a. udahan ah   b. Marilah kita berhenti   c. bubar-bubar   d. Ayolah berhenti

7. Jika Anda ingin pulang:
a. Saya akan pulang sekarang   b. mau pulang dulu nih
c. cabut ah   d. Permisi dulu ya

8. Kalau Anda memperkirakan sesuatu:
a. kayaknya sih...   b. sepertinya...

9. Kalau Anda tidak mau melakukan sesuatu:
a. Tidak mau   b. ngak mau   c. ogah   d. Tidak ingin

10. Kalau Anda ingin ikut pergi:
a. Bolehkah saya ikut serta?   b. ikut dong
c. Saya juga ingin ikut pergi            d. ikutan

11. Kalau Anda sedang mencari sesuatu:
   a. ilang ke mana nih barang?           b. Di manakah barang ini?
   c. Di mana tersembunyinya barang ini?  d. ngumpet di mana ini?

12. Kalau Anda sedang merasa muram:
    a. Saya sedang merasa muram           b. lagi bete nihc. suntuk banget...!
    d. Saya sedang tidak merasa gembira sekarang

13. Kalau Anda tidak setuju dengan sesuatu:
    a. ngak usah                        b. Tidak perlu
    c. Janganlah                       d. jangan digituin

14. Jika Anda memprotes suatu tindakan:
    a. Pake digituin segala              b. Kenapa harus dilakukan demikian?

15. Kalau Anda mempertanyakan suatu tujuan:
    a. Untuk apa pergi ke sana?         b. ngapain ke sana?

16. Jika Anda menanyakan tujuannya seseorang:
    a. mau ke mana?                    b. Sedang pergi ke manakah kamu?

17. Jika Anda menyarankan sesuatu:
    a. Lebih baik begini dari pada...     b. mendingan begini ketimbang...
    c. begini aja...                    d. Lakukanlah seperti ini...

18. Jika Anda menginginkan teman Anda datang:
    a. nanti dateng ya...              b. Datanglah kamu ke sini

19. Dengan keluarga Anda, Anda lebih memakai:
    a. bahasa Indonesia             b. bahasa daerah
    c. bahasa Indonesia dan bahasa daerah

20. Dengan teman-teman Anda, Anda lebih memakai:
21. Di tempat kerja Anda, Anda lebih memakai:
   a. bahasa Indonesia  
   b. bahasa daerah  
   c. bahasa Indonesia dan bahasa daerah  

22. Anda telah menjawab pertanyaan-pertanyaan di atas ini sesuai dengan:
   a. Yang menurut Anda adalah bahasa Indonesia yang baik dan benar  
   b. Bagaimana selayaknya orang Indonesia patut berbicara  
   c. Kenyataannya bagaimana orang Indonesia berbicara sehari-hari  

23. Apakah tanggapan Anda tentang pertanyaan-pertanyaan angket ini?
   a. lucu  
   b. menarik  
   c. konyol  
   d. aneh  

Tolong berikan alasan  

___________________________________________________________________________  
___________________________________________________________________________  

24. Dari skala 1 – 7, tolong berikan indikasi (dengan tanda ✔) apakah Anda lebih sering memakai bahasa Baku dalam interaksi sehari-hari atau bahasa sehari-hari.  

<table>
<thead>
<tr>
<th>Bahasa Baku</th>
<th>bahasa sehari-hari</th>
</tr>
</thead>
<tbody>
<tr>
<td>1------------</td>
<td>--</td>
</tr>
<tr>
<td>2------------</td>
<td></td>
</tr>
<tr>
<td>3------------</td>
<td></td>
</tr>
<tr>
<td>4------------</td>
<td></td>
</tr>
<tr>
<td>5------------</td>
<td></td>
</tr>
<tr>
<td>6------------</td>
<td></td>
</tr>
<tr>
<td>7------------</td>
<td></td>
</tr>
</tbody>
</table>

Terima Kasih  

_Ethics Application_: Halim Nataprawira, confirmed PhD student (1009261).  

_Research Instrument_: Survey questionnaire to study the usage of Indonesian language in the context of sociolinguistic diglossic theory.  

_Synopsis_: This survey collects information about people’s choice of language usage in everyday dialogue. The theoretical context is the sociolinguistic theory of diglossia, where
two variations of the same language (High and Low) exist. What differentiates diglossia from
register is that grammar, lexis and linguistic features (phonology, morphology and syntax)
changes between the High (H) and Low (L) variants. It is like having two sets of grammar for
the same language. As a conceptual analogy, in effect, the difference between H and L can be
likened to the difference between Swedish and Danish; Sicilian and Italian; Catalan and
Spanish, Standard German and Austrian German.

**Preface:**
The following translation, especially the multiple choice options of the questions, are
semantic, emphatic and impressionistic translation of the Indonesian. The H and L diglossic
difference can only be best captured in the context of register in English. I have translated H
Indonesian into an artificial formal-standard English register and L Indonesian into an
informal/colloquial English register.

In all reality, literal translation between English to Indonesian rarely works, let alone
attempting to translate the linguistic aspect of the diglossic variation, where there is a change
in grammar and linguistic aspects between H and L in Indonesian. As a demonstration, I shall
hypothetically create a diglossic linguistic situation in English. If we were to take a root verb
such as happy [+adjective], if English was diglossic, then the L phonetic variation of the
word happy might be hoopy. The morphological suffix in standard (H) English to
pronominalise happy is –ness = happiness [+Noun] (where it also undergoes an allomorphy
of ‘y’ → ‘i’). Let’s say the L pronominal suffix might be: -ba, so hoopyba = happiness. To
take it a step further, the negating prefix might change from un- to no-, thus nohoopyba = unhappiness.

Now if these phonological and morphological variations were to apply to just 40% of the
vocabulary of the English language (not to mention lexis that are totally different and
different syntax at work) when English is informally spoken in everyday context, then L
English diglossia would be quite a different language from standard English.
Questionnaire on the usage of everyday Indonesian

This questionnaire collects is a data collection instrument for my (Halim Nataprawira) PhD on the usage of Indonesian in everyday interactions.

The purpose of this questionnaire is to gauge the usage of everyday Indonesian in the main cities in Indonesia.

The questions in this survey differentiate between formal Indonesian and everyday Indonesian. Please answer the questions as honest and accurate as possible so that this research reflects the reality of Indonesian language use in everyday life.

This survey is voluntary. You are not obliged to answer. You do not need to continue with the survey if you feel hesitant about any of the questions.

The usage of everyday Indonesian that is targeted is vocabulary, expressions and informal speech that you would use when:

- Taking to family members (siblings, parents) at home
- Going shopping or eating out
- Socialising with school/uni friends or work colleagues
- Taking public transport
- Talking to relatives of the same age

Your gender: ________________ Your age: _______________
Education level: ________________ Occupation: ________________
Your home city: ________________
Date: ______________

Please circle the language that you would most likely use in the following situations:
1. If you want to ask something from a friend or family, you would say:
   a. Can I have some  
   b. Would you give it to me  
   c. just give me some

2. If you are asking a friend/colleague what they are doing:
   a. What are you doing?  
   b. what’s going on?  
   c. What are you up to?  
   d. What is it that you are doing right now?

3. If you wanted to invite friends to go out:
   a. Let’s go out  
   b. Why don’t we go out?  
   c. Would you like to go out?  
   d. Feel like going out?

4. If you wanted to ask a price to a seller:
   a. How much would this cost?  
   b. How much is this one?

5. When you are bargaining with a seller:
   a. Could you please reduce the price?  
   b. Can it be less?  
   c. You can do better  
   d. Please make it less

6. If you want to terminate an activity:
   a. Let’s quit  
   b. Let’s stop  
   c. break up  
   d. Let us stop now

7. If you would like to go home:
   a. I shall go home now  
   b. I’m off  
   c. I’m out of here  
   d. Please excuse me

8. If you are estimating something:
a. It’s kinda... b. It seems like...

9. If you don’t want to do something:
   a. I do not want to b. don’t wanna c. no way d. I’m not inclined

10. If you would like to come along:
   a. May I participate? b. can I please come
c. I would also like to go d. I’m coming

11. If you are looking for something:
   a. Where has it disappeared to? b. Where would it be?
c. Where would it be hiding? d. where is it hiding?

12. When you are in a bad mood:
   a. I’m in a bad mood b. feeling rotten c. I’m in a foul mood
d. I am not feeling very happy right now

13. If you don’t agree about something:
   a. nah b. no need c. do not d. don’t do it

14. If you are protesting something:
   a. What’s with that? b. Why would it have to be done as such?

15. If you are questioning a destination:
   a. Why would you want to go there for? b. why go there?

16. If you are asking someone where they are going:
17. When you are advising:

a. It would be better to...
   b. better off to...
   c. do it this way...
   d. Perform it as such...

18. If you would like to invite a friend:

a. Come around ok...
   b. Would you come later...

19. With your family, you would use:

a. Indonesian  
   b. regional dialect/language  
   c. Indonesian and regional dialect/language

20. With your friends, you would use:

a. Indonesian  
   b. regional dialect/language  
   c. Indonesian and regional dialect/language

21. In your workplace, you would use:

a. Indonesian  
   b. regional dialect/language  
   c. Indonesian and regional dialect/language

22. You have answered these questions according to:

   a. What you deem is correct Indonesian
   b. How Indonesians should be speaking
   c. The reality of how Indonesian is spoken

23. What is your opinion about this survey?
a. funny  b. interesting  c. ridiculous  d. strange

Please give reason

___________________________________________________________________________
___________________________________________________________________________

24. In a scale from 1 – 7, please indicate (with ✔) if you would use more formal Indonesian or informal Indonesian in your daily interaction.

<table>
<thead>
<tr>
<th>Formal Indonesian</th>
<th>Informal Indonesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Thank you
The alternative ‘kenapa’ is more commonly used than ‘mengapa’.

Historically, one of such famous Balinese kampong was located across on the southern end from where Monas in today.


Echols, John M. & Shadily, Hassan (1975) Kamus Inggris Indonesia. Pt Gramedia, Jakarta. Gramedia is a well known, well established Indonesian bookstore and publication company.