The Semantics of Utterance Particles
in Informal Hong Kong Cantonese
(Natural Semantic Metalanguage Approach)

Helen Hue Lam Leung
BA(Hons)/BComm ANU

School of Humanities, Languages and Social Science
Arts, Education and Law
Griffith University

Submitted in fulfilment of the requirements
of the degree of
Doctor of Philosophy

June 2016
Abstract

This study identifies the semantic invariants of some commonly-used Cantonese utterance particles in Hong Kong Cantonese. The particles are a distinctive and ubiquitous feature of informal, everyday Cantonese, occurring every 1.5 seconds on average (Luke 1990, 11). The particles are necessary for expressing speakers’ transitory attitudes, assumptions, or feelings connected with an utterance. Although they are not grammatically obligatory, conversation sounds unnatural when they are omitted. There are approximately 30 ‘basic’ particles, which can combine with each other to form ‘clusters’, resulting in roughly 100 variations. This number easily surpasses that of comparable particles in Mandarin, and is matched by very few, if any, other languages. Semantic analysis of Cantonese utterance particles is challenging because their meanings are extremely elusive, even to native speakers. The range of use of each particle is so varied and wide-ranging that some Cantonese speakers and scholars have concluded that the particles have no stable semantic content. Prior research on the particles has produced contradictory, vague, obscure or inaccurate descriptions.

This study demonstrates that particles have meaning, by using the Natural Semantic Metalanguage (NSM) approach to identify the semantic invariants, or ‘core’ meanings, of a selection of commonly-used utterance particles, namely laa1, wo3, gaa3, laa3, and zaa3. NSM expresses the meanings of words and concepts in reductive paraphrases called explications, where the language used is limited to a set of semantic primes (see e.g. Goddard 2008, Goddard and Wierzbicka 2002, 2014). Using this method, each particle’s meaning is identified and stated in versatile explications which are clear, accurate, translatable, and testable. The explications reliably explain each particle’s range of use in the Hong Kong Cantonese Corpus, which comprises 180 000 words of naturally-occurring Cantonese. One of the most significant findings is that explications for Cantonese utterance particles are typically short and simple. The results prove that the particles have stable and identifiable meanings.

In addition, the explications reveal the role of semantics in determining why particles can or cannot combine in particular ways. The particles selected for analysis occur in many common clusters, e.g. gaa3-laa1, gaa3-zaa3-wo3,
while other clusters are unacceptable, e.g. *laa1-wo3. The meanings of particle clusters are widely claimed to be the combined meanings of the particles of which they are made up, but there have been no serious attempts to verify this. To do so would first require accurate definitions of the individual particles. The explications proposed in this study shed light on this neglected area. It is found that where particle clusters are acceptable in speech, the combined explications reveal the meanings of the clusters. A semantic critique of sub-morphemic analyses of monosyllabic particles is also presented.

This study also considers the complexities of using NSM for Hong Kong Cantonese. If basic NSM assumptions are correct, any explication should be able to be expressed in simple and natural Cantonese, giving the same meaning as in any other language. This thesis identifies and evaluates Cantonese exponents of all the 65 proposed semantic primes, and explores some Cantonese-specific issues. Each particle explication is presented in English and Cantonese.
Statement of originality

This work has not previously been submitted for a degree or diploma in any university. 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 in the thesis itself.

Helen Hue Lam Leung
Acknowledgements

To begin with, I sincerely thank Cliff Goddard for his contributions to my PhD. He is knowledgeable, approachable, patient, and kind. He manages to push and reassure simultaneously – a unique ability I have been encouraged by many times. He has given me guidance not only on my thesis but on all things academic, for which I am also truly grateful. As well, I have benefited from the substantial experience of Andy Kirkpatrick, who gives swift feedback and straightforward advice.

I thank Kang Kwong Luke for generously allowing me the use of the Hong Kong Cantonese Corpus for my research. It has been essential to me in this project. I also thank John Wakefield, who kindly shared the audio files of the corpus with me.

I wish to acknowledge Zhengdao Ye, for encouraging and believing in me since before I began my PhD. If not for her, this thesis may never have been written. I should mention too that she gave me invaluable guidance and direction at the Australian National University where my research started.

Others who have helped along the way include Carol Priestley, Ulla Vanhatalo, Hilary Chappell, Anna Wierzbicka, Stephen Matthews, Virginia Yip, and C.T. James Huang.

All of the above have willingly offered help or shared ideas and experiences with me at various stages.

Last but not least, I thank my family and friends, in particular Mum, Dad, and Simon, for their love and support. To Mum I owe extra credit for always being my most patient and obliging native speaker consultant, from the very beginning.
Contents

Abstract i
Statement of originality iii
Acknowledgements iv
Contents v
List of tables x
List of abbreviations xi
Romanisation xii

Part One: Introduction 1

Chapter 1: The challenge of Cantonese utterance particles 3
  1.1 The semantics of particles 3
  1.2 The Cantonese language 8
  1.3 Cantonese utterance particles 11
    1.3.1 Cantonese utterance particles 11
    1.3.2 Particle combinations 17
    1.3.3 Choice of particles for this thesis 18
  1.4 Main research questions 20
  1.5 Previous scholarship on main research questions 21
    1.5.1 The meanings of Cantonese utterance particles 21
    1.5.2 Cantonese utterance particle combinations 31
    1.5.3 Summary 32
  1.6 Methodology used in this study: the Natural Semantic Metalanguage
     (NSM) 33
    1.6.1 Outline of NSM principles and practice 33
    1.6.2 Using NSM to describe particles 35
    1.6.3 Previous studies of Cantonese utterance particles using NSM 36
  1.7 Data: the Hong Kong Cantonese Corpus 40
  1.8 Thesis outline 43

Chapter 2: The Cantonese Natural Semantic Metalanguage 45
  2.1 Cantonese NSM 45
    2.1.1 Chinese NSM primes previously identified 45
    2.1.2 Cantonese exponents of NSM primes 46
2.1.3 Complexities of a Cantonese NSM 49
2.2 Substantive primes 50
2.3 Determiner and quantifier primes 54
2.4 Evaluator and descriptor primes 63
2.5 Mental predicates and speech-related primes 67
2.6 Primes of actions, events, and movement 73
2.7 Primes of location, existence, specification, and possession 75
2.8 Primes of life and death 75
2.9 Time-related primes 76
2.10 Place-related primes 81
2.11 Logical concept primes 85
2.12 Intensifier, augmentor, and similarity primes 86
2.13 Concluding remarks 89

Part Two: Semantic analysis of selected particles 91

Chapter 3: The semantics of particle laa1 93
3.1 Previous descriptions of laa1 94
3.1.1 Cantonese-English dictionary definitions of laa1 94
3.1.2 Previous academic studies of laa1 96
3.2 NSM semantic analysis of laa1 101
3.2.1 The NSM explication for laa1 101
3.2.2 Testing of explications and descriptions using the corpus 102
3.3 Concluding remarks 127

Chapter 4: The semantics of particle wo3 129
4.1 Previous descriptions of wo3 130
4.1.1 The particles wo3 and bo3 130
4.1.2 Classifications of wo3 as a question or a statement 131
4.1.3 Previous descriptions of the meaning and functions of wo3 134
4.2 NSM semantic analysis of wo3 139
4.2.1 The NSM explication for wo3 139
4.2.2 Testing of explications and descriptions using the corpus 142
4.3 Concluding remarks 170

Chapter 5: The semantics of particle gaa3 171
5.1 Previous descriptions of gaa3 171
5.2 NSM semantic analysis of *gaa3*

5.2.1 The NSM explication for *gaa3*  
5.2.2 Testing of explications and descriptions using the corpus  

5.3 Possible polysemy: *gaa3* in exclamations and questions

5.3.1 *Gaa3* in exclamations of ‘disbelief’  
5.3.2 *Gaa3* in questions

5.4 Concluding remarks

Chapter 6: The semantics of particle *laa3*

6.1 Previous descriptions of *laa3*

6.2 NSM semantic analysis of *laa3*

6.2.1 The NSM explication for *laa3*

6.2.2 Testing of explications and descriptions using the corpus

6.3 Concluding remarks

Chapter 7: The semantics of particle *zaa3*

7.1 Previous descriptions of *zaa3*

7.2 NSM semantic analysis of *zaa3*: testing possible key primes and components using the corpus

7.2.1 Possible key prime: *LITTLE~FEW siu2* (or *NOT MUCH~MANY m4 do1*)

7.2.2 Possible key prime: *SMALL sai3* (or *NOT BIG m4 daai6*)

7.2.3 Possible key prime: *A SHORT TIME jat1 zan6*

7.2.4 Possible key prime: *ONE jat1*

7.2.5 An early partial explication for *zaa3*

7.2.6 Possible components involving ‘not what you think’ and ‘true’

7.2.7 Possible component involving ‘it can’t be... like this’

7.2.8 Possible component involving ‘this someone can feel something because of this’

7.2.9 Possible component involving ‘it is not more’

7.2.10 The final NSM explication for *zaa3*

7.3 Concluding remarks

Part Three: Particle combinations

Chapter 8: The contractions hypothesis and the sub-syllabic morphemes hypothesis: a semantic critique

8.1 Overview of approaches to similarity between particles
8.2 The particle contractions hypothesis
8.2.1 The proposed gaa3 contraction
8.2.2 Contraction analyses for zaa3
8.2.3 The contractions hypothesis from an NSM perspective

8.3 The sub-syllabic morphemes hypothesis
8.3.1 The w-initial
8.3.2 The ‘mid level’ tone 3
8.3.3 The o rhyme
8.3.4 The ‘low level’/‘falling’ tone 4
8.3.5 The ‘low rise’ tone 5
8.3.6 The l-initial laa1 and l01
8.3.7 The -k coda and ‘intensifiers’
8.3.8 Sub-syllabic morphemes as ‘simplex particles’
8.3.9 Further observations

8.4 Other possible reasons for similar particles
8.5 An alternative approach: phonesthemes and sound symbolism
8.6 Concluding remarks

Chapter 9: The semantics of particle clusters
9.1 Previous descriptions of the meanings of particle clusters
9.2 Clusters made up of laa1, wo3, gaa3, laa3, and zaa3
9.2.1 Possible clusters and the mechanics of their combination
9.2.2 Non-occurring clusters
9.3 Clusters of two: gaa3-laa1 and gaa3-wo3
9.3.1 Gaa3-laa1
9.3.2 Gaa3-wo3
9.4 Clusters involving laa3
9.4.1 Gaa3-laa3
9.4.2 Laa3-wo3
9.4.3 Gaa3-laa3-wo3
9.5 Clusters involving zaa3
9.5.1 Gaa3-zaa3
9.5.2 Zaa3-wo3
9.5.3 Gaa3-zaa3-wo3
9.5.4 Summary
9.6 Reasons for the non-occurrence of certain particle clusters
9.6.1 Semantic reasons for the non-occurrence of *laa1-wo3 355
9.6.2 Semantic reasons for the non-occurrence of other clusters 357
9.6.3 Possible semantic reasons for rigid ordering in clusters 357
9.7 Concluding remarks 358

Part Four: Conclusions 361

Chapter 10: Review and prospects for further research 363
10.1 Main findings and contributions of the present study 363
10.1.1 Individual particles: laa1, wo3, gaa3, laa3, and zaa3 363
10.1.2 Particle combinations 367
10.1.3 The particles in use 368
10.2 Suggestions for further research and future directions 371
10.2.1 More particles 371
10.2.2 Questions 372
10.2.3 Phonesthemes and particle families 372
10.2.4 Particle clusters 373
10.2.5 Data 373
10.2.6 Testing of NSM explications with native speakers 374
10.2.7 Language learning 375
10.2.8 Culture 376
10.2.9 Similar particles in other languages 378
10.3 Concluding remarks 379

References 381
List of tables

**Table 1:** Onsets in the Jyutping system  xii
**Table 2:** Nuclei in the Jyutping system  xii
**Table 3:** Codas in the Jyutping system  xiii
**Table 4:** Syllabic nasals in the Jyutping system  xiii
**Table 5:** Finals in the Jyutping system  xiii
**Table 6:** Tones in the Jyutping system  xiii
**Table 7:** Initial consonants in Jyutping, IPA and Yale  xiv
**Table 8:** Final consonants in Jyutping, IPA and Yale  xiv
**Table 9:** Vowels in Jyutping, IPA and Yale  xv
**Table 10:** Tones in Jyutping, IPA and Yale (illustrated with the syllable u)  xv
**Table 11:** Frequency of particles selected for study in the Hong Kong Cantonese Corpus  19
**Table 12:** Cantonese semantic primes, with English equivalents  47
**Table 13:** Select phonesthemes of English  302
**Table 14:** Xi- phonestheme of Mandarin  303
**Table 15:** Phonesthemes in Mandarin words for ‘honesty and hypocrisy’  303
**Table 16:** Phonesthemes in Mandarin words for ‘shape and size’  304
**Table 17:** Possible particle clusters comprised of laa1, wo3, gaa3, laa3, and zaa3  311
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADV</td>
<td>Adverbial marker</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>CONT</td>
<td>Continuous aspect</td>
</tr>
<tr>
<td>DEL</td>
<td>Delimitative aspect</td>
</tr>
<tr>
<td>EXP</td>
<td>Experiential aspect</td>
</tr>
<tr>
<td>LP</td>
<td>Linking particle</td>
</tr>
<tr>
<td>PFV</td>
<td>Perfective aspect</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive aspect</td>
</tr>
<tr>
<td>PRT</td>
<td>Utterance particle</td>
</tr>
<tr>
<td>QPRT</td>
<td>Quantifying particle</td>
</tr>
<tr>
<td>VPRT</td>
<td>Verbal particle</td>
</tr>
</tbody>
</table>
This thesis uses the ‘Jyutping’ romanisation system developed by the Linguistic Society of Hong Kong. It is also known as ‘The Linguistic Society of Hong Kong Cantonese Romanisation Scheme’. Details of the Jyutping system, adapted from the Linguistic Society of Hong Kong (2015) website, are tabulated below and illustrated with Chinese characters. A variety of romanisation systems have been developed and used for Cantonese, with several different systems still in use. Further below, tables adapted from Matthews and Yip (2011, 461-463) show the Jyutping system alongside the International Phonetic Alphabet (IPA) symbols and the Yale system, another of the widely used Cantonese romanisation systems. The romanisation system used by the Hong Kong government for names matches none of these systems, and students in Hong Kong are unlikely to be taught any romanisation at all. Chinese characters are not used in this thesis because spoken Cantonese does not have a standardised written form on a par with ‘Standard Chinese’ or ‘Standard Written Chinese’, the standard written form of Chinese promoted and used in mainland China. For more on Cantonese and Chinese writing, see section 1.2.

The Jyutping system

### Table 1: Onsets in the Jyutping system

<table>
<thead>
<tr>
<th>Initial</th>
<th>Example Chinese Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>b (巴)</td>
<td>p (怕)</td>
</tr>
<tr>
<td>d (打)</td>
<td>t (他)</td>
</tr>
<tr>
<td>g (家)</td>
<td>k (卡)</td>
</tr>
<tr>
<td>gw (瓜)</td>
<td>kw (誇)</td>
</tr>
<tr>
<td>z (渣)</td>
<td>c (叉)</td>
</tr>
</tbody>
</table>

Null initials are not represented, e.g. ‘呀’ is only spelt as ‘aa’.

### Table 2: Nuclei in the Jyutping system

<table>
<thead>
<tr>
<th>Nucleus</th>
<th>Example Chinese Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa (沙)</td>
<td>i (詩/星/識)</td>
</tr>
<tr>
<td>yu (書)</td>
<td>oe (鋸)</td>
</tr>
<tr>
<td>a (新)</td>
<td>eo (詢)</td>
</tr>
</tbody>
</table>
Table 3: Codas in the Jyutping system

<table>
<thead>
<tr>
<th></th>
<th>p (濕)</th>
<th>t (失)</th>
<th>k (塞)</th>
</tr>
</thead>
<tbody>
<tr>
<td>m (心)</td>
<td>n (新)</td>
<td>ng (生)</td>
<td></td>
</tr>
<tr>
<td>i (西/需)</td>
<td>u (收)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Syllabic nasals in the Jyutping system

<table>
<thead>
<tr>
<th></th>
<th>m (唔)</th>
<th>ng (吳)</th>
</tr>
</thead>
</table>

Table 5: Finals in the Jyutping system

<table>
<thead>
<tr>
<th></th>
<th>i 思</th>
<th>ip 攪</th>
<th>it 洩</th>
<th>ik 習</th>
<th>im 閃</th>
<th>in 先</th>
<th>ing 升</th>
<th>iu 消</th>
</tr>
</thead>
<tbody>
<tr>
<td>yu 書</td>
<td>yut 雪</td>
<td>yun 孫</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u 夫</td>
<td>up 闔</td>
<td>uk 福</td>
<td>um 望</td>
<td>un 歡</td>
<td>ung 風</td>
<td>ui 灰</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e 些</td>
<td>ep 石</td>
<td>ek 石</td>
<td>em 洩</td>
<td>en 燈</td>
<td>eng 鄭</td>
<td>ei 灭</td>
<td>eu 與</td>
<td></td>
</tr>
<tr>
<td></td>
<td>eot 摔</td>
<td>eon 詢</td>
<td>eoi 需</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oe 鋸</td>
<td>oek 腳</td>
<td>oeng 疆</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o 可</td>
<td>ot 喝</td>
<td>ok 學</td>
<td>on 看</td>
<td>ong 康</td>
<td>oi 開</td>
<td>ou 好</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ap 汁</td>
<td>at 侄</td>
<td>ak 則</td>
<td>am 料</td>
<td>an 玲</td>
<td>ang 增</td>
<td>ai 擠</td>
<td>au 周</td>
</tr>
<tr>
<td>aa 渣</td>
<td>aap 集</td>
<td>aat 扎</td>
<td>aak 責</td>
<td>aam 站</td>
<td>aan 讚</td>
<td>aang 猛</td>
<td>aai 齋</td>
<td>aau 嘆</td>
</tr>
</tbody>
</table>

Table 6: Tones in the Jyutping system

<table>
<thead>
<tr>
<th></th>
<th>1 (夫/福)</th>
<th>2 (虎)</th>
<th>3 (副/霍)</th>
<th>4 (扶)</th>
<th>5 (婦)</th>
<th>6 (父/服)</th>
</tr>
</thead>
</table>

Tone marks appear at the ends of syllables.

Examples: fu1 (夫), fu2 (虎), fu3 (副), fu4 (扶), fu5 (婦), fu6 (父).
The Jyutping, IPA, and Yale systems

Table 7: Initial consonants in Jyutping, IPA and Yale

<table>
<thead>
<tr>
<th>Jyutping</th>
<th>IPA</th>
<th>Yale</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>p</td>
<td>b</td>
</tr>
<tr>
<td>p</td>
<td>ph</td>
<td>p</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
</tr>
<tr>
<td>d</td>
<td>t</td>
<td>d</td>
</tr>
<tr>
<td>t</td>
<td>th</td>
<td>t</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>g</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>k</td>
<td>kh</td>
<td>k</td>
</tr>
<tr>
<td>ng</td>
<td>η</td>
<td>ng</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>z</td>
<td>ts</td>
<td>j</td>
</tr>
<tr>
<td>c</td>
<td>tsh</td>
<td>ch</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>y</td>
</tr>
<tr>
<td>gw</td>
<td>kw</td>
<td>gw</td>
</tr>
<tr>
<td>kw</td>
<td>kwh</td>
<td>kw</td>
</tr>
<tr>
<td>w</td>
<td>w</td>
<td>w</td>
</tr>
</tbody>
</table>

Table 8: Final consonants in Jyutping, IPA and Yale

<table>
<thead>
<tr>
<th>Jyutping</th>
<th>IPA</th>
<th>Yale</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>ng</td>
<td>η</td>
<td>ng</td>
</tr>
<tr>
<td>p</td>
<td>p’</td>
<td>p</td>
</tr>
<tr>
<td>t</td>
<td>t’</td>
<td>t</td>
</tr>
<tr>
<td>k</td>
<td>k’</td>
<td>k</td>
</tr>
</tbody>
</table>

Matthews and Yip’s Yale romanisation, presented here, differs from the Yale system slightly in that the ‘high falling’ tone is not used. The words concerned are shown with a ‘high level’ tone, as these two tones are no longer distinctive in the speech of most Hong Kong Cantonese speakers today (Matthews and Yip 2011, 10-13).
Table 9: Vowels in Jyutping, IPA and Yale

<table>
<thead>
<tr>
<th>Jyutping</th>
<th>IPA</th>
<th>Yale</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>i (before ng, k)</td>
</tr>
<tr>
<td>yu</td>
<td>y</td>
<td>yu</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>u</td>
<td>o</td>
<td>u (before ng, k)</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>oe</td>
<td>œ</td>
<td>eu</td>
</tr>
<tr>
<td>eo</td>
<td>ø</td>
<td>eu (before n, t)</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a (with final consonant)</td>
</tr>
<tr>
<td>aa</td>
<td>a:</td>
<td>a (no final consonant)</td>
</tr>
<tr>
<td>aa</td>
<td>a:</td>
<td>aa</td>
</tr>
<tr>
<td>iu</td>
<td>iu</td>
<td>iu</td>
</tr>
<tr>
<td>eoi</td>
<td>øy</td>
<td>eui</td>
</tr>
<tr>
<td>ui</td>
<td>uy</td>
<td>ui</td>
</tr>
<tr>
<td>ei</td>
<td>ei</td>
<td>ei</td>
</tr>
<tr>
<td>oi</td>
<td>øy</td>
<td>oi</td>
</tr>
<tr>
<td>ou</td>
<td>ou</td>
<td>ou</td>
</tr>
<tr>
<td>ai</td>
<td>ei</td>
<td>ai</td>
</tr>
<tr>
<td>au</td>
<td>eu</td>
<td>au</td>
</tr>
<tr>
<td>aai</td>
<td>a:i</td>
<td>aai</td>
</tr>
<tr>
<td>aau</td>
<td>a:u</td>
<td>aau</td>
</tr>
</tbody>
</table>

Table 10: Tones in Jyutping, IPA and Yale (illustrated with the syllable u)

<table>
<thead>
<tr>
<th>Jyutping</th>
<th>IPA</th>
<th>Yale</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td>55</td>
<td>ù</td>
</tr>
<tr>
<td>u1</td>
<td>53</td>
<td>̀ù</td>
</tr>
<tr>
<td>u2</td>
<td>25</td>
<td>ú</td>
</tr>
<tr>
<td>u3</td>
<td>33</td>
<td>u</td>
</tr>
<tr>
<td>u4</td>
<td>21/11</td>
<td>̀uh</td>
</tr>
<tr>
<td>u5</td>
<td>23</td>
<td>̀uh</td>
</tr>
<tr>
<td>u6</td>
<td>22</td>
<td>uh</td>
</tr>
</tbody>
</table>
Part One:
Introduction
Chapter 1:
The challenge of Cantonese utterance particles

This thesis investigates the semantics of Cantonese utterance particles using the Natural Semantic Metalanguage (NSM) approach. The particles are a distinctive and ubiquitous feature of informal, everyday Cantonese, and are necessary for Cantonese speakers’ self-expression. Five commonly-used particles have been selected for this study, namely laa1, wo3, gaa3, laa3, and zaa32. Part One of this thesis introduces the challenges of Cantonese utterance particles and Cantonese NSM. Part Two identifies the invariant meanings of the selected particles. Clusters of two or three particles are also common in Cantonese, and Part Three considers the clusters composed of the five chosen particles, as well as various analyses according to which monosyllabic particles are composed of sub-units. Lastly, Part Four presents the conclusions, including the main findings of this study and future directions.

1.1 The semantics of particles

In many languages, particles are ubiquitous in ordinary speech, and play a very important role in everyday communication. Particles provide ways of expressing complex pragmatic meanings at minimal cost, expressing speakers’ attitudes towards addressees or situations, and revealing speakers’ assumptions, intentions, and emotions (Wierzbicka 2003, 341). Wierzbicka (1986, 519-520) describes particles as a vital part of linguistic machinery that help make human speech distinctly human, as opposed to the language of robots. Typical examples of English particles include well, just, even, and too. As mentioned, the foci of this thesis are the Cantonese utterance particles laa1, wo3, gaa3, laa3, and zaa3 (see section 1.3).

Despite their significance, relatively few studies have focused on what particles mean. Several reasons have been identified as causes of this neglect (Wierzbicka 1986, 519-521). First, most influential work in modern linguistic theory has been written in English by native speakers of English. English has an

---

2 The numbers indicate tones (see section 1.2).
unusually small number of particles, which makes it easy to overlook their role. But Cantonese utterance particles are a major feature of everyday Cantonese, and should not be neglected. Numerous different Cantonese utterance particles are found in speech, so they are difficult to ignore (see section 1.3).

Second, Wierzbicka states that there has been an a-semantic and a-pragmatic orientation in modern linguistics, strongly influenced by American structuralism and Chomskyan linguistics. But language is first and foremost to do with meaning. As Goddard (2011, 1) states, everything in a language conspires to realise the goal of expressing meaning in the fullest, richest, subtlest way. Semantic understanding is important in understanding any language, and exploring the meanings of words in a systematic way must be acknowledged as a vital and necessary part of linguistics (see Goddard and Wierzbicka 2014, 1-8). In fact, even in semantics, it is ‘difficult to point to a more grossly neglected area’ than that of particles (Wierzbicka 2003, 341). But understanding particle meaning is crucial to interaction, especially in Cantonese, and knowing what particles mean is necessary for semantic and communicative competence.

Third, much conventional grammatical description focuses on the sentence and often relies heavily on examples obtained from elicitation rather than from natural discourse (Goddard 2011, 163). Related to this, particles are usually grammatically optional, and do not interact with the major systems of grammar, making it is easy to overlook them and the role they play. These factors have also affected study of Cantonese utterance particles. As will be explained in section 1.7, to study them properly, it is important to use spontaneous, naturally-occurring speech, where they occur frequently.

Finally, particles are ‘extremely hard to deal with’, with most being ‘illogical’, ‘non-truth-functional’, ‘subjective’, and ‘generally rather messy’; their meanings are ‘often excruciatingly hard to state’ (Wierzbicka 1986, 520-521, 2003, 341). The same is pointed out by Goddard (2011, 162-163), who adds that they are particularly difficult to translate, and that they are often misunderstood and misused by non-native speakers. This applies to Cantonese utterance particles too. It will be shown throughout this thesis that previous descriptions of Cantonese utterance particles are often inaccurate, contradictory, and unhelpful. Some Cantonese textbooks and grammar books seem to avoid
providing definitions of utterance particles, while at the same time admitting that they are pervasive in speech, important for communication, and one of the most challenging features of Cantonese for learners. They state that ‘using particles appropriately is best learnt from practice and experience’ (Yip and Matthews 2001, 156).

**Common approaches to describing particles**

Where attempts have been made to describe particles, several approaches have traditionally been used. One approach is to compare the meaning or function of a particle with a particle or word in another language. This other language is often English. For example, we will see in Chapter 7 that many different Cantonese utterance particles are described as being like English *only* or *just*. These particles include *ze1, zek1, ze4, zaa3, zaa4, zaa5*, and *zaak1* (Kwok 1984, 53, Wakefield 2012b, 1, Fung 2000, 30). These Cantonese particles cannot all be used interchangeably, but describing them all as *only* or *just* ignores this fact. Moreover, English words like *only* and *just* are themselves fairly complex, despite also being fairly short and common words (see e.g. Aijmer 2002, 153-174, Wierzbicka 2003, 346-354), and it is unlikely that *only* and *just* are semantically equivalent even to each other, let alone to seven different Cantonese utterance particles. As Goddard (2011, 163) points out, particles almost never correspond perfectly with particles in other languages, even in number, but especially in meaning. Particles are usually highly idiosyncratic, and ‘untranslatable’ in the sense that no exact equivalents can be found in other languages (Wierzbicka 2003, 341). Furthermore, even if all these words were semantically equivalent, using a (complex) English word as a definition only helps English speakers.

More common in studies of Sinitic languages is a comparison with or ‘translation’ to Mandarin. Kwok (1984), for instance, identified Mandarin counterparts to many Cantonese utterance particles. For example, in explaining Cantonese *ge3*, she mentions Mandarin *de* (1984, 42), in explaining Cantonese *laa3*, she mentions Mandarin *le* (1984, 46), in explaining Cantonese *laa1*, she mentions Mandarin *la* (1984, 55), and in explaining Cantonese *maa3*, she mentions Mandarin *ma* (1984, 84-85). Kwok seems to assume that Cantonese particles and Mandarin particles have one-to-one correspondence and mean the
same thing. For instance, she transitions directly from explaining Mandarin *ma* to saying ‘an example of this in Cantonese would be...’. Notwithstanding the fact that these particles may indeed have similarities, this is problematic. Mandarin particles are highly unlikely to be fully equivalent to the Cantonese particles. Though Mandarin and Cantonese are related, they are very different and mutually unintelligible, especially in informal speech where Cantonese utterance particles are most abundant (Snow 2004, 2, 46) (see section 1.2). As will be explained in section 1.3, utterance particles in Cantonese greatly outnumber those in Mandarin. Additionally, this is not helpful for people who do not speak Mandarin, as the Mandarin counterparts were themselves not clearly defined by Kwok. It is counterintuitive to describe a frequently used, everyday word via one that is less well known and understood. For instance, Mandarin *le* is described as ‘inchoative’, ‘progress in story’, and ‘completed action as of the present’ (Kwok 1984, 46). Words like ‘inchoative’ are obscure even to native speakers of English, and the three descriptions of the same particle are contradictory, indicating different phases of an event.

Another common approach to describing particles is to use technical or semi-technical functional labels. For Cantonese utterance particles, some popular labels include ‘restrictive’, ‘realisation’, ‘noteworthy’, ‘assertive’, and ‘relevance’ (e.g. Fung 2000, Sybesma and Li 2007, Matthews and Yip 2011). However, this approach does not clearly explain meaning or give reliable guides to usage. Someone who is unfamiliar with the particle(s) in question is unlikely to grasp their meaning, or how to use and understand them, based on such labels. This approach also cannot escape the pitfall that multiple particles, even in the same language, are assigned the same labels, even though most particles are non-interchangeable.

Yet another approach, often used in conjunction with other approaches, has been to provide a list of examples and commentary on the various uses of the particles. This does not make explicit what the various uses have in common or how they differ, with many particles seemingly capable of being used in almost diametrically opposite ways, and the reader left to use their own linguistic intuition (Goddard 2011, 164). For example, the Cantonese *laat* is variously described in the literature as ‘advising’, ‘commanding’, ‘agreeing’, and ‘requesting’, just to name a few (Huang 1970, 414, Kwok 1984, 78-79, Lau 1977,
Other descriptions of the same particle *laa1* include ‘lack of completeness’ as well as ‘implying completion’, or ‘lack of forcefulness’ but also ‘urgency’ (Cowles 1965, 489, Kwok 1984, 55, 79, Meyer and Wempe 1947, 287). These descriptions are often accompanied by example sentences, but this mix of descriptions is more puzzling than illuminating, and leaves the reader confused and uncertain about the real meaning of *laa1*.

Even when various methods of describing particles are used in combination, as they often are, they rarely suffice to convey the meaning of a particle in a clear and complete way, with language learners unable to determine when a particle may or may not be used. The following quote from Wierzbicka (2003, 344) sums up these approaches.

The idea that the meaning of particles can be adequately elucidated by means of examples, synonyms and translation equivalents, would probably not be seriously defended by many linguists as a valid theoretical option, although the practice in question is still widespread. But the notion that the meaning of particles can be adequately elucidated by means of abstract formulae, totally unsubstitutable for the particles themselves, continues to be put forward as a theoretical program, and continues, I would add, to fail to produce empirically adequate clues to the use of the particles discussed.

As will be shown in this thesis, a much better approach to particles and semantics is to use NSM and reductive paraphrase (see section 1.6) to find explications that are clear, accurate, translatable, and substitutable in the context of the particle. Substitutability into natural examples from ordinary conversations helps in deducing the meanings of the particles, and provides an important test of the explications’ accuracy. Explications can be understood and judged by native speakers. Though meanings of particles may still be difficult to determine, the NSM approach avoids the major flaws of other approaches. This thesis largely relies on data from the Hong Kong Cantonese Corpus (see section 1.7). The next section, 1.2, briefly describes Cantonese, and following that, section 1.3 introduces Cantonese utterance particles specifically.
1.2 The Cantonese language

Cantonese is the primary spoken language of Hong Kong. It is the ‘essential community language’, used in high domains as well as low (Bolton 2012, 233). In 2011, roughly 90% of the population of Hong Kong aged over five (almost 6.1 million people) spoke Cantonese as their usual language (Census and Statistics Department 2012). In comparison, in the same year, 3.5% spoke English as their usual language, and only 1.4% spoke Mandarin as their usual language.3

Cantonese belongs to the Yue group of Chinese ‘dialects’, and there are close to 63 million speakers of Yue dialects worldwide, according to Ethnologue (Chinese, Yue 2016). It is the most widely known and influential variety of Chinese besides Mandarin (Snow 2008, Matthews and Yip 2011, 2). Differing Yue dialects are sometimes referred to as ‘Cantonese’ because Cantonese is the most well-known variety in the group, but this thesis uses the term ‘Cantonese’ to refer to the variety as spoken in Hong Kong, unless otherwise specified.

Cantonese is traditionally regarded as a Chinese dialect, not being afforded the status of a full-fledged language, even by its own speakers. Despite being the main language of the vast majority of Hong Kong people, and counter to linguistic definitions, Hong Kong people tend to believe that Cantonese is a vernacular dialect of Chinese. This may be for reasons of politics, education, and/or because Cantonese as it is spoken lacks a ‘formal’, standard written form (see below).

The standard official language of China, which will be referred to in this thesis as ‘Mandarin’, is the variety also known as ‘Modern Standard Chinese’, ‘Standard Mandarin’, or ‘Putonghua’ (it is also popularly referred to simply as ‘Chinese’). Cantonese and Mandarin are of course related, both being Sinitic languages, but they are mutually unintelligible. The two languages differ in phonology, vocabulary, and grammar. Cantonese is not just a different way of pronouncing Mandarin, and the perception that the Chinese ‘dialects’ share a common grammar is increasingly seen as misguided (Matthews and Yip 2011, 1-7). The relationship between Cantonese and Mandarin has been compared to

3 Though the status of ‘Hong Kong English’ as a variety is still debated, there has been a growing body of research on the topic (see e.g. Sung 2015, Setter, Chan, and Wong 2010, and Bolton 2002).
that of languages within the Indo-European language families: Cantonese differs from Mandarin to much the same extent as French differs from Spanish, or Swedish from German (Matthews and Yip 2011, 1-7).

The Hong Kong government promotes a policy called ‘biliteracy and trilingualism’ – being literate in Standard Chinese and English, and speaking Cantonese, English and Mandarin.4 Evans (2013) calls this an ‘ambitious goal’ and a ‘commendable if ill-defined aim’. Lam and McArthur (2004, 123) report that the majority of the Hong Kong population has not been very successful at achieving proficiency in either English or Mandarin, and there has been no build-up of momentum for one, let alone two, additional spoken languages.5 The statistics from the Census and Statistics Department given above support this claim.

Cantonese is also spoken in Guangzhou and areas around Guangzhou, with varieties of Cantonese also used in Chinese communities outside of China, such as in Singapore and Malaysia. It has a vital role as a lingua franca in Hong Kong, Macau, Guangzhou and overseas Cantonese communities (Matthews and Yip 2011, 1-3). The influence and popularity of Hong Kong Cantonese is strong, due in part to films, television programmes, and ‘Canto-pop’ music. The rapid economic development of the southern coastal districts of China, led by Guangdong province, also contribute to its growing prestige (Matthews and Yip 2011, 2-3). The importance of Hong Kong means that Cantonese as spoken in Hong Kong is considered the standard (Matthews and Yip 2011, 2-3, Baker and Ho 2010, xi).

This study considers only ‘everyday’, informal, spoken Cantonese. It considers only Cantonese as spoken in Hong Kong, as there may be slight differences in the lexicon and phonology of Cantonese as spoken elsewhere. Use of Cantonese utterance particles is reduced significantly in formal contexts, possibly because formal language is heavily influenced by Mandarin and the written language (of Mandarin). Mandarin has far fewer utterance particles or ‘final particles’ than Cantonese.

4 For language use, policy, and planning in Hong Kong, see e.g. Bolton (2012), Evans (2013), Li (2009), and Poon (2011).
5 As Lam and McArthur (2004, 122) point out, the Hong Kong government, in using the phrase ‘trilingualism’, has intentionally or unintentionally asserted the status of Cantonese as a language in its own right, and Beijing has not publicly objected.
Chinese writing/characters

This thesis uses the ‘Jyutping’ romanisation system developed by the Linguistic Society of Hong Kong. It is also known as ‘The Linguistic Society of Hong Kong Cantonese Romanisation Scheme’. A large variety of other romanisation systems have been developed and used for Cantonese, with several different systems still in use, such as the Yale system. The romanisation system used by the Hong Kong government for names is different again, and students in Hong Kong are unlikely to be taught any romanisation at all. For tables showing the Jyutping system alongside the International Phonetic Alphabet symbols and the Yale system, see pages xii-xv.

Chinese characters are not used in this thesis, partly because spoken Cantonese does not have a standardised written form on a par with ‘Standard Chinese’ or ‘Standard Written Chinese’, the standard written form of Chinese promoted and used in China since the early twentieth century. No form of written Cantonese is taught in schools or used in academic settings in any Cantonese-speaking community, although written Cantonese has developed a significant number of the attributes associated with ‘standard languages’ (Matthews and Yip 2011, 7, Snow 2008). When it comes to the written form, it is essentially the written form of Mandarin that is taught and learnt in Hong Kong. In writing, Cantonese and Mandarin are very similar, due not to the two languages being similar, but to the fact that there is a high degree of discontinuity between spoken and written language in Hong Kong. The situation has been compared by Snow (2004, 2) to that of Dutch speakers reading and writing in German. As mentioned, Cantonese is widely regarded by even its own speakers as a dialect of Chinese, and Mandarin, and especially written Chinese, are believed to be more ‘proper’. This may have implications for the use of a natural Cantonese metalanguage (see section 2.1.3).

Many features of spoken Cantonese and colloquial Cantonese terms, including utterance particles, do not have a traditional written form and are generally not found in writing. The exceptions to this are when the written form tries to imitate informal speech, such as in advertisements, gossip columns, comics, or informal personal communication between friends, such as online or via text messaging (James 2001, Matthews and Yip 2011, 5-6, Snow 2004).
Therefore, it would make little sense for a study of Cantonese utterance particles to concentrate on written Chinese.

Another reason not to use Chinese characters in this thesis is that it is not unusual for more than one character to be in widespread use for common Cantonese words, nor is it unusual for a single character to be in use for several different Cantonese words. Cantonese speakers often create makeshift characters. Even in dictionaries, the same utterance particle can be represented by different authors using different characters, or else the one character will be used in different texts for different particles. Sometimes it is impossible to know from the character which tone the particle would be pronounced in. For example, Chan’s (1955, 283) Cantonese-English dictionary gives a definition of a particle laa. Without romanisation, and combined with a vague and obscure definition, it is ambiguous whether this particle is laa1 or laa3. Distinctions such as this are very important because this study follows the precedent given by most scholars studying Cantonese utterance particles, including Yau (1965), Gibbons (1980), Kwok (1984), and Luke (1990), by considering tones to be integral to the lexical identity of a particle, rather than as being a morphological feature applied to ‘neutral’ or ‘toneless’ particles (cf. ‘sub-syllabic morphemes’ discussed in Chapter 8). Therefore, laa1, laa3 and laa4, as well as wo3, wo4, and wo5 etc. are all considered separate particles. Dictionaries and reference tools are usually influential in the process of standardisation, but no single dictionary has yet been accepted as the authoritative standard for Cantonese (Snow 2004, 58).6

1.3 Cantonese utterance particles

1.3.1 Cantonese utterance particles

Cantonese utterance particles attach to the ends of utterances, being like bound morphemes in that they cannot occur independently. Although they are not grammatically obligatory, they have important functions. They have no direct counterpart in English, where their functions are often conveyed by intonation

---

6 Lam and McArthur (2004) explored whether there should be Hong Kong-specific ‘biliterate and trilingual dictionaries’ for education and other purposes, and considered the lexicographical implications of the ‘biliterate and trilingual’ policy – they concluded that a biliterate and trilingual dictionary was not likely to eventuate.
patterns (see below). They sometimes correspond functionally to English question tags (Matthews and Yip 2011, 389-390). Cantonese utterance particles are also known as ‘sentence particles’, ‘sentence-final particles’, or ‘final particles’. This is slightly misleading, though, because the particles can occur at the ends of syntactic units which are not sentences, such as clauses, phrases, free-standing words, paragraphs, after the sentence topic, and at other natural breaks in the sentence (Gibbons 1980, 765, Luke 1990, 6-10, Matthews and Yip 2011, 393-394). For these reasons, this thesis uses the terminology of e.g. Gibbons (1980) and Luke (1990), and refers to the particles as ‘utterance particles’.

For Cantonese speakers, utterance particles are necessary for self-expression, being a way to indicate attitudes, assumptions, moods, intentions, feelings, or emotions. This may be because Cantonese is a tonal language with six clearly distinctive tones. Changing the tone of a Cantonese word changes its meaning, which means speakers cannot use intonation to the same extent as in English or other European languages. Much of the meaning typically conveyed by intonation in languages such as English is expressed via utterance particles in Cantonese instead (Yau 1980, 51, Matthews and Yip 2011, 389-390, Yip and Matthews 2001, 156, Wakefield 2011b). As Luke (1990, 6) puts it, it is often difficult to speak purely objectively, and since the potential for emotions to be expressed through intonation is limited in Cantonese, utterance particles assist by making emotions more recognisable.

To give an example of some utterance particles in Cantonese, example (1.1) has been taken from the Hong Kong Cantonese Corpus (see section 1.7) and presented below. It shows a teacher describing a student of hers whom nobody likes. This excerpt features the particles laa1, wo3, and gaa3, which are

---

7 Based on the position of the particles in sentences, Luke and Nancarrow (1997) divide the particles into three groups: initial particles (sentence-initial position), final particles (sentence-final position), and medial particles (at other places in the sentence).
8 The number of tones in Cantonese depends on analysis as well as variation and ongoing change (Matthews and Yip 2011, 27-37). In terms of analysis, checked syllables are traditionally counted contrastively, resulting in three additional tones, i.e. nine distinct Cantonese tones altogether. However, the three additional tones which occur before unreleased consonants may be seen as abbreviated counterparts of the three level tones which occur in other contexts. Therefore, Cantonese can be said to have six distinct tones. The system of six tones (and of nine tones) assumes that the ‘high level’ and ‘high falling’ tones are not distinctive.
9 Schubiger (1965) was apparently the first to point out the interrelationship between particles and intonation, in a comparative study of German and English, two non-tonal languages.
analysed in the coming chapters. It also shows use of the particle *ne1*, which is not analysed here. This example will be revisited in section 10.1.3.

(1.1)

\[
\text{Kei}4\text{taa}1 \quad \text{tung}4\text{hok}6 \quad \text{ne1,} \quad e6, \quad \text{kei}4\text{sat}6 \quad \text{ne1}
\]
\[
\text{other} \quad \text{student} \quad \text{PRT} \quad \text{eh} \quad \text{actually} \quad \text{PRT}
\]
\[
\text{dou1} \quad m4 \quad \text{zung}1\text{ji3} \quad \text{keoi5} \quad \text{gaa3}. \quad \text{Jan}1\text{wai6} \quad \text{ne1}
\]
\[
\text{also/all} \quad \text{not} \quad \text{like} \quad \text{him/her} \quad \text{PRT} \quad \text{Because} \quad \text{PRT}
\]
\[
\text{keoi5} \quad m4 \quad \text{do1} \quad \text{ceot1seng1} \quad \text{laa1}, \quad \text{ji4ce2} \quad \text{ne1}
\]
\[
\text{s/he} \quad \text{not much} \quad \text{speak} \quad \text{PRT} \quad \text{also/moreover} \quad \text{PRT}
\]
\[
\text{jau6} \quad \text{seng}4\text{jat}6 \quad \text{tau4sou3} \quad \text{jan4dei6} \quad \text{wo3}...
\]
\[
\text{also} \quad \text{always} \quad \text{complain} \quad \text{people/others} \quad \text{PRT}
\]

‘Other students *ne1*, eh, actually *ne1*, they don’t like him/her either *gaa3*. Because *ne1* s/he doesn’t talk much *laa1*, and *ne1* also always complains about other people *wo3*...’

The prevalence of Cantonese utterance particles in everyday, spoken Cantonese is immediately noticeable. Luke (1990, 11) found that they occur every 1.5 seconds on average in ordinary, continuous talk (although they occur less in formal situations). Gibbons (1980, 765) reported that in a radio discussion programme, 85% of the utterances used particles. Cantonese speakers widely agree that a conversation without any utterance particles sounds highly unnatural. Luke explains that if one were to imagine a real conversation without particles, ‘it would sound overly terse to some, hostile or perhaps funny to others, but in any case unreal, almost unintelligible’ (Luke 1990, 14). In the Hong Kong Cantonese Corpus, the most frequently used particle, *aa3*, is also, interestingly, the second most frequently occurring word overall. It comes second only to the word *hai6*, which means ‘is’ or ‘yes’. The second most frequently used particle, *gaa3*, is the 11th most frequently used word.

Cantonese has a very large number of utterance particles. Approximately 30 or more ‘basic’ or ‘independent’ (typically monosyllabic) particles were identified by Kwok (1984, 8). Yau (1980) counted 42 which occurred at least once in his recording, which ran for approximately 15 hours (with an additional 11 ‘bound’ particles which must occur with another particle). Some other scholars report even larger numbers; for example, Gibbons (1980, 764) reports that there are approximately 70 utterance particles which can stand alone. The particles can also be used together in ‘clusters’ of more than one particle, so that
the actual number of particles (basic and compound) in use in spoken Cantonese is said to be approximately 100 or more (Gibbons 1980, 764, Luke 1990, 1, Wakefield 2011b, 13, Yip and Matthews 2000, 131). In terms of sheer numbers, then, Cantonese utterance particles far outnumber their Mandarin counterparts, and is matched by very few, if any, other languages (Luke 1990, 1, Wakefield 2011b, 2). This rich collection of particles may mean that Cantonese uses utterance particles to express connotative meaning more than any other language (Wakefield 2011b, 80). The unusually large number of utterance particles, and their ability to combine into clusters, adds to their interest as a linguistic phenomenon.

The significance of the particles to Cantonese speakers can also be seen in that speakers use them even when not speaking Cantonese. The internet has become an essential part of everyday life for most Hong Kong people, and it has been observed that when writing online informally, Cantonese speakers often use utterance particles, even when writing in English (James 2001, Ho 2003, Wong 2009). This shows firstly that the meaning in those particles is so important for Cantonese speakers that they cannot be (easily/naturally) omitted, and secondly that those meanings are not easily translatable.

Some scholars have mentioned that use of utterance particles may vary between individuals, the sexes, and/or age groups. For example, Chan (2002) found gender-linked differences in the choice of particles and in their usage in sentence types (though her study was based on data from a television series filmed in Guangzhou in mainland China in the mid-1980s). Botha and Barnes (2013) considered the impact of age and gender on particles (though their study was based on Macau Cantonese). This thesis is not concerned with variations in particle use between sexes or age groups. Even if there are some differences in, say, relative frequencies of usage, there is no evidence to suggest that the particles investigated here are not used by all Cantonese speakers, and more importantly, there is no evidence to suggest that the particles are not understood by all speakers. This thesis is concerned with the particles’ meanings, not with sociolinguistic variation in their use. Any invariant meaning should stay constant and is separate from whether or not all speakers will want to convey those meanings.
Note that Cantonese also has verbal particles, which are not like utterance particles and are not the focus of this study. Verbal particles indicate ‘notions such as result (effect on an object) and phase of action (beginning, continuing or ending)’, and can be further divided by their functions into the categories of directional, resultative, quantifying, and adversative/habitual (Matthews and Yip 2011, 243-262). They are comparable in form and function to the particles of English phrasal verbs10. An example of the resultative verbal particle dou2, taken from Matthews and Yip (2011, 243), is given below. Aspect markers in Cantonese are also sometimes called ‘aspect particles’.

(1.2) 
Ngóh ngāam- ngāam sāu dóu chin 
I just-just receive PRT money

‘I’ve just received the money.’ (Matthews and Yip’s gloss and translation)

This section has introduced Cantonese utterance particles and shown that they are very important in Cantonese and for Cantonese speakers. Despite the significance and pervasiveness of utterance particles in Cantonese, it is generally agreed that their meanings, if they have any, are extremely elusive. Native Cantonese speakers themselves find it incredibly difficult to pinpoint what the particles mean. As will be shown throughout this thesis, accurate and comprehensive semantic analyses of Cantonese utterance particles are lacking. This study addresses the problem by using the NSM framework to pinpoint the invariant semantic content of some Cantonese utterance particles, and to test the idea that the meaning of particle ‘clusters’ is equal to the meaning of the individual particles combined.

Comparable particles in other Southeast Asian languages

Mandarin has fewer utterance particles (or sentence-final particles) than Cantonese, perhaps because Mandarin has only four lexical tones, not including the ‘neutral’ tone. Li and Thompson (1981, 238) listed six particles in Mandarin11, while Chao (1968, 795-814) listed 26 (and two additional ‘tonal

10 For example, English up denotes direction in pick up, but completion in eat up (Matthews and Yip 2011, 243).
11 They are le, ne, ba, ou, a/ya, and ma.
Particles in Mandarin are usually considered to be ‘toneless’ or in the ‘neutral tone’, and Mandarin speakers can more readily accommodate the pitch changes demanded by intonation patterns. Cantonese has a larger number and a richer variety of utterance particles, which are also able to occur together in more combinations. In Mandarin, only le combines with other particles (Matthews and Yip 2011, 389). Chappell (1991, 40-41) counted utterance-final particles roughly every 6 seconds in informal Mandarin conversation (compared to 1.5 seconds in Cantonese). Some other tonal languages, such as Vietnamese and Thai, also have comparable particles. The total number of particles found in Vietnamese ranges from 25 to 45, depending on the regional dialect (Tran 2010, 335-343). Peyasantiwong (1981) recorded roughly 40 final particles in Thai.

Like Cantonese, Singapore English has noticeable utterance final particles, and some Singapore English particles have even been suggested to have their origins in Cantonese. Lim (2007) considered the etymology of the Singapore English particles and suggested that Cantonese provided lor, hor, leh, meh, and ma, inclusive of tone. Additionally, one of the most studied Singapore English particles is la (sometimes spelt lah), and it has been suggested that a Cantonese particle laa is a possible source (Kwan-Terry 1978, Gupta 1992). Three Cantonese laa particles are in use, and it is not clear which would have been the source (hence omission of a tone here), although Besemer and Wierzbicka (2003, 25) noted a strong similarity between the meaning of Singapore English la and Luke’s (1990) analysis of Cantonese laa1. Other scholars have suggested that there is more than one Singapore English la, differentiated by tone, as is the case in Cantonese. Wong (2004) made a distinction of three different lexical tones for Singapore English la, while Kwan-Terry (1978) suggested that it has two lexical tones. This may have some correlation with Cantonese, where three separate particles laa1, laa3, and laa4 are generally listed in studies of particles, although the first two are far more salient, and occur more frequently than laa4. Some might also suggest that Cantonese laa has some relation to the Malay lah. However, this link seems more tentative, perhaps because Malay lah is seen as

12 There may be differences between the varieties of Mandarin, and some Mandarin particles may be polysemous.
being more closely linked with Singapore English la\textsuperscript{13}. This thesis does not explore the historical origins of these particles or whether they are related. No comparisons are attempted here, which would require a better understanding of the other languages.

1.3.2 Particle combinations

Several types of particle combinations are possible in Cantonese, or potentially possible. Some scholars have claimed that certain monosyllabic particles can be further broken down into meaningful units. There are two main sub-types of this kind of combination, termed in this thesis ‘particle contractions’ and ‘sub-syllabic morphemes’. The term ‘contraction’ is used to describe the claim that a given particle is made up of two or more other particles. For example, it has been claimed that $gaa3$ consists of two particles, $ge3 + aa3$. A similar but different idea is that of ‘sub-syllabic morphemes’, whereby particles are dissected phonologically into component parts that are not particles themselves, and that do not occur independently. For example, it has been claimed that $wo3$ can be analysed as $w- + o + \text{tone 3}$. A third possibility, for which this thesis uses the term ‘particle cluster’, refers to polysyllabic combinations of two or more particles, which can be heard being used one after the other. For example, $gaa3$-zaa3 is a cluster of two particles, $gaa3$ and $zaa3$. $Gaa3$-laa3-wo3 is a cluster of three particles, $gaa3$, $laa3$, and $wo3$. I use the term ‘particle combinations’ as an umbrella term for contractions, sub-syllabic morphemes, and clusters. Note that in the existing literature, some authors use the terms ‘combination’, ‘cluster’, and ‘contraction’ interchangeably, and may use them to include sub-syllabic morphemes.

Very large numbers of particle clusters have been counted in the literature. Yau (1980) reported that the 42 ‘bare’/‘independent’ particles and 11 ‘bound’ particles he found together created 164 particle clusters, all of which occurred at least once in his 15 hour recording. Kwok (1984, 8-11) listed a much smaller 74 combinations (mostly clusters, with a small proportion of contractions) derived

\textsuperscript{13} Besemeres and Wierzbicka (2003, 27-29) gave a brief comparison of the semantics of Singapore English $la$ and Malay $lah$, based on their own analysis of the former and Goddard’s (1994b) analysis of the latter. They found that the two are similar, although the formula proposed by Goddard could not fit all the examples used in their corpus.
from her list of 30 ‘basic’ particles. Matthews and Yip (2011, 394-395) point out that this is a small figure compared to the theoretically possible number of permutations given 30 basic particles.

Due to limitations of the Hong Kong Cantonese Corpus used in this study, there are no exact statistics on how often the particles occur in clusters in the corpus, or which clusters are the most popular. In searching for examples for the analysis in Chapter 9, it seems that gaa3-laai1 is one of the most frequently used, while laa3-wo3 and zaai3-wo3 are relatively less frequent. Unsurprisingly, clusters of three particles are also slightly harder to find than clusters of two particles, although they are not uncommon. Clusters of more than three particles are relatively rare, although they are not unheard of and have been reported in the literature (Gibbons 1980, 765, Matthews and Yip 2011, 394-397).

1.3.3 Choice of particles for this thesis

Where there have been attempts to describe utterance particles’ elusive meanings, linguists have often sought to analyse a large group, or the entire inventory, of known Cantonese utterance particles (Yau 1965, Gibbons 1980, Kwok 1984). Taking on such ambitious numbers all at once has led to individual particles being treated only briefly, all particles being considered under uniform criteria, and/or particles being described mainly in relation to other particles. Most descriptions are questionable when tested with additional examples. Studying the whole class of particles has arguably led scholars to jump to conclusions about individual particles before they have properly identified or analysed their range of use (Luke 1990, 17). Luke’s study had a much narrower focus, looking only at the particles laai1, lo1, and wo3. Since Luke’s study, more work on individual particles, or much smaller groups of particles, has been carried out.

This study will take a ‘micro’ approach, looking at only five Cantonese utterance particles: laai1, wo3, gaa3, laa3, and zaai3. These five particles were chosen for several reasons. Each is commonly-used, with laa1, wo3, gaa3, and laa3 being in the top six most frequently used particles, according to the Hong Kong Cantonese Corpus (zaa3 is the 17th most frequently used particle). The
five particles seem intuitively to a native speaker to be quite different from each other semantically, with the possible exception of laa1 and laa3. Including these, however, helps us determine whether particles that differ only in tone have any similarities. The five selected particles also allow comparison of particles which have various phonological similarities and differences in the three areas of initial, rhyme, and tone, intended to allow investigation of the sub-syllabic morphemes hypothesis discussed in Chapter 8. Additionally, various commonly-used clusters can be formed from these five particles (while some particles cannot combine with each other into clusters). They will be discussed in Chapter 9. The main focus in this thesis will be declarative utterances, although questions are mentioned where relevant. It is hoped that close analysis of this small set of particles may prove a useful starting point for investigation of larger numbers of utterance particles in the future.

Table 11 below shows the frequency of the five particles selected, according to the Hong Kong Cantonese Corpus. They are listed in order of their analysis in this thesis. The ‘particle ranking’ shows the ranking of the particle in a list of the most frequently occurring utterance particles. The ‘word ranking’ shows the position of the particle in a list of the most frequently used words overall. The ‘frequency’ column shows the number of tokens of the particle in the corpus. The eight particle clusters that can be formed using laa1, wo3, gaa3, laa3, and zaa3 are listed further below.

Table 11: Frequency of particles selected for study in the Hong Kong Cantonese Corpus

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Particle</th>
<th>Particle ranking</th>
<th>Word ranking</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Laa1</td>
<td>3rd</td>
<td>14th</td>
<td>1578</td>
</tr>
<tr>
<td>4</td>
<td>Wo3</td>
<td>5th</td>
<td>23rd</td>
<td>928</td>
</tr>
<tr>
<td>5</td>
<td>Gaa3</td>
<td>2nd</td>
<td>11th</td>
<td>1832</td>
</tr>
<tr>
<td>6</td>
<td>Laa3</td>
<td>6th</td>
<td>33rd</td>
<td>664</td>
</tr>
<tr>
<td>7</td>
<td>Zaa3</td>
<td>17th</td>
<td>144th</td>
<td>173</td>
</tr>
</tbody>
</table>

Resulting particle clusters:

Gaa3-laa1
Gaa3-wo3
1.4 Main research questions

The main research questions addressed in this thesis can be summarised as below. The context from which they were created is discussed in the next section.

1. Do Cantonese utterance particles have stable, invariant meanings?
   a. If so, can their meanings be stated clearly and precisely, using NSM, in English and Cantonese?
   b. Can the proposed meanings be tested and substantiated by evidence from a corpus of naturally-occurring, spoken Cantonese?
   c. What is the character and level of complexity of the explications?

2. To what extent can particle combinations be explained semantically?
   a. Are monosyllabic particles made up of smaller meaningful units?
   b. Do polysyllabic particle clusters have the same meaning as the combined meanings of the particles of which they are made up?

As well, this thesis necessarily addresses an important methodological preliminary, namely, whether it is possible to create a viable Cantonese NSM. This is investigated in Chapter 2, where it is established that Cantonese exponents of all the proposed semantic primes exist. Accordingly, all the explications of particles proposed throughout this thesis are given in both English and Cantonese.
1.5 Previous scholarship on main research questions

Some previous studies of Cantonese utterance particles are overviewed here, with a focus on the main research questions identified above. The methods used in these studies are evaluated, and it will be argued that the approaches to particles’ meanings have not been very satisfactory. Each chapter in Parts Two and Three includes additional reviews of the literature more specific to the topic of the chapter. Note that only research published in English has been reviewed here.

1.5.1 The meanings of Cantonese utterance particles

Do Cantonese utterance particles have meaning?

First and foremost in this study of Cantonese utterance particles is the question of whether or not the particles have meaning. This is a question which even native speakers of Cantonese are unable to immediately answer with confidence, and it is widely acknowledged that if Cantonese utterance particles have meaning, their meanings are extremely elusive. In one of the major works on Cantonese utterance particles, Luke (1990, 3-4) states simply ‘They have no semantic content’, listing this as the first of four distinctive features of Cantonese utterance particles. He explains that language studies in the Chinese tradition draw the distinction between shizi ‘full words’ and xuzi ‘empty words’, i.e. words that refer to ‘actual things and events’, and words that do not. He places utterance particles in the category of ‘empty words’, which would mean that they do not represent concepts, but instead help to organise language. Kwok (1984, 6) also classifies Cantonese utterance particles as ‘empty’, with the

---

14 The author is not suitably learned in academic Chinese writing, but this seems not to be a serious drawback, since prominent experts in the field, such as Luke (1990), also report that literature on Cantonese utterance particles (and even particles in Mandarin and other languages) is scant and not particularly useful. Luke states that existing descriptions lack empirical data and study an unmanageably large number of particles at once, such that individual particles are not sufficiently analysed. More recently, Wakefield (2011b, 18) agreed that published definitions are inadequate. Furthermore, as explained in sections 1.2 and 2.1.3, it is essentially the written form of Mandarin that is normally used in formal and academic contexts in Hong Kong. Thus, a case can be made that even research published in Chinese is not being presented in the true native language of Cantonese speakers.
distinction between ‘empty’ and ‘full’ being based on the ‘amount of semantic content’. 

It is understandable that Cantonese utterance particles have been labelled empty words which do not refer to ‘actual things and events’. However, it seems implausible that all words in any language can be fitted neatly into these two categories. Furthermore, Ye’s (2004) NSM study of Mandarin ‘emotional adverbs’, which in the Chinese linguistic tradition are also considered ‘empty words’, showed that the emotional adverbs were rich in semantic content. The distinction between ‘full’ and ‘empty’ words is not unlike the distinction drawn in Relevance Theory between words with ‘conceptual meaning’ (encoding a concept) and words with ‘procedural meaning’ (instructions regarding how to process a concept), a notion which Bordería (2008) and Sasamoto (2008), among others, have evaluated with regards to discourse markers.

One of the major attempts to describe the meanings of Cantonese utterance particles was carried out by Kwok (1984), who was interested in their meanings despite categorising them as ‘empty’. As an example, we can consider her description of the particle aa3, which was the most frequently occurring particle in her corpus. Kwok (1984, 45-46) states that when suffixed to statements, aa3 ‘does not add a great deal to the meaning of the sentence which carries it’. This description seems to support Luke’s assumption that particles have no meaning. She tells us that aa3 can occur at the end of a sentence or in a non-final position, and in declarative, interrogative, and imperative structures. This is not particularly illuminating either, and does not help to pinpoint the meaning of aa3. If anything, it confuses the picture by seeming to suggest that aa3 can occur almost anywhere. When aa3 is used in questions, it can be taken as the ‘neutral’ form, meaning that ‘it is used to refer to a form that is preferred except where there is “good reason” for choosing something else’ (Kwok 1984, 71-72). Kwok states again that aa3 ‘does not carry much semantic content, and does not add a great deal of meaning to the sentence, especially when it is weakly stressed’. Like its use in statements, its main function seems to be to

---

15 Peyasantiwong (1981) notes roughly 40 final particles in Thai and states that all scholars agree they have no lexical meaning of their own. He believes that they are far more complex than is generally understood, but felt unable to make concrete conclusions about their meanings because they often overlap and can be used in widely differing circumstances.
somehow make the sentence seem less abrupt. She makes no attempt at explaining what aa3 means when it is stressed, or how and why it makes a sentence seem less abrupt. More importantly, the meaning of aa3 has not been captured, and if one does not have a firm grasp on the meaning of aa3 in the first place, it is then impossible to know whether one wishes to convey a meaning not expressed by aa3.

Kwok’s description of the particle laa1 is equally perplexing (Kwok 1984, 78-79). She states that laa1 is the ‘neutral form’ for imperative sentences, but also that imperative sentences do not always take an utterance particle. Again, by ‘neutral’, she means that laa1 is ‘the form chosen except when one wishes to express some special meaning’, and again, this is ambiguous. Similarly to her treatment of aa3, she states that sentences with laa1 may involve the addressee, the speaker, or a third party. This merely draws attention to the unanswered question of when to use laa1, and does not capture laa1’s meaning. Despite Kwok’s apparent view that utterance particles have meaning, she seems unable to precisely state what these meanings are.

Others state that while the particles play a role in conversation, their meanings are not fixed, and/or are impossible to state. Botha and Barnes (2013), for example, studying Macau Cantonese, claimed that meaningful social information is conveyed using utterance particles, but that their social meanings are variable, not rigid or fixed. As mentioned, some textbooks and grammar books admit that Cantonese utterance particles are pervasive, important, and one of the most challenging features for learners; however, they state that ‘using particles appropriately is best learnt from practice and experience’ (Yip and Matthews 2001, 156). Besides not contributing to a reader’s understanding, such statements also support the idea that the particles have no ‘real’ meaning.

Is it possible that Cantonese utterance particles truly have no meaning? This is hard to believe, as it begs the question of why there are so many different Cantonese utterance particles, and why they are used in such abundance in normal conversation. A more likely possibility is that the particles do have meanings, but that they are difficult to state. When James (2000, 72, 2001, 15) asked his native Cantonese-speaking students for clarification of utterance particles’ meanings, the students, although very willing to help, were amazed at their own inability to explain what they understood intuitively. A frequent
response was that the words had no meaning, being used just for ‘emphasis’. Yet, a contrastive approach which attached different particles to the same utterance led the students to say that the utterances had different meanings, which suggests that the particles do have meanings. The question that arises then is: if these particles do have meanings, are their meanings identifiable in any consistent fashion?

**Are the meanings of Cantonese utterance particles identifiable?**

Although not the most reliable source, dictionaries can be a good starting point. Entries for utterance particles in Cantonese-English dictionaries usually consist of a description of the kinds of sentences that the particle can attach to, or translated ‘equivalents’ in the form of English verbs, sometimes accompanied by short examples. However, as explained in section 1.1, such methods of explaining particles’ meanings leave much to be desired. In the case of laa1, for example, descriptions are contradictory, use the same labels for multiple particles which are not truly synonymous, and are simply not very informative (see Chapter 3). Assessing dictionary definitions, Wakefield (2011b, 15) states that ‘some are helpful as a starting point, but many are inaccurate or misleading, and none are sufficient on their own for language learners, translators, or linguists’. In this thesis, Chapter 3 will be the only chapter to include discussion of dictionary descriptions, and it will become clear that they are not helpful.\(^\text{16}\)

Dictionaries are obviously not meant to contain in-depth semantic analyses of particles, but the problems just outlined also abound in scholarly work on Cantonese utterance particles. Why does it appear that the potential meanings of Cantonese utterance particles cannot be stated clearly? Kwok (1984) appears to be the first to have realised that each particle can be used in different contexts and situations. Luke’s (1990) separation of each particle into more than five sequences further promoted the idea that particles can behave differently in different contexts. One reason given by Luke (1990) for the contradictory nature of most prior definitions of Cantonese utterance particles was that no attempt had been made to capture the full range of uses of each

\(^\text{16}\) English-Cantonese dictionaries (e.g. Kwan et al. (2000)) do not appear to mention any utterance particles under the popular English descriptors found in the literature.
He believed that discrepancies existed because scholars had provided only partial pictures of the kinds of work each particle can perform.

Luke’s descriptions are amongst the most in-depth for individual particles. Although he studied far fewer particles, and his focus was much more narrow and specific than Kwok’s (1984), Yau’s (1965), or Gibbons’ (1980), he gave significantly more detailed analyses of each particle. Luke (1990) looked at the three particles laat, lo1, and wo3, and aimed to identify all the ‘sequences’ where each particle was possible. For example, the sequences identified for wo3 were ‘reportings and story-tellings’, ‘challenging a position’, ‘contact-establishments’, ‘disconfirmations’, ‘thankings’, ‘informings and remindings’, and ‘realisations’ (Luke 1990, 199-255). In each kind of sequence, Luke gave a thorough discussion of the particle, with examples. His aim was not to concisely explain what the particle meant, but quite the opposite. His emphasis was on showing that the same particle can be used in very different ways. It is easy to believe that Cantonese utterance particles have no meaning, if they appear to behave differently in different contexts. Luke’s acknowledgement of the wide range of use of each particle, and attention to detail, were important improvements towards explaining the particles comprehensively. Nonetheless, the vast range of contexts and sequential environments in which the particles occur, and their apparently myriad functions, makes them difficult to understand, and seems to confirm that they are inherently confusing and mysterious.

Luke’s (1990) approach fits the description of the widely-used ‘conversational’ and/or ‘discourse’ approach identified by Wierzbicka (1986, 524), which focuses on the role of particles in discourse or conversation. It is true that particles often seem to indicate, in a complex way, how the utterance that contains them is a response to, or a continuation of, some portion of the prior discourse, but as Wierzbicka argues, many particles can also be used in isolated sentences and have nothing to do with discourse cohesion or conversational structure (Wierzbicka 1986, 524). Luke (1990, 6-10) himself states that Cantonese utterance particles can be said to attach to a wide range of syntactic units, from whole paragraphs to free-standing words. Wierzbicka writes that a comprehensive theory of particles must be applicable to all particles, and that this approach does not provide a unified framework for the
study of all particles. Therefore, their role in the linguistic study of particles must be supplementary rather than central, and the key must lie elsewhere.

Nevertheless, the realisation that each particle can be used in different contexts and situations is very important. Luke’s work may have been one factor which influenced the gradual transition of studies of Cantonese utterance particles from quantitative to qualitative, and from looking at large numbers of particles to looking at a selection in more depth. This shift deserves an overview, to show that even though quantitative methods appear more objective and definitive, they have been unable to pinpoint any ‘core’ meanings of Cantonese utterance particles.

The first comprehensive and scholarly study of Cantonese utterance particles was carried out by Yau (1965). Yau conducted two main ‘tests’, both encompassing the whole range of 89 particles identified at the time. The ‘S-Q test’ was concerned with whether a statement with an utterance particle would be preserved as a statement, or be transformed into a question. Yau used questionnaires, determining the classification of each particle by calculating statistics from the participants’ answers.17 The three possible categories were ‘Q-type’, containing particles that demand a verbal confirmation, ‘S-type’, containing ‘affirming’ particles, and ‘S-Q-type’, which consisted of particles indeterminable between the two other categories. However, this S-Q test did not provide much description of meaning and is not particularly informative.

Yau’s S-Q test used only one criterion to assign each of the 89 particles into one of three categories. Each particle would seem to have the same property as many other particles, with the differences between unique particles in the same category being overlooked. Another potential problem with the S-Q test is related to his methodology. Example utterances were given to participants to show each particle’s typical use. As it has been established that the same particle can be used in a great variety of situations, particles may have

---

17 Participants were asked to place utterances with utterance particles on a five-point bipolar scale depending on whether the utterance it attached to would demand a verbal confirmation or provide affirmation (Yau 1965, 39-48). The five points were Q (obviously demanding a verbal confirmation), ½Q (less obviously demanding a verbal confirmation), S-Q (equally associated with Q and S or irrelevant to both), ½S (less obviously affirming) and S (obviously affirming). A number was arbitrarily assigned to each of the five positions of the scale. His calculations provided the mean, which indicated the opinion of the average respondent, and the standard deviation, which indicated the spread of the scores and the reliability of the test item, as well as a formula indicating significance when two particles scored similarly (Yau 1965, 58-62).
a wide range of ‘typical’ uses. Furthermore, the utterances were given independent of context, and it would have been possible for participants to misinterpret the examples and thus lead to less valid results, not to mention that, as explained, even native-speakers find particles difficult to explain.

The second test conducted by Yau, the ‘C-test’ (Yau 1965, 82-120), was used to determine which particles contain which of 12 ‘connotation concepts’. They were ‘coaxing’, ‘surprised’, ‘hesitating’, ‘fault-finding’, ‘patient’, ‘persuading’, ‘dissatisfied’, ‘conceited’, ‘reluctant’, ‘reminding’, ‘doubting’, and ‘politely urging’. This is more helpful than the S-Q test, but is still flawed in that 89 particles were restricted to 12 pre-determined connotation concepts. Each concept was inevitably assigned to more than one particle, and particles were assigned to more than one concept. The assigned concepts were not always compatible; for example, laat was labelled both ‘persuading’ and ‘reminding’. Equally problematic is that some particles, such as laa3, were found not to have ‘significant loading’ in any of the connotation concepts. In fact, only a small portion of the particles tested were found to have significant loading in any connotation concepts, which would seem to indicate that those particles have no meaning, and/or a defect in the test. The latter seems more likely, since it is improbable that the wealth of Cantonese utterance particles (Yau tested a total of 206 particles and clusters) fall into neat categories. In addition to these problems, the 12 labels represent fairly complex ideas to begin with, even in English (the language in which Yau was writing), which probably interferes with participants’ answers.

Another quantitative study was by Gibbons (1980), who also investigated the entire inventory of Cantonese utterance particles. Like Yau (1965), Gibbons appears to have believed that each particle could be straightforwardly categorised in fairly simple terms, and that the particles affected utterances in only one way. As with Yau’s study, all particles were considered under a single set of uniform criteria, and the broad classifications could not provide much information about particles’ complex meanings. Knowing that a speaker using a particular particle ‘strongly expects a response in terms of action’, for example, does not reveal much about the particle’s meaning. It is still unclear to readers what the speaker is trying to communicate. Furthermore, some descriptions were not very clear, such as when a particle is described as ‘weakly’ requiring a
Individual particles were mentioned only briefly, and described mainly in relation to other particles. Gibbons’ (1980) work appears to be concerned mainly with comparisons and placing particles in positions relative to each other. However, the meaning of each particle is lost as soon as it is considered in isolation. Yau and Gibbons both overlooked the fact that Cantonese utterance particles can attach to a wide range of different utterances with different effects.

The approach used by Yau (1965) and Gibbons (1980) is like the ‘scalar’ approach to the study of particles that Wierzbicka (1986, 527-528) criticises, where particles are placed in positions within a semantic ‘continuum’. Wierzbicka mentions in particular Tsuchihashi’s (1983) attempt to place Japanese particles along a continuum between declarative and interrogative, which is strikingly similar to Yau’s S-Q test. Wierzbicka describes these descriptions as being unhelpful from any practical point of view, saying that it seems naïve and simplistic to try to explain interrelations between particles by representing them in quantitative and graphical terms without first attempting to analyse them in qualitative, conceptual terms. Wierzbicka accepts that arranging a small number of particles along a single dimension may have some degree of plausibility, but Yau (1965) and Gibbons (1980) both considered all Cantonese utterance particles in their studies.

If Cantonese utterance particles have meaning, those meanings have not been identified clearly using the methods discussed so far. As Matthews and Yip (2011, 389) put it:

While these particles are one of the most intensively studied areas of Cantonese grammar, their functions remain elusive, and rather different conclusions are drawn according to [the approach taken]. While the function of a particle may seem clear enough in a specific context, it is extremely difficult to identify a common denominator underlying all its uses.

Despite this, it remains possible to believe that particles may have stable, invariant meanings, if it is possible to separate the invariant meaning of the particle from the meaning of the utterance to which it is attached.
Are the meanings of Cantonese utterance particles stable and invariable?

One of the main problems affecting existing descriptions of Cantonese utterance particles is that scholars have not differentiated between the particle’s meaning and the meaning of the utterance to which it is attached. One clear example is found in Meyer and Wempe (1947, 287), where they define laa1 as ‘implying completion; certainty, or urgency’, and they provide two examples. The first is zou6 hou2 laa1 [lit. ‘do complete/good laa1’] or simply ‘done’, which they translate as ‘it is finished’. The second is jat1-ding6 laa1 [lit. ‘definitely laa1’], which they translate as ‘certainly’. Note that ‘completion’ and ‘certainty’ would appear to be due to the meaning of the rest of the utterances, and not the particle laa1. Each utterance would clearly imply ‘completion’ and ‘certainty’ even without a particle. Taking into account the meaning of the utterance can lead people to believe that each particle has numerous meanings. One of the goals of this research must be to identify in particles that specific meaning which is invariable.

Similarly to this thesis, Kwok was interested in isolating the ‘core meaning’ of each particle, i.e. the ‘meaning shared by all occurrences of the particle in different contexts’ (Kwok 1984, 13-14). She understood that the same particle could behave somewhat differently, and that the more specific meanings in each case would depend on context. Kwok and Luke can actually be said to have had two very different ways of trying to describe Cantonese utterance particles accurately. Luke believed that the solution to the problem was to find the full range of uses of each particle and identify the various possible contexts, while Kwok aimed to find the ‘core’ meaning of each particle.

The particles in Kwok’s study were divided into two main groups, or four sub-groups, depending on the type of sentence they attached to\textsuperscript{18}. The particles were able to occur in more than one of her categories, and she gave short

\textsuperscript{18} The first main group contained particles and particle clusters which are ‘non-status changing’. This main group was sub-divided into three sub-groups. The first sub-group consisted of those particles and particle clusters which attached to statements without altering their grammatical status as statements. The second sub-group were those attached to questions without altering their grammatical status as questions. The third were those attached to commands and requests without altering their grammatical status as commands and requests. The second main group (and fourth sub-group) consisted of those particles and particle clusters which when attached to statements changed those utterances into questions (Kwok 1984, 20-21, 41-42).
analyses and descriptions of each particle in each type of sentence. Therefore, her work made the basic distinction that the particles act differently in different types of sentences. In this way, then, Kwok seemed to imply that the Cantonese utterance particles do mean different things in different contexts, and this created a certain tension with her aim of isolating core meaning. As she stated, ‘it is no easy task to isolate the inherent meaning from the peripheral meanings which depend on the circumstances in which the particle is actually used’ (Kwok 1984, 13-14).

Wakefield (2011b) also believed in the importance of isolating Cantonese utterance particles’ ‘core’ meanings, agreeing that descriptions of utterance particles are inadequate when they include meanings that come from the sentences or discourse contexts in which they occur. This would be similar to supposing that the plural morpheme ‘s’ in ‘cats’ means ‘more than one cat’, which would obviously be undesirable because it includes the meaning of ‘cat’ in the meaning of ‘s’ (Wakefield 2011b, 71-74). Wakefield saw this kind of faulty reasoning as the reason so many scholars have concluded that utterance particles have no meanings independent of context.

Wakefield speculated that each particle ‘contain[s] propositional and discourse-related deictic elements, which make their meanings appear to change from context to context’ (Wakefield 2011b, 70-71). He did not suggest that the explications of utterance particles should be free of references to discourse context, but rather that they should accurately represent such connections (while not including any meanings that are specific to particular sentences) (Wakefield 2011b, 4-5). By building in deictic elements, he was able to bring together the ideas that particles have invariant meanings, on the one hand, but can do very different things in different discourse contexts, on the other. As will be explained below, Wakefield’s ultimate goal was to equate Cantonese utterance particles with English intonation.

There is clearly still a need for thorough and accurate semantic analyses of Cantonese utterance particles. I agree with Wakefield (2011b, 18) that despite some useful studies, precise definitions of Cantonese utterance particles are still

---

19 For example, aa3 occurred in both the first and second sub-categories, and aat occurred in both the second and third.
lacking. Nonetheless, the existing discussion supports the idea of the importance of Cantonese utterance particles, the need for their research and analysis, and that scholars are aware of certain difficulties in studying them.

As a final point on the question of whether particles have invariant meaning, it can be helpful to briefly consider the particles of Singapore English, which are often accepted as behaving similarly to those of Cantonese (and, as noted earlier, it is even possible that some Singapore English particles are influenced by Cantonese utterance particles). In relation to Singapore English particles, Kwan-Terry (1978, 70-71) wrote as follows:

While it is true that the meaning of a particle depends in part on the context, it is difficult to maintain that a particle has no basic meaning of its own, for if that were the case, it would be near impossible for anyone to learn how to use it. It would seem to be more reasonable to hypothesise that each particle has one or more basic meanings which interact with the context to give rise to more specific meanings.

Kwan-Terry is not the only researcher into Singapore English particles who has suggested this. It has also been said that while the particles ‘combine with social and linguistic context to create a total meaning for the utterance’, the particles have a ‘consistent force’ (Gupta 1992, 35). This is what is referred to here as the invariant meaning. Gupta believed apparent differences in function resulted from the attachment of the particle to different sentences. The change associated with each particle is then an interaction between the particle's invariant meaning and a given context. Besemer and Wierzbicka (2003, 17) also explicitly support this approach, adding that it is only by assigning each particle its own ‘consistent force’ that one can hope to explain each particle’s unique range of use.

1.5.2 Cantonese utterance particle combinations

One of the research questions explored in this thesis is to what extent patterns of particle combinations in Cantonese can be explained semantically; this is the focus of Part Three. As explained in section 1.3.2, this thesis makes a distinction between putative combinations, including contractions, that are supposed to
result in monosyllabic particles (discussed in Chapter 8), and surface combinations of polysyllabic ‘particle clusters’ (discussed in Chapter 9).

Particle combinations are often claimed to contain the combined meaning of the elements of which they are comprised (Yau 1965, Gibbons 1980, Kwok 1984, 8-15, Yip and Matthews 2000, 131-132, Matthews and Yip 2011, 396-397, Wakefield 2011b, 13-14). However, this claim does not seem to have been rigorously and reliably tested. According to Wakefield (2011b, 13-14), all studies on Cantonese utterance particles appear to agree that this is the case, presumably based on native speaker intuition and assumption. Testing of such a claim would be a futile or impossible exercise if the meanings of the particles or morphemes that supposedly make up the combinations were not first accurately identified. NSM provides a new approach to the idea that the meaning of particle clusters is equal to the meaning of the individual particles combined (see also a preliminary investigation in HHL Leung (2013, 18-24)).

The specific issues regarding (supposed) particle combinations are complex, and best deferred until Part Three. Chapter 8 provides a detailed review of putative particle contractions and sub-syllabic morphemes, and introduces the idea that Cantonese utterance particles may have some qualities of or similarities with phonesthemes and sound symbolism. Chapter 9 looks at particle clusters using the results of Part Two, highlighting the advantages of using NSM and corpus data.

1.5.3 Summary

For a long time, the nature of particles’ meanings was not clearly understood. Some considered them to have no meanings, and others struggled to pinpoint their meanings in concrete ways. Traditional ideas and the influence of Mandarin played a part, as well as the overly ambitious nature of large-scale studies, and emphasis on quantitative methods rather than qualitative analysis. Generally speaking, more recent studies of Cantonese utterance particles have become more focused, investigating smaller numbers of particles, and research methods have become more descriptive, acknowledging the complexities and seemingly changing nature of the same particles.
The research questions outlined in section 1.4 address current gaps in our understanding of Cantonese utterance particles. While accepting that each particle can do an array of things, they still might have stable core meanings; hence the leading research questions are concerned with whether Cantonese utterance particles have identifiable, invariant meanings. It is only after finding those meanings that the research questions concerning particle combinations can be addressed. This idea that Cantonese utterance particles may have invariant meanings is still a fairly recent development. Obviously, testing this depends on having a reliable method of semantic analysis. This thesis uses the innovative NSM approach as its primary tool of semantic analysis.

1.6 Methodology used in this study: the Natural Semantic Metalanguage (NSM)

As we have seen, even those scholars who believe that Cantonese utterance particles do have meanings struggle to state exactly what their meanings are. A method is needed to pinpoint and clearly state any invariant meanings for any particles and combinations. This section addresses the methodology used in this study, namely the Natural Semantic Metalanguage. The Cantonese-specific NSM is explored in Chapter 2.

1.6.1 Outline of NSM principles and practice

NSM is a framework for constructing explications and cultural scripts using ‘semantic primes’: a set of commonly used, everyday words which are posited to represent the most basic meanings. The primes provide a finite inventory of primitive terms, the need for which was recognised as early as the 17th century (Goddard and Wierzbicka 2014, 10-11). There have been numerous studies of NSM by authors in many languages. Major works include Wierzbicka (1996), Goddard and Wierzbicka (1994, 2002, 2014), Peeters (2006), and Goddard (2008). A list of semantic primes is given in Chapter 2, which also proposes Cantonese exponents.

Semantic primes are, by definition, meanings that cannot be further broken down or paraphrased more simply, i.e. without circularity. The goal of NSM analysis is to decompose complex meanings into explications made up of
semantic primes via reductive paraphrase. NSM scholars believe that it is not helpful to have the meanings of everyday expressions, such as particles, expressed in obscure or esoteric terms which are themselves not defined well. Any explication should be framed in terms which are clearer and less complex than the expression being defined. Adopting reductive paraphrase to find meaning ensures meaning descriptions which are simple, non-circular, and testable. This approach allows for maximum clarity and cross-translatability, thus addressing the main problems with most other semantic methods.

Importantly, the semantic primes are expected to have semantic equivalents in all natural languages, which means that explications and cultural scripts can be readily translated into any other language to give an identical meaning. Empirical studies have confirmed that semantic equivalents of the primes exist in a wide range of geographically and typologically diverse languages, and even share a universal syntax. These languages include Amharic, Arabic, Danish, East Cree, Ewe, Finnish, French, Japanese, Korean, Koromu, Lao, Malay, Mandarin, Mbula/Mangaaba-Mbula, Polish, Russian, Spanish, Tok Pisin, and Yankunytjatjara (Goddard 2011, 68, Goddard and Wierzbicka 2014, 12). The NSM approach has been applied to the semantic description of a wide range of lexical and grammatical phenomena in many languages. Another advantage of using NSM is that, due to its cross-translatability, it can facilitate cross-cultural communication and language learning.

The explications in this thesis should, in particular, be understandable by native speakers of Cantonese. It would not make sense for the definition of a Cantonese word to not be available in Cantonese itself. If a definition were impossible to express in Cantonese, this would indicate that it represents an ‘outsider’ point of view and does not match the real thoughts of native speakers. To achieve an insider’s perspective, it is better for languages to be described in their own terms. Presenting explications in Cantonese ensures that they can really explain Cantonese words from a native speaker’s point of view. For this reason, the final proposed explications for all the particles and particle clusters in this thesis will be presented in parallel English and Cantonese versions. This, of course, requires that the Cantonese semantic primes are first identified. This is addressed in Chapter 2, which focusses on Cantonese NSM.
1.6.2 Using NSM to describe particles

Use of NSM and reductive paraphrase is ideal for the analysis of particles. The NSM approach has been applied to particles in various other languages, including Singapore English (Besemer and Wierzbicka 2003, Wong 2004, 2005, 2014), Mandarin Chinese (Chappell 1991), Mparntwe Arrernte (an Aboriginal language spoken in Central Australia) (Goddard 2011, 176-180), Malay (Goddard 1994a, 2001), Polish (Wierzbicka 2003, 371-389), French, (Waters 2010), Ewe (a Niger-Congo language) (Ameka 1991), Colombian Spanish (Travis 2005), Arabic (Rieschild 2011), English (Goddard 2011, 167-175, Wierzbicka 2003, 345-370), and Cantonese (Wakefield 2011b, 2012a, b, HHL Leung 2012, 2013) (see section 1.6.3). This also shows that particles offer a particularly fruitful field for semantic analysis.

These analyses demonstrate that many particles express bundles of thoughts, wants and feelings. They back Wierzbicka’s (2003, 342-343) theory, derived from the 17th century thinkers John Locke and Gottfried Leibniz, that utterance particles can be regarded as abbreviations for whole sentences, representing actions of the mind and having ‘illocutionary forces’. Wierzbicka states that illocutionary forces are bundles of assumptions, intentions, and ‘postures of the mind while discoursing’ (in Locke’s phrase). As such, the only possible way to represent these illocutionary forces accurately is to decompose them; i.e. to reconstruct the sentences they stand for.

Related to this is the fact that NSM explications can be written from the speaker’s point of view, using the primes ‘I’ and ‘you’ to shed light on the relationship between a speaker and addressee (Besemer and Wierzbicka 2003, 8, Wakefield 2011b, 75-76). This is very important because Cantonese utterance particles concern the interaction between the speaker and addressee. As Besemer and Wierzbicka (2003, 8) explain, ‘the only mode of analysis which would allow testing in context is one adopting a first person format: not a meta-comment on what the speaker is doing but an attempted explication of what the speaker is saying – that is, an explication adopting the speaker’s, not an external observer’s, point of view.’

Using NSM, this study will attempt to identify and isolate the invariant meanings of Cantonese utterance particles, using reductive paraphrases made of
semantic primes. The goal in each case is to find an NSM explication that can be applied to all instances of the particle under investigation. To begin with, it will be assumed that a single invariant meaning can be found for each utterance particle. If this is found to not be possible, only then should polysemy be considered. The presence of polysemy would not be surprising, as numerous words in numerous (if not all) languages have multiple meanings, but in the case of Cantonese utterance particles, setting out in advance with the assumption of multiple meanings for each particle would almost certainly result in varied and opposing definitions, which would not be very helpful.

As mentioned, it is significant that explications of particles using natural language can be tested by substitution in place of the particles they represent. Drawing on the work of Leibniz (1949), Wierzbicka (2003, 343) explains that if particles are abbreviations for whole sentences, ‘the possibility or otherwise of substituting the reconstructed sentence for the particle in question provides an empirical test of the adequacy of the proposed explication’. Testing by substitution is used throughout this thesis, in conjunction with the data from the Hong Kong Cantonese Corpus (see section 1.7). A range of possible explications are constructed for each particle, based on native speaker interpretation of numerous examples of the particle in use. By substituting the proposed explications into varied examples found in the corpus, we can see whether the proposed meaning is likely to have been intended by the speaker, and whether it makes sense in context. Consideration of plenty of examples helps to reject incorrect components, and the explication can be gradually fine-tuned until it can apply to all instances of the particle. Since the meaning of the particle should not include meanings specific to certain utterances or contexts, studying a variety of different examples is crucial. The metalanguage is very versatile and can formulate extremely precise explications, despite using simple words. (Chapter 7 in particular is laid out in a way that demonstrates the process of finding an explication by using NSM and the substitution tests.)

1.6.3 Previous studies of Cantonese utterance particles using NSM

A modified NSM was first applied to four Cantonese utterance particles by Wakefield (2011b, 2012a, 2011a) in the process of finding their equivalents in English intonation. His work had a different focus to the present study, as he
was mainly concerned with matching the meanings of Cantonese particles to intonation patterns in English. Wakefield started with the belief that Cantonese utterance particles could have exact English equivalents in the form of intonation contours, which he treated more or less as lexical items. He relied on bilingual speakers of English and Cantonese to ‘translate’ certain Cantonese utterances with particles into English, and recorded their intonation contours. He then examined the pitch contours of Cantonese-to-English audio translations, and found that four Cantonese utterance particles appeared to be comparable to specific English pitch contours. He reasoned that if language learners knew that an utterance particle were equivalent to a particular form of intonation in their native language, they would intuitively understand its meaning and function.

Though this is an interesting concept, there are several potential problems with some of Wakefield’s apparent assumptions. His basic starting point is the idea that Cantonese utterance particles are, to a significant degree, the counterpart to English intonation contours. Although this notion makes some sense intuitively, it is problematic to simply equate a specific English intonation pattern with a specific Cantonese utterance particle, because of several fundamental complications. Firstly, while lexical tone puts limits on Cantonese speakers’ use of intonation, use of intonation and other qualities of the voice are still possible to an extent and cannot be completely ruled out. It is possible to, for example, place more stress or emphasis on certain words, to shout, or even to turn a declarative sentence into a question without changing the words or structure of the sentence.

Secondly, many English speakers from different places and backgrounds use intonation very differently from one another, and even within one country there may be considerable variations. It is possible, and even likely, that different English speakers would produce different intonation patterns in the same context. In his analysis of the English equivalent of *lo1* (Wakefield 2011b, 111), Wakefield noted that one participant’s translation used a different pitch contour from the others, and that it seemed not to be a dialectal difference but a difference in personal speaking style related to personality. Though Wakefield stated that this participant’s utterances sounded like they carried the same
connotative meanings as the others’, Wakefield admitted that this was a complicating factor which he had not considered prior.

Finally, questions remain as to whether all utterance particles in Cantonese have equivalents in English intonation. Wakefield conducted in-depth study of four selected particles. Is it plausible that English will have distinct intonation contours for 30 or more particles? Wakefield (2011b, 250-251) himself recognised that this might not be possible. Furthermore, how can Wakefield’s claims be extended to include particle combinations? Presumably, intonation contours cannot be combined simultaneously, in which case an even larger number of intonation contours would need to be identified to account for the total number of particles and particle combinations. On such a large scale, the feasibility of finding equivalents is severely complicated.

In his process of analysis, Wakefield first produced explications for the particles he studied, using a modified NSM. Two examples from Wakefield (2011b, 107, 150) are given below. As will become clear in Part Two, these explications look rather different to the explications that will be proposed in the present thesis (for different particles). His use of P (and P2) for ‘proposition’ and D (and D2) for ‘discourse’ items were his innovation and are not usual in NSM work. They were intended to ‘link’ references in the explication with particular parts of the preceding context (as explained above), but give the explications an unwelcome technical quality and, especially when more than one P and/or D are involved, they can be hard to process. In the present thesis, anaphoric references are accomplished without any such supplementary devices.

P+lo1-equivalent intonation:
   a. you can know this (P)
   b. because you know something else (D)

P+aaimaa3-equivalent intonation:
   a. you can know this (P)
   b. because you know something else (D)
   c. I want you to think about this (P) now
   d. after this, you will not think this (D2)
   e. you will think something else (P2)
Another peculiarity of Wakefield’s work was a tendency to ‘embed’ certain particles within certain other particles. For example, he proposed that the meaning of the particle *lo1* is part of the meaning of *aa1maa3*. He also believed that the meaning of *zaa3* is entailed in and part of the meaning of *ze1* (Wakefield 2012b) (see Chapter 7). This opens the door for circularity.

Some NSM semantic analyses of Cantonese utterance particles have also been published by the author (HHL Leung 2012, 2013), which are earlier versions of some of the ideas presented in this thesis.

**Comparable particles in other Southeast Asian languages**

The NSM approach has been used in studying the Singapore English particles *la/lah, wut, meh, one, á, and lör* by Besemeris and Wierzbicka (2003) and Wong (2004, 2005, 2014). Malay *lah* and *pun* have been analysed using NSM by Goddard (1994b, 2001, 2011, 180-183). Chappell (1991) has studied Mandarin *me*. As mentioned, this thesis makes no attempt to explore the historical origins of these particles or to compare their meanings or uses. Nevertheless, the NSM explication of one particle each from Singapore English and Mandarin are given below as samples of typical particle explications.

Firstly, Wong (2014, 241) tells us that there are two senses of *là*, which can be described as follows.

It it like this *là*:

you do not think like this about it, ‘it is like this’
maybe you do not think that it can be like this
I think that it will be good if you think like this
I want you to think this now
you don’t have to think other things about it

Do this *là*:

you do not think that you want to do this about it
maybe you do not think that it can be like this
I think that it will be good if you do this
I want you to do this now
you don’t have to think other things about it

Secondly, Chappell (1991, 50, 56-57) proposes the following explications of the Mandarin *me*. They are written in an older version of NSM, and include some terms (such as ‘understand’ and ‘state of affairs’) which would not be allowable in current versions of the metalanguage. As well, the syntax of several
components is significantly more complex than would be allowed in current NSM.

The *me* of an obvious logical connection:

I say: Proposition X  
because I want you to understand why state of affairs Y happened  
I think you can understand why this state of affairs came to be  
I say X has to be the cause of Y  
I say this is true  
There’s nothing more to be said about it

The *me* of disagreement in face of a self-evident situation:

I say: Proposition X  
because I don’t think this state of affairs (Y) should have happened  
I say X is true  
because anyone can see it’s true  
There’s nothing more to say about it

[I feel something bad towards you/they just now  
because you/they should know that X is true too]

Since these are different languages, there would be little point in attempting close comparisons between these particles and Cantonese ones. However, a general point of difference between these explications and those to be proposed in the present thesis is that these explications are much longer.

1.7 Data: the Hong Kong Cantonese Corpus

The data and examples used in this study are mainly taken from the Hong Kong Cantonese Corpus, which consists of naturally-occurring, everyday Cantonese conversations (see Luke and Wong in press). The Hong Kong Cantonese Corpus contains about 30 hours of recording and 180 000 word tokens (as opposed to Chinese characters). The data is made up of informal, spontaneous speech either in ordinary settings among family, friends and colleagues, or from radio talk shows, and was recorded in Hong Kong between the late 1990s and early 2000s. There were some 100 Hong Kong Cantonese speakers, mostly aged in their 20s and 30s, with roughly equal female and male gender distribution. Such corpus data satisfies the criteria identified above. The majority of examples used in this thesis are from the ordinary settings among family,

---

20 The last component was placed in parentheses by Chappell as it is not an invariant feature of meaning.
friends and colleagues, rather than from the radio talk shows. The names mentioned in the conversations were changed in the transcription in the corpus.

All English glosses and translations which accompany the corpus data are my own. The data in the Hong Kong Cantonese Corpus is reported to have a four-line format: the original transcription in Chinese characters, the romanised version, a word-for-word gloss in English, and a free translation of each sentence. The version of the corpus that I have access to only shows the first two, i.e. the transcription in Chinese characters and the romanised version, so all of the word-for-word glosses in English and all of the English free translations in this thesis are my own work, and may differ slightly from that translated by others. I did not include the Chinese characters mainly because colloquial spoken Cantonese, including utterance particles in particular, does not always have standardised character forms (see section 1.2). Since I use a well-known romanisation system, this should not cause problems, and may be even more clear, for example, in notating differences between particles which differ only in tone. The corpus includes audio files.

Luke (1990) was the first to use and analyse real examples of Cantonese utterance particles from naturally-occurring, ordinary conversations. As he noted, invented examples are no substitute for spontaneous conversation (Luke 1990, 2), and his view is shared by Luke and Nancarrow (1997) and by Wakefield (2011b). Cantonese utterance particles are used primarily in informal or colloquial speech, so it is best if the data used to analyse them are from informal, everyday conversation.

Although Kwok (1984) did have a small corpus consisting of about two hours of telephone conversations, she acknowledged that she used this mainly to identify the particles and to describe their distribution. She instead preferred a more speculative and subjective approach in determining the meanings the particles convey, which then became the core part of her work (Kwok 1984, 3-4). Similarly, Yau (1965) used sixteen hours of recorded data, but only for the purposes of identifying the particles at the beginning of the study. He then relied on the intuitions of native speaker participants in his questionnaires.

The Hong Kong Cantonese Corpus was chosen as the data for this study for several reasons. Aside from featuring informal, ‘everyday’ speech of Hong Kong
Cantonese speakers, it is public and accountable, and has also been used by others in the field, such as Luke (1990) and Wakefield (2011). As well, other corpora of Cantonese were not accessible when the project was started, despite efforts to gain access. I agree with Luke (1990, 26) that the potential problems of the observer’s paradox are unlikely to affect corpus studies of Cantonese utterance particles, since it is unlikely that small linguistic details such as particles can be consciously monitored for extended periods of time. Furthermore, as mentioned, Cantonese speakers do not seem to be consciously aware of the role of utterance particles.21

When conducting the analyses, the data and examples were mostly read in order of their appearance in the conversations in the corpus. The examples ultimately chosen for inclusion in the thesis, particularly in Parts Two and Three, were selected to show as wide a range of different scenarios as possible. For example, in Chapter 3, the chosen examples of laa1 showed certainty in some cases, and uncertainty and approximations in other cases, especially as this had been a point of disagreement in the existing literature. This reflects one of the goals of the explications, which is to find an invariant meaning that is relevant and accurate in a wide range of contexts. Another consideration in choosing examples for discussion was that they were relatively self-explanatory, or at least not too difficult to explain, so that not too much background information was needed. This was partly for ease of exposition, and partly to ensure that the context could be accurately taken into account. For examples discussing individual particles, most of the examples selected showed the particle being used on its own, i.e. not in a cluster. This was done in case its meaning in the cluster was different from its meaning as an independent particle. The exception to this was Chapter 7, which laid out the semantic analysis differently, and also because zaa3 is the least frequently used of the particles investigated in this thesis, so that examples of it being used independently were slightly fewer. In a small number of cases, utterances were

21 Chan (2002) found that in two episodes of a Cantonese-language mainland television series filmed in Guangzhou, what was actually said in the television production was not identical to the corresponding script. The most noticeable difference was that the written scripts contained fewer utterance particles than the final television production. This indicates that the actors and actresses used utterance particles naturally and spontaneously, including them even when not scripted. This supports the idea that use of utterance particles is not fully consciously controlled.
excluded which featured potentially offensive language (even if apparently intended to be jokes or teasing among friends).

1.8 Thesis outline

This thesis consists of ten chapters grouped into four parts. Part One includes the current chapter and the next, which investigates Cantonese NSM.

Part Two includes five chapters and is the bulk of the semantic analysis of individual Cantonese utterance particles. Each chapter focusses on a different particle; they are laa1, wo3, gaa3, laa3, and zaa3. Each begins with an overview of the literature relevant to that particle, and then presents examples of the particle taken from the Hong Kong Cantonese Corpus and discussion of the proposed NSM explications. Often, several versions of explications or alternative components are discussed to give insight into the process of semantic analysis. The final proposed explications are presented in both English and Cantonese. Chapter 7, which investigates the semantics of zaa3, is presented differently, in a way that is intended to guide the reader through the process of finding an NSM explication.

Part Three explores combinations of Cantonese utterance particles. It consists of Chapters 8 and 9. The first assesses and critiques supposed combinations which result in monosyllabic particles. The second investigates polysyllabic particle clusters and draws on the particles and NSM explications discussed in Part Two.

Finally, Chapter 10 makes up Part Four, the conclusion to this thesis.
Chapter 2: The Cantonese Natural Semantic Metalanguage

The Natural Semantic Metalanguage (NSM) has been introduced in section 1.6. As mentioned, it is one of the goals of this thesis that the proposed explications should be understandable by native speakers in their own language. This chapter therefore presents a detailed investigation into the viability of a Cantonese-based NSM, addressing the methodological preliminary identified in Chapter 1. The main project is to identify Cantonese equivalents, also termed exponents, of all 65 primes. Although in the chapters that follow, only a subset of these are used in the explications of utterance particles, finding Cantonese exponents of all NSM primes contributes to the broad research program into NSM, reinforces its universality, and lays the groundwork for future NSM studies of Cantonese across all areas of the lexicon and grammar.

Section 2.1 starts by briefly outlining previous explorations into Chinese NSM. Following this, a table is provided of the 65 current proposed semantic primes, in English and Cantonese. Then, some Cantonese-specific issues and complexities are pointed out. The main body of this chapter (sections 2.2–2.12) considers each of the primes, and explores whether there are any complications in using them in Cantonese. This is the first attempt to investigate the full set of Cantonese semantic primes. Naturally, some primes were found to be more problematic than others, with particularly problematic Cantonese exponents including PLACE~WHERE, TIME~WHEN, VERY, and BAD. Other issues not encountered in English include, for example, the need to use classifiers in certain constructions.

2.1 Cantonese NSM

2.1.1 Chinese NSM primes previously identified

Full sets of Mandarin exponents have been proposed by Chappell (1994, 2002), with further discussion of Mandarin HAPPEN provided by (Tien 2009). As explained in section 1.2, however, everyday, spoken Cantonese is quite different
to Mandarin. For Cantonese, a limited number of primes related to time and space were proposed by Tong, Yell, and Goddard (1997, 249-251). Wakefield (2011b) also proposed nine Cantonese primes which he needed in his explications, and in a limited number of other cases he translated Chappell’s (2002) Mandarin exponents into Cantonese (Wakefield 2011b, 76). With 65 semantic primes currently recognised, this left a considerable number of primes with no identified Cantonese exponent. A complete list of Cantonese exponents was first published in HHL Leung (2012, 279-280), with a revised second version in HHL Leung (2013, 8-9). Several minor updates have since been made, and the latest inventory is presented in full below. Importantly, the previously published lists of Cantonese exponents were not accompanied by any discussion of the exponents and how they were chosen, or any potential complications. The present chapter is therefore the first attempt at a discussion of the full set of Cantonese NSM semantic primes. Some exponents proposed here match those of the scholars mentioned above, while others differ. The next section gives a table of the Cantonese and English semantic primes. Following that, section 2.1.3 considers some Cantonese-specific NSM issues.

2.1.2 Cantonese exponents of NSM primes

The list of 65 NSM semantic primes is given in Table 12, showing Cantonese and English exponents. The English exponents and their groupings follow the latest list published by leading NSM experts (Goddard and Wierzbicka 2016). Some Cantonese exponents are polysyllabic (the Chinese writing system can sometimes lead people to believe that all Chinese words are monosyllabic, but this is not the case). Some primes have allolexes, where words are semantically equivalent despite behaving differently grammatically. Following convention, allolexes are indicated by a tilde, e.g. English I~ME, MUCH~MANY. In Cantonese explications and cultural scripts, classifiers will sometimes be needed which are not shown in this table; the most likely classifiers are go3 and joeng6, as they are neutral and non-specific.

---

22 A full list was also provided by the author to the NSM webpage and is available at https://www.griffith.edu.au/humanities-languages/school-humanities-languages-social-science/research/natural-semantic-metalanguage-homepage/resources. It can be found alongside tables of semantic primes in many other languages, as well as other NSM resources.

23 See e.g. DeFrancis (1984).
Table 12: Cantonese semantic primes, with English equivalents

<table>
<thead>
<tr>
<th>Cantonese</th>
<th>English</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngo5</td>
<td>I~ME</td>
<td>Substantives</td>
</tr>
<tr>
<td>lei5</td>
<td>YOU</td>
<td></td>
</tr>
<tr>
<td>jan4</td>
<td>SOMEONE</td>
<td></td>
</tr>
<tr>
<td>je5</td>
<td>SOMETHING~THING</td>
<td></td>
</tr>
<tr>
<td>jan4</td>
<td>PEOPLE</td>
<td></td>
</tr>
<tr>
<td>san1 tai2</td>
<td>BODY</td>
<td></td>
</tr>
<tr>
<td>zung2</td>
<td>KINDS</td>
<td>Relational</td>
</tr>
<tr>
<td>bou6 fan6</td>
<td>PARTS</td>
<td>substantives</td>
</tr>
<tr>
<td>li1</td>
<td>THIS</td>
<td></td>
</tr>
<tr>
<td>tung4 jat1</td>
<td>THE SAME</td>
<td>Determiners</td>
</tr>
<tr>
<td>ling6 jat1~ling6 ngoi6</td>
<td>OTHER~ELSE</td>
<td></td>
</tr>
<tr>
<td>jat1</td>
<td>ONE</td>
<td></td>
</tr>
<tr>
<td>loeng5</td>
<td>TWO</td>
<td>Quantifiers</td>
</tr>
<tr>
<td>di1</td>
<td>SOME</td>
<td></td>
</tr>
<tr>
<td>dou1~cyun4 bou6</td>
<td>ALL</td>
<td></td>
</tr>
<tr>
<td>do1</td>
<td>MUCH~MANY</td>
<td></td>
</tr>
<tr>
<td>siu2</td>
<td>LITTLE~FEW</td>
<td></td>
</tr>
<tr>
<td>hou2</td>
<td>GOOD</td>
<td>Evaluators</td>
</tr>
<tr>
<td>m4 hou2</td>
<td>BAD</td>
<td></td>
</tr>
<tr>
<td>daai6</td>
<td>BIG</td>
<td>Descriptors</td>
</tr>
<tr>
<td>sai3</td>
<td>SMALL</td>
<td></td>
</tr>
<tr>
<td>zii (dou3)</td>
<td>KNOW</td>
<td></td>
</tr>
<tr>
<td>lam2</td>
<td>THINK</td>
<td></td>
</tr>
<tr>
<td>soeng2</td>
<td>WANT</td>
<td></td>
</tr>
<tr>
<td>m4 soeng2</td>
<td>DON’T WANT</td>
<td>Mental predicates</td>
</tr>
<tr>
<td>gok3 dak1</td>
<td>FEEL</td>
<td></td>
</tr>
<tr>
<td>gin3 (dou2)</td>
<td>SEE</td>
<td></td>
</tr>
<tr>
<td>teng1 dou2</td>
<td>HEAR</td>
<td></td>
</tr>
<tr>
<td>gong2~waa6</td>
<td>SAY</td>
<td>Speech</td>
</tr>
<tr>
<td>zii6</td>
<td>WORDS</td>
<td></td>
</tr>
<tr>
<td>zan1</td>
<td>TRUE</td>
<td></td>
</tr>
<tr>
<td>Cantonese</td>
<td>English</td>
<td>Notes</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>zou6</td>
<td>DO</td>
<td>Actions, events, movement</td>
</tr>
<tr>
<td>faat3 sang1</td>
<td>HAPPEN</td>
<td></td>
</tr>
<tr>
<td>juk1</td>
<td>MOVE</td>
<td></td>
</tr>
<tr>
<td>hai2</td>
<td>BE (SOMEWHERE)</td>
<td>Location, existence, specification</td>
</tr>
<tr>
<td>jau5</td>
<td>THERE IS</td>
<td></td>
</tr>
<tr>
<td>hai6</td>
<td>BE (SOMEONE/SOMETHING)</td>
<td></td>
</tr>
<tr>
<td>(hai6) ngo5 ge3</td>
<td>(IS) MINE</td>
<td>Possession</td>
</tr>
<tr>
<td>sang1 cyun4</td>
<td>LIVE</td>
<td>Life and death</td>
</tr>
<tr>
<td>sei2</td>
<td>DIE</td>
<td></td>
</tr>
<tr>
<td>si4 hau6</td>
<td>TIME~WHEN</td>
<td></td>
</tr>
<tr>
<td>ji4 gaa1</td>
<td>NOW</td>
<td></td>
</tr>
<tr>
<td>zi1 cin4</td>
<td>BEFORE</td>
<td></td>
</tr>
<tr>
<td>zi1 hau6</td>
<td>AFTER</td>
<td></td>
</tr>
<tr>
<td>noi6</td>
<td>A LONG TIME</td>
<td></td>
</tr>
<tr>
<td>jat1 zan6</td>
<td>A SHORT TIME</td>
<td></td>
</tr>
<tr>
<td>jat1 dyun6 si4 gaan3</td>
<td>FOR SOME TIME</td>
<td></td>
</tr>
<tr>
<td>jat1 haak1</td>
<td>MOMENT</td>
<td></td>
</tr>
<tr>
<td>dou6</td>
<td>PLACE~WHERE</td>
<td></td>
</tr>
<tr>
<td>li1 dou6</td>
<td>HERE</td>
<td></td>
</tr>
<tr>
<td>seong6 gou1</td>
<td>ABOVE</td>
<td></td>
</tr>
<tr>
<td>haa6 min6</td>
<td>BELOW</td>
<td></td>
</tr>
<tr>
<td>jyun5</td>
<td>FAR</td>
<td></td>
</tr>
<tr>
<td>kan6</td>
<td>NEAR</td>
<td></td>
</tr>
<tr>
<td>bin6</td>
<td>SIDE</td>
<td></td>
</tr>
<tr>
<td>leoi5 min6~jap6 min6</td>
<td>INSIDE</td>
<td></td>
</tr>
<tr>
<td>dim3 (zyu6)</td>
<td>TOUCH</td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>NOT</td>
<td></td>
</tr>
<tr>
<td>ho2 lang4~waak6 ze2</td>
<td>MAYBE</td>
<td>Logical concepts</td>
</tr>
<tr>
<td>ho2 ji5</td>
<td>CAN</td>
<td></td>
</tr>
<tr>
<td>jan1 wai6</td>
<td>BECAUSE</td>
<td></td>
</tr>
<tr>
<td>jyu4 gwo2 (...zau6...)</td>
<td>IF</td>
<td></td>
</tr>
</tbody>
</table>
2.1.3 Complexities of a Cantonese NSM

One of the main considerations to take into account when working with Cantonese NSM is the issue of diglossia and the perceived low status of Cantonese compared with Mandarin. These are not issues which arise for English NSM or Mandarin NSM. Two related issues are: How ‘formal’ or ‘colloquial’ should Cantonese explications be? And does it matter that everyday, spoken Cantonese does not have a standardised, written form (using Chinese characters)?

The position adopted here is that Cantonese NSM explications should, as much as possible, emulate ordinary, everyday Cantonese as usually spoken, i.e. it should not be presented in the formal register, which is more like Mandarin. This is in keeping with the basic principles of NSM; the ‘N’ stands for ‘natural’, after all, and exponents of semantic primes should be the simplest and most basic words of a language. Notwithstanding the fact that the official language policy of Hong Kong is to be biliterate in English and standard written Chinese, explications should be intelligible to the average native speaker of Cantonese without the need for formal education or familiarity with other languages or speech varieties. It would also be far from ideal to describe everyday, colloquial words, such as utterance particles, using less natural or less familiar words or syntax. This means that some of the Cantonese primes and explications may sound informal and colloquial to some educated Cantonese readers, and may seem odd in an academic context, but this is a necessary and worthwhile side effect to stay true to the goals of NSM.

There are, however, two further complications with the goal of using the most natural Cantonese words and syntax in the explications. First, due to the

---

24 These and other issues related to Cantonese NSM and Cantonese explications were presented by the author and discussed at the April 2013 semantic workshop held at the Australian National University, Canberra.
significant mismatch between written Chinese and everyday, spoken Cantonese, as explained in section 1.2, Jyutping is used in place of standardised Chinese characters. This may mean that if the average native speaker (unfamiliar with romanisation, since none is taught in schools) were given Cantonese explications, they may need to be read out to them or be transposed into ‘written Cantonese’ (developed by Cantonese speakers, sometimes by creating characters which would not be taught in schools).

Secondly, there is the problem of the native speaker’s impulse to include Cantonese utterance particles in the NSM explications themselves. This is understandable given the pervasiveness of particles in ordinary Cantonese, but this is clearly unacceptable because particles do not have simple, cross-translatable meanings, not to mention the danger of circularity. Cantonese explications therefore need to be phrased in language which aims to be natural but without utterance particles. Though more difficult, the particle explications developed in this thesis show that this goal is attainable. Admittedly, some explications and components sound unnatural due to the strict restrictions of the NSM, which are necessary to maintain simplicity and universality. Many of the components which sound odd are similarly odd in the NSM of other languages, but this ensures maximum clarity and cross-translatability.

Discussion of the proposed Cantonese exponents follows. Many are fairly straightforward, not requiring much discussion, while others are more problematic, and several possibilities are discussed. The primes are discussed in the same order as in Table 12 above. The examples in the remainder of this chapter are constructed.

2.2 Substantive primes

The substantive primes of NSM are necessary in a great many semantic explications and cultural scripts, and the particles analysed in Part Two are no exception.

**Ngo5 – I~ME and lei5 – YOU**

Unsurprisingly, the NSM primes I~ME and YOU will feature heavily in the particle explications to be presented in Part Two. The Cantonese exponent for
I–ME is ngo5. Ngo5 is used whether it is the subject or object of the sentence. The Cantonese exponent for YOU is lei5. Lei5 is the neutral, singular form, as is the prime YOU. There is a more ‘respectful’ or ‘polite’ form which is less commonly used in Cantonese (it is pronounced the same way).

Lei5 can also be pronounced nei5. This variation is due to a general shift from initial n to initial l, with initial l preferred by the younger generations and in less formal speech (see e.g. Matthews and Yip (2011, 21-22, 36)). This sound change also affects, for example, li1/nii1, the Cantonese exponent for THIS.

**Jan4:** *(jat1) go3 jan4 / jau5 (go3) jan4 – SOMEONE*

The Cantonese SOMEONE is jan4. This is the same Cantonese word as for the prime PEOPLE, and it usually depends on context which is meant.\(^25\) The generic or neutral classifier usually required before jan4 is go3, and the most unambiguous way of specifying a single SOMEONE, as opposed to PEOPLE, is to say jat1 go3 jan4 [ONE CL SOMEONE]. This is often favoured at the beginning of a sentence or if the SOMEONE has not already been introduced. Go3 jan4 [CL SOMEONE] is also acceptable in many situations and is understood to be a single SOMEONE. Other possibilities for SOMEONE include jau5 jan4 [THERE-IS SOMEONE] and jau5 go3 jan4 [THERE-IS CL SOMEONE], which frequently mean SOMEONE rather than ‘there is someone’. Jau5 jan4 used without the classifier is more likely to be interpreted as an unspecified or hypothetical SOMEONE. Jau5 jan4 can also mean PEOPLE (possibly unspecified or hypothetical people) or THERE ARE PEOPLE, although the latter expression does not seem to occur very commonly in explications. The options for Cantonese SOMEONE are hence *(jat1) go3 jan4 and jau5 (go3) jan4*. We can regard these as allolexes of SOMEONE.

\(2.1\)

\(2.1a\)

*(Jat1) go3 jan4 teng1 dou2 di1 je5.*

**ONE CL SOMEONE HEAR CL THING**

**SOMEONE HEARS SOMETHING.**

\(^{25}\) Jan4 can also be translated as ‘human(s)’ or ‘mankind’. It is unusual for jan4 to be used for non-human ‘beings’ such as God(s) or ghosts.
(2.1b)

\[
\begin{array}{llllll}
Jau5 & (go3) & jan4 & tenq1 & dou2 & di1 & je5.
\end{array}
\]

THERE-IS CL SOMEONE HEAR CL THING

SOMEONE HEARS SOMETHING.

For comparison, the Mandarin exponents for SOMEONE are yŏu rén and shéi (Chappell 2002, 248-249). The former is translated into Cantonese as jau5 jan4, and the latter as Cantonese seoi4, which is formal and rare in spoken, conversational Cantonese. It is used in written Chinese, but its unnaturalness in everyday, spoken Cantonese means that it does not qualify as a Cantonese prime.

**Je5: (jat1) joeng6 je5 / jau5 (joeng6) je5 / (di1) je5 – SOMETHING~THING**

The Cantonese exponent for SOMETHING~THING is je5. As to be expected, je5 in ordinary Cantonese conversation can be used with many different classifiers depending on the kind of THING referred to. In NSM explications, generic classifiers should be used, so the appropriate classifier for je5 SOMETHING~THING would usually be joeng6. THIS THING is therefore li1 joeng6 je5, THE SAME THING is tung4 jat1 joeng6 je5, SOMETHING ELSE is ling6 jat1 joeng6 je5, and THIS OTHER SOMETHING is ling6 ngoi6 li1 joeng6 je5. Joeng6 je5 may be preceded by jat1 ONE or jau5 THERE IS, just as with jan4 SOMEONE. When preceded by jau5, the classifier is not always necessary. For example, SOMETHING HAPPENS/ED could be expressed as jau5 je5 faat3 sang1 [THERE-IS SOMETHING HAPPEN]. Another possibility is (jat1) di1, a classifier which indicates plural and also means SOME. Jat1 ONE is optional before it. I WANT TO SAY SOMETHING would be ngo5 soeng2 gong2 (jat1) di1 je5 [I WANT SAY (ONE) CL:SOME SOMETHING]. A question using ‘what’ would use mat1 je5 [what thing] (sometimes contracted to sound like me1 e5).

It may be of interest that the exponents identified by Chappell (1994, 112-113) for Mandarin are unusual in spoken Cantonese, being more formal and more likely to be used in writing. For syntactic reasons, three allolexes are necessary for SOMETHING in Mandarin, but Chappell’s (2002, 249-250) examples for each can all be translated into Cantonese using je5 in one of the phrases listed above. A distinction between physical and non-physical things is
more prominent in Mandarin. Cantonese je5 is acceptable for both (the other option in Cantonese for non-physical things is si6).

**Jan4 — PEOPLE**

As mentioned, the Cantonese prime for PEOPLE is jan4, with context required to differentiate between PEOPLE and SOMEONE. PEOPLE often requires no classifier, or classifiers such as dit, which is also the exponent for the prime SOME. If the classifier go3 is used, it would mean SOMEONE. As mentioned, jau5 jan4 [THERE-IS SOMEONE/PEOPLE] can be interpreted as SOMEONE as well as PEOPLE (including unspecified SOMEONE/PEOPLE). This means that in certain phrases, there may be ambiguity between SOMEONE and PEOPLE, although other parts of the explication may be able to clarify.

**San1 tai2 — BODY**

The prime BODY is expressed in Cantonese using san1 tai2. This is standard for human bodies, but may also be used for animal bodies. It cannot, however, be used for dead bodies. *Sei2 (zo2 ge3) san1 tai2 [DIE (PFV LP) BODY] is unacceptable. To refer to a ‘dead body’ or ‘corpse’, a Cantonese speaker would have to say either si1 tai2 or sei2 si1.

**Zung2 — KINDS**

The Cantonese exponent for KINDS is zung2. Zung2 is usually sufficient on its own. The similar zung2 leoi6 implies something more like ‘species’. Just as Chappell (2002, 308-309) states for the Mandarin counterpart zhōng, zung2 behaves syntactically like a classifier and thus readily co-occurs with determiners and quantifiers.

**Bou6 fan6 — PARTS**

The Cantonese exponent for PARTS is bou6 fan6. The classifier usually required before bou6 fan6 is the generic go3. See dou1 ALL on p.59 for ALL PARTS.
2.3 Determiner and quantifier primes

\textit{Li1: li1 (joeng6 je5) – THIS}

The semantic prime \textit{THIS} is expressed in Cantonese by \textit{li1}. \textit{Li1} can also be pronounced \textit{nii}, \textit{lei1}, or \textit{yii} (Matthews and Yip 2011, 21-22). As with \textit{lei5} \textit{YOU}, this reflects a pronunciation shift and not a difference in meaning. A classifier is often required between \textit{THIS} and the noun; for example: \textit{li1 go3 jan4} [\textit{THIS CL SOMEONE}] \textit{THIS SOMEONE}, \textit{li1 di1 jan4} [\textit{THIS CL:PL/SOME PEOPLE}] \textit{THESE PEOPLE}, and \textit{li1 joeng6 je5} [\textit{THIS CL THING}] \textit{THIS THING}. \textit{Li1 dou6} [\textit{THIS PLACE}] \textit{THIS PLACE} behaves differently (see \textit{PLACE~WHERE} on p.81). \textit{Li1 THIS} will feature in all of the explications to be proposed in Part Two.

In many cases, where English explications use \textit{THIS} on its own, it is necessary in Cantonese to say \textit{THIS THING}, i.e. \textit{li1 joeng6 je5} [\textit{THIS CL THING}]. Two examples, (2.2) and (2.3), are shown below. In both, \textit{joeng6 je5} is required after \textit{li1 THIS}. In these cases, \textit{li1 joeng6 je5 THIS THING} may be recognised as an allolex of \textit{li1 THIS} in Cantonese. Occasionally, \textit{li1 go3} (recall that \textit{go3} is the other generic classifier) also functions naturally as \textit{THIS THING}, but this seems to be applicable to a smaller number of explications.

(2.2)

(2.2a)

*\textit{Ngo5 zou6 li1.}
\textit{I~ME DO THIS}

*I DO THIS.

(2.2b)

\textit{Ngo5 zou6 li1 joeng6 je5.}
\textit{I~ME DO THIS CL THING}

I DO THIS THING.

(2.3)

\textit{Li1 joeng6 je5 hou2 hou2 /m4 hou2.}
\textit{THIS CL THING ‘dummy’/VERY GOOD BAD}

\textit{THIS THING IS (VERY)\textsuperscript{26} GOOD/BAD.}

\textsuperscript{26} See \textit{VERY} on p.46.
In explications and cultural scripts where it might sound odd to say THIS THING, another option is to use gam2 joeng2 THIS-WAY/LIKE-THIS (see LIKE~AS~WAY on p.88). This may be necessary where THIS refers to, say, a series of events. For example, in the common NSM phrase BECAUSE OF THIS, use of BECAUSE OF THIS THING is often not ideal. Additionally, IT IS LIKE THIS, another common NSM component, generally sounds odd as IT IS LIKE THIS THING. Jan1 wai6 gam2 joeng2 [BECAUSE THIS-WAY/LIKE-THIS] or hai6 gam2 joeng2 [IS THIS-WAY/LIKE-THIS] is usually necessary instead. The NSM expression I DON’T WANT THIS (not to be confused with ‘I don’t want to have this’) would be highly unnatural as *ngo5 m4 soeng2 li1 [I DON’T WANT THIS] or *ngo5 m4 soeng2 li1 joeng6 je5 [I DON’T WANT THIS CL THING]. One must say, e.g. *ngo5 m4 soeng2 gam2 joeng2 [I DON’T WANT THIS-WAY/LIKE-THIS] I DON’T WANT IT TO BE THIS WAY/LIKE THIS or ngo5 m4 soeng2 li1 joeng6 je5 faat3 sang1 [I DON’T WANT THIS CL THING HAPPEN] I DON’T WANT THIS THING TO HAPPEN.27

**Tung4 jat1 – THE SAME**

The Cantonese exponent of THE SAME is tung4 jat1. Another option, jat1 joeng6, was not chosen because phrases like THE SAME SOMEONE become problematic – use of jat1 joeng6 would imply that this was the same kind of person, rather than the same person. Yet another option, tung4 joeng2, implies the same kind of way. Some example phrases are below.

\[ \text{(2.4)} \]

\[ \text{(2.4a)} \]

\[
\text{tung4 jat1 g03 jan4} \\
\text{SAME CL SOMEONE} \\
\text{THE SAME SOMEONE}
\]

\[ \text{(2.4b)} \]

\[
\text{tung4 jat1 joeng6 je5} \\
\text{SAME CL THING} \\
\text{THE SAME THING}
\]

---

27 The situation and problems discussed here for primes THIS and THE SAME are like those for the same primes in Chinese Hokkien.
Several possibilities exist for the Cantonese exponent of OTHER~ELSE. Tong, Yell, and Goddard (1997, 247) use ling6 jat1, while Wakefield (2011b) suggests ling6 ngoi6. Though they share the word ling6, the two have slightly different meanings. Jat1 is the Cantonese prime ONE, while ngoi6 means something like ‘outside’. Ling6 jat1 is ‘the other’, or ‘another’. Ling6 ngoi6 can be understood as ‘apart from’, or ‘besides’. Kei4 taa1 is another synonym, and seems to always mean ‘others’ in the plural. Additionally, I would point out that dai6 ji6 go3 is commonly used to mean OTHER, although in some cases it can mean ‘the second one’, as in its literal meaning.

Use of ling6 jat1 signals something singular in number; for example, SOMEONE ELSE and SOMETHING ELSE (singular) are ling6 jat1 go3 jan4 and ling6 jat1 joeng6 je5, respectively. Fortunately, ling6 jat1 can be followed by di1 which means SOME, to create ling6jat1 di1. Then, OTHER PEOPLE can be ling6 jat1 di1 jan4 and OTHER THINGS can be ling6 jat1 di1 je5. However, these can be interpreted as SOME OTHER PEOPLE/THINGS, instead of just OTHER PEOPLE/THINGS. More than one OTHER can also be expressed by ling6 ngoi6, ling6 ngoi6 di1, and kei4 taa1. OTHER PEOPLE has the additional possibilities ling6 ngoi6 di1/ge3 jan4 and kei4 taa1 jan4, and OTHER THINGS has the additional possibilities ling6 ngoi6 di1/ge3 je5 and kei4 taa1 je5. SOMEWHERE ELSE to refer to one single place is ling6 jat1 dou6 or ling6 ngoi6 jat1 dou6 (or ling6 jat1 go3 dei6 fong1). Another example, AT ANOTHER TIME, would best be translated as ling6 jat1 go3 si4 hau6 (/si4 gaan3), where ling6 jat1 and the classifier go3 help give the meaning ‘another’. Ling6 ngoi6 ge3 si4 hau6 (/si4 gaan6) is also possible but does not mean ‘another’, meaning something closer to ‘at other times’.
Ling6 jat1, kei4 taa1, and dai6 ji6 go3 cannot co-occur with li1 THIS. This other someone must be ling6 ngoi6 li1 go3 jan4, although this is not common either. This other thing is ling6 ngoi6 li1 (jat1) joeng6 je5. The plural these other people would be ling6 ngoi6 li1 di1 jan4, while these other things would be ling6 ngoi6 li1 di1 je5. This other part can be ling6 ngoi6 li1 go3 bou6 fan6 or ling6 ngoi6 li1 di1 bou6 fan, and this other kind can be ling6 ngoi6 li1 zung2.

The expression of the plural other kinds in Cantonese is problematic. The singular other kind ling6 jat1 zung2 is acceptable, but translations of the plural such as *ling6 jat1 di1 zung2, *ling6 ngoi6 di1 zung2, *kei4 taa1 zung2, and *kei4 taa1 di1 zung2, are all unacceptable. If expressing other kinds in plural, a Cantonese speaker must use the problematic gei2, roughly ‘quite/few/some’ (see siu2 LITTLE~FEW on p.63). Acceptable constructions include ling6 ngoi6 gei2 zung2 and kei4 taa1 gei2 zung2, roughly ‘a few other kinds’. This problem seems to be specific to kinds. For example, the singular other part is ling6 jat1 go3 bou6 fan6, and the plural other parts is ling6 jat1 di1 bou6 fan6 or kei4 taa1 bou6 fan6, with no need for gei2. The Cantonese exponent for other~else is proposed to be ling6 jat1~ling6 ngoi6 for now, though some issues are unresolved.

Jat1 – ONE

Jat1 is the Cantonese exponent for ONE. A classifier may be required after it. Jat1 may sometimes be used like an indefinite article.

Loeng5 – TWO

Cantonese has two distinct allolexes for the number TWO, namely ji6 and loeng5. Ji6 is used for counting and in compound numbers, while loeng5 is used to quantify nouns. A distinction also exists in Mandarin between èr and liǎng (Chappell 1994, 125). The Cantonese exponent for TWO is identified as loeng5 (even though ji6 is semantically equivalent) since the prime TWO is generally used to quantify nouns and therefore loeng5 is the most likely to be used in NSM explications. A classifier is often required afterwards. Some examples are given below. It is also possible to say loeng5 go3 dou1 [ TWO CL ALL] to mean ‘both’ (English has a word both which means ALL TWO).
(2.5)

(2.5a)
loeng2  go3  jan4
TWO   CL   PEOPLE
TWO PEOPLE

(2.5b)
loeng2  joeng6  je5
TWO   CL   THING
TWO THINGS

(2.5c)
loeng2  (go3)  bou6  fan6
TWO   (CL)   PART
TWO PARTS

(2.5d)
loeng2  zung2
TWO   KIND
TWO KINDS

(2.5e)
loeng2  go3  li1  di1  je5/jan4
TWO   CL   THIS   CL   THING/PEOPLE
TWO OF THESE THINGS/PEOPLE

(2.5f)
do1  loeng2  joeng6  je5
MORE   TWO   CL   THING
TWO MORE THINGS

(2.5g)
loeng2  go3  dou1
TWO   CL   ALL
ALL TWO / both

**Di1: (jat1/jau5) di1 – SOME**

The Cantonese exponent of *some* is identified as *di1*. *Jat1* or *jau5* are sometimes required before *di1*, but not always. Somewhat counter-intuitively, *jat1* is also the Cantonese exponent for *one*, mentioned above. *Jau5* is the Cantonese exponent for *there is*. 
Di1 is sometimes used as a classifier implying plurality. For example, jat1

di1 jan4 [ONE SOME/CL:PL PEOPLE] means SOME PEOPLE, as opposed to jat1 go3

jan4 [ONE CL SOMEONE] which means ONE SOMEONE. Ngo5 soeng2 gong2 (jat1)
di1 je5 is likely to be interpreted as I WANT TO SAY SOMETHING [I WANT SAY (ONE)

CL:PL THING], although it can also be interpreted as I WANT TO SAY SOME THINGS [I

WANT SAY (ONE) SOME THING]. A less ambiguous way of saying I WANT TO SAY SOME

THINGS would be to use gei2, as in ngo5 soeng2 gong2 gei2 joeng6 je5 [I WANT

SAY gei2 CL THING], but gei2 causes unwanted problems (see LITTLE~FEW on p.63). (Ngo5

soeng2 gong2 je5 [I WANT SAY THING] without di1 would be interpreted as I WANT TO

SAY THINGS, like ‘I want to speak’.)

Use of di1 SOME in *hai2 di1 dou6 [AT SOME PLACE] IN SOME PLACES is

unacceptable if dou6 is used for PLACE~WHERE. It would be acceptable with the

other possible exponent for PLACE~WHERE dei6 fong1, as in hai2 jat1 / jau5 di1

dei6 fong1 [AT ONE / THERE ARE SOME dei6 fong1] (see also PLACE~WHERE on

p.81).28

**Dou1~cyun4 bou6 – ALL**

Several possibilities were considered for the Cantonese exponent of ALL,

including dou1, so2 jau5, seng4, and cyun4 bou6. The most suitable exponent

for ALL appears to be dou1, although an exception needs to be made for ALL

PARTS. First, some explanations for dismissing the other options are given.

So2 jau5 can be translated literally as ‘whatever there is’. So2 jau5 can be

combined with, say, jan4 PEOPLE or je5 THING(s), but is less likely to be used in

phrases like (AT) ALL TIMES. It is possible that the meaning of so2 jau5 is closer
to that of English ‘every’, and therefore it does not combine well with TIME

because the Cantonese exponent for TIME, si4 hau6, does not indicate a specific

length or period of time (see TIME~WHEN on p.76). ALL TIMES implies something

more fluid, as opposed to ‘every time’, which implies frequency, or repetition of

something specific. Furthermore, Matthews and Yip (2011, 307) point out that

if the quantified phrase comes before the verb, then so2 jau5 must be

accompanied by another of the identified possibilities for ALL, namely dou1.

One must say so2 jau5 ... dou1 .... For example, ALL PEOPLE KNOW would not be

28 With regards to the primes SOME and ALL, Hokkien poses similar problems as Cantonese.
*so2 jau5 jan4 zi1 (dou3), but rather so2 jau5 jan4 dou1 zi1 dou3. For these reasons, so2 jau5 is not ideal as the exponent for ALL.

The Cantonese word seng4 also has a meaning similar to ALL. This word can be dismissed as the Cantonese exponent of ALL, however, because it usually refers to an ‘entire’ thing, or a ‘whole’. Therefore seng4 (go3) bou6 fan6 would not be ALL PARTS, but rather ‘the whole/entire part’. This meaning is completely different. Related to this meaning of ‘whole’, *seng4 jan4 [whole people] is not possible as ALL PEOPLE, although something like seng4 cyun1 jan4 [whole village people] makes sense as ‘a whole village of people’. Seng4 go3 jan4 [whole CL someone] with the classifier go3 for the singular person would mean ‘the whole person’. Seng4 is also optionally accompanied by dou1 and/or saai3, a quantifying verbal particle.

The next possibility, cyun4 bou6, can be combined with je5 SOMETHING–THING to form ALL THINGS in cyun4 bou6 (ge3) je5 [all LP THING], but does not combine to produce other necessary NSM phrases, such as ALL PLACES or ALL TIMES. One reason for this could be that cyun4 bou6 also has a meaning closer to English ‘whole’. Like so2 jau5, cyun4 bou6 is sometimes used with dou1 in sentences, as in cyun4bou6… dou1… It would seem, then, that cyun4 bou6 ought to be rejected as the Cantonese exponent for ALL. However, there is one exception, namely ALL PARTS, which will be discussed further below.

The best candidate for the Cantonese ALL seems to be dou1. It occurs after the noun and before the verb. However, use of dou1 is not entirely straightforward either. In some cases, it is necessary to reduplicate the word preceding dou1. This is sometimes the noun, for example, ALL PEOPLE is jan4 jan4 dou1 [PEOPLE PEOPLE ALL], (AT) ALL TIMES is si4 si4 dou129 [TIME TIME ALL], and ALL KINDS is zung2 zung2 dou1 [KIND KIND ALL]. Elsewhere, it is the classifier which is reduplicated, followed by the noun once. For example, ALL THINGS is not *je5 je5 dou5 [THING THING ALL], but rather, assuming the classifier joeng6 is appropriate, joeng6 joeng6 je5 dou1 [CL CL THING ALL]. ALL PEOPLE has this as an additional possibility, being able to be expressed as go3 go3 jan4 dou1 [CL CL PEOPLE ALL], where go3 is the appropriate classifier for a person. If the relevant noun is obvious, for example where THINGS/PEOPLE have already been

29 The hau6 in si4 hau6 is not necessary here.
identified as the subject, it is also possible to have only the reduplicated classifier and omit the noun completely, as in "ALL THINGS, joeng6 joeng6 dou1 [CL CL ALL]. In the case of "ALL PEOPLE, this results in "go3 go3 dou1 [CL CL ALL]. All three of these options for "ALL PEOPLE are acceptable and natural in Cantonese. Not all three of these possibilities may be acceptable for other nouns. As explained in the section on "PLACE~WHERE (see p.81), dou6 can be used as a noun as well as the classifier for 'place' nouns. It is unclear whether "dou6 dou6 dou1 "ALL PLACES" is a reduplication of the noun "PLACE" or a reduplication of the classifier for places. In any case, this is not a significant concern; it is possible to say "ALL PLACES" in NSM.

(2.6)

(2.6a)
jan4 jan4 dou1
PEOPLE PEOPLE ALL
ALL PEOPLE

(2.6b)
go3 go3 jan4 dou1
CL CL PEOPLE ALL
ALL PEOPLE

(2.6c)
go3 go3 dou1
CL CL ALL
ALL PEOPLE

In addition to meaning "ALL, dou1 can mean 'every' or 'each', or something like 'also', 'too', 'any', 'anyway', 'still', or 'even'. Perhaps because of this, it can seem less 'strong' or 'emphatic' than other words considered for "ALL, such as so2 jau5. Matthews and Yip (2011, 214) give an example where one person says "lei5 bou2 zung3 aa3 'take care', and another responds "lei5 dou1 hai6 [you dou1 is] 'you too'. Another example they give is "ngo5 dei6 dou1 lam5 zyu6 zau2 ge3 laa3 [we dou1 intend leave PRT PRT] 'we were going to leave anyway'. In these examples, dou1 does not mean "ALL. Sometimes it depends on context whether dou1 is used to mean "ALL or 'also', 'even', 'anyway', etc. The varied uses of dou1 can lead to ambiguity in some cases, but it should be remembered that when dou1 is used in NSM, its only available meaning is that of "ALL.
Consider now ALL PARTS in NSM. Bou6 fan6 parts (see p.53) cannot be combined with dou1 all to form ALL PARTS. Does this mean that Cantonese speakers are unable to express the meaning ALL PARTS? In fact, they can, but the answer lies in one of the aforementioned options for ALL. Cantonese speakers express the meaning ALL PARTS using cyun4 bou6. The use of cyun4 bou6 as ALL PARTS does not contradict the explanation above that cyun4 bou6 means something like ‘whole’ – it is simply that this meaning appears to be exactly that of ALL PARTS. Note that the bou6 in cyun4 bou6 is the same as the bou6 in bou6 fan6 parts. But although cyun4 combines with bou6 to form cyun4 bou6 all parts, this does not mean that cyun4 can combine naturally with the other primes to form e.g. ALL PLACES, ALL THINGS, ALL TIMES, or ALL PEOPLE. (Synonyms exist, like cyun4 jat6 ‘all day’ or cyun4 jan4 leoi6 ‘all people kind’, which refers more to the human species, like ‘all mankind’, but this is not relevant to NSM.)

In summary, it is proposed that the Cantonese exponent of ALL is dou1, with the exception of cyun4 bou6, which is needed for ALL PARTS. Cyun4 bou6 may still be used with dou1, as in cyun4 bou6 dou1..., although this is not essential.

**Do1 – MUCH~MANY**

MUCH~MANY is do1 in Cantonese. There is no systematic distinction in Cantonese between count and mass nouns. Do1 is symmetrical to siu2 LITTLE~FEW, and the two are commonly used together in the same phrase or sentence. The ‘dummy’ hou2 (see VERY on p.86) is often necessary, as in example (2.7).

(2.7)

\[
\text{Li1 dou6 jau5 hou2 do1 jan4.} \\
\text{HERE THERE IS VERY/‘dummy’ MUCH/MANY PEOPLE} \\
\text{THERE ARE (VERY) MANY PEOPLE HERE.}
\]

---

30 Alternatively, it is also possible (although unlikely) that ALL PARTS is itself unacceptable NSM, since ALL and PART might be seen as contradictory concepts. The acceptability of ALL PARTS may be discussed or re-evaluated in future research on other languages.
Siu2 – LITTLE~FEW

Siu2 in Cantonese has the meaning LITTLE~FEW. As mentioned, Cantonese does not make any distinction between count and mass nouns. LITTLE~FEW and SOME are two separate semantic primes, but since neither specify a precise number, the exact distinction between the two cannot be certain. We can only say that SOME seems to indicate more than LITTLE~FEW. A Cantonese word gei2, also mentioned under dit SOME (see p.58), causes some problems for both LITTLE~FEW and SOME, as gei2 sometimes sounds more natural. Gei2 can be translated as ‘few’ or ‘quite’. It seems that gei2 is somewhere in between or overlaps with both LITTLE~FEW and SOME, and therefore can be interpreted as either. Consider, for example, siu2 bou6 fan6 FEW PARTS, which is rarely used in Cantonese. Use of gei2, as in gei2 (go3) bou6 fan6, seems to be more natural than siu2 bou6 fan6 (to make matters worse, siu2 bou6 fan6 can be interpreted as A SMALL PART – see SMALL on page 66). Other ways of saying FEW PARTS in Cantonese include bou6 fan6 PARTS on its own, or jau5 bou6 fan6 [THERE-IS PARTS], as well as the reduplicated siu2 siu2, which means ‘a few’ or ‘a bit’.

2.4 Evaluator and descriptor primes

Hou2 – GOOD

The Cantonese exponent for GOOD is hou2. It is homophonous with hou2 VERY. Hou2 hou2 is natural as VERY GOOD (although a problem exists with VERY – see VERY on p.86). Hou2 GOOD will be necessary in some particle explications presented in Part Two.

(2.8)

\[
\begin{align*}
\text{hou2} & \quad \text{do1} & \quad \text{hou2} & \quad \text{jan4} \\
\text{VERY/’dummy’} & \quad \text{MUCH/MANY} & \quad \text{GOOD} & \quad \text{PEOPLE} \\
\text{(VERY) MANY GOOD PEOPLE}
\end{align*}
\]

M4 hou2 – BAD

The semantic prime BAD does not, at first, seem to have a straightforward Cantonese translation. Several different words seem to be used, depending on context. The possibilities include m4 hou2, sui1, chaai1, and waai6. Ultimately, m4 hou2 is chosen as the best exponent for BAD.
M4 hou2 can be used to describe almost anything, including people and things. Because of this, m4 hou2 seems to be the best option for the prime BAD. From a formal point of view, m4 hou2 is literally NOT GOOD, i.e. a combination of the primes m4 NOT and hou2 GOOD. However, m4 hai6 m4 hou2 (NOT IS NOT GOOD) is acceptable Cantonese for expressing the meaning NOT BAD, which seems to support that m4 hou2 is a fixed expression meaning BAD. As might be expected from the literal meaning of m4 hou2 as NOT GOOD, m4 hou2 is quite mild.

Hou2 is not only the Cantonese exponent for GOOD, but also for VERY. This does not seem to be a problem even if BAD is m4 hou2, because hou2 m4 hou2 [lit. VERY NOT GOOD] actually conveys the sense VERY BAD. This also supports m4 hou2 as a fixed expression meaning BAD. (Hou2 m4 hou2 can also be interpreted as GOOD NOT GOOD, in the ‘A-not-A’ format commonly used for asking a question in Cantonese, but this usage could never appear in an NSM explication. Therefore, any instance of hou2 m4 hou2 in Cantonese NSM will mean VERY BAD.) M4 hai6 hou2 hou2 [NOT IS VERY GOOD] is also acceptable as NOT VERY GOOD, and is not ambiguous with hou2 m4 hou2 VERY BAD. M4 hai6 hou2 m4 hou2 [lit. NOT IS VERY NOT GOOD] is acceptable as NOT VERY BAD, although it sounds a bit clumsy.

Another option for Cantonese BAD is sui1. As Vanhatalo, Tissari, and Idström (2014, 75-77) observe with reference to Finnish NSM, there is a difference between things being bad as in ‘low in quality’, and bad as in ‘evil or immoral’. Both could be the opposite of GOOD. They suggest that the semantic prime BAD refers to that which is ‘evil or immoral’. If this is the case, sui1 in Cantonese is a possible exponent for the prime BAD. When sui1 is used, there often seems to be some sort of subjective moral judgment, though it is not completely ruled out in contexts of something being ‘low in quality’. Sui1 is used mostly for people, and it is possible to call someone a sui1 jan4 (recall jan4 is SOMEONE, a person). When referring to people in spoken language, the word sui1 seems to be more common than the other possibilities for BAD. However, sui1 is less likely to be used to describe non-humans.

The Cantonese chaa1 also has a meaning like BAD. It is typically used for things (i.e. not people). It is often used as the opposite of hou2 GOOD, but it seems that chaa1 has the meaning of ‘bad’ as in ‘low in quality’. There may also
be some sort of level of degree implied; for example, somebody’s grades can be *chaat* as opposed to *hou2 GOOD*, possibly because a grade can be higher or lower on a scale. Grades cannot be *suit* because grades are more objective and there is no moral judgment. Weather can also be *chaat* – the opposite of ‘good (/fine) weather’, and a farmer’s poor harvest can also be referred to as *chaat*. If the semantic prime *BAD* must always be capable of being used about people, however, *chaat* is unlikely to be a suitable exponent for *BAD*.

Another possibility considered for Cantonese *BAD* is *waai6*. *Waai6* at first seems like a good exponent for *BAD* because, like *m4 hou2*, it can be used for people as well as things. At the same time, it feels ‘stronger’ than *m4 hou2*. *Waai6* seems to be used for a more objective category than, say, *chaat*. For instance, children are told not to do *waai6* things, and to beware of *waai6* people in the world (*waai6 jan4* would be roughly ‘evil or immoral’ people, though possibly milder than English ‘evil/immoral’, just as *BAD* is). This is almost like a classification of some things that are simply always bad, and some people that are simply bad people, as opposed to someone having done something bad once, in which case they might be labelled as *sui1*. When being used to refer to things, *waai6* usually means that something is faulty, broken, or rotten. For instance, a machine can be *waai6* (‘faulty/broken’), and a fruit can be *waai6* (‘rotten’). Hence it appears that *waai6* can be used both in the ‘evil or immoral’ context and in the ‘low in quality’ context.\(^{31}\)

Vanhatalo, Tissari, and Idström (2014, 75–77) suggest that NSM should sharpen the distinction between the ‘evil, immoral’ and ‘low in quality’ uses, proposing that the former might be the semantic prime *BAD* and the latter may be susceptible to being explicated (in which case, it could not be a prime). This possibility is set aside here.

The Cantonese exponent for *BAD* is proposed here to be *m4 hou2*, on the basis that it may be the most general and can be used for almost anything. It is acknowledged that *waai6* is another possible candidate, but based on native

\(^{31}\) *Waai6* is the Cantonese counterpart to Mandarin *huài*, identified as the Mandarin exponent for *BAD* by Chappell (1994, 142-143). Chappell stated that *huài* is not a symmetric opposite of *GOOD* because it is closer in meaning to ‘immoral’, ‘nasty’, or ‘evil’, than *GOOD* is to ‘saintly’. Later, Chappell (2002, 266) suggested that Mandarin *bù hào*, which translates into Cantonese as *m4 hou2*, is often substituted for *huài*, for reasons of ‘cultural appropriateness’ and in predicative uses.
speaker intuition, the author hesitates to propose that waai6 and m4 hou2 are semantically equivalent allolexes, mainly because waai6 is discernibly 'stronger' (though waai6 can be intensified by hou2 VERY, just as m4 hou2 can). It is possible that waai6 is strong enough to include a sense of something being dangerous/harmful or having relatively serious consequences, i.e. the meaning of waai6 may be closer to the meanings of 'evil' and 'immoral', as compared with the milder m4 hou2. If the meanings of the other possibilities considered for BAD were able to be explicated using m4 hou2, this would be convincing proof that the exponent proposed is the correct one. This is left for future research.

**Daai6 – BIG**

The Cantonese exponent for BIG is daai6. Daai6 can be used in most NSM explications, but an exception lies in BIG PLACE. As will be explained below, *daai6 dou6* is unnatural (see place~where on p.81). If dou6 is the Cantonese exponent for place~where, then we must say *li1 dou6 hou2 daai6* [this place very/‘dummy’ BIG], THIS PLACE IS (VERY) BIG. This allows us to avoid the problem of *daai6 dou6* BIG PLACE, but instead involves the use of the Cantonese ‘dummy’ hou2, homophonous to VERY (see very on p.86). This causes ambiguity between THIS PLACE IS BIG and THIS PLACE IS VERY BIG.

**Sai3 – SMALL**

The Cantonese word most suitable as the exponent for SMALL is sai3. It is the opposite of daai6 BIG, and the two are often used together in Cantonese, as in the expression daai6 sai3 [lit. big small] ‘size’. As with daai6 BIG, discussed above, *sai3 dou6* SMALL PLACE is unacceptable if place~where is dou6, although *lii dou6 hou2 sai3* THIS PLACE IS (VERY) SMALL is fine, bar the problem mentioned regarding ‘dummy’ hou2.

Another Cantonese word for SMALL is siu2, homophonous to siu2 LITTLE~FEW (although it uses a different written character). Some phrases, like a SMALL PART, are more likely to take siu2 with the meaning SMALL, rather than sai3. This means that siu2 bou6 fan6 can naturally sound like both a SMALL PART and FEW PARTS (see also siu2 LITTLE~FEW on p.63). Sai3 bou6 fan6 [SMALL
PART] is uncommon (it would be understood in speech, but is unacceptable in writing). This problem is left for the future.

2.5 Mental predicates and speech-related primes

Several of the primes discussed in this section will appear in the explications of Cantonese utterance particles later in this thesis, specifically zi1 (dou3) KNOW, lam2 THINK, soeng2 WANT, gok3 dAKI FEEL, and gong2~waa6 SAY.

Zi1 (dou3) – KNOW

The semantic prime KNOW is expressed in Cantonese using zi1 (dou3). There appears on the surface to be no meaning difference between zi1 and zi1 dou3. Zi1 is perhaps more common.

(2.9)
\[ Ngo5 \ zi1 \ hou2 \ do1 \ je5. \]
I KNOW ‘dummy’/VERY MANY THING
I KNOW (VERY) MANY THINGS.

(2.10)
\[ Lei5 \ m4 \ zi1 \ li1 \ joeng6 \ je5. \]
you not know this CL thing
YOU DON’T KNOW THIS.

Earlier work identified the Cantonese exponent of KNOW to have an allolexical variant sik1. A generalisation can be made that zi1 (dou3) is to do with knowing facts, and sik1 is to do with knowing people or skills. Similar distinctions are made in Romance languages, for example French savoir~connaitre, Spanish saber~conocer, Portuguese saber~conhecer, and Italian sapere~conoscere (Peeters (2006, 90-94)). In latest NSM work, the meanings in ‘know someone’ and ‘know how to do something’ have been explicated, so they are no longer regarded as allolexes (Goddard in press). This leaves zi1 (dou3) as the sole exponent for KNOW.

Lam2 – THINK

The best Cantonese exponent for THINK appears to be lam2. Lam2 usually seems to imply intentional or ‘active’ thinking, for example, when one is thinking about a problem or a puzzle. Two other Cantonese ‘think’ words, jing6
wai4 and gok3 dak1 (which appear to have very similar meanings to each other), are used to express a speaker’s general opinion. Another word ji5 wai4 is closer in meaning to ‘believed’ or ‘thought’, and implies that the speaker no longer thinks in that way.

Chappell proposed that the Mandarin exponent for THINK is xiăng. If translating between Mandarin and Cantonese via written Chinese, xiăng would correspond with Cantonese soeng2. However, Chappell’s (2002, 268-269) example sentences for Mandarin xiăng THINK would either be unacceptable or highly unusual in spoken Cantonese if soeng2 were used. Soeng2 is taken here to mean WANT (see below). Wakefield (2011b, 277) also believes lam2 to be the correct Cantonese exponent for THINK, commenting that it must be used in all Cantonese versions of the example sentences given in Chappell (2002) and that the Mandarin counterpart only holds the same meaning in writing.

**Soeng2 – WANT**

Wakefield (2011b, 274-275) identified jiu3 as the Cantonese WANT, explaining that when the example sentences given by Chappell (2002) are translated directly into Cantonese using the same Chinese characters, the Cantonese versions of the primes have the same valency properties as their Mandarin counterparts. Though this may be the case, as explained, written Chinese is quite different from spoken Cantonese. The Mandarin primes, translated into Cantonese, belong to the formal register and are therefore not appropriate for a Cantonese NSM (see section 2.1.3). In Cantonese, jiu3 has a meaning better translated as ‘need’ or ‘have to’.32

The Cantonese word soeng2 is more commonly used in everyday, natural speech, and I propose that this is the Cantonese exponent of the prime WANT.33 (Soeng2 jiu3 is also possible in Cantonese, although this is also associated with Mandarin, and more problematically, soeng2 jiu3 could be interpreted as ‘want to have’.). Interestingly, the Cantonese exponents for VERY and WANT, i.e. hou2 and soeng2, can be collated without problem, as in ngo5 hou2 soeng2… ‘I really want to…’. This is in contrast to the Mandarin exponents for VERY and WANT,

---

32 Perhaps needing something can be thought of as a ‘stronger’ kind of wanting something – this could explain why the two are often interchangeable, which could cause confusion.

33 Some arguments about the semantic prime WANT as a lexical and conceptual universal are presented in Goddard and Wierzbicka (2010).
which cannot be collated in the same way – *hěn yào is not possible (Chappell 2002, 307). (This is like *very want’ in English, although workarounds like ‘really want’ are possible).

\[(2.11)\]
\[
\text{Ngo5 soeng2 lei5 ji4gaa1 zi1 ngo5 dim2 lam2. I want you now know I how think}
\]
I WANT YOU TO KNOW WHAT I THINK NOW.

**M₄ soeng2 – DON’T WANT**

The prime DON’T WANT is one of the newer additions to the set of NSM primes, added (or rather, restored) to the prime inventory in Goddard and Wierzbicka (2014).\(^{34}\) The Cantonese exponent is simply the primes m₄ NOT and soeng2 WANT together, resulting in m₄ soeng2 DON’T WANT. Unsurprisingly, this will cause ambiguity between a negation of WANT and the prime DON’T WANT, just as would happen in English.

**Gok3 dak1 – FEEL**

The most suitable Cantonese word for the prime FEEL is gok3 dak1. Another ‘feel’ word, gam2 gok3, is more like a ‘feeling’ or like the other meaning of ‘feel’ in English, i.e. a physical feeling or sensation. Gok3 dak1 FEEL is more mental/emotional. Gok3 dak1 can also mean ‘think’ in the ‘opinion’ sense.

\[(2.12)\]
\[
\text{Ngo5 gok3 dak1 hoi1 sam1. I feel happy}
\]
‘I feel happy.’

**Gin3 (dou2) – SEE**

The Cantonese exponent of SEE is gin3 (dou2). Dou2 is a ‘verbal particle’ in the ‘resultative’ category. Resultative particles are generally used with transitive verbs to indicate an effect on the object, and they form compounds with the

\[^{34}\text{In early NSM work (Wierzbicka 1972, 1980) WANT and DON’T WANT were recognised as separate primes, but this insight was lost sight of in Goddard and Wierzbicka (1994) and much subsequent work. The basicness of DON’T WANT is very clear from its early appearance in child speech (expressed simply as ‘No!’), but the formal overlap with WANT confuses the picture. The issue cannot be discussed here for reasons of space, but NSM theorists are now convinced of the futility of trying to paraphrase DON’T WANT via WANT and NOT.}\]
verb. Dou2 can indicate ‘accomplishment or successful completion of an action’, and can be used with sensory verbs such as ‘see’ to form perception verbs which ‘denote actual perception of a stimulus’ (Matthews and Yip 2011, 149, 250-252).

Tai2 (dou2) is a close synonym of gin3 (dou2), and in most situations the two are interchangeable, although sometimes (not always) tai2 (dou2) is more ‘intentional’. Tai2 may have a meaning slightly closer to English look. Tai2 can be used in situations like reading a book, watching a movie, performance or football match, or watching over children playing. Tai2 haa5 [tai2 DEL] or lei5 tai2 [you tai2] can be used when telling someone to look at something. Nonetheless, both can be used when seeing something suddenly and unexpectedly, for example ‘when I was driving, I saw a kangaroo jump out unexpectedly’. The difference between gin3 and tai2 is not as distinct as that between English see and look. Confusingly, it is possible to tai2 gin3, with no obvious differences to tai2 or to gin3, although *tai2 gin3 dou2 is unacceptable.

(2.13)

Ngo5 gin3/tai2 dou2 di1 je5.
I see VPRt some thing(s)

I saw something / some things.

In (2.14) below, both gin3 and tai2 are acceptable, but gin3 sounds more neutral. Use of tai2 may imply, for example, to visit your grandmother if she is sick.

(2.14)

Ngo5 soeng2 heoi3 gin3/tai2 maa4maa4.
I want go see grandmother (paternal)

‘I want to go and see [my] grandmother.’

If talking about dreaming, if one specifies mung6 zung1 [dream in], one can either gin3 (dou2) or tai2 (dou2) something/someone. However, if see is to directly follow the word ‘dream’, only gin3 is possible.

---

35 Many resultative particles can exist independently as verbs or adjectives, while some do not occur independently. They may also be used in passive form (Matthews and Yip 2011, 250).
There is another synonym mong6, which is even closer in meaning to English look, and in some contexts stare, and more likely to be intentional. It might be used, for example, when telling someone ‘don’t look down!’ or to ‘see/look whether cars are coming’. One cannot mong6 a movie or a performance, but can mong6, say, a movie screen. Mong6 is not the exponent for see.

Gin3 is not compatible with seong6/haa6/zo2/jau6 ‘up/down/left/right’, although both tai2 and mong6 are. This may have similarities with ‘see up/down/left/right’ being less natural than ‘look up/down/left/right’.

I propose that gin3 (dou2) is the Cantonese exponent of see because it seems in some contexts to be more neutral than tai2 (dou2), which may be more like look. I acknowledge, however, that gin3 (dou2) and tai2 (dou2) are often interchangeable and that their exact differences are not as clear as those between English see and look.

**Teng1 dou2 – HEAR**

The Cantonese exponent for hear is teng1 dou2. The Cantonese word teng1 indicates hearing. Dou2 is the same as in gin3 (dou2) see above. Teng1 can be used when, for example, listening to music, or listening to the radio, but dou2 is included because teng1 on its own seems to imply listening intentionally. Teng1 dou2 may be either intentional or unintentional, as with hear in English.

**Gong2~waa6 – SAY**

Say can be expressed in Cantonese by gong2 or waa6. As Wakefield (2011b, 277-278) explains, gong2 takes NP objects, while waa6 takes clause complements. Therefore, the most simple and straightforward way to quote
someone or something is to use *waa6*. In some situations, one or the other is required/more natural, but in other cases, both *gong2* and *waa6* are possible and can be used interchangeably. Sometimes the same sentence can use both. It appears that there is no semantic difference between *gong2* and *waa6*. Some constructed examples are given below.

(2.16)

\[
\begin{array}{cccc}
\text{Keoi5} & \text{waa6} & \text{keoi5} & \text{soeng2} /\text{gin3 dou2} /\text{teng1 dou2} \\
\text{s/he} & \text{say} & \text{s/he} & \text{want saw heard}
\end{array}
\]

‘S/he said s/he wants/saw/heard...’ (*gong2* is unacceptable)

(2.17)

\[
\begin{array}{cccc}
\text{Jyu4 gwo2} & \text{lei5} & \text{gong2} & \text{do1 jat1 go3 zi6} \\
\text{if you say more one cl word}
\end{array}
\]

‘If you say one more word...’ (*waa6* is less likely)

(2.18)

\[
\begin{array}{cccc}
\text{Keoi5} & \text{deoi3 Mary} & \text{gong2, m4 hai6 deoi3 ng05 gong2} \\
\text{s/he to Mary say not be to I say}
\end{array}
\]

‘She said it to Mary, not to me.’

Using either *gong2* or *waa6*, it is not possible to express in Cantonese *someone says something not with words*. This phrasing has been used in English NSM, but in Cantonese, one can only *say* things with words. When someone expresses something not with words, Cantonese speakers must use *biu2 si6*, which has a meaning like ‘show (that)/express’. One can *biu2 si6* something either with or without words.

The Mandarin exponent for *say* is *shuō*. As with many other Mandarin words, in Cantonese, the corresponding *syut3* is usually used only in writing or formal speech. Cantonese *gong2 syut3 waa6* is possible and roughly means ‘say things’, but it is more formal and not as natural as *gong2 je5 [say thing]‘say things’.

**Zi6 – WORDS and zan1 – TRUE**

The semantic prime *WORDS* is expressed in Cantonese using *zi6*. The semantic prime *TRUE* is expressed as *zan1*. 

72
2.6 Primes of actions, events, and movement

**Zou6 – do**

The Cantonese DO is zou6.

**Faat3 sang1 – happen**

The best Cantonese exponent for the prime HAPPEN appears to be faat3 sang1. However, it is less commonly used than in English. In simple questions like English ‘what happened?’ Cantonese speakers may use faat3 sang1 but are far more likely to say something like mat1 je5 si6? [what thing/event] or zou6 mat1 je5? [do what]. The English ‘what happened to this thing?’ sounds quite unnatural in Cantonese. When faat3 sang1 is used in Cantonese, it is more than likely that something bad and probably noteworthy has happened. Typical examples might include a car crash or a building collapsing. Faat3 sang1 is unlikely to be used when something good has happened or is happening. Something happens to someone must be expressed by Cantonese speakers literally as ‘something happens on someone’s body’, just like in Mandarin.

The likelihood of faat3 sang1 being associated with bad or unfortunate events parallels the situation with the corresponding Mandarin word fāshēng. Chappell states that fāshēng is often used with adverse events because the majority of Mandarin verbs of happening tend to express misfortune, but believes that fāshēng itself is relatively neutral (Chappell 1994, 127-129). However, Tien (2009) believes that fāshēng is marked for ‘adversity’ and ‘serious mention’. He believes that it is not semantically neutral and therefore not semantically primitive. He proposes that the correct Mandarin exponents of HAPPEN are instead you3 and zen3(me)yang4/zhe4(me)yang4.

Let us consider Tien’s exponents for HAPPEN. Mandarin you3 corresponds to Cantonese jau5, polysemous with THERE-IS (you3 is also polysemous with THERE IS in Mandarin). In order for Mandarin you3 to express its distinctive HAPPEN sense, it ordinarily occurs with the complement shi4(qing2), which is one of the Mandarin exponents for SOMETHING~THING. It is necessary to point out that in Mandarin, shi4(qing2) necessarily means an event or non-physical thing; another allolex must be used for physical things. This distinction is not reflected in the Cantonese exponent for SOMETHING~THING, which is je5 (there
are Cantonese words for SOMETHING~THING corresponding to the Mandarin exponents and which differentiate between physical and non-physical things, but they have not been chosen for Cantonese NSM). The consequence of this non-distinction, combined with the polysemy with jau5 THERE-IS, means that Cantonese jau5 je5, which would mean SOMETHING HAPPENS/ED following Tien’s analysis, also means THERE-IS SOMETHING in Cantonese. (Jau5 has another polysemous meaning have in the ‘possession’ sense, which further complicates the picture.) The various other HAPPEN phrases which can be formed in Mandarin using these components also hinge on the explicit marking, via shi4(qing2) SOMETHING~THING, of the ‘subject’ being an event. It may be that the ‘event’ aspect of SOMETHING~THING is causing the HAPPEN interpretation, since events necessarily HAPPEN. While I do not deny that Cantonese jau5 je5 can be interpreted as SOMETHING HAPPENS/ED, the potential ambiguities in Cantonese make it more problematical to abandon faat3 sang1 as an exponent of HAPPEN.

The other Mandarin allolex, zen3(me)yang4/zhe4(me)yang4, does not translate directly into a suitable exponent for Cantonese HAPPEN either (in writing it is possible, but it is unnaturally formal in speech). The Cantonese dim2 (joeng2) can be used in place of zen3(me)yang4/zhe4(me)yang4 in some of Tien’s examples, but dim2 (joeng2) has a meaning closer to ‘how’.

I propose that faat3 sang1 is the Cantonese exponent for HAPPEN, but note that it often implies something bad has happened. Faat3 sang1 is rarely used for good or positive things, although it is possible. Whether the implications of an adverse or unfortunate event are part of the meaning of faat3 sang1 as such, or are the result of other factors, is left for future research.

**Juk1 — MOVE**

The prime MOVE is expressed in Cantonese using juk1.
2.7 Primes of location, existence, specification, and possession

*Hai2* – *BE (SOMEWHERE)*, *jau5* – *THERE IS*, and *hai6* – *BE (SOMEONE/SOMETHING)*

The Cantonese exponent of the locational prime *BE*, as in *BE SOMEWHERE*, is *hai2*. For example, *SOMEONE IS HERE* is *jau5 jan4 hai2 lii dou6*. *THERE IS* is expressed as *jau5* in Cantonese. For example, *THERE IS SOMETHING/SOMEONE IN THIS PLACE* is *jau5 (CL) je5/jan4 hai2 lii (jat1) dou6*. The Cantonese exponent for *BE (SOMEONE/SOMETHING)* is *hai6*.

**(Hai6) ngo5 ge3** – *(IS) MINE*

This prime has been updated several times in recent years, from *HAVE jau5*, to *BE (SOMEONE’S) hai6... ge3*, and now *(IS) MINE*, *(hai6) ngo5 ge3.*

2.8 Primes of life and death

*Sang1 cyun4* – *LIVE*

*LIVE* as in ‘SOMEONE LIVES FOR A LONG TIME’ is proposed to be *sang1 cyun4* in Cantonese. There is a very close synonym *sang1 wut6* (which shares the word *sang1*), but it is not chosen here because *sang1 wut6* may imply to live well, or be associated more with making a living, whereas *sang1 cyun4* is closer to ‘exist’, or having life, and may be more objective. As an example, in a sentence such as ‘after recovering from cancer, he is grateful to live’, Cantonese speakers would use *sang1 cyun4*, and *sang1 wut6* is unacceptable. Note, however, that it is often more natural in Cantonese to say that something ‘has life’ rather than ‘lives’. The English word ‘live’ also has a polysemous meaning of ‘dwells, resides in’, as in ‘someone lives in this place’. For this (non-prime) meaning, Cantonese uses *zyu6.*

---

36 For an analysis of *(IS) MINE* and the conceptual semantics of possession, see Goddard and Wierzbicka (2016).

37 A similar distinction exists in Finnish between *elää* ‘to live, to be alive’, and *asua* ‘to live, to stay, to reside’ (Vanhatalo, Tissari, and Idström 2014, 74).
**Sei² – DIE**

The Cantonese exponent for DIE is sei².

### 2.9 Time-related primes

Some primes relating to time are needed in the particle explanations given in Part Two.

**Si⁴ hau⁶ – TIME~WHEN**

Tong, Yell, and Goddard (1997, 246-249) identified the Cantonese exponent for TIME~WHEN as si⁴ (hau⁶). They note that there is an alternative si⁴ gaan³, but that it is more semantically complex, and implies a specific ‘length’ or ‘segment’ of time. Chappell (1994, 135-136) also identified shihou, the Mandarin counterpart to si⁴ hau⁶, as the correct exponent for TIME~WHEN in Mandarin. She believes, similarly, that shijiiän, or the Mandarin counterpart to si⁴ gaan³, refers to a specific period of time, as well as the concept of time. Chappell chooses shihou (si⁴ hau⁶) as the best Mandarin exponent because it is more ‘appropriately vague’. When WHEN is used, as in AT THE TIME WHEN..., Tong, Yell, and Goddard (1997, 248-249) propose that the Cantonese would be (dong³)... go² zan⁶ si⁴, where go² zan⁶ si⁴ means ‘at that time’, but I propose instead that dong¹... ge³ si⁴ hau⁶ ['at/when’... LP TIME] is more appropriate, and has the advantage of retaining si⁴ hau⁶.

I agree that there is a meaning difference between si⁴ hau⁶ and si⁴ gaan³, and si⁴ (hau⁶) is the preferable Cantonese exponent of TIME~WHEN. Although the two Cantonese ‘time’ expressions are often interchangeable, certain NSM phrases are better expressed using si⁴ hau⁶. For example, (AT) THIS TIME would best be translated as (hai²) li¹ go³ si⁴ hau⁶, and OTHER TIMES as ling⁶ jat¹ di¹ si⁴ hau⁶. MANY TIMES is hou² do¹ si⁴ hau⁶, while hou² do¹ si⁴ gaan³ is less common. In AT THE TIME WHEN..., si⁴ hau⁶ is again more common than si⁴ gaan³, as mentioned above.

However, the translation of some other typical NSM phrases into Cantonese shows that the situation is not that straightforward. For example, THIS OTHER TIME is expressed more naturally in Cantonese as ling⁶ ngoi⁶ li¹ go³ si⁴ gaan⁶ [OTHER THIS CL si⁴ gaan⁶], rather than as ling⁶ ngoi⁶ li¹ go³ si⁴ hau⁶.
[other this cl si4 hau6]. (At) the same time was translated by Tong, Yell, and Goddard (1997, 246-249) as tung4 (jat1) si4 hau6. This is acceptable, but may imply some period, such as the Qing Dynasty or World War Two. Si4 hau6 also seems slightly less common in the same time, with (hai2) tung4 jat1 (go3) si4 gaan3 [(at) same (cl) si4 gaan3] being perhaps more commonly heard, although si4 gaan3 implies a specific ‘length’ of time like a minute or an hour, while si4 hau6 is used for e.g. ‘morning’/‘afternoon’/‘when I was young’. In fact, the most natural and common way to say the same time in Cantonese is the shortened tung4 si4. Tong, Yell, and Goddard (1997, 246-249) also note that the basic Cantonese element required is si4, but si4 is not adequate on its own. For example, (At) this time, mentioned above, must be expressed as (hai2) li1 go3 si4 hau6, and *(hai2) li1 si4 is unacceptable. *Hai2 li1 go3 si4 gaan3 also sounds unnatural. (The only other acceptable synonym in this case is hai2 li1 go3 si4 doi6 which means something more like ‘in this generation/era’.)

Further complicating the problem is apparent polysemy of the English word ‘time’ – the ‘time’ in at this time is different from the ‘time’ in ‘it happened two times’. This distinction is also evident in Mandarin (Chappell 1994, 135-136), Bahasa Melayu (Goddard 2002, 136-139) and Koromu (Priestley 2012, 149-150). Tong, Yell, and Goddard (1997, 248) note that Cantonese ci3 would be required for the second, i.e. what they call ‘frequency time’, but they add that si4 hau6 and ci3 are not allolexes. As pointed out by Vanhatalo, Tissari, and Idström (2014, 72-77), it is not always clear whether we are dealing with polysemy of the universal language-independent NSM primes or polysemy of the English exponents of the NSM primes.

No candidate for Cantonese time seems perfect, and it is unclear whether si4 hau6 and si4 gaan3 should be considered allolexes (at least, they seem to have some overlapping meaning), or whether we should ignore that si4 gaan3 sometimes sounds more natural. If the prime time is intended to be polysemous and can be used both in at this time and it happened two times, this would mean the Cantonese exponent also needs to be valid for both uses, bringing back the possibility of ci3 being required. Alternatively, if only one meaning of time is appropriate in NSM, all NSM researchers should be more attentive about not using time in the other sense. For now, we follow Tong, Yell,
and Goddard (1997) in proposing \textit{si4 hau6} for Cantonese \textit{time}, but note that the matter requires more consideration.

**Ji4 gaa1 – NOW**

The prime \textit{NOW} is expressed in Cantonese as \textit{ji4 gaa1}. Tong, Yell, and Goddard (1997, 249-251) also believe \textit{ji4 gaa1} to be the correct exponent for \textit{NOW}, but importantly, show that it is not possible to combine the primes \textit{BEFORE} or \textit{AFTER} directly with \textit{NOW} in Cantonese (see below).

**Zi1 cin4 – BEFORE and zi1 hau6 – AFTER**

Tong, Yell, and Goddard (1997) identified \textit{zi1 cin4} as the Cantonese exponent for \textit{BEFORE}, and \textit{zi1 hau6} as the exponent for \textit{AFTER}. Other possibilities, \textit{ji5 cin4} and \textit{ji5 hau6}, for \textit{BEFORE} and \textit{AFTER} respectively, were considered by Tong, Yell, and Goddard (1997, 249-251), but ultimately dismissed.\footnote{For Mandarin, Chappell (1994, 2002) believes that \textit{yīqián} and \textit{yīhòu}, the counterparts to Cantonese \textit{ji5 cin4} and \textit{ji5 hau4} respectively, are the best exponents for \textit{BEFORE} and \textit{AFTER}.} Both \textit{zi1 cin4} and \textit{zi1 hau6} can combine with \textit{li1 THIS}, \textit{si4 hau6} (or \textit{si4 gaan3} \textit{TIME}), and the classifier \textit{go3} to form \textit{li1 go3 si4 hau6 (/si4 gaan3) zi1 cin4/zi1 hau6 \textit{BEFORE}/\textit{AFTER \textit{THIS \textit{TIME}}}. Tong, Yell, and Goddard (1997, 249-251) showed that it is not possible in Cantonese to combine \textit{zi1 cin4 \textit{BEFORE}} or \textit{zi1 hau6 \textit{AFTER}} directly with \textit{ji4 gaa1 \textit{NOW}}. In subsequent NSM work, the combinations ‘before now’ and ‘after now’ have been eliminated and alternatives have been found. Tong, Yell, and Goddard suggested phrasings such as \textit{IT ISN’T HAPPENING NOW}; \textit{IT HAPPENED \textit{BEFORE}/\textit{AFTER \textit{THIS \textit{MOMENT}}}. I would additionally suggest the possibility of \textit{BEFORE}/\textit{AFTER \textit{THIS \textit{MOMENT}} \textit{li1 jat1 haak1 zi1 cin4/zi1 hau6} \textit{[THIS \textit{MOMENT \textit{BEFORE}/\textit{AFTER}] \textit{in some circumstances}}. Chappell (1994, 138) points out that it is very common cross-linguistically to find semantic change moving in the direction from spatial to temporal meanings, so it may be of note that \textit{cin4} means ‘front’ and \textit{hau4} means ‘behind’.

**Noi6 – A LONG TIME**

Two possibilities were considered for the Cantonese exponent of the prime \textit{A LONG TIME}, namely \textit{noi6} and \textit{coeng4 si4 gaan3}. Tong, Yell, and Goddard (1997) identified \textit{noi6} as the most appropriate. \textit{Noi6} may also be pronounced \textit{loi6}. 
Noi6 can combine with hou2 VERY to form hou2 noi6 A VERY LONG TIME, although there is ambiguity between this and A LONG TIME when the ‘dummy’ hou2 is used, since the two are identical phrases. Noi6 can also be combined with both BEFORE and AFTER, although the ‘dummy’ hou2 is again required (if very is not used). Therefore, a Cantonese speaker may say hou2 noi6 zi1 cin4 to mean either A LONG TIME BEFORE or A VERY LONG TIME BEFORE, and hou2 noi6 zi1 hau6 to mean either A LONG TIME AFTER or A VERY LONG TIME AFTER. In Mandarin, a ‘dummy’ intensifier hĕn is also required with jiŭ A LONG TIME in its positive form (Chappell 2002, 294).

One of the main attractions of the other option coeng4 si4 gaan3 is that it can be divided into coeng4 [long] and si4 gaan3 [time], and is symmetrical with dyun2 si4 gaan3 [lit. short time]. However, it is slightly more formal than noi6, and perhaps less idiomatic. Furthermore, coeng4 si4 gaan3 cannot combine with BEFORE or AFTER to create A LONG TIME BEFORE or A LONG TIME AFTER (although it can combine with hou2 VERY). Therefore, we chose to overlook the attractive symmetry between A LONG TIME and A SHORT TIME – besides, dyun2 si4 gaan3 was ultimately not proposed to be the exponent for A SHORT TIME, for reasons explained below. In any case, as far as one can tell, noi6 and coeng4 si4 gaan3 are semantically the same, so they may be allolexes.

**Jat1 zan6 – A SHORT TIME**

Tong, Yell, and Goddard (1997) identified jat1 zan6 as the most appropriate exponent for A SHORT TIME. Since MOMENT was not yet a prime when their paper was published, the author considered the possibility that jat1 zan6 might be the exponent for MOMENT instead, in which case dyun2 si4 gaan3 [lit. short time] would be the natural replacement for A SHORT TIME. Reasons for considering jat1 zan6 for MOMENT and dyun2 si4 gaan3 for A SHORT TIME include the fact that jat1 zan6 cannot be combined with hou2 VERY (nor fei1 soeng4 ‘extremely’/‘unusually’ or gei2 ‘quite’ etc.) to form A VERY SHORT TIME. *VERY MOMENT is understandably unacceptable, since MOMENT is already a maximally short period of time, whereas A VERY SHORT TIME seems like it should be possible. Dyun2 si4 gaan3 can be combined with very to create hou2 dyun2 si4 gaan3 A VERY SHORT TIME.*
However, as explained below (see p.80), *jat1 haak1* was eventually chosen as the Cantonese exponent for *moment*, thereby allowing *jat1 zan6* to be reconsidered for a *short time*. Ultimately, *jat1 zan6* was reselected as the prime for a *short time*, because as with *coeng4 si4 gaan3* [lit. long time] mentioned above, *dyun2 si4 gaan3* is slightly more formal than *jat1 zan6*, and perhaps less idiomatic. Furthermore, *dyun2 si4 gaan3* cannot combine with *before* or *after* to form *a short time before* or *a short time after*. *Jat1 zan6*, on the other hand, can combine with *zi1 hau6* *after* to form *jat1 zan6 zii hau6* [a *short time after*] a *short time after*. *Jat1 zan6 zii cin4* [a *short time before*] sounds strange in Cantonese and is not commonly used, although it would be understood.

It is possible that *jat1 zan6* and *dyun2 si4 gaan3* are semantically equivalent, although it is also possible that *jat1 zan6* implies a shorter time period and/or that it is more vague. That *jat1 zan6* implies a shorter time than *dyun2 si4 gaan3* would be supported by the fact that *jat1 zan6* cannot co-occur with *hou2* *very*, while *dyun2 si4 gaan3* can. In line with the proposal that *jat1 zan6* means a *short time*, it could also be that co-occurrence of *very* and a *short time* is no longer needed, following the addition of *moment* to the list of semantic primes. If so, explications and cultural scripts which currently include a *very short time* may need to be revised.

**Jat1 dyun6 si4 gaan3** — *for some time*

*Jat1 dyun6 si4 gaan3* [one period time], appears to be the best Cantonese exponent for *for some time*. It happens to correspond to Chappell’s (2002, 294) proposal of *yŏu yīduàn shíjiān* for Mandarin.

**Jat1 haak1** — *moment*

The Cantonese *jat1 haak1* is proposed to be the exponent for *moment*. Another possibility *jat1 zan6* was first considered. The shared component *jat1* means *one*. *Zan6* is also a specialised classifier occurring with brief events, but *jat1 zan6 haak1* is not possible, unlike *jat1 dou6 dei6 fong1* (see PLACE~WHERE on p.81).

As mentioned, *jat1 zan6* was proposed by Tong, Yell, and Goddard (1997) as the Cantonese exponent of a *short time*. However, this was before *moment*
was included in the list of semantic primes, and it seemed possible, after
MOMENT was added, that the length of time specified by jat1 zan6 was closer to
MOMENT than to A SHORT TIME. This was supported by the fact that jat1 zan6
does not combine with hou2 VERY (see jat1 zan6 A SHORT TIME on p.79).

Ultimately, jat1 haak1 is proposed as the Cantonese exponent for MOMENT,
because MOMENT is supposed to be as short as possible, and jat1 haak1 seems
even shorter than jat1 zan6. Furthermore, li1 jat1 haak1 is possible as THIS
MOMENT, while *li1 jat1 zan6 is unacceptable. Li1 jat1 haak1 THIS MOMENT can
also combine with BEFORE and AFTER, so li1 jat1 haak1 zii cin4 [THIS MOMENT
BEFORE] BEFORE THIS MOMENT and li1 jat1 haak1 zii hau6 [THIS MOMENT AFTER]
AFTER THIS MOMENT are both possible. Like jat1 zan6, jat1 haak1 cannot occur
with hou2 VERY, which makes sense because MOMENT should indicate a time
period so brief that it cannot be intensified; *VERY MOMENT is also impossible in
English. This is unlike, say, A LONG TIME, which can logically be intensified as A
VERY LONG TIME. MOMENT occurs in the semantic explication proposed for wo3
in Chapter 4.

2.10 Place-related primes

Dou6 – PLACE~WHERE

PLACE~WHERE is expressed in Cantonese as dou6. This was also identified by
Tong, Yell, and Goddard (1997, 253-254) as the best Cantonese equivalent for
PLACE~WHERE. Hai2 BE (SOMEWHERE) is usually present. ‘In the place where...’
can be expressed as hai2... go2 dou6 [BE... THAT PLACE], and a statement like ‘I
know where she is’ can be expressed as ngo5 zii keoi5 hai2 bin1 dou6 [I know
she is/at which place]. Questions using ‘where’ are naturally expressed using
(hai2) bin1 dou6? [(at) which place]. Note that ‘who’ questions use bin1 go3, i.e.
the same bin1 ‘which’, with go3, the generic classifier which is used for people.
‘Who’ questions cannot be formed using bin1 with the noun jan4
SOMEONE/PEOPLE. This means that there is no symmetry with PLACE~WHERE, if
we take dou6 to be the noun. In fact, although dou6 can function as a noun in
colloquial Cantonese, it is used in writing and more formal contexts as the
classifier for places, and it is perhaps for this reason that bin1 dou6 is possible.
Also unusually, *li1 dou6 THIS PLACE does not require a classifier, probably for the same reasons.

Another possibility, *dei6 fong1, also considered by Tong, Yell, and Goddard (1997, 253-254), was reconsidered for several reasons. *Daai6 dou6 BIG PLACE and *sai3 dou6 SMALL PLACE are unnatural, whereas daai6 dei6 fong1 and sai3 dei6 fong1 are more acceptable. Also, phrases like A KIND OF PLACE, A PLACE OF ONE KIND, and SOME PLACES are unacceptable with dou6. With dei6 fong1, however, it is possible to say things like A KIND OF PLACE and SOME PLACES. Additionally, use of dei6 fong1 as PLACE~WHERE would better distinguish THIS PLACE from HERE. Assuming that PLACE~WHERE is dou6, then *li1 dou6 THIS PLACE would sound the same as *li1 dou6 HERE, whereas use of dei6 fong1 as PLACE~WHERE would allow THIS PLACE to be expressed by *li1 CL dei6 fong1 [THIS CL PLACE]. As mentioned, dou6 can be used as a classifier for places, so a Cantonese speaker may say jat1 dou6 dei6 fong1 [ONE CL dei6 fong1], where dou6 is the classifier for dei6 fong1. IN THE PLACE WHERE... can be expressed as hai2... ge3 dei6 fong1 [BE... LP dei6 fong1], and ‘I know where she is’ can also be expressed as ngo5 zi1 keoi5 hai2 bin1 go3 dei6 fong1 [I know she is/at which CL place]. Questions using WHERE can be expressed by hai2 bin1 go3 dei6 fong1 [at which CL place] or hai2 mat1 je5 dei6 fong1 [at what place], although bin1 dou6 may be more common.

Ultimately, dou6 remains the better exponent for PLACE~WHERE, for several reasons. Dou6 seems more abstract than dei6 fong1. One reason is that as Tong, Yell, and Goddard (1997, 253) point out, dei6 fong1 implies a bounded place. This would mean that the natural-sounding daai6 dei6 fong1 and sai3 dei6 fong1 might not be BIG PLACE and SMALL PLACE respectively, but rather, something like ‘big space’ and ‘small space’. The more ‘vague’ option seems more appropriate. A more likely explanation for the unacceptability of *daai6 dou6 BIG PLACE and *sai3 dou6 SMALL PLACE is that BIG PLACE and SMALL PLACE are simply not ‘good’ or acceptable phrases in NSM. This could easily be managed by using, instead, THIS PLACE IS BIG OR THIS PLACE IS SMALL, both of which are acceptable in Cantonese using dou6 as PLACE~WHERE. In order to show that *li1 dou6 THIS PLACE is comprised of two separate primes which together give a different meaning to *li1 dou6 HERE, we can insert other primes in between *li1 THIS and dou6 PLACE~WHERE. Li1 jat1 dou6 [THIS ONE PLACE] is an
acceptable and natural way of saying THIS PLACE and helps to differentiate it with HERE. (*Li1 jat1 dou6 and li1 dou6 usually mean the same thing, unless the jat1 ONE is stressed to emphasise that there is one – just like li1 go3 jan4 [THIS CL SOMEONE] and li1 jat1 go3 jan4 [THIS ONE CL SOMEONE] both mean THIS (ONE) SOMEONE.*) These two places can be expressed in Cantonese as li1 loeng5 dou6 [THIS TWO PLACE]. On the other hand, loeng5 TWO does not combine with li1 dou6 HERE.

Still, it is unclear why phrases such as A KIND OF PLACE, A PLACE OF ONE KIND, and SOME PLACES sound odd in Cantonese using dou6 for PLACE–WHERE. This is likely to have an impact on explications which refer to places. It might be that dei6 fong1 is an allolex of dou6, which would allow Cantonese NSM to express otherwise unnatural phrases such as A KIND OF PLACE, and SOME PLACES. The problem in some cases might be to do with PLACE not having a defined area or clear boundaries. The solution is left for future research.

**Li1 dou6 – HERE**

Cantonese HERE is expressed by li1 dou6. As with lei5/nei5 YOU, the shift from initial n to initial l in Cantonese affects li1 dou6. Not only is the pronunciation ni1 dou6 possible, lei1 dou6 and yi1 dou6 may also be used (see e.g. Matthews and Yip (2011, 22)).

*Li1 dou6 HERE* is not to be confused with *li1 (jat1) dou6 THIS (ONE) PLACE*, the difference being that *li1 dou6 HERE* is one word, whereas *li1 dou6 THIS PLACE* is two words (see PLACE–WHERE above). *Li1 dou6* was also identified by Tong, Yell, and Goddard (1997, 255) as the Cantonese prime HERE, where it was also argued to be fully lexicalised with the meaning HERE. They point out that unless the verb is motional, *li1 dou6 HERE* is normally preceded by the locational marker hai2 ‘at’.

**Seong6 gou1 – ABOVE and haa6 min6 – BELOW**

Tong, Yell, and Goddard (1997, 257-259) first found ‘a confusing number of near-synonyms’ for ABOVE, including seong6 gou1, seong6 min6, and seong639.  

---

39 In many cases, seong6 min6 appears to be semantically equivalent to seong6 gou1 ABOVE, but seong6 min6 is vague and ambiguous between ABOVE and ‘on’. Tong, Yell, and Goddard (1997, 257-259) point out that seong6 min6 can be used, for example, in (hai2) coeng4 seong6 min6
There were also many apparent overlaps in usage between the possible exponents of ABOVE and ‘on’ (in the sense of physical contact), which was being considered as a possible NSM semantic prime at the time\(^{40}\). They concluded that soeng6 gou1 invariably and unambiguously indicates ‘aboveness’, and this thesis proposes that soeng6 gou1 is the Cantonese exponent of ABOVE.

The best candidate for Cantonese BELOW was found by Tong, Yell, and Goddard (1997, 259) to be haa6 min6. They also considered haa6 dai1, haa6 dai2, and dai2 haa6, which seem to be closer in meaning to ‘under’, ‘underneath’, or ‘beneath’. The situation regarding the primes ABOVE and BELOW seems to be just as complex in other Sinitic languages such as Mandarin and Hokkien.

\textit{Jyun5 – FAR and kan6 – NEAR}

Cantonese FAR is jyun2, and Cantonese NEAR is kan6. Tong, Yell, and Goddard (1997, 255-256) identified the same exponents. As they note, the reference points may be people, things, or locations. Lei4 may be used to mark the reference point. The ‘dummy’ hou2 is necessary if there is no degree modifier. Jyun2 corresponds to Mandarin yuan, and kan6 corresponds to Mandarin jìn\(^ {41}\) (Chappell 2002, 298-299).

\textit{Bin6 – SIDE}

As put forth by Tong, Yell, and Goddard (1997, 256), bin6 is the Cantonese exponent for SIDE.

\(^{40}\) In the end, Tong, Yell, and Goddard could not find an unambiguous exponent of ‘on’ in Cantonese, and it has since been decided that ‘on’ is not a viable semantic prime. A better candidate for a prime of ‘physical contact’ was found in TOUCH, which was added to the list of NSM primes later and has a more straightforward Cantonese exponent (see p.41).

\(^{41}\) It may be worth pointing out that Mandarin has a special allolex for NEAR when the reference point is a person, as in ‘stand near me’. Cantonese does not require an allolex; ‘stand near me’ in Cantonese can be expressed as kei5 kan6 ngo5 [stand near me].
**Leo5 min6~jap6 min6 – INSIDE**

Tong, Yell, and Goddard (1997, 257) suggest that the Cantonese exponent for INSIDE is *leo5 min6*. I additionally propose *jap6 min6*, as the two have no obvious differences.

**Dim3 (zyu6) – TOUCH**

The Cantonese exponent for TOUCH is proposed to be *dim3*. *Dim3* sometimes sounds unnatural on its own, with the implication that the touching was only for one instant. *Zyu6* is the marker of the continuous aspect in Cantonese, and *dim3 zyu6* is more natural in many sentences than *dim3* on its own. It indicates that something is ‘touching’.

### 2.11 Logical concept primes

Many of the semantic explications given in Part Two for Cantonese utterance particles will feature the following primes for logical concepts.

**M4 – NOT**

The Cantonese *m4* has a general meaning of negation. It can combine with other primes to form, for example, *m4 hou2* NOT GOOD, *m4 soeng2* DON’T WANT, *m4 zi1* (dou3) DON’T KNOW, or *m4 ho2 ji5* CAN’T. Note that DON’T HAVE is *mou5* and not *m4 jau5*.

**Ho2 lang4~waak6 ze2 – MAYBE**

MAYBE is expressed in Cantonese as *waak6 ze2* or *ho2 lang4*. Wakefield (2011b, 274-275) proposed only *ho2 lang4* as the Cantonese MAYBE, but this was based on a translation from Mandarin. The exact difference (if any) between *waak6 ze2* and *ho2 lang4* is unclear, and may require more in-depth analysis; for now, they are considered to be semantically equivalent to each other.

**Ho2 ji5 – CAN**

The Cantonese exponent of CAN is proposed to be *ho2 ji5*. It encompasses ‘possible’ as well as ‘permissible’ things. Another option *dak1* was considered, which can be used in *zou6 dak1* CAN DO, *juk1 dak1* CAN MOVE, etc. *Dak1* also includes both ‘possibility’ and ‘permission’, but cannot be used in all situations,
such as in IT CAN BE LIKE THIS. In all of these, ho2 ji5 is acceptable. There is another synonym lang4 gau3, which was identified as the Cantonese exponent of CAN by Wakefield (2011b, 274) based on translation of Chappell’s (2002) Mandarin primes, but lang4 gau3 is more formal and is better translated as ‘able’.

**Jan1 wai6 – BECAUSE and jyu4 gwo2 (...zau6...) – IF**

The Cantonese exponent for the prime BECAUSE is jan1 wai6. The exponent for IF is jyu4 gwo2. At times, zau6, roughly ‘then’, is necessary, to produce jyu4 gwo2... zau6... ‘IF... then...’. If zau6 is present, jyu4 gwo2 may sometimes be omitted if the meaning of IF is clear. This is similar to colloquial Mandarin, where ‘if-markers’ are often omitted if another method is used to ensure understanding (Chappell 1994, 130-132, 2002, 303-304).

2.12 Intensifier, augmentor, and similarity primes

This is the final category of NSM semantic primes. Do1 (di1)~zoi3 MORE and gam2 joeng2~ci5 LIKE~AS~WAY will be used to explicate the utterance particles in Part Two.

**Hou2 – VERY**

The most obvious option for Cantonese VERY is hou2. This is also the Cantonese exponent for GOOD (see p.63). As mentioned, hou2 hou2 is acceptable as VERY GOOD. More problematic, however, is that hou2 can also be what Tong, Yell, and Goddard (1997, 251-252) refer to as a ‘dummy’; this thesis uses the same label. This ‘dummy’ hou2 is virtually obligatory with predicative adjectives unless the adjective is modified by a genuine degree adverb such as fei1 soeng4 ‘extremely’ or gei2 ‘quite’. In this function, hou2 is ‘weaker’ than English VERY. Matthews and Yip (2011, 179) call this hou2 a ‘default intensifier’, noting that Cantonese predicative adjectives do not require the copular verb ‘be’. They give the following example:

(2.19)

`Lei5 go3 zai2 hou2 gou1.`

you CL son hou2 tall

‘Your son is (very) tall.’ (not *hai6 hou2 gou1)
The translation of (2.19) may be either ‘your son is tall’ or ‘your son is very tall’. This ambiguity would also exist in NSM. *Hou2 do1 je5* may be either *MANY THINGS* or *VERY MANY THINGS*. Similarly, *SOMEONE IS IN A PLACE FOR A LONG TIME* is expressed in the same way as *SOMEONE IS IN A PLACE FOR A VERY LONG TIME*. The solution to this problem is unknown at present. Tong, Yell, and Goddard (1997, 251-252) suggested that *VERY* is a stressed *hou2*, but they note that this is a partial violation of Goddard’s (1994b, 13) ‘Strong Lexicalisation Hypothesis’. In any case, this would not be an ideal solution. Sometimes speakers may stress or emphasise the *VERY* (as opposed to the ‘dummy’) in normal speech, but this is neither a requirement nor a foolproof way to differentiate between the two. In many, if not most, situations, speakers will say *hou2* in the same way whether they mean *VERY* or not.

Another word *fei1 soeng4* was considered for the Cantonese exponent of *VERY*, and is also mentioned by Tong, Yell, and Goddard (1997, 252). While this would avoid the ambiguity caused by *hou2*, the use of *fei1 soeng4* is not an ideal solution. *Fei1 soeng4* is much stronger than *VERY* and is more likely to be translated as ‘extremely’ (the two characters literally mean ‘unusually’).

The word *han2* was also considered. Its Mandarin counterpart *hên* is the Mandarin exponent for *VERY*, and is far more fitting than any other Mandarin intensifier. Because Chinese writing is fairly standardised, an educated Cantonese speaker would certainly understand *han2* to mean *VERY* in writing. However, *han2* is not used in spoken Cantonese. One might even say that *han2* is not a Cantonese word, but a word used in Mandarin and in writing (since written Chinese is different from spoken Cantonese).

Therefore, *hou2* is proposed as the best Cantonese exponent for *VERY*. The problem of ambiguity caused by the ‘dummy’ *hou2* remains unsolved for now. Fortunately, this does not affect the explications proposed in this thesis for utterance particles.

**Do1 (di1)~zoi3 – MORE**

The Cantonese exponent for *MORE* is *do1 (di1)~zoi3*. *Do1* is the Cantonese exponent for *MUCH~MANY*, and in the phrase *MANY MORE PEOPLE*, the Cantonese would be *do1 hou2 do1 jan4*, as in (2.20). In some cases, Cantonese *zoi3* is preferred, especially when used in the sense of *ANYMORE*. Both *do1 (di1)* and
zoi3 are used in the explications proposed for utterance particles later in this thesis.

(2.20)
\[
d o 1 \quad j a n 4 \\
\text{MANY PEOPLE}
\]

(2.21)
\[
d o 1 \; d i 1 \; j a n 4 \\
\text{MORE PEOPLE}
\]

(2.22)
\[
h o u 2 \; d o 1 \; j a n 4 \\
\text{VERY/‘dummy’ MANY PEOPLE}
\]

(2.23)
\[
d o 1 \; h o u 2 \; d o 1 \; j a n 4 \\
\text{MORE VERY/‘dummy’ MANY PEOPLE}
\]

Gam2 joeng2~ci5 – LIKE~AS~WAY

The Cantonese exponent for LIKE~AS~WAY is proposed to be gam2 joeng2~ci5. The joeng2 in gam2 joeng2 is often optional but its omission sometimes causes a slight meaning change, so gam2 joeng2 is generally favoured. Ci5 means LIKE but seems to be used less frequently in explications, as many cases can use THIS WAY/LIKE THIS satisfactorily. Some explications in Part Two will feature gam2 joeng2.

(2.24)
\[
N g o 5 \; g a m 2 \; j o e n g 2 \; l a m 2 : \; ‘ \ldots ’ \\
1 \sim \text{ME THIS-WAY/LIKE-THIS} \; \text{THINK} \; ‘ \ldots ’
\]
I THINK LIKE THIS: ‘\ldots’

\[\text{See also li1 THIS on p.10.}\]
(2.25)  
\[ \text{Ng05} \ (m4) \ \text{soeng2} \ \text{lei5} \ \text{gam2 joeng2} \ \text{zou6}. \]
\[
1\sim\text{ME} \ \text{NOT} \ \text{WANT} \ \text{YOU} \ \text{THIS-WAY}/\text{LIKE-THIS} \ \text{DO}
\]
I (DON'T) WANT YOU TO DO IT THIS WAY/LIKE THIS.

(2.26)  
\[ \text{Jyu4 gwo2} \ \text{ngo5} \ \text{gam2 joeng2} \ \text{zou6}\ldots \]
\[
\text{IF} \ \ 1\sim\text{ME} \ \text{THIS-WAY}/\text{LIKE-THIS} \ \text{DO}
\]
IF I DO IT THIS WAY/LIKE THIS...

(2.27)  
(2.27a)  
\[ \text{Ci5} \ \text{ngo5} \ \text{ge3} \ \text{jan4} \]
\[
\text{LIKE} \ \ 1\sim\text{ME} \ \text{LP} \ \text{SOMEONE/PEOPLE}
\]
SOMEONE/PEOPLE LIKE ME

(2.27b)  
\[ \text{Ng05} \ \text{gam2 joeng2} \ \text{ge3} \ \text{jan4} \]
\[
1\sim\text{ME} \ \text{THIS-WAY}/\text{LIKE-THIS} \ \text{LP} \ \text{SOMEONE/PEOPLE}
\]
SOMEONE/PEOPLE LIKE ME

2.13 Concluding remarks

This has been the most thorough investigation of Cantonese NSM to date. This chapter has addressed the methodological preliminary presented in Chapter 1, and shown that Cantonese exponents of all the NSM semantic primes are identifiable. For some primes, alternative exponents are acknowledged, but overall, all 65 primes have Cantonese exponents, and they can occur together in the mini-syntax of NSM, reinforcing its universality. It was decided that Cantonese NSM should reflect ordinary Cantonese speech as much as possible within the rules of NSM, and the exponents identified in this chapter establish that this is feasible. They also illustrate that Cantonese and Mandarin are quite different at the informal, spoken level. The Cantonese exponents identified here will reappear in the coming chapters, where explications are presented in English and Cantonese to demonstrate that the two can be used in parallel in explications.

Part One has addressed what Cantonese utterance particles are, the research questions and where they came from in the context of the existing literature, as well as the data and methodology which will now be used to
address the main research questions. Part Two will focus on semantic analyses of the five particles selected for study (laa1, wo3, gaa3, laa3, and zaa3), with one chapter for each particle.
Part Two:
Semantic analysis of selected particles
Chapter 3:
The semantics of particle laa1

This is the first of the five chapters in Part Two of this thesis, which present semantic analyses of the utterance particles laa1, wo3, gaa3, laa3, and zaa3, respectively. The present chapter investigates the semantics of the particle laa1. Laa1 is one of the most salient and frequently used particles in Cantonese. In the Hong Kong Cantonese Corpus, laa1 is the 3rd most frequently used particle (after the particles aa3 and gaa3). Laa1 occurs in the corpus a total of 1578 times, making it the 14th most frequent word in the corpus overall. Particles are often used even when Cantonese speakers are speaking English, and laa1 is, significantly, the most frequently used Cantonese utterance particle in Wong’s (2009, 102) data from MSN Messenger and ICQ chats. Non-Cantonese speakers, such as English speakers and even Mandarin speakers, often imitate Cantonese speakers by adding a laa sound to the ends of their sentences (whether they mean to use laa1, laa3, or laa4 is unknown; it is more likely that they are unaware there is more than one laa particle). As mentioned, this thesis follows the precedent of previous work on Cantonese utterance particles, including that of Yau (1965), Kwok (1984), and Luke (1990), by considering the tones as lexical. In other words, the particles laa1, laa3 and laa4 are all considered to be different from each other. The particle laa3 will be looked at in Chapter 6. (Laa4 is a much less commonly-used particle, used only 24 times in the Hong Kong Cantonese Corpus, and will not be considered in this thesis.)

This chapter begins by providing an overview of previous work on laa1, and problems with these descriptions will be discussed. Some common problems include descriptions which are contradictory, vague, inaccurate, or obscure. One clear, testable, and cross-translatable NSM explication will be proposed for the invariant meaning of laa1. The explication will be shown to be supported by examples taken from the corpus, which help to examine and justify the explication. The wide range of different examples of laa1 allows us to observe its use in an array of real and natural situations, and shows that the proposed explication is able to account for all of its uses. This chapter also uses the examples to show the problems with previous descriptions. The analysis
shows that *laa1* does have a stable meaning, and the NSM explication is able to expose this meaning clearly. It is found that *laa1* indicates that the speaker believes the hearer to know how the speaker thinks about something, and that the speaker feels they do not need to say more about it.

NSM explications for *laa1* were previously published in HHL Leung (2012, 2013). Those earlier versions have since been slightly modified (see section 3.2.1). The semantic analysis given in those papers was significantly condensed compared to the analysis presented in this chapter. This chapter also features an updated literature review. Section 3.1 overviews previous descriptions of *laa1*. To begin with, definitions of *laa1* found in Cantonese-English dictionaries will be discussed in section 3.1.1. Following this, scholarly work on Cantonese utterance particles will be examined in section 3.1.2. Section 3.2 presents the NSM semantic analysis of *laa1*.

### 3.1 Previous descriptions of *laa1*

#### 3.1.1 Cantonese-English dictionary definitions of *laa1*

Typical problems with dictionary definitions of particles include that they are contradictory, uninformative, describe different particles in the same way, only focus on certain aspects of the particle, and frequently mistake utterance meaning for particle meaning. These points will be expanded upon briefly here, and the remaining chapters in this thesis will not consider dictionary definitions.

The first problem with dictionary definitions of particles is that many ideas seem contradictory. For example, Huang (1970, 414) and Lau (1977, 480) both describe *laa1* as simultaneously ‘commanding’ and ‘requesting’, though these can be thought of as opposing speech acts. Both also state that *laa1* indicates ‘agreement’, which is again different. Cowles (1965, 489) and Meyer and Wempe (1947, 287) both describe *laa1* as implying ‘urgency’ and ‘completion’, which overlooks the fact that ‘urgency’ implies something has not yet been completed. They also state that *laa1* indicates ‘certainty’.

Secondly, dictionary definitions are generally not very informative, because they are not specific enough. For example, Chik and Ng Lam (1994, 63,
have two entries for laa1 particles (with differing Chinese characters), both simply described as ‘a phrase-final particle’, and O’Melia’s (1941, 83) entry for laa1 states ‘final, declarative, imperative’. This does not reveal much about meaning, and furthermore, many of the labels used for laa1 are used for multiple other particles. For example, O’Melia’s description of the particle laa3 is exactly the same, save for the addition of ‘emphatic’ – itself a vague and unhelpful description – placed in parenthesis. Neither definition is satisfying, and the difference between laa1 and laa3 is not clear.

Thirdly, the same definitions are given for multiple particles even if they are not interchangeable or even synonymous. Lau (1977, 480) defines laa1 in this way: ‘expresses idea of requesting, commanding, or advising at end of imperative statements; expresses idea of an agreement of some kind having been reached’. If we look at Matthews and Yip’s (2011, 391) list of utterance particles, we see that the descriptor ‘requests’ is also used for aa1, ‘advice’ is also used for laa3, and ‘agreement’ is also used for lo1.43 However, aa1, laa3, and lo1 are not truly synonymous with laa1. Replacing laa1 with any of these particles would either change the meaning or tone of the utterance, sound unnatural, or not make sense.

Finally, another of the main problems with previous descriptions of laa1 has been the lack of focus on the particle’s ‘core’ or invariant meaning. The example of Meyer and Wempe (1947, 287) was used in Chapter 1. They provide two examples to show that laa1 implies completion and certainty: zou6 hou2 laa1 [lit. ‘do complete/good laa1’], meaning ‘done’, which they translate as ‘it is finished’, and jat1-ting6 laa1 [lit. ‘definitely laa1’], which they translate as ‘certainly’. It should be obvious that this is due to the lexical content of the utterances themselves, which convey completion and certainty with or without the particle laa1. Of course, dictionaries are not intended to be in-depth semantic analyses of utterance particles, and so they are not further reviewed in the following chapters. The next section focusses on previous academic studies of laa1. Many of them have the same problems.

43 This, despite Matthews and Yip’s list not even including all Cantonese utterance particles – their aim was to show the main variants of the particles which appear to form groups.
3.1.2 Previous academic studies of laa1

In one of the earliest studies of Cantonese utterance particles, Yau’s S-Q test categorised laa1 as a Q-type particle, or one ‘demanding a verbal confirmation’ (Yau 1965, 39-68). Other aspects of Yau’s work have been critiqued in Chapter 1, but here it can be added that many examples from the corpus show that laa1 can be used outside of situations that demand a verbal confirmation. In Yau’s C-test, laa1 was found to be associated with the ‘connotation concepts’ of ‘coaxing’, ‘persuading’, ‘reminding’, and ‘politely urging’ (Yau 1965, 82-120). This is more helpful than the results of the S-Q test, but still not very helpful, because many other particles were found to be associated with the same connotation concepts.

The second study of Cantonese utterance particles was carried out by Gibbons (1980). Gibbons differentiated between a particle la with a high level tone and a particle la with a high falling tone, although both are transcribed as ‘laa1’ in the Jyutping system. Gibbons (1980, 768, 770) categorised both of these as ‘mands’ requiring a response in terms of action, using a meta-category from Lyons (1977, 745)44. This corresponds somewhat with Yau (1965, 39-68), but Yau’s conclusion was that laa1 ‘demands verbal confirmation’, whereas Gibbons believes it requires a response in terms of action. In Gibbons’ study, la with a high level tone was given a ‘strength’, or ‘degree of expectation of a response’, of 2 on a scale with 1 being the weakest and 3 being the strongest. La with a high falling tone was given a ‘strength’ of 3 (Gibbons 1980, 770).

Gibbons’ descriptions can be useful when viewed as parts of a wider study of all particles and where these descriptions can be considered relative to other Cantonese utterance particles, but they are less helpful when one wants to find out the meaning of an individual particle, and circularity is a big problem. For example, it may be useful for some to know that laa1 is categorised in the same class (‘mand’) as the particle le3, but that le3 has a ‘strength’ of 1, which is weaker than both la with a high level tone and la with a high falling tone (Gibbons 1980, 770). However, this is not helpful if one does not already know the meaning of le3 – which is described in relation to the other particles studied,

44 Lyons (1977, 745) explains that ‘mand’ is ‘a general term to refer to commands, demands, requests, entreaties etc.’, and that mands are a subclass of directives.
in particular the two *laaa1* particles. Furthermore, as with many other descriptions of Cantonese utterance particles, descriptions such as ‘desire a response in terms of action’ are inadequate and do not give readers a clear sense of when and how a particle is used.

Kwok (1984) writes that when suffixed to statements, *laaa1* is similar in function to *la* in Mandarin, which is a particle of ‘lively enumeration’. It indicates ‘a certain lack of definiteness, a lack of finality or completeness’ (Kwok 1984, 55). When used in imperative sentences, Kwok (1984, 78) believes *laaa1* to be the ‘neutral form’, being chosen as the suffix to commands and requests ‘except when one wishes to express some special meaning’. She states that when used in imperative sentences, *laaa1* also denotes ‘a lack of forcefulness’, ‘so that the sentence is more of a request than a command’ (Kwok 1984, 79).

It should be apparent that these descriptions are quite ambiguous. The statement that *laaa1* is the ‘neutral form’ for imperative sentences is perplexing given that Kwok also states that imperative sentences do not always take an utterance particle (Kwok 1984, 78). Likewise, the idea of *laaa1* being used ‘except when one wishes to express some special meaning’ is very ambiguous. Without knowing the meaning of *laaa1*, it is equally impossible to know when one wishes to convey a meaning not expressed by *laaa1*. As explained in Chapter 1, Kwok’s tendency to link Cantonese utterance particles with a Mandarin counterpart is not helpful.

Furthermore, it will be shown in this chapter that Kwok’s definition cannot be applied to all instances of *laaa1*, despite her intention to isolate a ‘core’ meaning for each particle. Kwok (1984, 55-57) explains that *laaa1*’s lack of definiteness may be reinforced by words like *dou2*, which means ‘about’ or ‘approximately’, or *waak6ze2*, which means ‘maybe’ or ‘perhaps’. An example used by Kwok is ‘*sei3 jyut3 dou2 laa1*’, meaning ‘around April *laaa1*’, reproduced in Kwok’s gloss and translation in (3.1) below. She states that this ‘shows the idea of something which is approximate and not definite’. It seems, though, that Kwok has made the common mistake of confusing the meaning of *laaa1* with the meaning of the utterance to which it is attached. In a sentence with *dou2* or *waak6ze2*, the words *dou2* and *waak6ze2* themselves convey indefiniteness and approximation. The example below would still be translated into English as ‘around April’, even without *laaa1*.  

97
‘Well, around April, I suppose.’ (Kwok’s gloss and translation)

The literature is not in agreement about this, as laa1 has been described as conveying ‘lack of definiteness’, ‘certainty’, ‘completion’, and when appended to imperatives, can be very direct and forceful (Kwok 1984, 55-57, Cowles 1965, 489, Meyer and Wempe 1947, 287, Fung 2000, 103). Numerous examples of laa1 presented in this chapter will show that the feeling of something being approximate and not definite is not part of the invariant meaning of laa1. These ideas will be discussed further in section 3.2.2 as examples from the corpus are given.

The next major work on Cantonese utterance particles was carried out by Luke (1990), who separated the uses of laa1 according to the main kinds of sequences in which utterances with laa1 tend to occur: ‘reportings and storytellings’, ‘listings and instructions’, ‘understanding checks’, ‘suggestions’, ‘agreements’, and ‘pre-closings’. Though a significant advance on previous work, Luke’s analysis still does not give sufficient information on the semantics of laa1. Knowing that laa1 can sometimes occur in lists and sometimes in pre-closings, for example, does not tell the reader much about what laa1 actually means. In fact, in each of the contexts identified by Luke, there is no requirement to use laa1, or even any particle. In other words, while it may be true that laa1 is used in, for example, ‘understanding checks’, not all cases of understanding checks require the use of laa1, or any other particle. Furthermore, some of the identified sequences imply that the same particle is used in very different ways, while conversely, some examples may not fit easily into any of the identified sequences. As an example, consider (3.2), taken from the Hong Kong Cantonese Corpus, where the speaker says she does not want her friends to spend too much money on her for her birthday. This example does not fit clearly into any of the sequences listed above.

(3.1)

\[ sei\ jyt \quad dou \quad la \]
April approximate la

‘Well, around April, I suppose.’ (Kwok’s gloss and translation)

(3.2)

\[ Fai3si6 \quad jiu3 \quad nei5dei6 \quad po3fai3 \quad laa1. \]
not-bother need you (PL) spend-too-much-money PRT

‘I don’t want to bother [make] you guys to spend too much money laa1.’
Luke believes that his overarching label of ‘common ground establishment’ encompasses the large variety of work laa1 performs in conversation. Laa1 provides ‘a linguistic resource with which conversation participants can, at particular points in an unfolding interactional scene, document to each other their assumption that what is being talked about is something known-in-common’ (Luke 1990, 112). This is actually quite close to the account presented in this chapter, albeit that, as mentioned in Chapter 1, Luke (1990, 3-4) himself states that Cantonese utterance particles have no semantic content.

Grouping all l-initial particles into a ‘family’, Fung (2000) proposed that all the described senses of the particles laa1, laa3, laa4, laak3, lei4, lo1, lo3, lo4, lok3, le3, le4, and le5 are derived from one core semantic feature, namely ‘realisation of state-of-affairs’. She believes that all other senses can be inferred from this core semantic feature combined with contextual information (Fung 2000, 74). At the same time though, Fung states that not all of the l-initial particles encode all the semantic features of l- (Fung 2000, 83). This is rather puzzling, and brings into question the plausibility of the ‘sub-syllabic morphemes’ approach (this is explored in more detail in Chapter 8). Fung (2000, 96-99) claims that a speaker using laa1 assumes that the hearer has knowledge of the state-of-affairs in question. This is similar to Yiu’s (2001, 140-141) description ‘obviousness’ for laa1.

In marking directives, Fung (2000, 82) claims that laa1 is used for ‘persuasion’. This is reminiscent of Yau’s (1965) finding, mentioned above, that laa1 is associated with the connotation concept of ‘persuading’. Fung’s accompanying example, though, appears to suffer from the common problem of taking the utterance meaning to be the particle meaning. Her example is given below as (3.3). An utterance like ‘do believe me’ could be described as ‘persuasive’, even if laa1 were omitted.

(3.3)  
Nei seon ngo laa1.  
you believe me FP45  
‘Do believe me.’ (Fung’s gloss and translation)

45 Her abbreviation ‘FP’ stands for ‘final particles’.  
99
Fung claims that in this example, the speaker does not declare the realisation of state-of-affairs in the real world, but instead envisages the realisation of the state-of-affairs in the potential world (Fung 2000, 82-83). The speaker supposedly considers the hearer to have the potential to let the state-of-affairs be realised, and the hearer is persuaded to bring the state into reality.

Regarding the use of laa1 in lists, Fung (2000, 100-101) refers to this as laa1’s ‘enumeration function’, but she acknowledges that enumeration is still conveyed after omission of laa1, instead coming from utterances containing parallel constructions. Fung posits that in such utterances, laa1 performs its usual function, which is to indicate that every state conveyed in the proposition has been realised. This has some similarities with the analysis of laa1 in lists as given by Luke (1990, 69-74), as well as the analysis here which explains the meaning of laa1 using NSM. This will be discussed further below (see e.g. example (3.18)). One point of concern, though, is that Fung believes it is the high tone of laa1 that somehow facilitates the enumeration function – she states that the high tone is regarded as a listing tone signalling incompleteness and non-finality. This is questionable, considering that the particle aa3 is also commonly used in between list items. The reason laa1 (and aa3) is more compatible with lists is more likely to be because of semantics.

Matthews and Yip appear to have been influenced by many of the descriptions already discussed. They state that laa1 is used for ‘requesting, seeking common ground’ (Matthews and Yip 2011, 391). They include laa1 in the group of ‘imperative and persuasive’ particles which are typically used in ‘giving directions, including not only straightforward commands, but also suggestions, requests and advice’ (Matthews and Yip 2011, 403-404). Specifically, laa1 is said to often accompany invitations and requests. They also write that laa1 will ‘often serve to fill a pause, for example in lists’ (Matthews and Yip 2011, 393).

The bulk of this chapter will discuss the NSM semantic analysis of laa1. Since previous descriptions of laa1 have varied so much and are at times contradictory or inaccurate, finding the invariant meaning of the particle is not as straightforward as simply summarising all the previous accounts or finding where the previous accounts overlap with each other. The proposed NSM
explication is given in section 3.2.1. Following this, a large number and range of examples of laa1 will be discussed in section 3.2.2.

3.2 NSM semantic analysis of laa1

3.2.1 The NSM explication for laa1

The latest explication proposed for laa1, labelled [3A], is presented below in both English NSM and Cantonese NSM. Numerous real, naturally-occurring examples of laa1 from the Hong Kong Cantonese Corpus were used to determine and to test this explication, with some of these examples given throughout this chapter to demonstrate the explication’s validity. It will be shown that the explication can be used to fully and clearly explain numerous instances of laa1, in a range of situations and contexts. The proposed explication also has some links with previous descriptions of laa1.

[3A] Final explication for laa1:

\[
\text{you now know how I think about this} \\
\text{I can not-say more}
\]

\[
\begin{align*}
\text{lei5} & \quad \text{ji4gaa1} & \quad \text{zi1dou3} & \quad \text{ngo5} & \quad \text{dim2} & \quad \text{lam2} \\
\text{you} & \quad \text{now} & \quad \text{know} & \quad \text{I} & \quad \text{how} & \quad \text{think} \\
\text{li1} & \quad \text{jat1} & \quad \text{joeng6} & \quad \text{je5} \\
\text{this} & \quad \text{one} & \quad \text{CL} & \quad \text{thing} \\
\text{ngo5} & \quad \text{ho2ji5} & \quad \text{m4} & \quad \text{zoi3} & \quad \text{gong2} \\
\text{I} & \quad \text{can} & \quad \text{not} & \quad \text{more} & \quad \text{say}
\end{align*}
\]

Several comments about the explication may be helpful here. The ‘you’ in the explication refers to the person to whom the utterance is directed. The ‘I’ refers to the speaker him/herself. ‘This’ refers to the topic the speakers are talking about, especially the utterance to which the particle is attached. It should be noted that the last part of the explication is ‘I can not-say more’, and not ‘I cannot say more’. ‘Not’ and ‘say’ have been hyphenated above to highlight this. Roughly, ‘I can not-say more’ means ‘I don’t have to say more’ or ‘I don’t need to say more’, but these phrasings are not acceptable NSM.\(^46\) The component ‘I can

\(^{46}\)For semantic analyses of have to, have got to, and must, see Goddard (2014a). Using have to in explications causes translation problems in languages like Russian and Yankunytjatjara.
not-say more’ is acceptable in NSM, including in the Cantonese version of this explication.

Note in relation to the component ‘you now know how I think about this’, that the ‘knowledge’ may go beyond the explicit content of the utterance laa1 is attached to. Nonetheless, the speaker’s message is that this utterance is enough in context for him/her to be confident that the hearer ‘now knows how I think about this’. Whether the hearer does know in reality is irrelevant (although most of the time, it is true that the hearer understands). Each speaker’s language reflects his or her own (purported) perception of things, which is not necessarily the reality. It is also not a requirement that the speaker immediately stop talking – the explication only states that the speaker is pointing out that it is not necessary to elaborate on what is (perceived to be) understood.

To help show the development of the explication over time, the first published explication is reproduced below, labelled [3B]:

[3B] Early explication proposed for laa1 (HHL Leung 2012, 256):

a. I say this because I want you to know what I think
b. I think like this now: ‘you know what I think about this’
c. because I think like this now, I can not-say more

As may be evident when compared with [3A], explication [3B] was modified to reduce unnecessary elements and thereby ‘streamline’ the explication. This reflects an important realisation: that shorter, simpler explications act more like particles themselves. The smaller number of components means that they are more readily ‘attachable’ to a wider range of utterances. Nevertheless, the overall intended meaning is essentially the same.

3.2.2 Testing of explications and descriptions using the corpus

This section presents a range of roughly 20 examples of laa1 taken from the Hong Kong Cantonese Corpus (some excerpts feature more than one instance of the particle). As explained in section 1.7, they have been chosen to support

Instead, have to can be captured in an explication. Have to includes CAN’T NOT, therefore it is not surprising that CAN NOT SAY as used in the explication of laa1 is roughly don’t have to say.

47 Before the publication of [3B], a component considered in place of the first line (a) was ‘I say this now because I think this’. An earlier component considered in place of the second line (b) of [3B] was ‘I think like this now: ‘you know what I think about this now/after this’. The first Cantonese explication of laa1 was not published until HHL Leung (2013, 18).
discussion of the different uses or interpretations of the particle in everyday conversation. This also allows testing of the explication by allowing substitution into real utterances, demonstrating that the explication is a plausible model of what laa1 means. This section also makes references to the previous descriptions of laa1 discussed in section 3.1, as well as NSM components of earlier explications. The subsections that follow are for expository convenience only and are not intended to indicate that laa1 has multiple meanings or that its different interpretations are unrelated. We start with the use of laa1 in ‘persuasiveness’ contexts.

**Uses of laa1 in ‘persuasiveness’ contexts**

Example (3.4) below is the beginning of one of the recorded conversations in the Corpus. It shows two instances of laa1 being used. The first is given by speaker A and is labelled (3.4a). According to explication [3A], laa1 is being used to mean ‘you now know how I think about this, I can not-say more’. As the two speakers were conversing prior to this excerpt, they obviously have knowledge of what they were discussing before, and speaker A believes that B knows what she is referring to. Hence, the first part of the explication. (The exact reasoning for why the hearer is expected to know how the speaker thinks would not be necessary to detail in the explication, as it would then become too narrow and unable to cover the whole range of uses of laa1.) Since this is the case, speaker A does not need to, for example, repeat the earlier conversation or explain anything else. In other words, speaker A does not need to say more, and indeed, B understands and picks up the conversation again.

(3.4)

A:  
*Batjyu4 gai3zuk6 tau4sin1 gong2 ge3*  
let’s/why-not continue earlier speak LP  
*je5 laa1. (3.4a)*  
thing(s) PRT  
‘Let’s continue the conversation from earlier laa1.’

48 In the translation given here for this sentence, the English formula ‘let us’, used often by English speakers to sound less imposing, seems to have reduced the feeling of persuasiveness (see Wierzbicka (2006, 183-203) for English ‘let-constructions’). Another possible translation might have been ‘why not continue the conversation from earlier?’ although the use of ‘why not’
B:

Hou2 aa1 hou2 aa1 hou2 aa1. Gam2 hai6

PRT good PRT good PRT so yes

laa1, (3.4b) gong2 faan1 ngo5dei6 gan6fong3

PRT speak back we recent-situation

aa1.

PRT

‘Good aa1, good aa1, good aa1. So yes laa1, going back to talk about our recent situation aa1.’

This utterance by speaker A in (3.4a) could be described as ‘persuasive’, which, as we have seen, is a commonly used description of laa1 (Yau 1965, 82-120, Fung 2000, 82, Matthews and Yip 2011, 403-404). Yau also adds the synonym ‘coaxing’. Yet the English verb persuade is not very fitting because it implies that the addressee is less willing than the speaker to do something, that the speaker anticipates some resistance, and because of this, that the speaker has to say a number of things in an attempt to change the addressee’s mind. This is not the case in (3.4a), and we can see that speaker B responds by stating her agreement more than once. These aspects of persuade are incompatible with laa1.49 According to the proposed explication, and supported by (3.4a), a Cantonese speaker using laa1 assumes (or appears to assume) that the addressee understands how the speaker thinks, and that there is no need to say more.

We can now focus on speaker B’s use of laa1, labelled (3.4b). The explication can be used to explain laa1 here as well. Speaker B’s use of laa1 conveys that speaker B thinks A knows how she thinks. Speaker A is expected by B to remember their earlier conversation and to know what speaker B is referring to. The reason for B making this assumption is irrelevant to the meaning of laa1, but it may be because A mentioned it in the previous turn, because B knows A was present, because it was not very long ago, etc. In any case, it seems obvious to B that A knows how she thinks. Next, there is the

---

49 For a more detailed comparison of the meaning of laa1 with the meaning of the English persuade, see HHL Leung (2012, 259-260).
sense that because A supposedly knows how B thinks, B does not need to say more about it. This corresponds with the second component of the explication, ‘I can not-say more’. Notice again, that this is different from ‘I cannot say more’, and equally it is different from ‘I will not say more’. Using laa1 does not require that the speakers stops talking, and B in this example carries on.

Speaker B’s use of laa1 in (3.4b) is clearly not persuasive, since it is already known that speaker A would like to continue the previous conversation. It also does not demand a verbal confirmation, as Yau (1965, 39-68) claimed. Here, speaker B is in agreement. As seen above, ‘agreement’ has also been offered as one of the attributes of laa1, for instance in Huang (1970, 414) and Lau (1977, 480). Luke (1990, 98-102) also identified ‘agreements’ as one of the sequential contexts in which laa1 occurs, drawing attention to the role that laa1 plays in displaying mutual agreement and common understanding. This is a valuable observation, as Luke refrains from saying that the particle itself indicates agreement, and maintains that laa1 can help carry out a wide range of interactional tasks. Observe that an utterance like ‘good, good, good, so yes...’ clearly conveys agreement even when laa1 is not used. The proposed explication of laa1 is compatible with situations where agreement is being conveyed, and laa1 can be used in utterances which convey agreement, but to say that laa1 means agreement would be to make the common mistake of confounding the meaning of the utterance with the meaning of the particle. A number of other examples of laa1 in this chapter show that ‘agreement’ is not part of its invariant meaning (see e.g. (3.13), (3.15)).

It may be interesting to note that one of the earliest versions of the explication for laa1 included the component ‘I think you can know what I think’ in place of the current first line ‘you now know how I think about this’. However, ‘I think you can’ made the speaker seem too uncertain, and was rejected in favour of a component in which the speaker seemed more confident in his/her assumption.59 This was supported by instances of laa1 like speaker B’s in (3.4b). Since speaker A has already expressed her opinion, and the two speakers are in agreement, it is unlikely that B would only think that A could know what she was thinking. Many other examples from the corpus also

59 The component which was used instead in explication [3B] was ‘I think like this now: “you know what I think about this”’. 105
support this change. The related idea of ‘certainty’ will be looked at more closely further below, with reference to the use of *gang2hai6* ‘of course’.

One description of *laa1* more applicable to (3.4b) is Luke’s (1990) idea of ‘common ground establishment’. Speaker B can be interpreted as signalling to A that there is mutual understanding or agreement, and that they are thinking the same thing. We can discuss Luke’s description further by considering example (3.5). Example (3.5) is similar to (3.4) in that the speaker using *laa1* is referring to some prior knowledge. In the first turn of (3.5) shown below, the conversation stops temporarily as speaker B answers the phone, and upon resuming, *laa1* is used when referring to an earlier point in the original conversation.

(3.5)

B:

Ei3, *jau5 din6waa2 aa3, ngo5dei6 zaam6ting4 eh* have phone[call] PRT we pause

*haa5 sin1... Hou2 laak3, ngo5dei6 ho2ji5 gai3zuk6.* DEL first good PRT we can continue

‘Eh, I have an incoming call *aa3*, we’ll pause for a while first... Okay *laak3*, we can continue.’

A:

*M6.*

mm

‘Mm.’

B:

*Gam2 tau4sin1 gong2 dou3 ne1, zau6 waa6 tung4 so* previously speak to PRT then say with

*di1 jan4 lyun4lok3 *laa1.*** CL people contact PRT

‘So before we were talking about *ne1*, contacting people *laa1.*’

We can understand the use of *laa1* in (3.5) more clearly by considering the proposed explication. The first part of the explication ‘you now know how I think about this’ indicates that B now thinks A knows how B thinks about contacting people. In this example, B reasonably expects the addressee to have prior knowledge and understanding of the first part of their conversation. Because they should be in mutual understanding as to what B is thinking about,
B can not say more – this is shown in the second part of the explication, ‘I can not-say more’. This might be taken to mean that B does not need to repeat the earlier part of the conversation about contacting people.

The use of laa1 in (3.5) is highly compatible with Luke’s (1990) description of laa1 as being used to establish common ground. According to Luke (1990, 56), there can be organisational problems where speakers have to sustain mutual orientation as to what they are doing, where they are in the report, and what to do next. The validity of Luke’s view can be seen in example (3.5). When B restates the previous topic after having answered the phone, this may be interpreted as a brief summary of their earlier conversation or as an introduction to what will be said next, which helps with ‘organisation’ and ‘orientation’. Following this line of thought, Luke explains that in ‘reportings and story-tellings’, laa1 is sometimes ‘used to segment an extended reporting into chunks’. We can interpret the first ‘chunk’ of conversation in (3.5) as that which occurred before the phone was answered. The second ‘chunk’ would be the part of the conversation to take place after the phone has been answered.

The important point is that this organisation and mutual understanding in (3.5) corresponds with the proposed explication for laa1. According to the line ‘you now know how I think about this’, speaker B expects speaker A to know what they were discussing before about contacting people. Once they have mutually ‘organised’ their conversation, the second line of the explication ‘I can not-say more’ indicates that speaker B believes nothing more has to be said about the earlier part of the conversation. Through this organisation, the use of laa1 has helped speaker B ‘announce’ (or in this case, re-announce) the topic she wants to talk about, i.e. contacting people.

Luke (1990, 56-57) admits, though, that laa1 is only sometimes used for such segmentation, saying that ‘laa1 is not the only particle that can be used to segment an extended reporting into chunks; neither is it a necessity that some particle be used’. Therefore, one could argue that identifying sequential contexts in which laa1 can occur, as Luke has done, does not give sufficient information as to what laa1 means or how it should be used. It is not as helpful as having laa1’s invariant meaning explicated using NSM.
A related use of laa1 suggested by Luke (1990, 63-64) is that laa1 helps to announce a topic on which extended talk is about to be delivered. Through the use of laa1, the topic introducer displays his/her assumption that the addressee can know what the topic is. This allows the speaker to secure an extended slot for its delivery. This does not seem to be due to laa1 here. Although in (3.5) speaker B does continue talking with a longer turn, the same sentence without laa1 would probably still allow for an extended slot. Another indication can also be found in the rest of the conversation. Before answering the phone, speaker B already had many extended turns talking at length, while A’s responses were much shorter. Speaker A’s responses included saying things like ‘yes’, or asking questions so that B could continue the narrative. When B says they can continue after having answered the phone, A merely says ‘mm’ instead of contributing anything more significant or informative. This is likely because it has already been established in other ways that speaker B has a lot more than A to say on the topic. This idea that laa1 introduces something for a longer turn has not been included in the proposed explication for laa1.

Other previous descriptions of laa1 also appear not to be valid in light of example (3.5). The utterance in (3.5) is not ‘persuasive’, ‘commanding’, or ‘requesting’, and does not imply ‘urgency’ or ‘completion’. Furthermore, laa1 here does not, however strongly or weakly, require a response in terms of action, or demand a verbal confirmation. This shows that Gibbons’ (1980) label of ‘mand’ and Yau’s (1965) categorisation of laa1 as a ‘Q-type particle’ that demands a verbal confirmation are both unable to be applied to all cases of laa1. Descriptions of laa1 more applicable to example (3.5) are Kwok’s (1984, 55) ‘lack of finality or completeness’, since the speaker is continuing something that was unable to be finished previously, Yau’s (1965, 82-120) ‘reminding’, since the speaker is restating the topic, and Luke’s (1990) ‘common ground establishment’, since B can be said to be checking that they are thinking of the same thing. Examples like (3.5) led to consideration of NSM components like ‘I know that you know it’, but these were later rejected as they did not apply to all examples (see e.g. (3.10) and (3.11)).
Is ‘certainty’ part of the meaning of *laa1*?

Next, it will be beneficial to examine utterances which convey ‘certainty’, a description of *laa1* given by Cowles (1965, 489) and Meyer and Wempe (1947, 287). Such utterances occur very frequently in the corpus. As explained, the proposed NSM explication aims to be general enough to cover all uses of *laa1*, regardless of the content of the rest of the utterance. Since ‘certainty’ is not part of the proposed meaning of *laa1*, testing the explication against some utterances which convey certainty will reveal whether or not it is applicable in such common situations. Consider example (3.6), which is a response to someone talking about their sister’s pet guinea pigs smelling very badly.

(3.6)

```
Nei5 dou1 m4 tung4 keoi5 cung1loeng4
you even not with it shower
```

`gang2hai6  laa1.
of-course  PRT`

‘You don’t even give it [the guinea pigs] showers/washes, so of course [they smell bad] *laa1*.’

The *laa1* in (3.6) can be considered against the explication proposed. The speaker thinks it is obvious that guinea pigs that do not get washed become smelly. Since this is a fairly logical conclusion to come to, the speaker assumes the addressee understands what she thinks about this. Because the link between hygiene and smell/dirtiness is so ‘obvious’ or logical, B does not feel the need to explain further, and does not say anything more in this turn. Thus the explication ‘you now know how I think about this, I can not-say more’ can be applied to this example.

It would not be surprising for B’s utterance in (3.6) to be described in English as conveying ‘certainty’ or ‘persuasiveness’. ‘Persuasiveness’ has been discussed above in relation to (3.4), but the idea of ‘certainty’ can be examined further. The use of *gang2hai6* needs to be highlighted. *Gang2hai6* means ‘of course’, or literally ‘definitely is’. It therefore seems reasonable to assign the feeling of ‘certainty’ to *gang2hai6* and not *laa1*. The same sense of certainty is felt even in the free English translation of (3.6), and even without *laa1*. The word *dou1*, which is very roughly glossed as ‘even’, also contributes to the feeling of certainty here. It has a complex meaning that is not easily translatable.
into English, but seems to add some ‘strength’ to the speaker’s position. Without *gang2hai6* (and perhaps *dou1*), the utterance does not sound very strong or certain. This supports the idea that certainty is not part of the meaning of *laa1*.

Examples (3.7) and (3.8) similarly show examples of *laa1* being used with *gang2hai6* ‘of course’ and statements that seem to point out very obvious or reasonable things. This is in line with Yiu’s (2001, 140-141) description ‘obviousness’ for *laa1*. Neither demand verbal confirmation (Yau 1965, 39-68) or require a response in terms of action (Gibbons 1980, 768, 770). In (3.7), speaker E is saying that a school’s principal would obviously claim that the school is a very good one. Since this is to be expected, E can use *laa1* to express ‘you now know how I think about this, I can not-say more’.

(3.7)

M:

\[\begin{array}{llllll}
Daan6hai6 & go3 & haau6zoeng2 & sing1cing1 & waa6 & keoi5 \\
but & CL & school-principal & claim & say & s/he \\
nit & gaan1 & hai6 & hou2 & hou2 & ge3 & hok6haau6. \\
this & CL:building & is & very & good & auxiliary & school \\
\end{array}\]

‘But the school principal claims that this is a very good school.’

E:

\[\begin{array}{lll}
Keoi5 & sing1cing1 & ze1. & Gang2hai6 & laa1, \\
s/he & claim & PRT & of-course & PRT \\
keoi5 & haau6zoeng2 & aa1 & maa3. \\
s/he & school-principal & PRT & PRT \\
\end{array}\]

‘S/he claims *ze1* [that’s all]. Of course *laa1*, s/he is the school principal *aa1-maa3*.’

(3.8)

A:

\[\begin{array}{llllll}
Wai3 & dit3 & zo2 & lok6 & ngaa4i4 \\
hey & fall & PFV & down/drop & cliff \\
dou1 & mei6 & sei2? & \\
still/even & not-yet & die \\
\end{array}\]

‘Hey [he] fell off a cliff but still hasn’t died?’
Example (3.8) is from part of a conversation in which two speakers are discussing a drama. There is some confusion over whether or not one of the characters has died, and speaker B says that the plot will ‘of course’ say that the character survived, because otherwise there will be no more drama/story. Again, this seems to be fairly reasonable in the eyes of the speaker, who is pointing it out as something that should be obvious. The speaker is expressing ‘you now know how I think about this, I can not-say more’.

Example (3.9) shows laa1 being used with a synonym of gang2hai6, namely dong1jin4, which is also translatable as ‘of course’. M wanted to speak to one of E’s younger siblings, Ziling, but E reveals that his family has gone on a holiday. M questions whether this means Ziling is not available, and E says ‘of course
laa1’. E is saying ‘you now know how I think about this, I can not-say more’ presumably for several reasons: M should know that Ziling is part of the family, he has already stated that they are away in the previous turn, and it is reasonable that his younger siblings would be with his parents on a family trip. From his use of dongjin4 ‘of course’, we know that E thought this was quite obvious. It makes sense that he would feel he did not have to say more about it.

As has been shown, gang2hai6 ‘of course’ and its synonyms are often used in the same utterance as laa1. Their frequent collocation can easily lead some to conclude that laa1 conveys certainty. A quick look at the corpus data shows that, in fact, in the vast majority of cases, gang2hai6 is accompanied by laa1, although a much lower proportion of laa1 examples are accompanied by gang2hai6.51 In other words, laa1 is a popular choice when speakers want to convey certainty, but laa1 also frequently occurs in a wide range of other situations – certainty is not part of the invariant meaning of laa1.

While laa1 itself may not mean ‘certainty’, there must be something in the meaning of laa1 that is highly compatible with such words and phrases. The proposed explication can help explain this too. The explication ‘you now know how I think about this, I can not-say more’ corresponds well with a sense of something being obvious, clear, or to be expected, and therefore, from the speaker’s point of view, not in need of further explanation. If something is obvious or clear, it is compatible with gang2hai6 ‘of course’ and its synonyms, even in English.

The numerous examples of laa1 with gang2hai6 ‘of course’ and its synonyms also support the decision, mentioned above, to omit the early component ‘I think you can know what I think’. Again, this would make the speaker seem tentative and unsure, although many examples show that the speaker can be confident in what is being said. When using laa1, the speaker seems to always assume that the addressee knows what s/he thinks. The current component ‘you now know how I think about this’ is more fitting with phrases like gang2hai6 ‘of course’. At the same time, it can explain those utterances that do not contain a word like gang2hai6.

51 Unfortunately, the current corpus seems unable to determine exact figures for how often gang2hai6 ‘of course’ or its synonyms are collocated with laa1.
Does laa1 convey a ‘lack of definiteness’?

Recall that Kwok (1984, 56) states that when suffixed to statements, laa1 conveys a ‘lack of definiteness’, and may be reinforced by words like dou2 ‘about/approximately’ or waak6ze2 ‘maybe/perhaps’. As evidenced by the naturally occurring examples above which convey certainty, indefiniteness cannot be part of the ‘core’ meaning of laa1. But although sentences with ‘about’ or ‘maybe’ might seem different from previous examples where the speaker was very sure or certain about something, and although Kwok’s analysis appears misguided, it is true that laa1 can attach to utterances where speakers seem unsure. Kwok’s example (3.1) does sound natural to a native speaker. Real, naturally-occurring examples of laa1 from the corpus provide more context, and allow us to gain a better understanding of how laa1 is being used in these cases. Example (3.10) below from the corpus shows an utterance very similar to Kwok’s example (3.1). Just like example (3.1), laa1 here is used with dou2 ‘about/around/approximately’. Notably, in such cases, the proposed explication for laa1 is still valid. Here, speaker B asks speaker A about the university admissions process. B has been through the process before, while A has not.

(3.10)

B: ‘Oh, when do you choose [universities] gaa3, [through] JUPAS [Joint University Programmes Admissions System]?’

A: ‘Wow! You should know better than me wo3.’

B: ‘I forgot lo3.’

A: 

\[
\begin{array}{llllll}
Ngo5 & nam2 & sap6jyut6leng4 & dou2 & \text{laa1}.
\end{array}
\]

I think around-October around/approximately PRT

‘I think around October or so laa1.’

B: ‘You mean Form Seven?’

The proposed explication for laa1 is substitutable in example (3.10). Speaker A is using laa1 to say ‘you now know how I think about this’ – that it is around October – and ‘I can not-say more’. (This apparent confidence on the part of the speaker does not of course guarantee that the addressee does know how the speaker thinks, so the fact that in this case B asks for some further information
does not at all invalidate the proposed explication.) Thus the proposed explication is valid when laa1 is used with dou2 ‘around/approximately’ or where there is a ‘lack of definiteness’. This example helped to reject an earlier considered component ‘I know that you know it’. In this example, B clearly does not know when the universities need to be chosen.

Example (3.11) below shows two instances of laa1 being used with dou2 ‘about/approximately’. Speaker J is describing her work and her company to speaker A. Prior to the excerpt given here, J explained that the company has two divisions, one concerned with batteries and one concerned with rubber keypads. She works in the keypad division. Here, she explains that the battery division accounts for approximately two thirds of the company, while the rubber keypad division which she works in accounts for approximately one third.

(3.11)

J:
... Gam2 din6sam1 go2bin1 zau6 kei4sat6 zau6 so/then battery that-side then actually then zim3 go2 gaan1 gung1si1 bei2gaau3 account/take-up that CL company comparatively daai6bou6fan6 lo1. big-portion PRT ‘... And then battery that side actually accounts for a comparatively big portion of the company lo1.’

A:
Hai6 aa4? is/yes PRT ‘Yes aa4?’

J:
Caa1m4do1 saam1fan6ji6 dou2 laa1. (3.11a) roughly two-thirds about/approximately PRT Gam2 ng05 go2bin6 go2go3 division zau6 so/then I that-side that-CL division so/then zim3 saam1fan6jat1 dou2 account/take-up one-third about/approximately laa1. (3.11b)
PRT
‘Around two thirds laa1. And then on my side my division accounts for about one third laa1.’

In both (3.11a) and (3.11b), the speaker is giving a rough idea of how big each division within her company is. She is saying ‘you now know how I think about this, I can not-say more’, because she is giving facts (albeit approximations) which should be sufficient for A. It appears from the conversation that the two are friends and are not in, for example, an employer/employee relationship where J needs to report on specific company matters. One would assume that A does not need exact figures, and we know from the rest of the corpus that speaker A does not ask J for any more details about the structure of the company. The proposed explication is valid for these examples. These examples support the rejection of the earlier considered component ‘I know that you know it’, since it is clear that J does not know.

A few turns later, in the same conversation, speaker A asks J what brand the batteries are. Speaker J responds that she is not very sure, but guesses that the brand is Golden Power. This is shown in example (3.12).

(3.12)

J:

Ngo5  m4.  ng05 mou5  lau4ji3  keoi5
I  not  I  not-have  pay-attention-to  they/it

diu  din6sam1  w03.  Golden  Power  gwaa3,  gaan1
CL  battery  PRT  Golden  Power  PRT  CL:building

je5  giu3  Golden  Power  ng05  gu2  hai6  laa1.
thing  call  Golden  Power  I  guess  is  PRT

‘I don’t – I didn’t pay attention to their batteries w03 [laughter]. [I guess] ‘Golden Power’ gwaa3. The thing [company] is called ‘Golden Power’ so I guess that’s it laa1!’

In this example, the speaker is again speaking without certainty or definiteness, but the explication proposed is still substitutable. Laa1 here means that J thinks A now knows how J thinks of this, and that J does not need to say more. In other words, speaker A should understand that J is not sure, and hence there is not much more to say. Following the excerpt given in (3.12), speaker A moves on to ask a question confirming whether J’s company manufactures the batteries themselves. This shows that speaker A understood what J meant.
The potential for laa1 to be used with utterances conveying certainty, as well as utterances like in (3.10), (3.11) and (3.12) where the speaker is uncertain, highlights that laa1 can be used in a wide range of contexts. It demonstrates that the content of the utterance preceding laa1 is not particularly relevant. What is most relevant is that the speaker needs to have some expectation of the addressee’s understanding, that the addressee will know what the speaker thinks and that the speaker will not have to say more. The proposed explication appears to be valid in all situations where laa1 is found.

Are speakers always in agreement?

Let us now consider whether speakers are in ‘agreement’ when laa1 is used. Example (3.13) is a conversation between three people. One of them, labelled C, owns a rabbit, which sharpened its teeth on a bicycle seat. Another speaker, A, believes that the rabbit is doing this because A is not taking care of it properly. (Speaker B does not say much.) This leads to a disagreement. This is contrary to the observations by Huang (1970, 414), Lau (1977, 480) and Luke (1990, 98-102) that laa1 is used in situations where speakers are in agreement. Example (3.13) is similar to previous examples which featured gang2hai6 ‘of course’, but uses laa1 with hou2 ming4hin2 ‘very clear’.

(3.13)

C: ‘It was grinding its teeth. Like a carrot aa3.’
A: ‘Grinding its teeth? Your rabbit?’
C: ‘The rabbit aa3.’
A: 

Laa4 hou2 ming4hin2 nei5 ziu3gu3 dak1 keoi5
PRT very clear you take-care ADV it
m4 hou2 laa1. (3.13) Tou5ngo6 aa3 keoi5.
not good PRT hungry PRT it
‘Look, clearly you didn’t take good care of it laa1. Hungry aa3 it was.’
C: 

M4hai6 wo3. Keoi5 hai6 niidi1 tin1sing3 lai4
not-is PRT it it these nature come
gaa3 wo3, hai2 dou6 mo4 zim1 di1 ngaa4.
PRT PRT here grind sharp CL teeth

‘No wo3. This is their nature gaa3-wo3, to sharpen their teeth.’

A: ‘Then you should give – then you should give it normal things to grind its teeth on lo1.’

B: ‘Yes lo1.’

C: ‘Yes lo1. My bicycle lo1.’

A: Hou2 ming4 hin2 nei5 ziu3 gu3 dak1 keoi5 very clear you take-care ADV it
m4 gau3 hou2 laa1. (3.13b) not enough good PRT

‘Clearly you didn’t take good enough care of it.’

C: ‘No aa3. Very good.’

As the examples of laa1 in (3.13) show, despite the speaker’s assumption that the addressee will know what he thinks (‘you now know how I think about this’), the other person can disagree. There is no reason to believe the speaker thought the addressee would disagree, and so this does not require the explication to be changed. Recall that the explication indicates that the speaker believed the addressee would know how the speaker thinks, but the explication does not predict or restrict what the addressee’s response can be. This is similar to Fung’s (2000, 80) point that the speaker’s assumptions about the world are not necessarily in line with the actual state of the world. The explication is therefore still applicable. It should also be remembered that in (3.13), there is a third person, B, present. The ‘you’ in the explication could refer to B, and the explication would also still be valid. It seems speaker B does know how A thinks, because B indicates agreement. Due to the limitations of the corpus, it is impossible to tell whom speaker A is really addressing.

It may be interesting to explain the rejection of another component here. Initially, a component such as ‘I want you to think the same way’ was considered for laa1. It is feasible that in example (3.13), A wanted C to think that C had not taken good care of the rabbit, as it was hungry. This would have been more ‘persuasive’. However, this component is questionable when we see that A
responds to C’s disagreeing ‘no’ and explanation that this is the nature of rabbits by then saying that C should give the rabbit something normal to sharpen its teeth with. In other words, after being challenged or corrected, A quickly changed his stance from assuming the rabbit was underfed, to suggesting the rabbit was not given something appropriate to grind its teeth on. It seems that if laa1 included a component like ‘I want you to think the same way’, the speaker would not have given up on the idea of the rabbit being hungry so quickly or easily. Example (3.14), from a conversation about holidays and leave from work, is similar, and also supports the rejection of such a component.

(3.14)

A: 

\[\text{Jyu4gw02 heoi3 jing1gwok3 ng05 seng4jat6 gok3dak1.}\]

\(\text{if go England I always feel}\)

\[\text{Zik1hai6 nei5 heoi3 au1zau1 nei5 gang2hai6 heoi3 - meaning you go Europe you of-course go}\]

\[\text{heoi3 jat1 go3 jyut6 dai2 laa1.}\]

\(\text{go one CL month cheap/good value PRT}\)

‘If going to England, I always feel. I mean if you’re going to Europe you of course go –going for one month is more worth it laa1.’

B: ‘No aa3. The worst thing is you need to find someone to act in your position aa3.’

A: ‘Yes wo3…”

The explication proposed can be applied to example (3.14). Again, the current component ‘you now know how I think about this’ fits more accurately than the rejected component ‘I want you to think the same way’. In (3.14), speaker A says (again with gang2hai6 ‘of course’) that if going to Europe, it is more worth it to stay for a month. Speaker B disagrees, and gives a reason for this, perhaps because B knows that A expects her to know what A thinks. As in (3.13), A then changes her view, and agrees with B. Thus we can see that it is not unusual for a speaker to change his/her stance when corrected or questioned. This provides further evidence that English definitions like ‘commanding’ or ‘persuading’ are inaccurate, as these words imply some sort of rigidity whereby the speaker will not change his/her position. This change in view might be seen as being due to a ‘lack of forcefulness’, as claimed by Kwok (1984, 79) to be a feature of laa1, but
this was not found in this analysis to be part of the invariant meaning of laa1.
In fact, examples like (3.15) contradict this claim.

(3.15)
A:  Nei5 go2 jat6 dim2 aa3 tai2 maan6lyun4
you that day how PRT watch Manchester United
gok3dak1? Hai6 m4hai6 gok3dak1 m4 jing1goi1
feel is not-is feel not should
bong1 keoi5 ne1 haa6nin2?
help them PRT next-year
‘How were you feeling that day aa3 watching Manchester United? Do you feel that you shouldn’t help them ne1 next year?’

B:  M4hai6, haa6nin2...
not-is next-year
‘No, next-year…’

A:  Zyun2tau4 lei6mat6pou2 laa1. (3.15a)
turn/switch/transfer-to Liverpool PRT
‘Transfer to [support] Liverpool laa1.’

B:  Haa6nin2 zyun2tau4 –,
next-year turn/switch/transfer-to
gai3zuk6 gai3zuk6 zici4 maan6lyun4.
continue continue support Manchester-United
‘Next year transfer – continue continue support Manchester United.’

[1 min 45 secs later, after continued discussion about football:]

B:  Goi2tau4 maan6lyun4 laa1! (3.15b)
turn/switch/transfer-to Manchester-United PRT
‘Transfer to [support] Manchester United laa1!’

Example (3.15) shows two speakers discussing football. Speaker A is a supporter of Liverpool, while speaker B is a supporter of Manchester United. In this excerpt, A tells B to change his support to Liverpool instead, using an utterance with laa1 attached. This could be perceived as a forceful utterance,
and therefore goes against the descriptions of Kwok (1984, 55, 79) that laa1 conveys a lack of forcefulness or a lack of finality. Later on in the same conversation, B exclaims to A to switch to Manchester United, again using laa1 in a way which could equally be considered forceful. This is more compatible with Fung’s (2000, 103) claim that laa1-appended imperatives can be very direct and forceful. It could be that each speaker wanted a response in terms of action, as suggested by Gibbons (1980, 768, 770), but we know from other examples in this chapter that this is not an invariant meaning of laa1. The description ‘persuasive’ and the rejected NSM component ‘I want you to think the same way’ would also have worked very well with such utterances. Nonetheless, we have seen from other examples in this chapter (e.g. (3.13) and (3.14)) that these are not part of the invariant meaning of laa1 either.

The current explication ‘you now know how I think about this, I can not-say more’ still works with utterances such as those in (3.15). The speakers can be interpreted as saying, roughly, that the other person knows that they support a certain team and would like the other to also support that team. They also convey that there is no need to say more since, in the eyes of passionate football fans, it is obvious that their team is the best, and they are being playfully dismissive of what the other thinks. We know from the rest of the conversation (and from the traits of typical football fans) that each speaker, when telling the other to change their support to the other team, knows that the other is highly unlikely to actually do so. The conversation is light-hearted and the two speakers are having fun teasing the team and players that the other likes. This also shows again that use of laa1 does not prevent the hearer from disagreeing. The speakers’ uses of laa1, particularly the second line of the explication, can be seen as attempts to make it seem as if they have ‘won’ the argument and the discussion is over.

Despite these examples of disagreements, the proposed explication for laa1 can also be tested with (3.16), in which two speakers agree with each other. Speaker C is telling M about her recent trip to Guilin as part of a tour group. In one place, C was being pushed to buy Chinese medicine, and there were lots of people heavily promoting the products. The two agree that there was no need to buy Chinese medicine there as they cannot be sure of their trustworthiness, and
there are good doctors in Hong Kong. In all three uses of laa1 in (3.16), the proposed explication can be substituted.

(3.16)
C:
mm mm mm mm of-course not-good buy

laa1 (3.16a) sing4joek6. M4 zit mat1je5 lai4 ge2.
PRT medicine not know what come PRT

‘Mm. Mm. Mm. Mm. Of course it’s not good to buy laa1 medicine. Don’t know what it is ge2.’

M:
Gang2hai6 laa1. (3.16b) Hoeng1gong2 gam3 hou2
of-course PRT Hong Kong such good
jisang1, m4sai2 laa1, (3.16c) sai2 - sai2 mat1
doctor not-need PRT need need what
heoi3 - heoi3 - heoi3 go2dou6 aa3? Ci1sin3, hai6
g0 go go there PRT crazy is
m4hai6 gam2 gong2 aa3?

not-is this say PRT

‘Of course laa1. Such good doctors in Hong Kong, no need laa1, what’s the need – need to go – go – go there aa3? Crazy, wouldn’t you say aa3?’

C: ‘So how many days were you there for aa3? At that time?’

M: ‘Five days aai-maa3.’

The interesting part of this excerpt from their conversation is how quickly the speakers transition into the next topic. When speaker C asks ‘Crazy, wouldn’t you say?’ this appears to be for effect or emphasis and is not a question that C expects to be answered seriously. This is signalled by various things including M’s previous turn in which M has already made clear her stance, or perhaps both speakers’ use of gang2hai6 ‘of course’, but also, it would seem, their recurrent use of laa1. At least partly through their use of laa1, both speakers have indicated to each other that they think the same thing and that nothing more needs to be said. Mutual understanding and ‘common ground establishment’ (Luke 1990) have occurred. This is proven by the quick way that the topic is brought back to that of holidaying in Guilin. In an almost ‘abrupt’ way, speaker M asks how many days C was there for. This does not sound
abrupt in Cantonese and C simply moves on and answers this question, which suggests that it sounded natural to C as well. This contradicts Yau’s ‘S-Q test’ in which he found that laa1 was a ‘Q-type’ particle ‘demanding a verbal confirmation’ (Yau 1965, 39-68). It also goes against Kwok’s (1984, 55) description that laa1 indicates a ‘lack of finality or completeness’.

Use of laa1 in ‘listing’ contexts

Example (3.17) shows a speaker giving a lengthy explanation about cheap ways to travel. The speaker is explaining that an airline promotional offer allows travel from any European city to Hong Kong via London, with the flight from the European city to London being free. The speaker gives a few examples of modes of transport which the other person can save on, and ends the list with laa1. This excerpt shows the end of the explanation, where the speaker is concluding that this way you can save money that would have otherwise been spent on a boat, a train, or ‘whatsoever’ to get to London. Note that the listing could also have been accomplished without laa1, and laa1 does not indicate listing in itself.

We can test the proposed explication of laa1 with its use in (3.17). The first line of the explication, ‘you now know how I think about this’, indicates that the speaker thinks the hearer understands and will themselves be able to think of modes of transport that could be saved on. The second line of the explication, ‘I can not-say more’, explains why the speaker stops listing – since it is assumed the other person can understand and can continue the list themselves, there is no need to say any more.
While example (3.17) shows laa1 being used at the end of a list, example (3.18) shows the particle laa1 being used in between the items of a list (as with the particle aa3 in (3.17)\(^{52}\)). In (3.18), the speaker is recounting who was present at a dinner. Again, laa1 means ‘you now know how I think about this, I can not-say more’. After naming Elton and then Raymond, the speaker reasonably believes the hearer to know who those two people are, and he feels he does not need to say more. Laa1 thus helps the speaker to ‘finish’ or ‘conclude’ these items before adding to the list. Again, the list could have been achieved even without the particle, as the particle does not in itself indicate that items are being listed.

(3.18)

\[
\begin{align*}
\text{Cin4maan5} & \quad \text{sik6faan6} & \quad \text{ge3} & \quad \text{si4hau6} & \quad \text{ne1,} & \quad \text{gam2} \\
\text{previous-night} & \quad \text{eat-dinner} & \quad \text{that time} & \quad \text{PRT} & \quad \text{so/then} \\
\text{zau6} & \quad \text{jau5} & \quad \text{aa3} & \quad \text{aa16ji5deon6} & \quad \text{laua1, (3.18a)} \\
\text{then have prefix} & \quad \text{Elton} & \quad \text{PRT} \\
\end{align*}
\]

\[
\begin{align*}
\text{Raymond} & \quad \text{laua1, (3.18b)} & \quad \text{zyu1zai2} & \quad \text{tung4maai4} & \quad \text{ng05} \\
\text{Raymond} & \quad \text{PRT} & \quad \text{‘Piggy’} & \quad \text{and} & \quad \text{I} \\
\text{sei3} & \quad \text{go3} & \quad \text{sik6faan6} & \quad \text{gam2joeng2.} \\
\text{four CL} & \quad \text{eat-dinner} & \quad \text{like-that} \\
\end{align*}
\]

‘That night when we had dinner ne1, so there was Elton laa1, Raymond laa1, Piggy and I, the four of us eating dinner, like that.’

Luke (1990, 69-71) reported that one of the most likely sequential contexts to find laa1 in was in lists. In particular, Luke draws attention to the ‘one-at-a-time’ manner in which such lists proceed, and that using laa1 after an item constitutes a proposal that it has been fully delivered. This is in line with the meaning of laa1 proposed here. Luke (1990, 70-72) also writes that laa1 displays the speaker’s intention to move on to the next item and allows for the hearer to acknowledge each item, although he accepts that this is not strictly necessary. Luke does not claim that these things happen when laa1 occurs in other contexts, and they are not included in the proposed explication. Related is

\(^{52}\) On the surface it seems that the difference between laa1 and aa3 when occurring after items on a list could be that laa1 is used when the speaker thinks the hearer should be able to guess, know, or have some familiarity with the items already (this is in line with the proposed explication). When aa3 is used in lists, the hearer’s expected familiarity with the items may be lower. However, this requires much more in-depth analysis, especially of aa3.
that neither instance of laa1 here demands verbal confirmation (Yau 1965, 39-68) or requires a response in terms of action (Gibbons 1980, 768, 770).

A different view was presented by Kwok (1984, 55), whose description ‘lack of finality or completeness’ extends to the idea that in lists, laa1 indicates that what is stated is not the total, and that the list is not meant to be complete. Interestingly, this may have been an adequate description of example (3.17), which does not feel complete, but appears not to reflect what is happening in example (3.18). The speaker in (3.18) uses laa1 after naming Elton and again after naming Raymond. In each case, the speaker is indicating that the hearer knows how he thinks and that he does not need to say more about that particular item (person). Kwok’s description cannot be applied to all cases of laa1 found in lists. Laa1 in (3.18) is also not serving purely ‘to fill a pause’, as Matthews and Yip (2011, 393) reported that laa1 often does when it is used in a list. The explication and analysis presented in this chapter shows that laa1 has a meaning of its own.

Laa1 in ‘pre-closing’ contexts

Laa1 can also be used in ‘pre-closing’ contexts. For example, laa1 is very often used with syun3 to create ‘syun3 laa1!’ This is quite noticeable in the corpus. Syun3 means something like ‘let it go’ or ‘forget it’ and is used when one wants to drop something and move on. This appears to be particularly compatible with the second line of the explication. In (3.19), A and B are discussing A’s falling out with another friend. A wants to stop talking about it. Instances of laa1 such as in (3.19) seem to express something like ‘you now know how I think about this’, i.e. ‘I want you to let it go’, and ‘I can not-say more’.

(3.19)

\[
\begin{align*}
\text{Hei3!} & \quad M_4 \quad \text{hou2} \quad \text{gong2} \quad \text{laa1 (3.19a)} \\
\text{hei not good say} & \quad \text{PRT} \\
\text{niudii} & \quad \text{je5.} \quad \text{Syun3} \quad \text{laa1! (3.19b)} \\
\text{these thing let-it-go} & \quad \text{PRT}
\end{align*}
\]

‘Hey [sigh]! Stop talking about these things laa1. Let it go/forget it laa1!

The proposed explication can also explain the use of laa1 in (3.20). Two people, A and B, are having a conversation when somebody, C, enters (immediately prior to this excerpt, he presumably signals his arrival and A and B call for him
to ‘come in’). It appears that C is some sort of IT worker who has come to fix a machine. C asks if A and B can enter the network, and when they reply that they could in the morning, C concludes that this must mean that the machine works. After the brief exchange, A and B indicate that C can now leave by thanking him and saying goodbye, with B using ‘hou2 laa1’, or ‘good/complete laa1’.

(3.20)
B: ‘Oh, Mr. Lee.’
C: ‘I’ve been coming over to help Jane, um, fix that machine. Can you enter the network here?’
B: ‘[We] could before, in the morning.’
A: ‘Yes, could enter in the morning, yes.’
C: ‘Oh, that means it’s okay [it works] aa3. Yes.’
B: ‘Yes aa3. Haven’t tried in the afternoon.’
A: ‘Oh. Thank you.’
B: Hou2 laa1, m4goi1 saai3.
good/complete PRT thank-you all
‘Good laa1, thanks a lot.’
A: ‘Bye bye!’

The proposed explication ‘you now know how I think about this, I can not-say more’ again makes sense here substituted for laa1. Applied to this example, B’s use of ‘hou2 laa1’ indicates that B believes that C now knows how B (and A) thinks – that the job is completed and C is not needed. This leads to the second part of the explication, which indicates that there is nothing more to be said because all parties understand that the job is done. The last line of the explication is demonstrated particularly well in this example because B (and A) are indicating that C can leave. C appears to have understood this too.53

53 Unfortunately, the data does not indicate explicitly whether C then left, but as he does not appear any more in the conversation after that, it is most likely that he did. Native speaker intuition and the surrounding utterances such as thanking, which can also be used in English to indicate the end of a conversation, also suggest it would have been unnatural for him to stay after that.
To link example (3.20) back to some previous studies of laa1, we can see that this fits some, but not all, of the descriptions. It does convey a sense of ‘completion’, as suggested by Meyer and Wempe (1947, 287) and Cowles (1965, 489), but does not imply ‘urgency’, as they also suggest. It does not indicate ‘coaxing’, ‘persuading’, ‘reminding’, or ‘politely urging’ (Yau 1965, 82-120). It is also in clear contrast with Kwok’s (1984, 55) description of ‘lack of finality or completeness’. The existing description of laa1 most relevant to (3.20) can be found in Luke’s section on ‘pre-closings’.

According to Luke (1990, 102), laa1 has a ‘pervasive presence in pre-closing sequences’, or near the ends of conversations. Indeed, laa1 is very common even at the ends of telephone conversations. This makes sense if we consider the proposed explication, because laa1 has the meaning component ‘I can not-say more’. Luke (1990, 109-110) explains that when a ‘pre-closing favourable environment’ has been constructed, preparatory work can be made towards conversational disengagement. One of the main ways to achieve this environment is to produce a ‘contentless’ signal, like ‘hou2 laa1’ in (3.20). Luke explains that this ‘records the speaker’s assumption that mutual understanding and agreement obtains, but adds nothing new to what has been said so far in the conversation, and, in so doing, proposes to yield the turn’.

The three requirements identified by Luke (1990, 110) as necessary for ‘pre-closing initiators’ can be related to (3.20), and also closely correspond with the explication proposed. The first requirement is to express the assumption that whatever needed to be dealt with in the conversation has been dealt with to the satisfaction of both parties. This is consistent with the first line of the proposed explication, ‘you now know how I think about this’. The second requirement is to signal that there are no further matters to raise. This is consistent with the second line of the proposed explication ‘I can not-say more’. Lastly, with the first two requirements satisfied, the next activity identified by Luke is closing. These three requirements have all been fulfilled in (3.20), and furthermore, have parallels with the NSM explication proposed for laa1. Since there is understanding that the matter of the machine has been settled and that there are no more matters to be raised, nothing more needs to be said and C leaves. It is important to note, however, that although Luke labels utterances such as ‘hou2 laa1’ as ‘contentless’ and ‘adding nothing new to what has been
said so far in the conversation’, it has been demonstrated in this chapter that such utterances are far from contentless.

3.3 Concluding remarks

This chapter has investigated the semantics of the particle laa1. It has offered a detailed appraisal of the previous descriptions and has shown that they are mostly insufficient in explaining the meaning of the particle. Although they do give helpful insights into certain aspects, they cannot explain all instances of laa1 found in the Hong Kong Cantonese Corpus. Some descriptions are vague, contradictory, inaccurate, limited to certain contexts, or attribute the meaning of an utterance to the meaning of the particle. The most insightful previous descriptions are Luke’s (1990) idea of ‘common ground establishment’, and Fung’s (2000, 96-99) claims about knowledge of the state-of-affairs in question.

Using NSM, an explication was proposed in this chapter for the ‘core’ or invariant meaning of laa1. It was demonstrated that laa1 does have semantic content, which can be explicated in NSM as ‘you now know how I think about this, I can not-say more’. Laa1 was found to indicate the speaker’s expectation that the addressee shares some relevant knowledge or understanding, and that more does not need to be said. A range of real, naturally-occurring examples of laa1 found in the corpus were examined, where the validity of the explication was demonstrated by its capacity to consistently explain the meaning of laa1.

The content of the utterance to which laa1 attaches seems not to be very important to the meaning of laa1, and so laa1 appears to be able to attach to a wide range of utterances and to change meaning depending on context, for example, conveying at times certainty, as well as uncertainty. This may help explain why utterance particles like laa1 are sometimes perceived as having no meaning. Another reason for this is that the explication is relatively short compared to the average NSM explication. In hindsight, this is to be expected because the particle has a meaning that is relevant in many contexts and can easily attach to a vast variety of utterances, not to mention other particles. The explication helps shed more light on its regular use in Cantonese conversation, despite the lack of an equivalent word in English. The next chapter investigates the semantics of the particle wo3, using the same approach.
Chapter 4:
The semantics of particle wo3

This chapter looks at the particle wo3. In the Hong Kong Cantonese Corpus, there are 928 instances of wo3, making it the 5th most frequently used particle, after aa3, gaa3, laa1 and lo1. It is the 23rd most frequently used word in the corpus. In Yau’s data, wo3 is the 6th most frequently occurring utterance particle (Yau 1965, 17-25). Wo3 does not occur in combination with laa1, but does frequently combine with gaa3, laa3, and zaa3 to create the clusters gaa3-wo3, laa3-wo3, and zaa3-wo3, as well as gaa3-laa3-wo3 and gaa3-zaa3-wo3. These are discussed in Chapter 9.

This chapter, like the preceding one, looks first at the existing literature on wo3, then moves on to the NSM semantic analysis. The proposed explication is given in section 4.2, accompanied by discussion of examples of wo3 from the Hong Kong Cantonese Corpus. Substitution of the explication into the examples demonstrates the validity of the explication. The examples also show how previous descriptions, including earlier proposed explications, are not adequate. An NSM explication for wo3 was first proposed in HHL Leung (2013, 21-22), but without any supporting discussion or examples other than with regards to the cluster gaa3-wo3.

In the past, wo3 was often studied in comparison to particles such as wo5, wo4, and bo3 (e.g. W-m Leung 2009, 2010b, a, Matthews 2007). This thesis does not analyse the semantics of wo5, wo4, or bo3 in depth because of their low frequency in the corpus: wo5 occurs six times, wo4 occurs nine times, and bo3, sometimes considered to be the same particle as wo3, only occurs 10 times.54 Wo3’s alleged breakdown into ‘sub-syllabic morphemes’ and its possible connection to wo5, wo4, and bo3 are discussed further in Chapter 8.

54 In the corpus, bo3 occurs slightly more frequently than wo4 or wo5, but wo4 and wo5 receive disproportionately more attention in the literature (see Chapter 8).
4.1 Previous descriptions of wo3

4.1.1 The particles wo3 and bo3

Early studies of particles considered a particle bo3 which some claim to be wo3. For example, Kwok does not look at wo3, instead including a section on the particle bo3, which she claims is in free variation with and often heard as wo3 (Kwok 1984, 12). (However, Kwok claims that laa3-wo3 and laa3-bo3 both exist, which confuses the picture and calls into question whether wo3 and bo3 are really the same.) Examples of bo3, taken from Kwok, are shown and discussed in section 4.2.1, labelled as (4.6) and (4.7). According to Kwok, bo3 attaches to statements and is used ‘to remind the hearer to take something into special consideration’ (Kwok 1984, 64). It ‘may have the general meaning of “drawing someone’s attention to something”, or “asking someone to take a certain matter into consideration”’ (Kwok 1984, 14). When suffixed to hai6 ‘yes’, hai6 bo3 is used to indicate the speaker’s agreement, while suggesting that the idea has not occurred to that person before (Kwok 1984, 65-66).

However, because of bo3’s low frequency in the Hong Kong Cantonese Corpus, occurring 10 times as opposed to wo3’s 928, it does not seem to be in free variation with wo3. According to W-m Leung (2010b, 15-20), wo3 is primarily used to convey contrast, realisation, reminder, hearsay, and emphasis, while bo3 mainly conveys contrast and perhaps realisation. Despite the apparent overlap, she claims that in modern Hong Kong Cantonese, wo3 and bo3 are different particles with different meanings and functions, and are not interchangeable as previously assumed (W-m Leung 2010b, 21).

W-m Leung (2010a, 86), who studied the use of wo3 over time, acknowledged that most previous scholars considered wo3 a phonetically weakened variant of bo3, and that the focus of discussion was often bo3. Little attention was paid to wo3 until Luke (1990) (W-m Leung 2010a, 86, W-m Leung 2010b, 13). W-m Leung’s study found that the frequency of wo3 was apparently less than that of bo3 in the 1940s, but by the 1970s wo3 had become more frequent and common than bo3. This is supported by Yau (1980, 38-39), who reported that in about 15 hours of recordings provided by 104 participants in the mid-sixties, wo3 occurred about 20 times as often as bo3. By the 1990s,
bo3 was rarely heard in conversations and was seen almost exclusively in the company of other particles, in particle clusters (W-m Leung 2010a, 94). The decline of bo3 may be because the functions of wo3 expanded, gradually taking up and replacing the functions of bo3 (W-m Leung 2010b, 21, 2010a, 94). According to W-m Leung, the main functions of wo3 in the 1940s were realisation, reminder, and hearsay, with the first two having originally belonged to bo3 – this was the early sign of the merging of these two particles. By the 1970s, wo3 could also show contrast – another original function of bo3.

Yau (1965, 53-66, 112-116, 1980, 50) additionally reported that wo3 had the connotation concepts ‘conceited’ and ‘reluctant’, while bo3 had the connotation concept of ‘politely urging’. This seems not to support the claim that wo3 and bo3 are in free variation. Gibbons (1980) looked at bo3 and not wo3, stating that bo3 indicated the (assumed) newness of information to the hearer, showing that the speaker believes the hearer should already know something, and was often used in reminding.

Luke (1990, 199) acknowledged that bo3 was sometimes used as a variant of wo3, but stated that wo3 was ‘by far the most frequently found form in the Cantonese of contemporary Hong Kong’. Luke did not find in his data that the variation depended on any of the ‘usual social parameters’ or contextual factors.

4.1.2 Classifications of wo3 as a question or a statement

In the earlier studies of Cantonese utterance particles, a common way of categorising the large number of particles was based on whether a particle was used in statements or questions. However, the literature seems uncertain of whether wo3 is used with statements or with questions. Yau (1965, 53-66) categorised wo3 as Q-type (‘obviously demanding a verbal confirmation’) or S-Q-type (indeterminable between ‘obviously affirming’ and ‘obviously demanding a verbal confirmation’, or irrelevant to both) depending on the interpretation used55 (bo3 was classed as S-Q-type). This categorisation is

---

55 Wo3 is a Q-type particle according to Yau’s ‘nominal interpretation’, which was based on the opinion of the majority of respondents, being a means of classification and not an arrangement along a continuum (under this interpretation, wo3 contrasts with bo3, which is S-Q-type). Wo3 belongs in the S-Q category according to Yau’s ‘numerical interpretation’, which is a measurement that relies on each particle’s mean and standard deviation, calculated from an arbitrarily assigned digit (under this interpretation, bo3 was also found in the S-Q category).
contradicted by Yau’s own example sentences (see discussion of example (4.3) and (4.4) below)\(^{56}\). According to Kwok (1984, 41), *bo* is used with statements without altering their grammatical status as statements. Matthews and Yip (2011, 391) describe *wo* as ‘informative (noteworthiness)’, implying that they would also treat this particle as being used with statements.

So what kind of sentence is *wo* really used with? In the Hong Kong Cantonese Corpus, the vast majority of *wo* examples are used with statements. Other examples show *wo* being used in sentences which are somewhat like imperatives (being used to, say, ‘suggest’ or ‘remind’, e.g. example (4.14) below). While there are no exact statistics due to limitations of the corpus, it is relatively difficult to find examples of *wo* in questions. Instead, there are some examples in the corpus which show *wo* being used in statements which are then followed up with a question about that statement. Perhaps this is where some of the confusion stems from. See examples (4.1) and (4.2) below, for instance. In (4.1), A and B are talking about going to New Zealand. This excerpt starts with B answering a question about New Zealand’s Cantonese name, and then A starts to talk about the scenery, and they talk about New Zealand attractions. The particle *wo* is followed closely by ‘isn’t it?’ which could perhaps be argued to be either the same or a separate sentence.

(4.1)

B: ‘Eh, some people call New Zealand like New, directly translated to “New Zealand” [nau5saailaan4]. Some call it “New Zealand” [samisailaan4]. Actually it’s the same, the same place.’

A:

\[
\begin{array}{llllll}
\text{Hai6} & \text{lo1}. & \text{Hou2ci5} & \text{gei2} & \text{leng3} \\
\text{is} & \text{PRT} & \text{(Very) like} & \text{quite} & \text{pretty/beautiful} \\
\text{wo3} & \text{diu} & \text{fung1ging2,} & \text{hai6} & \text{m4hai6} & \text{aa3?} \\
\text{PRT} & \text{CL} & \text{scenery} & \text{is} & \text{not-is} & \text{PRT} \\
\end{array}
\]

‘Yes *lo1*. Seems quite beautiful *wo3*, the scenery, isn’t it *aa3*?’

---

\(^{56}\) We can also consider particle clusters. Yau (1980, 47) states that in most cases, the particles, and those particle clusters that have that particle as their last component, fall into the same type in the S-Q test. This seems to be true for *wo* in the numerical interpretation, but does not hold true consistently in the nominal interpretation. According to Yau’s nominal interpretation, clusters such as *gaa3-wo3* and *zaa3-wo3* are in the S-Q-type as opposed to *wo*’s Q-type, although clusters such as *laa3-wo3* are categorised as Q-type.
In (4.2), we see another example of *wo3* being used in a statement which is then followed up with a question about that statement. As the speaker says she does not know what results the addressee achieved, an answer about the results can be expected from the addressee. The statement has an expectation of an answer, just like a question. This is made explicit by the speaker’s follow-up question immediately afterwards. However, this does not make ‘I don’t know what results you got *wo3*’ an interrogative.

(4.2)

B:

```
Nei5 go2zan6si4 kei4sat6 haau2 mat1je5
you at-that-time actually test what
sing4zik1 ngo5 dou1 m4 zit wo3.
results I even not know PRT
```

In fact, there may be a problem with asking whether *wo3* is used in statements or questions in the first place. Luke (1990, 255-256) suggested that either native speaker intuitions about *wo3* are hazy and uncertain, or that native speakers find it hard to agree on *wo3*’s ‘grammatical function’. It would seem that both of Luke’s suggestions are true. Luke found that previous research variously characterised *wo3* as a marker of strong assertions, confirmation-seeking statements, questions, and sometimes assertions and sometimes questions. He went on to reason that perhaps there is something fundamentally wrong with asking whether *wo3* is or is not a question particle. In Luke’s
opinion, this was an unsolvable and even misleading question (Luke 1990, 257-258). Often, there is nothing in the morphology or word order of Cantonese utterances on the basis of which their grammatical status can be determined. The same utterance, in different contexts, can be understood very differently. Far from compensating for this, utterance particles considered in isolation do not provide any answers – context is needed. Luke (1990, 256-257) stated:

The difficulties with approaches which rely primarily on intuitions as linguistic data, coupled with traditional assumptions about the relationship between sentence types (declarative, interrogative, and imperative) and functional categories (assertions, questions, and commands/requests) – for example, the assumption that declarative sentences are (normally) used to make assertions, and interrogatives to ask questions – show up most dramatically in previous treatments of this particle.

4.1.3 Previous descriptions of the meaning and functions of wo3

As mentioned above, Yau calculated that the ‘connotation concepts’ for wo3 were ‘conceited’ and ‘reluctant’ (Yau 1965, 112-116, 1980, 50). Yau does not explain who or what is supposed to be being conceited/reluctant (is the speaker feeling these things, or is the speaker pointing out somebody else being conceited/reluctant?). Overall, these are not good explanations of wo3.

Moreover, when we look at real, naturally-occurring examples of wo3 from the Hong Kong Cantonese Corpus, we see that the vast majority of examples do not indicate conceitedness or reluctance. Both examples (4.1) and (4.2) above have shown this. Numerous other examples of wo3 in this chapter will also show that there is no evidence of conceitedness or reluctance being conveyed by or even with wo3, and therefore Yau’s results appear to be incorrect. In fact, both ‘conceited’ and ‘reluctant’ seem unable to adequately explain even Yau’s own example sentences of wo3. His two examples are reproduced here, labelled (4.3) and (4.4), with English gloss added but with Yau’s original English free translation (Yau 1965, 318).
You don’t have to mention the result of the investigation’ (Yau’s free translation, my gloss)

At the moment we are discussing the traffic problem, don’t you see’ (Yau’s free translation, my gloss)

It is difficult to see how either of Yau’s examples are either ‘conceited’ or ‘reluctant’. A noticeable problem is that there is no context for either of these examples. Let’s construct a plausible scenario for example (4.3) and assume that the speaker who produces this utterance is speaking to a person X, who is about to read a report to a group of people, and the speaker is telling this person X ‘you don’t have to mention the result of the investigation [when you read the report to these people]’. It seems to me that the most obvious reason why the speaker would use the particle wo3 is if s/he thought that person X would mention the results unless told not to. In other words, it seems that the speaker believes person X will mention the results, and therefore feels the need to instruct person X not to. ‘Conceited’ and ‘reluctant’ are not labels that come to mind when trying to explain this scenario. An imaginative person could believe that the speaker did not want to disclose the results because s/he was too conceited, or because s/he was reluctant to share information, but it is highly unlikely that this is conveyed through the particle wo3. We could omit the wo3 at the end of the utterance, and this interpretation would still be possible.

As with (4.3), Yau provides no context for his other example (4.4), although we can construct another plausible scenario for it. Let’s say that some people are discussing a traffic problem, but one person has misunderstood or gone off-topic. At this point, this speaker says ‘at the moment we are discussing the traffic problem, don’t you see?’ It is impossible to tell whether this speaker is calm and merely trying to help or redirect the other person, or whether this speaker is frustrated, impatient, and pointing out the other person’s mistake in an angry way. Perhaps in the latter case, we could hypothesise that the speaker was indeed conceited. As for ‘reluctant’, this does not appear to be a good
description of this scenario. Perhaps we could construct a different backstory where the speaker is being asked about something else, which they are reluctant to talk about, and therefore this speaker points out that they are talking about traffic, by way of diverting the conversation to traffic to avoid the other topic. There is no way of knowing what situation Yau had in mind when constructing this example, but more importantly, we can see that it requires a purpose-built backstory, and ‘conceited’ and ‘reluctant’ are by no means adequate descriptions in all cases of wo3 (this becomes even more apparent when we look at more examples from the Hong Kong Cantonese Corpus).

Yau’s free translation of (4.4) provides more insight into his intuition about wo3. We can pretend that the particle wo3 at the end of this utterance were omitted, and we were left with the sentence ‘[we/they are] now discussing traffic’. The fact that Yau has added ‘don’t you see’ in his free translation is interesting because it is not present in this truncated utterance and therefore can only correspond with what was left out: wo3. In other words, it seems as if Yau is equating the wo3 in this example sentence with ‘don’t you see’. From the feeling of the Cantonese utterance we know that this is most likely a rhetorical question which the speaker does not expect a genuine answer for (the Cantonese utterance is not a question). It could be that the speaker is pointing out what s/he deems as obvious or as something that should already be known. This is plausible in this utterance, and more helpful in understanding wo3, although it would be far-fetched to suggest that ‘don’t you see’ has connotations of conceitedness or reluctance.

Not only do Yau’s own examples of wo3 not connote conceitedness or reluctance, but (4.3) and (4.4) are both declarative statements, at least in Cantonese. This further calls into question Yau’s claims that wo3 is either a Q-type or S-Q-type particle (Yau 1965, 53-66). His examples support neither his own S-Q test results nor his own C-test results.

An example of wo3 similar to (4.4) exists in Matthews (2007, 6). Again, the interrogative part of the English free translation, by Matthews, appears to be due to the particle wo3, as the same utterance without wo3 would be a declarative statement. Matthews’ observations about wo3, as well as wo4 and wo5, are discussed further in Chapter 8.
Luke (1990) studied wo3 using a Conversation Analysis point of view, as he did with laa1 (see Chapter 3). Luke (1990, 257) reported that ‘the state of our knowledge concerning wo3 is nothing short of total confusion’, with disparate and contradictory accounts of the same particle. He himself reports seven contexts in which wo3 is used, namely ‘reportings and story-tellings’, ‘challenging a position’, ‘contact-establishments’, ‘disconfirmations’, ‘thankings’, ‘informings and remindings’, and ‘realisations’. According to Luke, use of wo3 is closely tied to different kinds of sequential positions. For example, after a turn which seeks agreement, confirmation, or contact, wo3 would serve to mark a dispreferred response, but after a turn in which some general rule or position is stated, wo3 would serve to challenge or undermine a position. Within the context of a story or report, wo3 is used to highlight the unusualness or remarkableness of the action or event being reported, and after an information offer, wo3 can be heard as remembering or realising something (Luke 1990, 262).

The conclusion Luke arrived at was that wo3 ‘can be characterised in the most general way as a linguistic device with which objects, persons, events, situations, etc. can be highlighted for noteworthiness, unexpectedness, or remarkableness’ (Luke 1990, 262). Luke offers much valuable insight into wo3, but his explanation of the particle often depends on fairly subjective ideas about what is ‘noteworthy’, ‘remarkable’, etc. For example, in the section on ‘informings and remindings’, Luke (1990, 233) states that a speaker can either provide information in response to an inquiry, or can volunteer information on his/her own initiative. Volunteering information is a means of highlighting the prominence of the information and presenting it as noteworthy information. This makes sense, but since a vast amount of information is volunteered in conversations, an extraordinarily large amount of information would be regarded by Luke as ‘noteworthy’. Possibly, most people believe that most of what they have to say is ‘noteworthy’. 
From the point of view of someone who believes utterance particles have no meaning, Luke’s analysis of *wo3* would appear to be solid and comprehensive. Luke (1990, 258) states that the question ‘what does this particle really mean?’ is a misleading question, explaining that in his view, an attempt to pin down what *wo3* means has led to descriptions which are confined too narrowly to one particular use, with that being called the particle’s intrinsic meaning. This may be true, but the problem may be due to faulty conclusions drawn in previous work. As for Luke’s many different contexts and functions of *wo3*, they could perhaps be criticised as being too numerous and varied, and as suggesting that the same particle changes its role continuously without explanation and is used in very different, unstable ways. In this thesis, I argue that *wo3* does have stable semantic content, and demonstrate this via NSM explications.

Another scholar who has identified several functions for *wo3* is W-m Leung (2009, 1-2). As mentioned, she believes that *wo3*’s function now indicates ‘realisation’, ‘reminder’, ‘showing contrast’, ‘unexpectedness’ and ‘noteworthiness’. As can be seen, there is overlap between the functions identified by W-m Leung and Luke. Revisiting example (4.1), most of the functions are already shown to be unjustified, at least in terms of an invariant. In that example, speaker A’s comment about the scenery in New Zealand is not a realisation or reminder that the scenery is beautiful, nor is there any reason to believe that either speaker finds this unexpected. There does not appear to be any contrast or noteworthiness, unless perhaps these are to do with the new topic being introduced. It is not very likely, but still possible that the contrast which is shown is meant to be the correct name for New Zealand, contrasted with the idea of beautiful scenery. It could be that this small shift in topic is supposed to be noteworthy. Alternatively, it could be that the speaker is asking a question, therefore it needs to be somehow emphasised and marked as noteworthy to catch the attention of the hearer so that the hearer can answer. Whether something is noteworthy or not can be a very subjective matter so it is difficult to say whether the speaker believed her utterance to be noteworthy. We can re-examine these supposed functions of *wo3* when more examples of *wo3* are presented.
According to Matthews and Yip (2011, 391-392, 406), \textit{wo3} is categorised as an ‘evidential particle’, which is informative and indicates something new and noteworthy. It emphasises newsworthy information, and is also used in reminders. This overlaps again with the descriptions of \textit{wo3} outlined above. There may be minor internal inconsistency in that the same particle is reported to indicate new information, as well as reminders (where the information is by definition previously known and therefore not new). This may be because scholars believe the same particle has seemingly opposing functions. Alternatively, perhaps we can speculate that information can be forgotten, and when one is reminded, that information can be described as ‘new’ again. The common thread, therefore, can be said to be that the person was not thinking about that particular thing at that particular time (see below). Next, the NSM explication proposed for \textit{wo3} will be presented.

4.2 NSM semantic analysis of \textit{wo3}

4.2.1 The NSM explication for \textit{wo3}

The NSM explication [4A] below explains the proposed meaning of the particle \textit{wo3}. More than 20 typical examples of \textit{wo3} from the corpus are provided in the remainder of this chapter to demonstrate the explication’s validity.

[4A] Final explication for \textit{wo3}:

\begin{quote}
\textbf{you aren’t thinking about this at this moment}
\textbf{it is good if you think about it}

\textit{leit\textsubscript{5} li1 jat1haak1 m4 hai6 lam2 gan2}
\textit{you This moment not is think PROG}

\textit{li1 jat1 joeng6 je5}
\textit{this one CL thing}

\textit{jyu4gwo2 leit\textsubscript{5} lam2 haa5 zau6 hou2}
\textit{if you think DEL\textsuperscript{57} then good}
\end{quote}

\textsuperscript{57} It is acceptable to delete the delimitative aspect marker \textit{haa5}, but this explication sounds like a command without it, being abrupt and therefore unnatural. This is improved marginally by including \textit{haa5} (an alternative would have been to use the reduplicated \textit{lam2 lam2}). Note, though, that the English explication may also sound unnatural or like a command to some. Importantly, both conform to the rules of NSM.
An earlier explication of \textit{wo3} was published in HHL Leung (2013, 21-22). The difference between that and [4A] is that ‘at this moment’ in the first line was previously ‘now’. This was changed because the period of time the explication refers to was felt to be more immediate and shorter – only a moment.\footnote{The only other change is that ‘are not’ in the first line of the previous explication was contracted to ‘aren’t’ here, which has not changed the meaning of the explication.} Still earlier versions of this explication included, for instance, ‘I am thinking about something now, you aren’t thinking about this something now’, or ‘I am thinking about something, you are thinking about something, these two things aren’t the same’. There was also the possibility of something like ‘I want you to think about this now, I don’t want you to think about something else’. Ultimately, explication [4A] was chosen, which seems to convey the meaning of \textit{wo3} the best.

This proposed explication for \textit{wo3} is compatible with some of the previous descriptions of \textit{wo3} discussed in section 4.1. The first line ‘you aren’t thinking about this at this moment’ encompasses in a straightforward way \textit{wo3}’s ability to code new information as well as reminders. The second line of the explication ‘it is good if you think about it’ shows the noteworthiness/newsworthiness of the information or the purpose of the reminder. Someone who thinks ‘you aren’t thinking about this at this moment, it is good if you think about it’ could be imagined to add ‘don’t you see?’ to a sentence.

Explication [4A] for \textit{wo3} is also consistent with Kwok’s descriptions of \textit{bo3} (which will be taken to be the same as \textit{wo3} for the purposes here). As explained, Kwok believed the particle \textit{bo3} could draw someone’s attention to something, or ask or remind someone to take something into consideration (Kwok 1984, 14, 64). If such were the case, presumably the addressee was not already paying attention to that thing. This can be seen in the explication ‘you aren’t thinking about this at this moment, it is good if you think about it’. Kwok (1984, 64-65) provides two examples of \textit{bo3} from her corpus, which she feels show that \textit{bo3} is used to remind the hearer to take something into special consideration. These are reproduced below as examples (4.6) and (4.7). Kwok’s examples show more typical and natural examples of \textit{wo3} than Yau’s (see section 4.1, examples (4.3) and (4.4)) and therefore it is relatively easy to picture...
a relevant scenario. Her examples can be used to test the NSM explication proposed.

(4.6)

\[
\begin{align*}
\text{k\text{\textcircled{\text{ao}}} \text{bat} \text{ dzo jip} \text{ hau jau tsin}} & \\
\text{he \ graduate \ aspect \ marker \ later \ have \ money} & \\
\text{sin wa:n dak bo} & \\
\text{before \ return \ can \ bo3} & \\
\end{align*}
\]

‘You must take into consideration the fact that you have to wait until he graduates and is able to make some money before he is able to pay you back.’ (Kwok’s gloss and translation)

The first part of the explication explains that the addressee is not thinking about this at this time, i.e., in example (4.6), that the addressee is not thinking about the fact that this person can only return the money after he graduates and earns money. Kwok describes this as ‘reminding’, but one could also imagine a scenario where this is the first time the hearer has considered that that person needs to graduate first. The explication allows for both possibilities. The second part of the NSM explication states that it is good if the addressee thinks about this, i.e. it is good if the addressee thinks about the fact that s/he will have to wait until this person graduates. Kwok describes this as ‘taking something into special consideration’. This also corresponds with the NSM explication. Note that the Cantonese utterance does not include any statement such as ‘you must take into consideration the fact that…’. That Kwok felt the need to include this in her English free translation could well be because she is picking up the meaning from wo3, specifically the second line of the proposed explication ‘it is good if you think about it’. Kwok describes her second example, labelled (4.7), as one speaker trying to encourage the other to paint the walls of his flat by himself, although with a mild warning.

(4.7)

\[
\begin{align*}
\text{nei jiu bun hoi sa:i di je bo} & \\
you \ need \ move \ away \ all \ those \ thing \ bo3 & \\
\end{align*}
\]

‘But I must tell you that you will have to move everything away first.’ (Kwok’s gloss and translation)

From the first part of the NSM explication, we can speculate that the speaker in (4.7) believes the addressee was not thinking about this. It is difficult to imagine a scenario where a speaker would use wo3 if the speaker thought the
hearer was already thinking about the thing being spoken about. The second line of the explication ‘it is good if you think about it’ seems to have been interpreted by Kwok to be a mild warning in this example. According to Kwok (1984, 14), in a particular context, meanings specific to the situation can be added to bo3’s general meaning, such as ‘a note of warning’ and so on. This is in line with NSM explications of particles’ core or invariant meanings – the core meaning does not change, but extra meanings, attitudes, feelings etc. are often notable in context. ‘It is good if you think about it’ can be interpreted generally as ‘drawing someone’s attention to something’, or ‘asking someone to take a certain matter into consideration’, with the feel of ‘warning’ coming about in this specific example (and therefore not to be included in a definition of wo3’s invariant meaning). Note again that the original Cantonese example does not include anything like ‘But I must tell you that...’ and it is likely that Kwok is getting this from the particle.

4.2.2 Testing of explications and descriptions using the corpus

This section tests the explication against examples of wo3 from the Hong Kong Cantonese Corpus. The examples are also used to evaluate existing descriptions of wo3, introduced above, as well as some alternative NSM explications and why they were rejected. The examples show a wide range of scenarios in which wo3 can be used. The subheadings separate rough groups of examples for expository convenience and should not be taken as separate categories of wo3. Many of the examples could actually fit into multiple groups. Many examples feature more than one use of wo3.

Answering questions in a ‘negative’ or ‘unexpected’ way

Examples (4.8a) and (4.8b) are used in negative responses to B’s questions. Speaker A seems to understand that B wants to start up a conversation about movies and that his negative responses are not what B is hoping for. Therefore he indicates ‘you aren’t thinking about this, it is good if you think about it’ to signal that he is aware of this inconsistency, and to express to B that he does not have much to contribute to the conversation.
Examples (4.8a) and (4.8b) are somewhat like 'disconfirmations', where according to Luke (1990, 227-231), wo3 can highlight a discrepancy between a prior assumption and some facts or evidence following a confirmation-seeking turn. However, B's questions are not exactly confirmation-seeking, and so wo3 seems to indicate a discrepancy between A's actual answer and the answer that A perceives B to be expecting. Explication [4A] provides a better explanation. An alternative NSM component considered earlier is 'I want to say something about this'. This would have been acceptable for many examples of wo3, but
examples like (4.8a) and (4.8b) disconfirm it, because the speaker here does not seem to want to say anything more.

In (4.8c), speaker B uses *wo3* as she introduces what she really wants to talk about: the new *Batman* movie. Speaker B expresses ‘you aren’t thinking about this at this moment, it is good if you think about it’ because A is evidently not thinking about the *Batman* movie, though B wants to discuss it. Also, as this is new information which speaker B is volunteering, she can mark it as something new in the conversation, i.e. something that speaker A is not thinking about. This is similar to some examples below, like examples (4.13d), (4.15), and (4.20), where the speaker provides some additional information to support what they are saying. This could be related to Luke’s idea of volunteering information in ‘informings and remindings’, which highlights some information as ‘noteworthy’ (Luke 1990, 233). In this case, the ‘noteworthiness’ appears to be that B wants to discuss the *Batman* movie. Speaker A appears to understand B’s intention here and asks B about the movie.

(4.9)

B:

*Nei5 m4 geng1 taa13 san1fu2*

you not scared too difficult/painful

*aa3? Nei5 zi1 nei5 gei2 gam3...*

PRT you know you how-many so

‘You aren’t scared it’ll be too difficult *aa3?* You know you’re so...’

A:

*Ngo5 haau2leoi6 ze1 maa3. Ngo5 mou5 waa6*

I consider PRT PRT I not-have say

*jat1ding6 wui5 duk6 *wo3 ji4gaa1 zan1hai6.*

definitely will read/study PRT now truly

‘I’m [only] considering [it] *ze1-maa3*. I didn’t say I’d definitely study it *wo3*, really now.’

B:

*Jyu4gw02 ne15 hai6 m4 duk6 j11 ge3waa2...*

if you is not read/study medicine in-that-case

‘If you don’t study medicine then...’

In example (4.9), speakers A and B are talking about what A will apply to study at university. Speaker A says that he is considering a lot of options, one of
which is medicine. B asks whether (/suggests that) medicine will be too
difficult, and A repeats that he is only considering it. Speaker A’s use of wo3
‘you aren’t thinking about this at this moment, it is good if you think about it’
helps to point out and to remind B that he has not made a proper decision yet.
It implies that B forgot that A said only that he is considering it, and it would be
good for B to think about this. It ‘downplays’ the medicine option and in a way
defends against or at least deflects B’s questioning. Speaker B understands and
goes on to discuss a scenario where A is not studying medicine.

We can also consider example (4.9) with regards to some previous
descriptions of wo3. It is possible to imagine that speaker A’s utterance could
be translated into English as ‘I didn’t say I’d definitely study it, don’t you see?’,
similarly to Yau’s translation of example (4.4). ‘Don’t you see’ suggests that the
speaker believes the hearer is not currently thinking properly about something,
and is also a way to encourage the reader to think about something more. This
expresses something similar to the explication: ‘you aren’t thinking about this at
this moment, it is good if you think about it’. According to Matthews and Yip
(2011, 391-392, 406), wo3 is informative and reminds or indicates something
new, noteworthy, and newsworthy. In example (4.9), there is an informative
reminder of what speaker A has said previously, namely that he is (only)
considering studying medicine. It could be considered noteworthy as it
addresses B’s concerns that medicine is too difficult. However, this information
is not new, and therefore not particularly newsworthy. Example (4.9) does not

(4.10)

B:
Nei5 dei6 gei2 si4 gaan2 gaa3 JUPAS?
you when choose PRT JUPAS
‘When do you choose [universities] gaa3 JUPAS [i.e. through the
universities’ admission system]?’

A:
Waa3! Nei5 jing1 goi1 nei5 cing1 co2 gwo3
wow you should you clear more
ngo5 wo3.
I PRT
‘Wow! You should know better than me wo3.’
In example (4.10), speaker A and speaker B are still discussing A’s options for universities and admission (this is later in the conversation after (4.9)). The Joint University Programmes Admissions System, or JUPAS, is a system in Hong Kong used for applying to universities. Speaker B is already studying at a university, therefore when B asks A about the JUPAS process and deadline, A exclaims that B should know better than him. The particle wo3 helps to point out how strange B’s question. From speaker A’s perspective, B seems to have ‘forgotten’, or is not thinking about the fact that she went through the process herself, so A is ‘reminding’ her and signalling that she should reflect on this. In the next turn, B explains that she really has forgotten. The proposed explication ‘you aren’t thinking about this at this moment, it is good if you think about it’ is substitutable in place of wo3 here. Note that ‘at this moment’ in explication [4A] works better than ‘now’ from the earlier explication in HHL Leung (2013, 21-22).

Considering Yau’s (1965, 112-116, 1980, 50) connotation concepts for wo3, we can consider whether wo3 connotes conceitedness or reluctance. It is difficult to see how this example shows conceitedness. Perhaps reluctance could be argued for in that speaker A does not answer straightforwardly in the first instance, but given he provides the answer when speaker B explains that she has forgotten, it is more likely that he is simply pointing out the strangeness of the question coming from speaker B. In terms of the contexts identified by Luke (1990), the best fit for example (4.10) might be ‘informings and remindings’. According to Luke (1990, 233-242), information can be presented in a way that points to its own noteworthiness, and can therefore be interpreted as informing, reminding, or both. The NSM explication helps to clarify what the speaker’s thoughts are, more so than just labelling what the utterance might be doing.
This use of wo3 in example (4.10) can be seen as similar to that in example (4.9) because speaker A is in both cases using wo3 in a response to a question from B. In both cases, A feels that B’s question is somewhat unusual, and B is saying that A should know something already but is not thinking about it.

(4.11)

A:

\[
\text{Matije5 paai4zi2 aa3?}
\]

what brand PRT

‘What brand aa3?’

J:

\[
\text{Matije5 paai4zi2 aa3? Ngo5 m4 - ngo5 mou5}
\]

what brand PRT I not I not-have

\[
\text{lau4ji3 keoi5 di1 din6sam1 wo3.}
\]

pay-attention they/it CL battery PRT

\[
\text{Golden Power gwaa3. Gaan1 je5 giu3}
\]

Golden Power PRT CL:building thing call

\[
\text{Golden Power ng05 gu2 hai6 laa1!}
\]

Golden Power I guess is PRT

‘What brand aa3? I don’t – I didn’t pay attention to their batteries wo3 [laughter]. [I guess] ‘Golden Power’ gwaa3. The thing [company] is called ‘Golden Power’ so I guess that’s it laa1!’

Example (4.11) is another in which wo3 is used in a response to a question, but the use of wo3 is slightly different. Speakers A and J are talking about J’s work. J works in a company with two divisions. She is involved in the division which handles rubber keypads, while the other division produces batteries. Speaker A asks her what brand the batteries are. J responds, laughing a little, that she did not pay attention to their batteries, and speaker A laughs while J explains. Speaker J uses wo3 when she says she did not pay attention to the batteries because she knows this is contrary to what A is expecting. Speaker A expects J to know what brand the batteries her own company produces are, and since J has not noticed, J uses wo3 to indicate ‘you aren’t thinking about this at this moment’. She also expresses ‘it is good if you think about it’, since this is the answer to her question, is unexpected and funny, and possibly considered noteworthy. This example was also used in Chapter 3.
This use of wo3 in a response to a question is different from examples (4.9) and (4.10), where wo3 and the whole response was used to point out something ‘wrong’ with the question being asked. Here, speaker J is not suggesting that A’s question is unusual in any way. On the contrary, speaker A’s question is quite reasonable and it is J who realises that her own answer is unexpected. She admits she does not know something which she should know, and is able to laugh at herself. Thus she does not seem ‘conceited’. Example (4.11) could possibly be argued to be ‘reluctant’, although J is not so much reluctant to as unable to, and she does take a guess, thus ‘reluctant’ is not the best description. Additionally, ‘reluctant’ has a slight connotation of something being bad or serious, and we know from the laughter and the speaker’s tone of voice that this was not considered a bad or serious issue. This means that Yau’s (1965, 112-116, 1980, 50) connotation concepts are again challenged.

One of the contexts Luke found for wo3, ‘reportings and story-tellings’, may be relevant to example (4.11). Luke (1990, 200-216) states that in giving a report or telling a story, the use of wo3 indicates that actions or events are unusual, extraordinary, or unexpected, setting up an explicit or implicit contrast with a norm or initial expectation. Though J’s response in (4.11) might not be labelled as a report or telling of a story, the rest of Luke’s observations appear to be true – wo3 highlights that J’s utterance is unusual and unexpected. Luke’s observations are compatible with explication [4A].

(4.12)

A: ‘Hey did you know Jenny ne1, besides ALO [Assistant Labour Officer] ne1, also got an offer for EO [Executive Officer]. Did I tell you aa3?’

B: ‘No aak3.’

A: ‘No aa4? That day Lily told me gaa3.’

B:

Daan6hai6 keoi5 zou6 m4 zou6 aa3?

but she do not do PRT

‘But will she do it [accept the job offer] aa3?’

A:

Ngo5 m4 ziu wo3, daan6hai6 ziu3

I not know PRT but according
Example (4.12) above shows wo3 again being used as a response to a question, but though it is ‘unexpected’, it is fairly ‘straightforward’. Speakers A and B discuss their friend ‘Jenny’. They both knew that Jenny had received a job offer, but speaker A now brings up the news that Jenny has received a second job offer, i.e. now has two offers. Speaker B asks whether Jenny will take it, but speaker A does not know, although she makes a guess. In response to B’s question, and especially because in the preceding turns, it seemed like A knew a lot about Jenny’s situation, wo3 serves to inform B that the truth is contrary to what he expects; speaker A does not know. From A’s perspective, B believes that A knows what Jenny will do, but A does not, hence ‘you aren’t thinking about this at this moment’. It is good for him to think about it because A does not know, but like A, he can guess based on the two available jobs.

The ‘unexpectedness’ in (4.12) is different from the ‘unexpectedness’ of examples (4.9), (4.10) and (4.11). Common amongst the other three examples is that the question asked is not answered with a ‘straightforward’ or ‘simple’ answer. In (4.9) and (4.10), the wo3-suffixed utterance helps to ‘avoid’ answering the question in a straightforward way. In (4.11), the wo3-suffixed utterance answers the question in an unexpected way which elicits laughter. In this example, the wo-suffixed utterance is not meant to be funny, although the wo3 is still unexpected because in the previous turns A seems to know a lot about the situation.

The same comment about reluctance can be made as with examples (4.10) and (4.11). The speaker that uses wo3 is not ‘reluctant’ to give the answer. She simply does not know the answer. But given that she makes a guess to try to
answer as best she can, it would be wrong to say she was ‘reluctant’. It is also not the case that speaker A is making a ‘realisation’ or giving a ‘reminder’ here, as believed by W-m Leung (2009, 1-2). It is highly unlikely that A is signalling that she has just made a realisation that she did not know what Jenny is going to do. Nor is A reminding B that she does not know, since they have not talked about this before. The NSM explication proposed ‘you aren’t thinking about this at this moment, it is good if you think about it’ seems more likely. Example (4.13) below shows four instances of wo3.

(4.13)

B:

\[\text{Hai2 mei5gwok3 mai6 hou2do1 go2di1 go2di1 at USA there-are’ very many those those giu3zou6 ceot1 - go2di1 factory outlet ne1. called out - those factory outlet PRT}\]

‘In the US [you know] there are many those, those called out – those factory outlets ne1.

A:

\[\text{Mou5 \textbf{wo3}, (4.13a) mou5 mat1 not-have PRT not-have any dim2 heoi3 \textbf{wo3}. (4.13b) how go PRT} ‘Didn’t \textbf{wo3}, didn’t really go \textbf{wo3}.’

B:

\[\text{Nei5 mou5 heoi3 haang4 me1? you not-have go walk PRT} ‘You didn’t go walking [shopping] me1?’

A:

\[\text{Mou5 \textbf{wo3}, (4.13c) go3 dou6jau4 dou1 mou5 not-have PRT CL tour-guide even not-have daai3 ngo5dei6 heoi3. bring us go} ‘No \textbf{wo3}, the tour guide didn’t even take us there.’

B:

\[\text{Go2dou6 maa5 je5 dai2 di1 \textbf{wo3}. (4.13d) There buy thing cheap some PRT} ‘It’s cheaper to buy things there \textbf{wo3}.’
In example (4.13), speakers A and B are talking about A having been to the United States. B’s statement at the beginning of the excerpt implies that B wants to talk about the factory outlets in the US and seems to expect that A visited some. Speaker A understands this and uses wo3 suffixed to her negative answer in (4.13a–c). Speaker A is indicating ‘you aren’t thinking about this at this moment, it is good if you think about it’ because A knows that B is expecting her to have something to say about the factory outlets, and wants her to understand that she did not go and therefore does not have much to say about them. The unexpectedness of A’s answer can also be seen in B’s question in the third turn of this excerpt, as well as A’s volunteering of extra information after (4.13c) about the tour guide not taking her to the outlets, to help B understand.

It was considered, during the development process that led to the current explication, that a possible component could be something like ‘I want you to think like this’, or something similar that would suggest that the speaker wants the hearer to change their mind about something. For instance, in (4.13a) and (4.13b), speaker A could be trying to change B’s assumption that she went shopping at a factory outlet. In (4.13c), speaker B seems to be trying to convince A that factory outlets are cheaper and that she should have gone. However, these components were not valid for all examples of wo3 found. For instance, (4.8a-b) show the answers to neutral questions ‘Have you watched any movies?’ in which the speaker does not have an existing idea that needs to be changed. Example (4.12) also shows the answering of a question, where the speaker does not need to try to change the hearer’s mind about anything.

‘Disconfirmations’, one of the contexts identified for wo3 by Luke, is relevant here, (Luke 1990, 227-231). He states that following a confirmation-seeking turn, wo3 can highlight a discrepancy between a prior assumption and some facts or evidence which contradict or cast doubt on the validity of that assumption. This appears to be what is happening in (4.13a–c). Explication [4A] is compatible with such ‘disconfirmations’.

At the end of example (4.13), speaker B uses wo3 as well, when explaining that products are cheaper at factory outlets. This is labelled (4.13d). She is providing information which, with wo3 suffixed, acts more like a suggestion (or possibly a reminder, depending on whether the hearer could have known this before). B seems to be expressing something like ‘why didn’t you go shopping at
a factory outlet to get cheaper products’, ‘you should have gone shopping at a factory outlet to get cheaper products’ or ‘next time you should go to a factory outlet to get cheaper products’. Wo3 helps her indicate this because wo3 means ‘you aren’t thinking about this at this moment, it is good if you think about it’.

From the perspective of Luke’s sequential contexts, example (4.13d) would probably fit under the category of ‘informing and reminding’, with the information being volunteered, i.e. offered by the speaker on her own initiative (as opposed to being given in response to an inquiry). Recall that Luke (1990, 233) believes that volunteering is a means of highlighting the prominence of some information, and presenting it as noteworthy (although this depends on a subjective judgment of noteworthiness, since much is volunteered in conversations). Luke also explains that wo3 can accompany a sequentially misplaced information offer. For example, where one should indicate information receipt, one could provide an information offer – this is like example (4.13d). The (mis)placement of a turn highlights the information for special salience and noteworthiness (Luke 1990, 234-235). Explication [4A] is compatible with Luke’s observations, but uses ‘it is good if you think about it’ to indicate what Luke calls ‘noteworthiness’.

Similarly to Luke, Matthews and Yip (2011, 391-392, 406) report that wo3 is informative and reminds or indicates something new, noteworthy, and newsworthy. These examples of wo3 could be described as mildly informative in that A is informing B that she did not go to any factory outlets, and B is informing A that it is cheaper to shop at factory outlets. (4.13c) is just speaker A repeating what she said before, and is not a genuine ‘reminder’ of what she has just said in the previous turn. But as she is repeating that she did not go, (4.13c) cannot be considered to signal ‘new’ information either. As for being noteworthy or newsworthy, again, this is rather subjective.

**Wo3 in suggesting and reminding**

As we can see from (4.13d), answering questions is not the only role of wo3. It can also be used with statements which help to make suggestions or reminders. The following example is similar. In example (4.14) speaker B learns that speaker A is planning to watch the new Batman movie the coming Tuesday. Speaker B responds by informing/suggesting/warning/reminding A that he
should buy the tickets earlier. This utterance is suffixed with the particle \textit{wo3} ‘you aren’t thinking about this at this moment, it is good if you think about it’. Speaker B fears that A is not thinking about buying the tickets early. It is good if he thinks about this, presumably so that he can organise them early and avoid disappointment on Tuesday as the tickets may be difficult to come by at that stage. Speaker A’s response to this is consistent with his understanding this and agreeing that buying tickets early would be better.

(4.14)\textsuperscript{59}

B:

\begin{verbatim}
... \textit{Nei5 jau5mou5 nam2 zyu6 heoi3 }
you have-not-have think CONT go
tai2 \textit{aa3 Batman?}
watch PRT Batman
\end{verbatim}

‘... Are you planning to go and watch \textit{aa3}, Batman [the movie]?’

A:

\begin{verbatim}
\textit{Jau5. Dang2 sing1kei4ji6 sin1 laa1.}
have wait Tuesday first/before PRT
\end{verbatim}

‘Yes. Wait until Tuesday first \textit{laa1}. [I’ll watch it on Tuesday.]’

B:

\begin{verbatim}
\textit{O3, gam2 daan6 nei5 jiu3 jyu6zou2}
oh so/then but you need plan-earlier/prepare
\textit{maai5 ding6 fei1 \textit{wo3}.}
buy already ticket PRT
\end{verbatim}

‘Oh, but then you have to buy tickets early/beforehand \textit{wo3}.’

A:

\begin{verbatim}
\textit{Gam2aa6 hai6.}
so/like-that is
\end{verbatim}

‘That’s true.’

This use of \textit{wo3} does not signal anything particularly unusual about the preceding turn, as in examples (4.9) and (4.10) where the preceding question was not answered, nor is the utterance unexpected or funny as in example (4.11). On the contrary, there is nothing particularly unexpected or uncalled for in this excerpt. Being quite similar to example (4.13d), this example might also

\textsuperscript{59} This is not the same conversation as (4.8), which was also about the \textit{Batman} movie.
fit under Luke’s category ‘informings and remindings’, signalling something ‘noteworthy’. Wo3 does not purely indicate the giving of information (that you need to buy tickets early). By looking at the context we can see that the information provided is intended as a helpful suggestion or warning. The NSM explication proposed for wo3 is still valid and applicable to this example and shows how versatile the particle is.

Other NSM components previously considered for wo3 included ‘I think you don’t know this’ or ‘I want you to know this’, where the thing being referred to as ‘this’ could be either a fact or some opinion of the speaker’s. As with other rejected components, this seemed plausible according to native speaker intuition, and could be an intended meaning in some, but critically not all, situations in which wo3 is used. Such components would have seemed fitting for examples like (4.13) and (4.14), as well as (4.2) (discussed earlier), among others. But consider (4.1), which features an utterance like ‘the scenery there seems quite beautiful wo3, isn’t it?’ It does not make sense to suggest ‘I think you don’t know / I want you to know that the scenery there is beautiful’, and then subsequently ask whether the scenery there is beautiful or not.60

(4.15)
A: ‘Where did we go to eat that time? Hot pot, wasn’t it?’
J: ‘[I] don’t remember.’
A: ‘[We] went to Siu3gyun1’s home.’
J: ‘We did?’
A: ‘Went to Siu3gyun1’s home for hot pot.’
J: ‘That was Siu3gyun1’s home?’
A: ‘Yes.’
J: ‘Are you sure?’

60 Example (4.2) is like (4.1) in that it shows wo3 being used with a statement immediately followed by a question. Stating that you do not know something is a good way to ask a subsequent question about it. In an utterance like ‘I don’t know what results you got wo3. Five As and then what?’ these early components could have been substituted as ‘I think you don’t know / I want you to know that I don’t know what results you got’. This would have made the question about what results he achieved a predictable and natural follow-up.
Example (4.15) shows w03 being used more like a reminder, pointing out something that has been overlooked or forgotten. Speaker A and J talk about a hot pot meal celebration they had previously at a friend’s house, which J has forgotten about. Speaker A says that their friend’s house was the one in Yuen Long (a place in Hong Kong), and w03 is used with this utterance. Speaker A is saying that speaker J is not thinking about this at this moment – this has been made obvious by J who has forgotten about the event – and that it is good if J thinks about it – perhaps the information about the house being in Yuen Long will help her recall the event. Actually, we can see from the next turn that J does remember a dinner in Yuen Long, although she had forgotten that that particular dinner was to celebrate the event they are talking about. The NSM explication works well for this example too. The latest explication [4A] using ‘at this moment’ seems to work better here than the earlier version which used ‘now’. Examples like this also help to reject alternative components such as ‘you didn’t think like this before’, because it is clear that the speaker believes the hearer should have known it or thought about it before. W-m Leung’s (2009, 1-2) description ‘reminder’ is very applicable in this example, and also overlaps with descriptions by Luke (1990) and Matthews and Yip (2011, 391-392, 406). Nonetheless, as we see from the other examples from the corpus, ‘reminder’ is not the sole or invariant meaning of w03.

In many examples, one had the feeling that it might be bad for the addressee not to have previously known whatever was being spoken about. This prompted consideration of evaluative components such as ‘you didn’t know this before, this is bad/not good’, which might have made sense for examples like (4.9) and (4.10). Other examples, like (4.11) and (4.12), might have included a similar component that indicates the speaker (as opposed to the addressee) did
not know something, and that this was bad/not good. Nonetheless, it was quickly found that a negative evaluation is unnecessary. For instance, in an utterance like ‘the scenery there seems quite beautiful wo3, isn’t it?’ in example (4.1), there is no reason to indicate ‘this is bad/not good’. Some other examples further below will also demonstrate that the negative evaluation is not a viable semantic component.

In example (4.16), speakers M and E are discussing where their group of friends will meet at an upcoming event. Both M and E live in the New Territories in the north of Hong Kong, but they have a friend ‘Elton’ who prefers to meet in places like Wan Chai District on Hong Kong Island, which is in the south of Hong Kong (and requires crossing the harbour to get to). M and E are half joking and half serious about suggesting places in Hong Kong to meet which are further and further north. In this excerpt, they suggest that the furthest they will go to meet Elton is Tsim Sha Tsui District (which is in Kowloon, between the New Territories and Hong Kong Island). They say they do not want the meeting place to only be convenient for Elton and not the rest of the group.

(4.16)

M:

*Ngo5dei6 dou1 hai6 heoi3 zim1saa1zeoi2keoi1* we even/still is go Tsim Sha Tsui District

*go2di1 laa1.*

those PRT

‘We’ll (again) be going to Tsim Sha Tsui District or surrounding areas laa1.’

E:

*E6.*

eh

‘Eh.’

M:

*Cin1kei4 m4hou2 heoi3 waan1zai2keoi1* under-no-circumstances not-good go Wan Chai District
wo3. (4.16a)  Fai3si6  zau6
PRT  not bother  suit / for their sake/convenience
aa3  aai6ji5deon6.
name-prefix  Elton

‘Under no circumstances go to Wan Chai District wo3.  [I] can’t be
bothered / don’t want to [go] for the convenience of Elton.’

E:
Hai6  lo1.  Ngo5  seng4jat6  gok3dak1
is  PRT  I  always  feel
zau6  keoi5  m4hai6  gei2
suit / for their sake/convenience  he  not-is  very/quite
hou2  wo3. (4.16b)
good  PRT

‘Yes lo1.  I always feel that [us] doing things in a way that suits him is not
very good wo3.’

M:
Mou5  lei5jau4  gam3  do1  jan4
not-have  reason  so/that  many  people
zau6  keoi5  jat1  go3  gaa1  maa3.
for their sake/convenience  he  one  CL  PRT  PRT

‘No reason / it doesn’t make sense for so many people to do things in a way
that suits / is convenient for just one person (him) gaa1-maa3.’

M’s use of wo3 in (4.16a), that under no circumstances should they arrange to
meet in Wan Chai District, is like a stern statement or even mild warning.  She
uses wo3 ‘you aren’t thinking about this at this moment, it is good if you think
about it’, probably because E has hesitated a little after her previous comment,
and also to emphasise the sense of warning to M, that he can think about what a
bad idea it would be to go to Wan Chai District and know not to do it.  (Clearly a
one word description like ‘warning’ is not adequate to explain wo3 or to identify
its invariant meaning, even if it helps to understand this specific example.)

E’s use of wo3 in (4.16b) helps express ‘you aren’t thinking about this at
this moment, it is good if you think about it’, because he is further pointing out
that they all always do things that suit Elton, and that he does not like this.  The
two speakers then continue to vent their frustration at how they always do
things Elton’s way, for the next 20 seconds or so, even talking over each other to
talk about how much they do not want to travel so far.  (And they continue to
talk about Elton even after the topic of travelling that far is over). We can see here that E uses \textit{wo3} to ‘argue’ his point even though he is actually agreeing with M. Use of \textit{wo3} helps them to bring in more points to support their ‘argument’, introducing points which may have been overlooked so far. It shows that they feel fairly strongly about what they are saying and want the other person to consider and understand them. This use of \textit{wo3} does not fit well with any of the previous descriptions of \textit{wo3} in the literature, besides perhaps ‘noteworthiness’.

(4.17)  
\begin{verbatim}
M:         \text{Ngo5 gau3 zyu6 zo2 sap6 gei2}
I even/also live PFV ten a few/something  
jaa6 nin4 dou1 hai6 zyu6 soutuk1cyun1 lok3.
twenty year also/still is live So Uk Estate PRT
'I also lived for ten or twenty years at So Uk Estate \textit{lok3}.'
\end{verbatim}

\begin{verbatim}
E:            \text{Daan6hai6 nei5 ji4gaa1 hai6gam2 zik1hai6, e6,}
but you now keep/constantly meaning eh  
\text{gaan3m4zung1 jau6 bun1 haa5 uk1}
every-now-and-then again/also move DEL house  
gam2joeng2 \textbf{wo3. (4.17a)} Bun1 jyun4 jau6
like-that PRT move finish again/also  
bun1 gam2joeng2 \textbf{wo3. (4.17b)}
move like-that PRT
\end{verbatim}

[Laughter] ‘But now you keep I mean, eh, moving every now and then \textit{wo3}. Like moving again and again \textit{wo3}.’

M: ‘Moving again and again?’

E: ‘Yes.’

M: ‘I want to move far away from you \textit{aat-maa3}.’

In example (4.17), the speakers are talking about how they have lived in the same places for over 10 or 20 years. E makes a comment about M who has moved several times recently. This example further supports the rejection of previously considered components such as ‘I think you don’t know this’ or ‘I want you to know this’. It is unlikely that E’s use of \textit{wo3} means something like ‘I think you don’t know this’ or ‘I want you to know this’, as obviously M would
already know about her own moving. Also, as we can see after E’s comments, M makes a joke that she was moving away from E; this also confirms that it is true. Explication [4A] makes much more sense. M has said she lived in the same place for a long time and E is bringing up the fact that she has moved several times, saying ‘you aren’t thinking about this at this moment, it is good if you think about it’ to counter her statement that she has lived in the same place for a long time (it seems M was talking about having lived at a previous place for a long time, before she started moving around).

(4.18)

A: ‘Don’t you like Gigi Leung [a Cantopop singer] gaa3-mei1?’ [‘I thought you liked Gigi Leung.’]

B:

Teng1 haa5 hou2 laa3.
listen DEL good/enough PRT
‘Listening a bit is good/enough [and nothing more] laa3.’

A:

Teng1 haa5 keoi5 di1 go1?
listen DEL her CL song
‘[You mean] listening to her songs?’

B:

Teng1 haa5… Gam3 gou1.
listen DEL so/that tall
‘Listening… [She’s] so tall.’

A:

M4 gwaan1 ne15 si6 wo3.
not related you thing PRT
‘That has nothing to do with you wo3.’ [Laughter]

In the above example, (4.18), speaker B comments that he does not like a Cantopop singer very much (or as much as speaker A thinks he does). Speaker B mentions that she is too tall, and speaker A remarks that that has nothing to do with B, and laughs. Speaker A is teasing him with a light-hearted put-down. Her use of wo3 indicates that he is not thinking about this and that it is good if he thinks about it. It helps her emphasise her joke by implying that he has no chance and no relationship with the singer. It is funny and ‘unexpected’, like
example (4.11) above, but instead of poking fun at herself, this speaker pokes fun at the other person.

Relevant here are Luke’s observations about challenging a position. He states that wo3 is commonly used to somehow undermine or challenge the truth, validity or generality of a rule proposed in a prior turn, often pointing to aspects of a situation which have been overlooked (Luke 1990, 216-223). This fits well in this example, as A is pointing out something which B has overlooked (that B has no chance with the singer) and challenges his view that the singer’s height is relevant.

**Giving information with wo3**

In example (4.19), speakers A and B are discussing their friend ‘Raymond’, who earns $20 000 a month. They are having a slight disagreement, as speaker B says this is not a lot, but A is questioning B, because she does not agree that it is not much. Speaker B then points out that Raymond’s boss earns $130 000 a month. This helps his argument because this is significantly more than what Raymond is earning. B’s use of wo3 points out to A that A was not thinking about how much Raymond’s boss earns and that it would be worthwhile for A to take that into consideration.

(4.19)

B:

**Wai3! Keoi5 lou5sai3 sap6saam1maan6 wo3.**

‘Hey! His boss [earns] 130 000 [dollars] wo3.’

In example (4.20), speakers A and J are talking about J having been to Shenzhen with her boss. Her boss returned to Hong Kong after a meeting and did not stay overnight, so J had to stay there by herself. Speaker A laughs that J is scared of ghosts, and J uses wo3 to point out the fact that she was alone, which is her reasoning/defence for being scared, and she asks (somewhat rhetorically) if A would be scared in that situation. Wo3 again indicates ‘you aren’t thinking about this at this moment, it is good if you think about it’. It is similar to, for instance, (4.15) above, in that it is supplying additional information (in this case, that she was alone in the room), but this example is
not a reminder. After the wo3-suffixed utterance, speaker A appears to change from initial surprise to agreement.

(4.20)
A: ‘Then were you scared aa3?’
J: ‘Scared [yes] aa3! Actually I was most scared of ghosts.’
A: Geng1 guai2?!
scared ghost
‘Scared of ghosts?!’ [Laughter]
J: Wai3 jat1 go3 jan4 hai2 gaan1
hey one CL person at CL:building
fong2 wo3 m4 geng1 ge3 me1?
room PRT not scared PRT PRT
‘Hey one person [just me] in the room wo3 not scary [wouldn’t you be scared] ge3-me1?’
so/like-that is also/even strange/unfamiliar
‘That’s true. It’s [the place is] also strange/unfamiliar.’

Prior to the excerpt shown in example (4.21) below, speakers A and B have been talking about cheap flights. Speaker A starts talking about going to Disney World in Orlando, using wo3 to say firstly that it is the 25th anniversary, and secondly that it is good value to go. This utterance introduces a new topic. It is unlikely that A wants to say it is bad that B did not know this before, and this supports the view that the negative evaluation considered above is not part of the meaning of wo3. The explication [4A] ‘you aren’t thinking about this at this moment, it is good if you think about it’ makes good sense here because speaker A is starting a new topic and wants to discuss it with her friend.

(4.21)
A:
Ngo5 teng1 pang4jau5 gong2 waa6 heoi3, e6,
I listen friend say say go eh
Orlando go2go3 gau6 – go2go3 dik6si6nei4 ne1, Orlando that old that Disney PRT
jaa6ng5 zau1nin4 wo3. (4.21a) Hou2
twenty-five anniversary PRT very
daiz waan2 wo3. (4.21b)
worth/good-value play PRT

‘I heard a friend talking about going to, eh, that old Orlando – that Disney World ne1, twenty-fifth anniversary wo3. Really worth/good-value going (playing) wo3.’

In (4.22), speaker B has been telling A about his results at school in a very matter-of-fact and nonchalant way. He states it so nonchalantly that it seems he was not thinking of his high achievement, and B wants to compliment him and encourage him to think about how good this is. Example (4.22) can be interpreted as B praising A and ‘reminding’ or ‘informing’ him that his result is very impressive. Wo3 here helps to reinforce and highlight how surprised/impressed B is by A’s result in maths. She may also be trying to praise and compliment A by expressing awe at his achievement. Again, ‘you aren’t thinking about this at this moment, it is good if you think about it’ seems very fitting.

(4.22)

B: ‘So in the end did you get an A in maths aa3?’
A: ‘Yes aa3.’

B:
O3, waa3 gam2 dou1 hou2 sai1lei6 wo3. oh wow so/then even/also very impressive PRT
Maths haau2 A zi3siu2 jiu3 gau2sap6gei2 maths test A minimum/at-least need ninety-something
fan1, gau2sap6ng5 fan1 ji5soeng6 gaa3 wo3. point/mark ninety-five point/mark above PRT PRT
‘Oh, wow then that’s actually very impressive wo3. To get an A in maths [you] need at least ninety-something points, ninety-five points or above gaa3-wo3.’

A:
Caa1m4do1 laa1.
not-much-difference PRT
‘Approximately/about that much laa1.’
Example (4.23) is part of a conversation about whale-watching in Australia. For some time, speaker B has been talking about going whale watching. Speaker A states, rather dismissively, that she is not interested (recall that laa1 indicates the speaker thinks the other person knows how they are thinking, and that they can not-say more, and ‘syun3 laa1’ is a fairly common phrase). Speaker B then states that she wants to go whale watching, adding the utterance particle wo3, or ‘you aren’t thinking about this at this moment, it is good if you think about it’. This wo3 helps speaker B is point out something which she feels A is not paying attention to or not understanding. It seems here that B thinks A could have known this before; A should have known that B wants to see whales due to what B has said earlier in the conversation. B feels it is necessary to draw A’s attention to this, and to (continue to) direct the conversation towards this. Speaker A’s next utterance reinforces the fact that she now realises this is how speaker B felt, and could be interpreted as an agreement to discuss it more.

(4.23)
B: ‘But I don’t know which month is best for whale watching laa1.’

A: Syun3 laa1. Ngo5 deoi3 king4jyu4 dou1
let-it-go PRT I towards whale also/even
mou5mat1 hing3ceoi3.
not-much interest
‘Let it go (/forget it) laa1. I don’t even have much interest in whales.’

B:
Aa3, ngo5 soeng2 tai2 haa5 wo3.
ah, I want see DEL PRT
‘Ah, I want to see wo3.’

A:
Hai6 me1?
yes/is PRT
‘Really me1?’

The particle wo3 in example (4.23) (as well as some other examples, for instance (4.8c)) seems to be helping with topic organisation. According to Luke (1990, 235-237), wo3 is sometimes found in turns that are ‘topic initial in the sense that it is a proposal which may contribute to the generation of further talk,
if and when it is retrospectively established as a next topic’. Wo3 gives a turn of talk a ‘noteworthy’ character. If wo3 means ‘you aren’t thinking about this at this moment, it is good if you think about it’, this is compatible with a speaker signalling to the hearer that s/he wants the hearer to engage in a conversation about what has just been said, and we can see why it might be appropriate for such topic organisation.

Using wo3 when remembering and realising

Often, wo3 is used when one remembers or realises something. In these cases, the proposed NSM explication works if we think about the speaker talking to themselves or thinking out loud, where ‘you’ in the NSM explication would be directed at oneself. The following excerpt, (4.24), shows wo3 being used as if talking to oneself. Speakers A and B are talking about football. Earlier in the conversation, B had said that Sheringham was ‘a piece of rubbish wearing a number 10 shirt’. After a few turns, A returns to this remark and B eventually realises he was wrong about Sheringham wearing number 10.

(4.24)

A: ‘Who aa3? Sheringham wears number ten me1? Number ten shirt me1? Beckham net then?’

B: ‘Number seven lo1.’

A: ‘Changed to number seven aa4?’

B: ‘Correct.’

A: ‘Why ge2?’

B: ‘He was always number seven gaa3-wo3.’

A: ‘Beckham was number ten gaa3-wo3 last year.’

B: ‘Hey! Well his England shirt is number seven gaa1 maa3.’

A: ‘Oh.’

B:

Gam2, ji2 m4hai6 wo3, (4.24a) si2gouisi1 cat1
So, huh not PRT Scholes seven
hou6 gaa3 wo3. (4.24b) Bik1haam4 gei2do1
number PRT PRT Beckham how-many
‘So, [voice becomes quiet] huh, no wo3, Scholes is number seven gaa3-wo3. What number is Beckham aa3? [Returning to normal voice] I forgot tim1. [Jokingly] Tell you later!’

A: ‘Beckham in – you mean in the England national team?’

B: ‘Yes aak3.’

A: ‘Number seventeen aa3!’

B: ‘Number seventeen me1?’

A: ‘Yes aa3.’

B: ‘Number seven, wasn’t it aa3?’

A: ‘Seventeen!’

B: ‘Number seventeen is McManaman aa5!’

A: ‘He wore McManaman’s football shirt aa3, how can you still talk!’

B: 
Aai5! Hai6 wo3! (4.24c) Zan1hai6 aa1, dou1 ai yes PRT really PRT also/still aam1 ge3, dou1 aam1 ge3. correct PRT also/still correct PRT

‘Oh! Yes wo3! Really aa1, you’re right ge3, you’re right ge3.’

[A and B then both laugh and continue their joking and banter.]

The occurrences of wo3 in (4.24) by B appear to be directed at B himself, especially as, in (4.24a) and (4.24b), he lowers his voice and sounds puzzled when he says them, before raising his voice back to the normal level he uses in the rest of the conversation. Then, in (4.24c), B finally realises that he is mistaken and again uses wo3 directed at himself. Wo3 here signals that B was mistaken. He realises that he was not thinking about the correct shirt numbers (‘you aren’t thinking about this at this moment’), and that it is good if he thinks about it to recall the correct numbers (‘it is good if you think about it’). Based on what B is saying, as well as the way his voice changes when he realises his
mistake, it is clear that B is directing these wo3-suffixed utterances not to speaker A, but to himself.

Kwok states that when suffixed to hai6 ‘yes’, hai6 bo3 is used to indicate the speaker’s agreement, while suggesting that the idea has not occurred to him before (Kwok 1984, 65-66). Given that hai6 ‘yes’ on its own already indicates the speaker’s agreement, we can postulate that Kwok’s own intuitions point to the particle as doing the additional work of suggesting that the idea has not occurred to the hearer before. Again this fits in well with the proposed NSM explication, ‘you aren’t thinking about this at this moment, it is good if you think about it’.

The use of wo3 in example (4.24) has some similarities with ‘realisations’ as described by Luke. Amongst other things, wo3 can indicate ‘being reminded’ (Luke 1990, 242-255). He believes, however, that the use of wo3 may still be directed at the hearer. The recipient of some information can retrospectively give that information a special status of being something that the recipient should have known or something which has led to a realisation or remembrance. If we look at example (4.24c), Luke appears to believe that speaker B is retrospectively marking the previous turn, i.e. A’s remark that Beckham wore McManaman’s football shirt, as something that B should have known, and which has led to B realising his mistake.

This is plausible with (4.24c), but when we look at (4.24a) and (4.24b), this is slightly problematic in that A appears to agree with B that Beckham was number 10. Perhaps it could be that the necessary information does not have to be in a clear way in the preceding turn; the multiple preceding turns and A’s apparent confusion may have been enough to spark B’s realisation. In this case, we would have to say that B is retrospectively marking the multiple previous turns as special or ‘noteworthy’. However, when B says ‘Scholes is number seven gaa3-wo3’ in (4.24b), it would seem that wo3 is attached to the utterance that B is producing at that time, rather than retrospectively attaching to any previous utterance that A might have said. Furthermore, given that B lowers his voice and sounds puzzled when he produces (4.24a) and (4.24b), then returns to his normal voice, it still seems likely that he was speaking to himself. In any case, the explication seems to be valid here.
Some of the descriptions used by W-m Leung (2009, 1-2) can be applied to this example of wo3. For example, it is being used at a time of ‘realisation’ as the speaker realises his mistake, and is also ‘showing contrast’ between what he thought and what is true. Although the wo3-suffixed utterances do signal the speaker remembering something, they are not being used to give a reminder. The wo3-suffixed utterances here can perhaps be argued to be signalling that a reminder has just been given by the other speaker, but this seems less likely with (4.24a) and (4.24b), for the same reasons as above. W-m Leung also stated that wo3 indicates ‘unexpectedness’, which is somewhat plausible in this example if we interpret wo3 as signalling that the speaker did not expect that he has remembered the wrong number for the various football players, but this is better described by the label ‘realisation’.

As speaker B clearly recognises his own mistake in these examples of wo3, this again goes clearly against the idea that use of wo3 conveys conceitedness. If he was being conceited, he might not have admitted he was wrong. Nor is he being reluctant to admit he is wrong, as when he realises his mistake, the change in his voice shows when it happens, and he immediately admits his mistake out loud. It is simply taking him this many turns to realise his mistake. The nature of their relationship (they are clearly friends) and the light-hearted nature of the conversation and banter gives no reason for him to be reluctant to admit he is wrong. Yau’s (1965, 112-116, 1980, 50) connotation concepts clearly do not apply to all (or even most) examples of wo3, although the NSM explication proposed does clearly explain what wo3 is adding to the utterances.

(4.25)  
M: ‘... I have to go to work on Wednesday aa3.’  
E: ‘Yes aa4?’  
M: ‘Yes aa3.’  
E:  
\text{Dim2gaai2} \quad \text{ge2?}  
why \quad \text{PRT}  
‘Why ge2?’
M:

E2, m4hai6 wo3! Lai5baai3saam1 ngo5 zan1hai6
eh not-is PRT Wednesday I really
hou2ci5 m4sai2 faan1 aa3 go2jat6.
like not-need go PRT that-day

‘[Loudly] Eh, no wo3! [Pause, then returning to normal voice] On Wednesday I think I really don’t need to go to work aa3 on that day.’

Example (4.25) is similar to (4.24). In example (4.25), speakers E and M are discussing where they will go on Wednesday. Speaker E asks about M working on Wednesday, and at first, M states that she needs to go to work on Wednesday. However, she soon realises that she was mistaken, and suddenly exclaims loudly and in a more high-pitched voice ‘Eh, no wo3!’, then pauses before returning to her normal voice and saying that she actually might not have to work on Wednesday. It appears that M is saying this to herself as she realises her mistake. Her use of wo3 indicates something like ‘I forgot’, or ‘I was not thinking about this’, and ‘it is good if I think about it’. Again, use of ‘at this moment’ in explication [4A] works better than use of ‘now’ from an earlier explication, as it is a shorter, more precise period in time, and shows more immediacy.

Again, we can consider this example against W-m Leung’s (2009, 1-2) descriptions of wo3, with similar observations as with example (4.24). Example (4.25) can be described as ‘realisation’ and ‘showing contrast’ as M realises that she actually does not need to work on Wednesday, in contrast to her thinking she did need to work on Wednesday. This instance of wo3 is produced as the speaker remembers something, but is not being used to give a ‘reminder’.

At one point during the investigation, it was considered possible that wo3 indicated that the hearer could or should have known something before that point in time. This seemed to be right according to native speaker intuition (and could be an intended meaning in some but not all situations). For example, in utterances like ‘the scenery there seems quite beautiful wo3, isn’t it?’ (like in example (4.1)), it was plausible that wo3 indicated something like ‘you know this’, ‘you could have known this before’, ‘I think you know about this’, ‘I think you knew this before’, etc. It would have made sense for utterances like ‘I said I’m considering studying medicine, I didn’t say I’d
definitely do it *wo3* (like in (4.9)), ‘You should know better than me *wo3*’ (like in (4.10)) or ‘Remember, we went to the house in Yuen Long *wo3*’ (like in (4.15)). In utterances directed at oneself, for example when one forgets something and then remembers again or makes a realisation about it (like in (4.24) or (4.25)), they could be saying something like ‘you [I] knew this before’.

However, investigation of a larger number of real examples of *wo3* showed that this was not a stable and invariant meaning of *wo3*. As the NSM explication aims to find the invariant meaning of the particle, such components could therefore not be included. For example, an utterance like ‘I’ve never paid attention to my own company’s batteries *wo3*’ (like in example (4.11)) is unlikely to have *wo3* attached if *wo3* means something like ‘you know/knew this’. There would have been no reason to expect the other person to already know information like that, especially given the other person has just explicitly asked what brand the batteries are. In example (4.8), there is another instance of *wo3* being used to answer an explicit question, this time the question being whether that person has watched any movies lately, with the response being ‘no *wo3*’. Again, the particle here is unlikely to indicate that the person asking the question should already have known this. Also in example (4.8), one person gives new information with *wo3* attached, in an utterance like ‘yesterday I went to watch the *Batman* movie *wo3*’, and again, there is no reason for the speaker to think that the hearer already knew this. As there are cases like this of *wo3* being used with utterances which provide new or sought-after information, it must not be part of the invariant meaning of *wo3* that the speaker thinks the hearer should have already known something.

Note that some considered components like ‘I think you don’t know this’ or ‘I want you to know this’ contrast with other considered components such as ‘you know this’, ‘you could have known this before’, or ‘I think you know about this’. This type of contrast shows that particles like *wo3* can be used in very different scenarios with very different implications. As a result, native speaker intuitions about the meaning of particles is usually hazy, conflicting or inaccurate, and examination of a large number of examples taken from naturally-occurring examples is necessary to find the particles’ core meanings. Finally, consider example (4.26).
(4.26)  
M:  
\[ \text{jì2 neì5 jau5 zek3 san1 biu1 wo3.} \]  
\[ \text{eh you have CL new watch PRT} \]  
\[ \text{‘[Loudly] Eh [returning to normal voice] you have a new watch wo3.’} \]  

In example (4.26), M makes a comment to E about something which E surely already knew about, namely that he is wearing a new watch. This is similar to examples like (4.17). Rather than signalling something like ‘I think you don’t know this’, or ‘you didn’t know this before, this is bad’, mentioned above, wo3 is being used by M to draw attention to something new in the conversation that she wants to talk about (possibly to compliment E). Therefore ‘you aren’t thinking about this at this moment, it is good if you think about it’ makes more sense. This is M’s ‘realisation’ during a conversation about something else.

4.3 Concluding remarks

This chapter has investigated the Cantonese utterance particle wo3 and proposed that it has an invariant meaning, which can be expressed in the NSM explication ‘you aren’t thinking about this at this moment, it is good if you think about it’. This meaning is readily expressible in Cantonese. There have been notable studies of wo3 (sometimes considered the same as bo3) carried out in the past, with some previous descriptions being compatible with the explication proposed here. However, there are still problems with previous descriptions, and the explication is more helpful. The proposed explication appears to be valid for wo3 as used in a very wide range of situations and contexts, rather than only certain types of utterances. This is supported by evidence from the Hong Kong Cantonese Corpus, which also helped to reject inferior or unsuitable components.

The particle wo3 is discussed further in Chapter 8, which considers whether Cantonese utterance particles can be further broken down into multiple morphemes, and whether there are ‘families’ of semantically related particles. Wo3 is often grouped with wo4 and wo5, which will be discussed in that section, as well as the alleged meanings of wo3’s initial, rhyme and tone (as well as potential coda). The next chapter investigates the semantics of gaa3.
Chapter 5:  
The semantics of particle gaa3

Gaa3 is the 2nd most frequently used utterance particle in the corpus, and the 11th most frequently used word overall. It is the most frequently used of the particles analysed in this thesis, occurring in the corpus 1832 times. Surprisingly, the existing literature on gaa3 appears to be the most scant of the particles covered in this thesis. Where it was studied, gaa3 was usually analysed as part of a large group of particles, and/or as a contraction of the particles ge3 and aa3. Use of NSM and corpus data here helps to demonstrate that gaa3 has a meaning which can be simply and clearly stated, and which can be tested and supported by substitution in place of gaa3 in real conversations. The proposed explication covers the meaning of gaa3 in statements only.

As in previous chapters, section 5.1 begins with an overview of prior work on gaa3. It will be shown that accurate and comprehensive semantic analysis of gaa3 is lacking. Then, in section 5.2, the proposed NSM explication for gaa3 is explained and supported by real, naturally occurring examples from the corpus. This process also shows that numerous other, seemingly similar explications are actually not supported by the available evidence, and demonstrates again that NSM is capable of differentiating between very small variations in meaning. A shorter version of this analysis was published in HHL Leung (2013). Section 5.3 briefly overviews gaa3’s possible polysemous uses.

5.1 Previous descriptions of gaa3

Gaa3 has attracted noticeably less attention from linguists than the other particles investigated in this thesis, despite having the highest frequency of the five in the Hong Kong Cantonese Corpus. Well-known studies of Cantonese utterance particles, such as Gibbons (1980), Kwok (1984), and Luke (1990), do not explore the semantics of gaa3. One possible reason for the neglect of such a commonly used particle is that its meaning or function is too broad and general, which makes it difficult to pinpoint. Another is that it is widely accepted to be a
contraction of two other particles, so scholars may have felt it unnecessary to study \textit{gaa3} as a particle in its own right.

Scholars such as Kwok (1984), Fung (2000), and Matthews and Yip (2011) believe that \textit{gaa3} is a contraction of \textit{ge3} + \textit{aa3}. The origin of this pervasive idea is unclear, and Kwok does not explain why she believes this. Matthews and Yip (2011, 391-392) state that the contraction is obligatory and that the combination \textit{ge3-aa3} does not occur, but Kwok (1984, 46) presents examples of \textit{ge3-aa3} as a disyllabic particle. As for the particles \textit{ge3} and \textit{aa3}, according to Kwok (1984, 42-43, 45, 71), \textit{ge3} indicates that the speaker believes something is true and strengthens the force of the assertion, while \textit{aa3} does not carry much semantic content, mainly causing a sentence to sound less abrupt. Matthews and Yip (2011, 391-392) define \textit{ge3} as ‘affirmative: “this is the case”’ and \textit{aa3} as a ‘softening statement or question’. Fung (2000, 168-171) is largely concerned with comparing \textit{gaa3} with \textit{ge3} and with \textit{aa3}, as well as a ‘\textit{g}-initial family’.

The idea that monosyllabic particles may be broken down into smaller meaningful units raises many questions, which will be revisited in Chapter 8. There does not seem to be convincing proof that \textit{gaa3} is a contraction of \textit{ge3} + \textit{aa3}, and in any case, the possibility that it is a contraction or that it belongs to a larger family of \textit{g}-initial particles should not prevent us studying \textit{gaa3} as a particle in its own right, especially given its high frequency in speech. Here, we take the view that the classification of \textit{gaa3} as ‘basic’ or a ‘contraction’ is irrelevant to the analysis to find its core meaning.

Fung (2000, 168-171) states that sentences suffixed with \textit{gaa3} ‘remind the hearer of situations that should be known but may have been overlooked or neglected by the hearer... \textit{gaa3} assumes that the hearer has no knowledge of a situation that should have been known and is a given (as opposed to a new) situation’. She generalises that all utterance particles with the velar stop as the onset, including \textit{gaa3}, mark a situation that is part of the presumed knowledge of the participants in the exchange (Fung 2000, 136-138). Yet many real, naturally occurring examples of \textit{gaa3}, as will be provided in this chapter, show it being used to introduce new information, to teach someone something, or even as part of a question. Moreover, her description could be used to describe many examples of \textit{wo3} which we have seen in Chapter 4 (recall that \textit{wo3} can be
used for ‘reminders’ and ‘realisations’). Matthews and Yip (2011, 391-392) state that *gaa3* ‘has the effect of seeking confirmation of a statement’.

Yau appears to be the only scholar to have considered *gaa3* as a ‘basic’ particle, i.e. not as a contraction or as a member of a particle ‘family’. As explained, Yau studied all the Cantonese utterance particles he observed (and therefore none in much depth), using uniform quantitative methods for all. Yau (1965, 66) categorises *gaa3* as an ‘S-type’ particle, meaning that is used in utterances which are obviously affirming and which do not demand a verbal confirmation. This is at odds with Matthews and Yip’s (2011, 391-392) statement that *gaa3* ‘has the effect of seeking confirmation of a statement’. In fact, *gaa3* is used at the end of both statements and questions. Yau (1965, 112-116) also found that *gaa3* was associated with the connotation concept of ‘conceited’. However, ‘conceited’ is itself a complex English concept, and furthermore, it is not a meaning applicable to all occurrences of *gaa3*, as will be shown by many of the examples further below.

As can be seen, existing descriptions of *gaa3* are scarce and brief. The following section presents the NSM semantic analysis of *gaa3*. A large number of seemingly similar explications were tested and ultimately rejected, and some of these are also discussed. Real examples of *gaa3* from the corpus are provided to support the final explication and to show why some of the other considered explications were dismissed. They also show that the descriptions reviewed above are unsatisfactory.

### 5.2 NSM semantic analysis of *gaa3*

#### 5.2.1 The NSM explication for *gaa3*

Analysis of naturally-occurring examples reveal that the invariant meaning of *gaa3* in statements can be stated in NSM as follows: ‘it is good if you know this’. This simple explication explains the use of *gaa3* in a wide variety of situations, including in providing information, responses to questions, giving instructions, boasting, persuading, teaching, or even giving warnings. The examples considered below show how the proposed explication covers a wider range of uses of *gaa3* than alternative explications, as well as the existing descriptions of
gaa3 mentioned in section 5.1. Explication [5A] is given in English NSM and in Cantonese NSM, with gloss in English.

[5A] Final explication for gaa3, as used in statements:

   it is good if you know this

   jyu4gwo2 lei5 zi1dou3 zu6 hou261
   if you know then good

‘It is good if you know this’ may appear to be a vague and short definition, but if one remembers that this is the second most frequently used particle and the eleventh most frequently spoken word in Cantonese overall, it makes sense that it has a fairly simple meaning, which is applicable to a wide range of utterances and contexts. Furthermore, the simplicity and wide applicability of gaa3 may be a factor in the large number of particle clusters in which gaa3 appears.

The short length of the explication was not anticipated or planned at first, but emerged as a result of following the data. The explication does not specify who the content of the utterance is good for, and can be compared with ‘it is good for me if you know this’ and ‘it is good for you if you know this’. In the following, examples of gaa3 will be used to help explain this analysis, and also to justify the exclusion of alternative explications such as ‘I want you to think about this’, ‘it is good if I say something about this’, ‘I think you didn’t think about this before’, and many more.

5.2.2 Testing of explications and descriptions using the corpus

Roughly 20 examples of gaa3 in statements from the corpus are presented here (questions are considered further below). As in previous chapters, the subheadings below are not meant to indicate that these are different meanings of the particle, and some examples could have been placed under multiple subheadings.

61 This Cantonese explication does not directly specify THIS, as is the case in the English explication. It would be possible to add li1 joeng6 je5 [THIS CL THING] in Cantonese, but this makes the explication sound less natural, and the idea of THIS is understood without it.
Gaa3 used in giving information

To start with, consider example (5.1) below, which is taken from a conversation about going whale-watching in Australia. This excerpt shows four examples of gaa3 produced by speaker B providing information about whale-watching, labelled (5.1a–d). Example (5.1a) introduces the topic of whale watching, (5.1b) and (5.1c) introduce the idea that timing is important, and (5.1d) introduces the tours and the fact that they do not guarantee whale sightings. In addition to providing information, these utterances with gaa3 attached are presumably perceived by the speaker to be interesting or useful facts. The explication ‘it is good if you know this’ can be substituted in place of each instance of gaa3, and explains what the speaker is saying.

(5.1)

A: ‘Yes lo1. Seems quite beautiful wo3, the scenery, right aa3?’

B: ‘Mm, yes aa3.’

A: ‘Yes lo1.’

B:

Jau5 hoi2tyun4 ding6hai6 jau5 king4jyu4 t'ai2
there-are dolphin or there-are whale see

gaa3. (5.1a) Jau5 dii gun1tyun4, ho2jji5
PRT there-are some viewing-tour can
ceot1hoi2.
go-out-to-sea

‘You can see dolphins or whales gaa3. There are some viewing tours, you can go out to sea.’

A:

Nei5 dou1 m4 jau4seo1i2.
you also/even not swim

‘You don’t even swim.’

B:

Daan6hai6 jii3 - m4hai6 aa3, go2di1 jiu3 gaap3
but need - not-is PRT those need combine
aam1 si4gaan3 gaa3. (5.1b) Jiu3 t'ai2
correct time PRT need look/see
‘But even if you join those tours – it’s still not necessarily definitely guaranteed that you’ll definitely be able to see [whales]. I mean their what’s-it-called “whale observation tour”, or something like that. It’s not definitely guaranteed that you’ll see [whales] lo1. I mean they will have – have a boat to go out to sea.’

We can use these examples to evaluate the previous descriptions of gaa3. The four gaa3-suffixed utterances in (5.1) would probably not be described as ‘conceited’, nor does speaker B appear to be seeking confirmation of anything (Yau 1965, 112-116, Matthews and Yip 2011, 391-392). Speaker B is not asking questions, and has fairly long turns compared to A. Speaker B is clearly the one who knows about whale watching, and in contrast, speaker A makes it clear throughout the conversation that she does not know anything about it. There is no need for ‘softening’ or to sound ‘less abrupt’, as Matthews and Yip (2011, 391-
and Kwok (1984, 45, 71) point out that aa3 does. It seems to be true that the speaker assumes the hearer has no knowledge of the situation, but the accuracy of Fung’s (2000, 168-171) description depends on whether the speaker believes that that knowledge should have been known. The conversation in the corpus suggests that this is not the case, as B brings it up on her own and shows no sign of having expected A to know anything about it. B carries on for some time even though A does not know anything about it and even does not seem very interested (A actually states in a later part of the conversation that she is not very interested, as B continues talking about whales). It is unlikely that gaa3 is being used to indicate that speaker A should have known about whale-watching. More applicable to (5.1) would be Yau’s (1965, 66) categorisation of gaa3 as a particle which does not demand a verbal confirmation.

(5.2)

A: ‘Hey did you know Jenny ne1, besides ALO [Assistant Labour Officer] ne1, also got an offer for EO [Executive Officer]. Did I tell you aa3?’

B: ‘No aak3.’

A: Mou5 aa4? Go2 jat6 aa3 - aa3
not-have PRT that day name-prefix name-prefix
Lily gong2 bei2 ngo5 teng1 gaa3.
Lily say give me listen PRT
‘No aa4? That day Lily told me gaa3.’

Example (5.2) shows another typical example of gaa3. This conversation also features in Chapter 4. Speaker A uses gaa3 at the end of her utterance to express ‘it is good if you know this’, perhaps so that B does not mistake the information about Jenny to be very reliable, or perhaps just because it would be interesting since Lily is a mutual friend. Observe that descriptions of gaa3 such as ‘conceited’ or ‘seeking confirmation’ are inaccurate. There is also no reason to believe that the information is presumed knowledge, as Fung (2000, 136-138) believes that all g- particles indicate, since speaker A explicitly asks whether B knew, and B replies ‘no’.

Examples like (5.2) might have been explicated as ‘it is good if I say something about this’, since it seems that speaker A may be trying to be clear about how reliable her information is. However, ‘it is good if I say something
about this’ was not a meaning found to be shared in all examples of *gaa3*. For example, consider next a different usage of *gaa3* in (5.3). This example shows two people, A and B, talking about A’s exam results. It includes three occurrences of *gaa3* produced by A, where he is bragging about his exam results. In this context, it seemed ‘it is good if I say something about this’ was not as appropriate as the current proposed explication. A speaker who brags could plausibly attach ‘it is good if you know this’ to their utterances, since presumably they would think it good that the hearer knows of their accomplishments.

(5.3)

B:  
Wai3 jing1man2 dou1 haau2 B aa4?  
hey English also/even examine B PRT

‘[You] even got a B in English aa4?’

A:  
B saam1 *gaa3 (5.3a)* ngo5 di1 jing1man2.  
B three PRT I CL English

Loeng5 go3 - ngo5 jau5 loeng5 zoeng1 paper  
two CL I have two CL paper

A *gaa3. (5.3b)*  
A PRT

‘I got a B3 [the better subdivision of B] *gaa3*, in English. Two – I got As in two papers *gaa3*.’

B:  
Gam3 sai1lei6? Bin1 zoeng1 aa3?  
so/that impressive which CL PRT

‘So impressive / that good? Which [papers] aa3?’

A:  
Ngo5 zok3man2 tung4maai4 oral dou1 hai6  
I writing and oral also/all is

A *gaa3. (5.3c)*  
A PRT

‘I got As in both writing and oral *gaa3*.’

The use of *gaa3* in example (5.3) may have the connotation of ‘conceited’, as Yau (1965) described. Speaker A seems very proud of his English results, and
speaker B is obviously very impressed by A’s results too. In examples (5.3a) and (5.3b), speaker A is not just answering B’s questions, but also giving additional information that makes himself sound even better. The uses of gaa3 may be highlighting or drawing attention to this ‘noteworthy’ information which he believes it is good if B knows. In (5.3c), he again uses the particle gaa3 to help highlight what he is saying, which might have sounded less conceited or proud had he omitted gaa3, ‘it is good if you know this’.

The explication ‘it is good if you know this’ does not itself imply conceitedness, but is compatible with situations where a speaker is being conceited. As we have seen, ‘conceited’ is not an invariant meaning of gaa3, and so should not be included in the explication (examples of gaa3 being used when the speaker is being unsure or being humble or modest are shown further below). These utterances would likely sound somewhat conceited even without gaa3 attached at the end. Speaker A’s tone of voice also indicates that he is boasting or proud and he places emphasis on his grades ‘B3’ and ‘A’.

At an earlier stage of this study, an alternative explication considered was ‘I want you to know this’. This would have been appropriate in the examples above, including providing interesting or helpful information or introducing a new topic, and especially in cases of boasting or bragging. However, ‘I want you to know this’ was rejected because sometimes the speaker does not personally or ‘actively’ want the hearer to know something.

**Gaa3 used in answering questions**

The next few examples show gaa3 being used in responses to questions. Example (5.4) shows a typical example. Speaker B questions whether a certain university has the option to study accounting. Using gaa3, speaker A expresses ‘it is good if you know this’, because B clearly wants to know, and A is answering B’s queries. The rejected component ‘I want you to know this’ would incorrectly imply that the speaker is always the one who wants the hearer to know something. This is not to say that the speaker would rather the hearer not know, but it seems that the person who most wants to know is the person who asked the question – the hearer. The current explication ‘it is good if you know this’ is appropriately neutral. This would be different from ‘it is good for me if you know this’ and ‘it is good for you if you know this’.
(5.4)
B: ‘The University of Science and Technology offers accounting me1? They only have very few...’
A: ‘Have [yes]. [They] have accounting.’
B: ‘Yes me1?’
A:
Koi5 jau5 business go2bin6 di1 je5
it/they have business that-side CL things
duk6 gaa3.
read/study PRT
‘They [the university] have courses to do with business gaa3.’

Example (5.4) again calls into question Fung’s assertion that gaa3 is used to ‘remind the hearer of situations that should be known but may have been overlooked or neglected by the hearer’ (Fung 2000, 168-171). There is no apparent reason for A to believe B should know this. This example is also unlikely to be described as ‘conceited’, nor is it seeking confirmation (Yau 1965, 112-116, Matthews and Yip 2011, 391-392).

(5.5)
M:
Nei5 hoeng1haa2 bin1syu2 aa3?
you hometown/country-home where PRT
‘Where is your hometown/country home aa3?’
E:
Ngo5 - ngo5 hoeng1haa2 hai2 ni1go3
I I hometown/country-home at this
hoi2fung1 gaa3. Zi1 m4 zi1 bin1dou6
Haifeng PRT know not know where
lai4 nei?
come PRT
‘My – my hometown/country home is this Haifeng gaa3. Do you know where it is nei?’

Example (5.5) above similarly shows that Fung’s description of gaa3 is not valid for all examples of gaa3. In this example, M asks where E’s hometown/country home is. This is something which people generally do not expect their friends to
know (in Hong Kong this often involves knowing where one’s parents/grandparents are from). When E answers, he actually asks whether M even knows where that place is. There appears to be no expectation that the hearer, M, should have known at all. E’s utterance does seem to show that he expects a verbal response, but this should be attributed to his follow-up question ‘do you know where it is?’ Again, the rejected component ‘I want you to know this’ would imply that the speaker, E, is the one who wants the hearer, M, to know where his hometown is. In this excerpt it is M who asks E, and it seems obvious that it is M who wants to know. The proposed explication is valid again because E is answering M’s question and is expressing ‘it is good if you know this’.

(5.6)
A:  
Gam3 gan2jiu3 aa4 sap6ji6jyut6.
so serious PRT December

‘So serious aa4 in December?’

B:  
Hai6 aa3, sei3sap6gei2 dou6 gaa3 (5.6a) dou1.
is/yes PRT forty-something degree PRT even

‘Yes aa3, forty-something degrees gaa3.’

A:  
G02 zan6 si4 teng1 jan4 gong2 ne1,
that CL time listen people talk/say PRT
waa6 ou3zau1 go2 dou6 ne1, hou2
talk/say Australia that place PRT very
gwa12 do1 wui1j1ing1 gaa3 w03. (9.7c)
devil/ghost (vulgar) much/many flies PRT PRT

Hai6 m4 hai6 aa3.
is not is PRT

‘Some time ago I heard people say ne1, say that in Australia ne1, very “devil” many flies gaa3-w03. Right aa3?’
B:

\[ \text{Hai6 aa3. Keo15 waa6 hou2 daai6 zek3} \]

is/yes PRT he say very big CL

**gaa3. (5.6b)**

PRT

‘Yes aa3. He [my husband] says they’re very big gaa3.’

Examples (5.6a) and (5.6b) above are also responses to questions. It is good if the addressee (speaker A in both these cases) knows that it can be forty-something degrees and that speaker B’s husband says the flies are very big, because the addressee has indicated that they want to know. This, of course, does not mean it is necessarily good that Australia experiences forty-something degree summers with lots of very big flies; it is just good that speaker B is answering A’s questions and providing helpful/relevant information. The particle gaa3 ‘it is good if you know this’ is valid for examples like (5.6a) and (5.6b), and can make the speaker seem to be more convincing or persuasive. (The use of gaa3-wo3 from speaker A above will be discussed in Chapter 9.)

The attitude of speaker B in example (5.6) could be compatible with Fung’s (2000, 168-171) claim that gaa3 ‘reminds’ the hearer of things that should be known but which may have been overlooked or neglected by the hearer. While ‘reminding’ is not exactly right, it is possible that speaker B in this excerpt has judged the situation in Australia as something which should have been known, but which speaker A seems to have no (or not enough) knowledge of. Yau’s finding that gaa3 is obviously affirming and does not demand a confirmation is also endorsed by these examples, although ‘conceited’ still seems an unusual description. Notice also that in situations like example (5.6b), where a speaker is responding to a question about flies, there is no need for ‘softening’ or to sound ‘less abrupt’, as Matthews and Yip (2011, 391-392) and Kwok (1984, 45, 71) state that aa3 does. Again, though, the current proposed explication ‘it is good if you know this’ makes sense in these examples.

Since most examples of gaa3 show it being used with the statement of very concrete, factual information, this led to the consideration of a possible component ‘this is true’. In the examples above, the speaker generally seems quite sure of what they are saying, but it was considered that even if the statement was not always completely true, or if the speaker was not always
completely sure, the speaker still seems to want to come across as being very
definite and sure. Ultimately, this explication was rejected. Particles generally
seem to express something about the speaker’s feeling or attitude, and it was not
clear why a speaker would constantly strengthen their statements in this way. It
was considered that this could be a cultural tendency for Cantonese speakers to
want to sound very sure about things, but this still did not make sense for all
examples. Furthermore, some examples exist where the speaker is apparently
unsure of something but still uses the particle gaa3.

**Can the speaker be unsure?**

In example (5.7) below, two speakers are discussing probations. Speaker A is
unsure about her own probation, saying ‘I don’t know’ and ‘I think’, but uses
gaa3 in the utterance anyway. The speaker A in example (5.8) also seems
unsure, using hou2ci5 ‘seems like’, as B asks him a question about school
examinations in previous years. This shows that gaa3 does not necessarily have
to be used in situations where the speaker is (or is trying to appear) very certain
about what they are saying. The proposed explication is still applicable. It is
also possible to say, for example, ho2lang4 hai6 gaa3 ‘maybe gaa3’ or m4 zit
pe4 gaa3 ‘don’t know gaa3’ (see also examples (5.11) and (5.12)).

(5.7)

J:

```
Nei5 ne1? Nei5 hai6 gei2 go3 jyut6 aa3?
```

‘You ne1? How many months is it [the probation] for you aa3?’

A:

```
Ngo5 - ngo5 m4 zit jau5mou5 probation,
I I not know have-not-have probation
```

```
ngo5 nam2 ngo5 mou5 probation gaa3.
I think I not-have probation PRT
```

‘I – I don’t know if I have probation, I think I don’t have probation gaa3.’

(5.8)

A: ‘The Certificate of Education Examination also/only counts your best
six subjects ze1.’

B: ‘Yes me1?’
A: ‘Yes.’

B:
\[\text{Gam}2 \ \text{ngo}5 \ \text{go}2\text{zan}6\text{si}4 \ \text{ne}1?\]
then \ I \ \text{that-time} \ \text{PRT}

‘What about in my time \text{ne}1?’

A:
\[\text{Nei}5 \ \text{hou}2\text{ci}5 \ \text{dou}1 \ \text{hai}6 \ \text{gaa}3.\]
you \ \text{seem-like/probably} \ \text{also} \ \text{is} \ \text{PRT}

‘I think you were the same \text{gaa}3.’

In example (5.9), A and B are discussing dictionaries. A does not know which is best and B is recommending some to her. B says here that one of the dictionaries is that of Lu Gusun, and that A has probably used it. It seems that this is just a guess, though. This use of \text{gaa}3 fits Fung’s report that \text{gaa}3 assumes the hearer has no knowledge of a situation that should have been known. This is indeed one situation in which \text{gaa}3 can occur, but it is not invariant. The proposed explication ‘it is good if you know this’ is deliberately broader and more widely-applicable.

(5.9)

B:
\[\text{Ling}6\text{ngo}i6 \ \text{jat}1 \ \text{bun}2 \ \text{ne}1 \ \text{zau}6 \ \text{hai}6 \ \text{luk}6\text{guk}1\text{syun}1\]
other \ one \ \text{CL} \ \text{PRT} \ \text{then} \ \text{is} \ \text{personal-name}
\text{ge}3. \ \text{Ngo}5 \ \text{nam}2 \ \text{nei}5 \ \text{jing}1\text{goi}1 \ \text{jau}5 \ \text{jung}6 \ \text{gaa}3.\]
POSS \ I \ \text{think} \ \text{you should have} \ \text{use} \ \text{PRT}

Another one [dictionary] \text{ne}1 \ is \ \text{Luk}6\text{Guk}1\text{Syun}1’s \ [\text{Lu Gusun’s}]. \ I \ \text{think} \ \text{you should have used}/\text{are using} \ \text{it}] \text{gaa}3.’

It was also previously considered that the explication for \text{gaa}3 could be about thinking, and that it would use the NSM prime \text{THINK} rather than \text{KNOW}. Another alternative explication considered was ‘I think you didn’t think about this before’. This explication reflects the fact that \text{gaa}3 is often used when giving someone new or unexpected information, as already seen in some examples above. Examples which this explication would not fit well include example (5.10). Given that (5.10) is a response to a fairly specific, closed-ended query, it is unlikely that speaker X thought speaker Y had not thought about it before. Similar explications such as ‘I think you haven’t thought about this
before’, ‘I think you don’t know this’, and ‘I think you didn’t know this before’ were also considered.

(5.10)

Y:

\[\begin{align*}
\text{Gam2} & \quad \text{nei}5 & \quad \text{mai}6 & \quad \text{jiu}3 & \quad \text{tai}2 \\
\text{so/then} & \quad \text{you} & \quad \text{then} & \quad \text{have-to} & \quad \text{read} \\
\text{hou}2\text{do1} & \quad \text{jik6bun}2 & \quad \text{lo}1.
\end{align*}\]

very-many translated-texts PRT

‘So/then you have to read very many translated texts lo1.’

X:

\[\begin{align*}
\text{M6.} & \quad \text{Dou1} & \quad \text{jing1goi1} & \quad \text{hai}6 & \quad \text{gaa3}. \\
\text{mm} & \quad \text{also} & \quad \text{should} & \quad \text{is} & \quad \text{PRT}
\end{align*}\]

‘Mm. It should be like this [supposedly] gaa3.’

An NSM component that would seem to address the issue highlighted by the rejected component ‘I think you didn’t think about this before’ is ‘I know you want to know something like this’. However, this kind of component is clearly not a good solution. It would be too specific and would not cover the wide range of uses of gaa3. This explication may seem acceptable for some instances of gaa3, particularly responses to questions like examples (5.3c), (5.4), and (5.6b), but it does not explain examples like (5.1a–d). Not all declarative uses of gaa3 are responses to questions.

Other earlier think explications that were considered and ultimately rejected include ‘I want you to think about this’ and ‘it is good if you think about this’. These seem simple enough for the varied uses of gaa3, and would be compatible with many examples. (5.6a) and (5.6b), for example, would seem to be able to be explained by ‘I want you to think about this’ since the speaker is providing new information. ‘It is good if you think about this’ seemed plausible too, since the speaker who uses gaa3 is likely to think that the response will be helpful/interesting/good for the person asking the question.

Even so, examples (5.11) and (5.12) clearly invalidate the possibility of using explications like ‘I want you to think about this’ or ‘it is good if you think about this’. It is unlikely that a speaker would want to express ideas like ‘I want you to think about the fact that I don’t know’ or ‘it is good if you think about me not knowing’. Thinking about something implies that someone dwells on
something for a relatively longer period of time than would be plausible in such situations. Since examples (5.11) and (5.12) are acceptable and were found in the corpus, it was concluded that formulations with THINK were unlikely to provide an invariant meaning of gaa3. The utterances in these examples act almost like disclaimers and the current explication ‘it is good if you know this [i.e. that I don’t know]’ seems to work better.

(5.11)
A: 
\[Heoi3\ gwo3\ gei2\ ci3\ zek1,\]
go \ EXP \ how-many \ time \ PRT
\[nei5\ lou5gung1\ Orlando\ go2dou6?\]
you \ husband \ Orlando \ that-place

‘How many times has your husband been to Orlando zek1?’

B: 
\[Ngo5\ m4\ zii\ gaa3,\]
I \ not \ know \ PRT
\[gam2\ peng4\ mai6\ heoi3\ lo1.\]
so/like-that \ cheap \ so-then \ go \ PRT

‘I don’t know gaa3, [he] just goes if it’s cheap lo1.’

(5.12)
A: 
\[Gei2si4\ dou2\ jau5dak1\ tai2\ zek1\]
what-time \ approximately \ can/able-to-have \ see \ PRT
\[gam2\ niidi1,\ zii\ m4\ zii\ daai6koi3?\]
so/this \ these \ know \ not \ know \ approximately/roughly

‘Approximately when can we see these [whales] zek1, do you know approximately?’

B: 
\[Ngo5\ m4\ zii\ gaa3.\]
I \ not \ know \ PRT.

‘I don’t know gaa3.’

Although the current proposed explication includes the prime GOOD, the explication does not require that the thing being spoken about is inherently good. Examples (5.11) and (5.12) both involve the speaker answering a question by saying that they do not know something. The assumption is that the person asking the question would evaluate it as good to know the answer to their
question (in these cases, the answer is that they do not know). Examples (5.11) and (5.12) further support the rejection of ‘I want you to know this’, discussed above. These are also examples of *gaa3* that challenge descriptions like ‘conceited’ (Yau 1965, 115), ‘reminding’ (Fung 2000, 170), or ‘seeking confirmation’ (Matthews and Yip 2011, 391-392).

**Gaa3 in ‘warnings’**

The current proposed explication ‘it is good if you know this’ is better than ‘it can be good if you know this’, because the latter makes the speaker seem unsure. For example, when *gaa3* is used in responses to questions, where someone has indicated explicitly that they want to know something, it makes sense to state that it is good for the person asking the question to know the answer. Another revealing situation in which *gaa3* is often used is when teaching children. The constructed examples labelled (5.13a–c) help to demonstrate this (no children were recorded in the corpus). These examples show some typical statements often directed at children, and support the rejection of ‘it can be good if you know this’ in favour of ‘it is good if you know this’.

(5.13)

(5.13a)

*M4* *ho2ji5 gam2joeng2 gaa3.*

not can like-this PRT

‘(You) can’t be like this *gaa3’ / ‘(You) can’t do that *gaa3.*’

(5.13b)

*Gam2joeng2 zou6 hai6 m4 ngaam1 gaa3.*

like-this do is not correct/right PRT

‘Doing this / behaving like this is wrong *gaa3.*’

(5.13c)

*Go2dou6 hou2 ngai4him2 gaa3.*

that-place very dangerous PRT

‘That place is very dangerous *gaa3.*’

We can reconsider some previous claims about *gaa3* here. These utterances do not seek confirmation, although they may or may not expect a verbal response. Fung (2000, 168-171) stated that sentences suffixed with *gaa3* ‘remind the hearer of situations that should be known but may have been overlooked or
neglected by the hearer... *gaa3* assumes that the hearer has no knowledge of a situation that should have been known and is a given (as opposed to a new) situation’. Whether or not Fung’s report is substantiated by utterances such as (5.13a–c) depends again on whether or not the child (hearer) is expected by the speaker to already know certain things. It could be that the speaker is informing the child for the first time, but it could also be that the child has been told before, but has ‘disobeyed’ and the speaker thus deems reminding to be necessary. In any case, we can see that Fung’s claims would not be applicable to all statements, and therefore there is a limit to its helpfulness.

Substitution of other utterance particles such as *laa1* or *wo3* would make examples (5.13a–c) sound very strange. The particle *laa1* includes the component ‘you now know how I think about this’, which is unreasonable to expect of young children. Further, the component ‘I can not-say more’ explains why the speaker may sound impatient or as if s/he does not want to talk about it. These utterances, with *laa1* substituted, probably sound unacceptable to most Cantonese speakers. The particle *wo3* has the meaning ‘you aren’t thinking about this at this moment, it is good if you think about it’. In these utterances it would imply that the child was not thinking or was not aware that the behaviour was wrong, and that the child should reflect on this and think about what they have done / are doing. Interestingly, the utterances would sound less odd if the combination *gaa3 wo3* were used (rather than *wo3* alone). This may be because the speaker in these cases would most likely want to express ‘it is good if you know this’, regardless of whether *wo3* were used (see also Chapter 9).

On one hand, the utterances in (5.13) may seem ‘softer’ than the same utterances without *gaa3* attached. The proposed explication ‘it is good if you know this’ may help explain why utterances with *gaa3* can sometimes seem to be ‘softened’ (Matthews and Yip 2011, 391) and ‘sound less curt and abrupt’ (Kwok 1984, 71). Utterances without the particle *gaa3* might sound as if the speaker does not want the addressee to know things, which can make it seem like the speaker is frustrated, annoyed, or impatient. On the other hand, the meaning ‘it is good if you know this’ can be used to help convince/persuade or even threaten/warn. What seems to be more influential on the level of ‘softness’ of the utterances in (5.13) is how the utterance is delivered, which can affect the
utterance drastically even if the particle remains in place. An angry parent shouting these utterances would have a very different effect from someone who was calmly teaching. *Gaa3* is not necessarily ‘soft’ in all cases, and ‘softening’ or ‘less curt’ are not part of *gaa3*’s invariant meaning. The explication reflects and allows for both situations.

(5.14)

\[
\begin{align*}
\text{Cin1kei4} & \quad m4hou2 & \quad zing2 & \quad go3 & \quad daan6gou1 \\
\text{under-no-circumstances} & \quad \text{do-not} & \quad \text{make} & \quad \text{CL} & \quad \text{cake} \\
\text{ceot1lai4} & \quad ji3 & \quad ng05 & \quad gaap3 & \quad \text{cin2} & \quad \text{gam2joeng2}. \\
\text{come-out} & \quad \text{need} & \quad I & \quad \text{share} & \quad \text{money} & \quad \text{like-that} \\
\text{Ngo5} & \quad \text{sat6} & \quad \text{sei2} & \quad \text{dou1} & \quad m4 & \quad \text{gaap3} & \quad \text{gaa3}. \\
\text{I} & \quad \text{definitely} & \quad \text{die} & \quad \text{also/even} & \quad \text{not} & \quad \text{share} & \quad \text{PRT} \\
\text{Ngo5} & \quad \text{gong2ming4} & \quad \text{aa3}. \\
\text{I} & \quad \text{say-clearly} & \quad \text{PRT} \\
\end{align*}
\]

‘Don’t get a cake under any circumstances and make me share money [contribute money towards it]. I’ll definitely not share/contribute even if I die *gaa3*. I’m telling you aa3.’

Example (5.14) from the corpus shows another kind of warning, which is almost a (‘friendly’) threat. Two adults are making fun of their friend, whom they are a bit fed up with. One speaker produces this utterance and they laugh. This ‘warning’ is given in a half-joking way. The proposed explication ‘it is good if you know this’ is substitutable in place of *gaa3* here and helps explain what the utterance means. *Gaa3* is being used here to emphasise that she will definitely not put money towards the cake, and it is good if her friend knows this.

**‘Humble’ uses of *gaa3***

In contrast to examples of boasting like in (5.3), *gaa3* can also be used when being more ‘humble’ or ‘modest’. In example (5.15), speaker E is making a point about not having moved house. We can tell, especially by E’s tone of voice, that this is presented as a good thing. At the same time, K and E are saying they are not as rich as their friend Lily, who has moved many times recently. This could be perceived as negative, but it appears that K and E are actually saying that Lily is weird and unusual for moving so often. Therefore it might be able to be said that this example shows conceitedness in that the speakers are proud of not
having moved. Observe that an example such as (5.15) does not have an ‘effect of seeking confirmation’.

(5.15)

K:

Bin1 hok6 ci5 Lily gam3 jau5 cin2

who/how learn like Lily that-way have-money

seng4 jat6 ho2 ji5 bun1 uk1 gaa3 aa1?

always can move house PRT PRT

‘Who would be rich like Lily to always move house gaa3-aa1?’

E:

Mai6 hai6 lo1. Ngo5 zyu6 zo2 sap6 gei2

so/then is PRT I live PFV ten something

jaa6 nin4 dou1 hai6 zyu6 niidou6 gaa3.

twenty year all/even is live here PRT

Bin1 dou6 ci5 nei5 go2 di1...

where like you those

‘Yes/exactly lo1.’ I’ve lived all ten or twenty-something years here gaa3.
Not like you...’

Example (5.16) shows again that ‘conceited’ is not a good description. In this example, speaker J is complimenting speaker A, when A ‘rejects’ this compliment. Her response conveys surprise and that her friend is crazy to think she is that good. Her use of gaa3 expresses something like ‘it is good if you know that this is crazy / that I will get a job so quickly’.

(5.16)

J:

... nei5 niidou gam3 jau5 sat6 lik6, m4 -

you these so/that have competence/strength not

m4 sai2 geng1, jatt g03 lai5 bai3

not-need scared one CL week

ho2 ji5 wan2 dou2 gung1.

can find arrive job

‘You are so competent, don’t – don’t need to be scared, in one week you’ll be able to find a job.’
Example (5.17) below shows another typical example of gaa3 in use. In this example, speaker A is asking about J’s work. J says that there are no other ‘fresh grads’ (people who have recently graduated) in her workplace, and that it is quite lonely (although she does not sound very embarrassed). She attaches gaa3 to her utterance. J seems to be saying ‘it is good if you know this’ because A has just asked her about it directly, and she is making conversation.

(5.17)
A:
Kau5 mou5 di1 tung4 nei5 jat1 cai4
have-not-have some with you together
hai6 fresh grad go2 di1 jan4 aa3?
is fresh grad those people PRT
‘Are there any people who are also fresh grads aa3?’

J:
Mou5 aa3, dak1 ngo5 jat1 go3,
not-have PRT have/only I one CL
dou1 gei2 gui3 duk6 gaa3.
also/even quite lonely PRT
‘No aa3, only me, it’s quite lonely gaa3.’

Being lonely is not a trait associated with being conceited, and this example is very different from those like (5.3), where a speaker is boasting and could be described as conceited, as reported by Yau (1965, 112-116). The two examples are also different in that the thing being spoken about is not GOOD. That is, doing well in exams would be considered a good thing, whereas being lonely is generally considered not to be a good thing. Recall that the proposed explication indicates only that it is good if the hearer knows something, not that the thing itself is inherently good. From the corpus data, it seems safe to say that in example (5.17), there is no expectation that A already knows anything about it. This again contradicts Fung’s (2000, 168-171) claims that when gaa3 is used, the hearer is expected to have been able to know something about the
situation. She is also not seeking confirmation, as claimed by Matthews and Yip (2011, 391-392).

Fung’s observations are not completely invalid, of course, and the next two examples from the corpus support her view. For instance, in example (5.18), it is possible that the speaker is trying to remind the hearer of a situation that should be known but may have been overlooked or neglected, as Fung (2000, 168-171) states. Here, speaker A gives a somewhat shaky defense about her relationship with their mutual friend. Speaker B comments that their relationship is not good, and A states that their relationship used to be very good. It appears that B is aware of this, as he states in the third turn of this excerpt that that was a very long time ago. Hence it could be that A is trying to remind B of how things were in the past. The explication ‘it is good if you know this’ makes sense and can be substituted into each of the two utterances. It helps explains why A uses gaa3 to ‘defend’ herself and help convince B that they used to have a very good relationship.

(5.18)

B:

`Nei5 loeng5 go3 gwaan1hai6 gam3 m4 hou2 gam2.`
you two CL relationship so not good like-that

‘You two have such a bad relationship like that.’

A:

`Ngo5dei6 ji5cin4 di1 gwaan1hai6 hou2 hou2`
we before CL relationship very good

`gaa3. (5.18a)`

PRT

‘We used to have a very good relationship gaa3.’

B: ‘Before? Wow! [That was] a very long time ago laa3-wo3. How many years ago?”

A: ‘Three years.’

B:

`Saam1 nin4 cin4 laat!`
three year before PRT

‘Three years ago laat!’
As a final point, consider that, according to the explication, *gaa3* would not be used in situations where the thing being spoken about is not good for the hearer to know. This seems to hold true. A constructed sentence like ‘I don’t want to tell you this now *gaa3*’ would be very strange in Cantonese (unless we imagine a complicated and unusual scenario, where it is good for the hearer to know that the speaker does not want the hearer to know something). *Gaa3* also seems unlikely to be used for breaking bad news to someone who will feel very bad, such as that someone they are close to has just died, is very sick, or has had a terrible accident. In these cases, the hearer would probably not find this news good to know.

This section has overviewed *gaa3* as used in statements, and has discussed many diverse examples found in the corpus. The data appears to support the proposed explication, and furthermore, shows that it is much more comprehensive than previous descriptions of *gaa3*, despite being short. This section also reviewed alternative explications which were considered at earlier points in the research process, and has shown that the current proposed explication is the most suitable definition of *gaa3*. The next section discusses *gaa3*’s possible polysemous meanings.

### 5.3 Possible polysemy: *gaa3* in exclamations and questions

As explained in Chapter 1, the initial assumption when seeking an NSM explication for a particle (as well as other words) is that there is one meaning that explains all uses of the particle. If identifying a single meaning that can apply to all the uses is found not to be possible, only then should polysemy be considered. Of course, polysemy is a linguistic fact and needs to be posited for many words at some point, but if multiple meanings were sought at the onset, this would open the way for numerous varied and opposing definitions before the search for the more economical solution was exhausted.
The present investigation has found that ‘it is good if you know this’ is able to cover the various uses of $gaa3$ in statements, but in many instances of $gaa3$ in questions, there are more complications (see section 5.3.2). There is also an interesting example of $gaa3$ being used in exclamations (section 5.3.1). It would be difficult to make the explication broader than it currently is, while still ensuring that the explication adds value, since it is already intentionally broad and simple. This indicates that the particle $gaa3$ may be polysemous. If so, this is a significant finding because polysemy has never been posited for $gaa3$ before. If there were two or even three meanings for $gaa3$, and if they could be explicated, the explications would still provide a much more comprehensive and probably more accurate account of the semantics of $gaa3$ than can currently be found in the small handful of works which mention it. More serious analysis of the potentially polysemous meanings of $gaa3$ is not feasible in this chapter due to the small number of relevant examples in the corpus, but a brief overview is given below (some examples are constructed). The other particles analysed in this thesis do not appear to be polysemous; $gaa3$ may have the broadest range of use, supported by the fact that it is the most frequently occurring of those studied in this thesis.

Importantly, polysemy in $gaa3$ could potentially have substantial implications for the particle $aa3$, since $aa3$ is the only particle in the Hong Kong Cantonese Corpus which has a higher frequency than $gaa3$. $aa3$ is therefore presumably the utterance particle with the broadest meaning, or largest number of polysemous meanings. In fact, $aa3$ occurs in the corpus double the amount of times that $gaa3$ does (3674 times, compared with 1832 times). $Gaa3$'s frequency is actually much closer to the next most frequently occurring particle, $laa1$ (occurs 1578 times). As described in Chapter 3, $laa1$ appears to have an invariant meaning which remains stable in all observed contexts.

5.3.1 $Gaa3$ in exclamations of ‘disbelief’

Sometimes, expressions which seem to show some kind of disbelief or shock can use $gaa3$. For example, ‘$gam2\ dou1\ jau6\ gaa3$’ [like-this even have PRT] or ‘$gam2\ dim2\ dak1\ gaa3$’ [like-this how can PRT] both roughly mean ‘how can it be like this $gaa3$’. They are typically used rhetorically, where someone may be showing annoyance, anger, or disbelief, and in these situations $gaa3$ may be
behaving differently to the examples considered above. One similar example found in the Hong Kong Cantonese Corpus is shown below. Out of five days of A’s holiday, she only had two days to ‘play’ (i.e. for fun activities). J is expressing shock or annoyance, or empathising with her friend. This use of gaa3 appears not to have been described before.

(5.19)
A:

Gei1bun2soeng6 ne1 waan2 ne1 hai6 dak1
basically PRT play PRT is only
loeng5 jat6 gam3 do1 zaa3.
two day like-this many PRT
‘Basically ne1 play [/do things] ne1 was only for two days zaa3.’

J:

Ce2! Gam2 gaa3.
exclamation like-this PRT
‘Ce2! Like that gaa3.’

It is possible that gam2 gaa3 [like-this PRT] is like a fixed expression and could be an abbreviated form of gam2 dou1 jau6 gaa3 / gam2 dim2 dak1 gaa3, roughly meaning ‘how can it be like this gaa3’, as mentioned above. This may be like ‘how can?’ in Singapore English, which may be taken to mean roughly ‘how can it be like this?’ It is also possible that utterances like gam2 gaa3 may be acting like ‘discursive’ uses of interjections (see e.g. Goddard 2014b) which can refer to things ‘out of context’. For instance, consider a situation where something disgusting happens to someone, who is called A. If another person, B, is told about what happened to A, B can respond with ‘yuck’ (or similar) even though it did not happen to B. The situation could even be imaginary or hypothetical. Note that the exclamation Ce2! in J’s utterance, which expresses bad feelings, seems to be being used discursively (cf. (5.16), which features the exclamation Waa3!). Taking all this into account, it seems possible that J in example (5.19) is expressing, roughly, ‘how can it be like this?!’ as if the thing

62 For the discourse of can in Singapore English, see Wong (2014, 139-179).
63 Discursive uses of interjections were found by Stange (2016, 123-135) to be very common in the speech of adult British English speakers, accounting for 20% and 23% of uses of ugh! and yuck! respectively.
that happened to A (i.e. only having two days for activities) is a bad thing that happened to J herself.

Gaa3-suffixed utterances which behave like this appear to be relatively limited. It is possible that gaa3 in these expressions retains its meaning ‘it is good if you know this’, perhaps as the speaker wants to express shock, annoyance, anger, or other feelings which the situation calls for, and wants the hearer to understand what they are feeling or to understand the situation. The speaker may be doing something like complaining or showing dissatisfaction. Alternatively, this use of gaa3 may be polysemous. Since this seems to depart from the usual use of gaa3, and far fewer examples of it were documented in the corpus, this is not pursued further.

5.3.2 Gaa3 in questions

Gaa3 can be used in questions, either as an individual particle, or as part of a particle cluster. One example of such a particle cluster is gaa3-me1. Me1 can attach to several different particles, including also laa3 to form laa3-me1, and zaa3 to form zaa3-me1. The particle me1 seems to turn statements into questions, and seems to be used with questions only. (Questions are also possible using the particle gaa4⁶⁴; for discussion of particle ‘families’ and the sub-syllabic morphemes hypothesis, see Chapter 8.)

In the case of gaa3, the distinction between statements and questions is fairly clear. This is different from wo3. As noted in Chapter 4, wo3-suffixed utterances are generally statements, although some are followed immediately by questions, which can sometimes lead those wo3-suffixed utterances to seem like questions. This does not appear to be as common with gaa3, and the gaa3-suffixed utterances are generally more distinct between statements and questions.

Since the starting point of this research was that there could be one invariant meaning identified for each utterance particle, at an earlier point in the analysis, it was thought that one explication of gaa3 could accommodate for gaa3 in both statements and questions. The use of the prime SOMEONE was

⁶⁴ It seems possible to the author that gaa4-suffixed questions generally invite more gaa3-suffixed responses than other kinds of questions.
considered in place of YOU, to create the explication ‘it is good if someone knows this’. This would have allowed the meaning to remain the same while being able to be directed at different people (either the speaker or the addressee). Ultimately, this possibility was rejected because a component using SOMEONE allows the interpretation that this SOMEONE is a third party other than the speaker and addressee(s).

It was also considered that the proposed explication ‘it is good if you know this’ could remain unchanged, and when gaa3 was used in a question, ‘this’ would refer to the answer to the question. In other words, the speaker could be expressing something like ‘it is good if you know the answer to my question’. Consider example (5.20), which is from the whale watching conversation which also gave the excerpt in (5.1). In both (5.20a) and (5.20b), speaker A is worried and unconvinced about going out to sea to see whales. Speaker A’s questions can be interpreted either literally or rhetorically; this did not seem to matter, if gaa3 expressed something like ‘it is good if you know this [the answer to my question]’. Note that these examples of gaa3 suffixed to questions also cause problems for the descriptions of gaa3 by those like Yau (1965, 66) and Fung (2000, 136-138).

(5.20)

B:
... Jan1wai6 mei6 gin3 gwo3 gam3 daai6
because not-yet see EXP so/that big
tiu4 king4jyu4 aa1 maa3.
CL whale PRT PRT
‘[I want to see whales] ... Because [I] haven’t seen such a big whale aa1-maa3.’

A:
Ngaau5 m4 ngaau5 jan4 gaa3? (5.20a)
bite not bite people PRT
‘Do [the whales] bite gaa3?’

B:
Gam2 ngo5 nam2 ngo5 nam2 zik1hai6
Well/then I think I think meaning
‘Well I think, I think, I mean you observe from a distance gaa3-aa1-maa3.’

A:

Jau5mou5 ngai4him2 gaa3? (5.20b) Daai6lou2.

have-not-have danger PRT big-man

‘Is there danger / is it dangerous gaa3? “Dude”.’

Similar interpretations of gaa3 as indicating ‘it is good if you know this [the answer to my question]’ also seemed plausible for (5.21), where a speaker asks her friend whether it is cheaper to shop overseas, and (5.22), where the speaker asks whether the hearer knows a person at their school.

(5.21)

Hai6 m4 hai6 dai2 hou2
is not is cheap/good-value very

d01 gaa3 go2 bin1? Maai5 je5.
much PRT that side buy thing

‘Is it much cheaper/better value gaa3 over there? Shopping.’

(5.22)

Nei5 sik1 m4 sik1 keoi5 gaa3?
you know not know s/he PRT

‘Do you know him/her gaa3?’

As these are yes/no questions in the form ‘A-not-A’, it may be useful to consider the semantics of yes/no questions, which have been decomposed by Goddard (2003) using NSM. Based on his analysis and explication, it seems that the question structure has meaning and that it may be possible for (5.22), as an example, to be explicated as in [5B] below. Components (a-e) follow Goddard’s (2003, 4) explication and represent the meaning of the yes/no question without an utterance particle attached. Component (f) is the meaning of gaa3 as explicated in [5A]. Component (c) is specific to example (5.22). For different questions, component (c) can be substituted with relevant information. Note that [5B] is suggested without detailed investigation and based on the examples presented in this section only; a more complete and in-depth analysis is not possible here.
Tentative explication for yes/no questions suffixed with *gaa*3, using (5.22) as an example

Do you know him/her *gaa*3?

a. I want to know something
b. I think maybe it’s like this:
c. you know him/her
d. I don’t know
e. I want you to say
f. it is good if you know this

A more puzzling aspect of *gaa*3 as used in questions presents itself when those questions are suffixed with the particle *aa*3 instead (*aa*3 can also be used in both statements and questions). For example, the differences between the constructed questions in (5.23) are not explained by the explication [5A].65 Ideally, the explication for *gaa*3, perhaps along with an explication for *aa*3, would be able to help determine why these differences occur.

(5.23)

(5.23a)

*Lei5 zou6 me1 aa3?*
you do what PRT
‘What/how are you doing?’

(5.23b)

*Lei5 zou6 me1 gaa3?*
you do what PRT
‘What do you do (for a living) / what is your role?’

It may be that these differences are caused by ambiguity in the *zou6* expression. It seems to have two meanings, roughly ‘do something’ and ‘do something for a living’.

It might be the case that the first implies something relatively trivial or short-term, and the second implies something relatively important or serious. This might then affect the choice of particle, if the particle *aa*3 has a meaning more compatible with the first, and *gaa*3 has a meaning more compatible with the second. Of course, this is not yet very satisfactory without an explication of *aa*3 and consideration of more examples.

65 I am grateful to a helpful reviewer of an earlier paper for pointing out these pairs of *gaa*3- and *aa*3- suffixed questions to me.

66 *Zou6 me1 [do what]* is also used to mean ‘what happened’ or ‘what is happening’.
Substitution of $aa3$ in place of $gaa3$ does not cause this complication in all questions. Such substitution in (5.20a) in particular sounds odd, if not unacceptable, but most other examples shown above of $gaa3$ being used in questions could be replaced by $aa3$ acceptably. Some, like (5.22), even seem largely unchanged. This would seem to support the suggestion that the unusual contrast in (5.23) is due to the $zou6$ expression.

Consider, though, example (5.24), which was taken from the corpus, and is about choosing university programmes. This utterance, suffixed with $gaa3$, could refer to choices which have been made, or which are to be made in the future (from context in the corpus we know that the choices were yet to be made in this instance). Substitution for $aa3$ in this example would suggest that the speaker believes the selection has not yet happened, and therefore $aa3$ seems to be more restrictive. The constructed utterances in (5.25) are comparable, featuring the same question format and use of $gei2si4$ ‘when’.

(5.24)

O5 nei5dei6 gei2si4 gaan2 $gaa3$ JUPAS?
oh you(PL) when choose PRT JUPAS

‘Oh when do you [when will you] choose $gaa3$ JUPAS [programmes on the Joint University Programmes Admissions System]?

(5.25)

(5.25a)

$Lei5$ $gei2si4$ lei4 $aa3$?
you when come PRT

‘When will you come?

(5.25b)

$Lei5$ $gei2si4$ lei4 $gaa3$?
you when come PRT

‘When will you come / when did you come?’

It might be possible that explication [5A], proposed for $gaa3$ in statements, can explain the difference between (5.25a) and (5.25b), since ‘it is good if you know this’ implies that there is something to know already, i.e. that either the decision has already been made, or that the person has already arrived. It is possible that the meaning of $aa3$ is only compatible with a situation where the person has not yet come; this requires semantic analysis of $aa3$, as well as more examples.
For now, the meaning of *gaa3* in questions is not very clear. It may be that the apparent complications discussed here are due to ‘external’ factors in the utterances which affect the particle, and/or it may be that *gaa3* is polysemous. More detailed investigation is left for future research, which would need to look at more examples (including constructed ones, since corpora are highly unlikely to contain examples of minimal pairs) and more kinds of questions. A semantic analysis of *aa3* would also be valuable.

5.4 Concluding remarks

This chapter has proposed an NSM explication for the Cantonese utterance particle *gaa3*, which plausibly captures its invariant meaning when used in statements. Previous descriptions of *gaa3* are scarce, which is strangely at odds with *gaa3*’s high frequency in the corpus. Aside from the widespread belief that *gaa3* is a contraction, a possible reason for the relative lack of scholarly attention is that *gaa3*’s range of uses is so very broad that it may seem difficult, if not impossible, to pinpoint any meaning for it. However, it has been shown that, so far as statements are concerned, the proposed explication ‘it is good if you know this’ explains the use of *gaa3* much more consistently than previous descriptions. This deliberately brief and vague explication is compatible with a wide range of contexts and situations, but the discussion also showed that various alternative explications, many of them equally brief and vague, were not compatible with the attested range of use. Possible polysemy and the use of *gaa3* in questions were left as open issues for further research. The assumption that *gaa3* can be decomposed into *ge3 + aa3* will be discussed in Chapter 8.
Chapter 6:  
The semantics of particle *laa3*

*Laa3* is one of the most salient and frequently used particles in Cantonese. In the Hong Kong Cantonese Corpus, it is the 6th most frequently used particle, and the 33rd most frequently used word in the corpus overall. *Laa3* is used even when Cantonese speakers communicate in English; it is the 2nd most frequently used Cantonese utterance particle in Wong’s (2009, 102) data from MSN Messenger and ICQ chats. As noted in Chapter 3, non-Cantonese speakers, such as English speakers and Mandarin speakers, often imitate Cantonese speakers by adding a *laa* sound to the ends of their sentences. *Laa3* frequently occurs with other Cantonese utterance particles, such as *gaa3* and *wo3*. Out of 664 instances of *laa3* in the Hong Kong Cantonese Corpus, more than 200 of these are as part of *gaa3-laa3*. *Laa3-wo3* is also common, occurring 60 times.67

As in the preceding chapters, this chapter starts with an overview of previous descriptions of *laa3*. It then moves on to the NSM semantic analysis, which includes the proposed explication for the particle and discussion of corpus examples, existing literature, and rejected NSM components.

6.1 Previous descriptions of *laa3*

Yau (1965, 39-68) categorised *laa3* as an ‘S-type’ particle, i.e. one that affirms and does not demand a verbal confirmation, and this is supported by the corpus data. Use of the particle *laa3* does not turn a declarative statement into an interrogative one (although *laa3* is occasionally used in questions). On his ‘connotation test’, *laa3* was found to have ‘significant loading’ in none of the 12 ‘connotation concepts’68 (Yau 1965, 82-120). This could be read as suggesting that *laa3* has no meaning. However, in Yau’s own 15 hours of recorded conversation (Yau 1965, 13-25), *laa3* occurs independently 154 times, and its frequency would be even higher if its occurrence in particle clusters were

67 *Laa3-wo3* is one of only two disyllabic particles identified as ‘basic’ by Kwok (1984, 8), although her reasoning for this is unclear.


203
included in the count. If laa3 has no meaning, it is difficult to understand why participants use the particle repeatedly. It is understandable that such a large-scale test as Yau’s may have had some limitations, and it is probable that the result for laa3 reflects a problem with the test, rather than a conclusive indication that the particle has no meaning. It could be that laa3 has some slight correlation with several of Yau’s connotation concepts, but that this was not strong enough to register significantly, thereby giving an ambiguous result, and/or it is possible that laa3 is linked with connotations other than the 12 chosen by Yau. If more connotation concepts were studied, or if Yau’s test were open-ended and asked participants to choose their own concepts which they thought were most appropriate, a different result would have been likely.

According to Gibbons (1980, 770-772), laa3 is an assertion, with an assertion strength of 2 (on a scale of 1 – 3). Laa3 also indicates ‘newness’. More specifically, laa3 ‘expresses instability or change, and tends to be used together with other particles which clarify the type of change’. The second part of this statement is unclear if we take into account Gibbons’ statement that la occurs in the cluster gelabo, where both ge and bo supposedly mean that something should already be known. (He claims that gelabo expresses that a new idea is being firmly stated.)

As noted earlier, there is a tendency to try to match Cantonese particles with particles in Mandarin, and laa3 is often matched with Mandarin le. Kwok (1984, 46) explains that they have similar functions. Li (2006, 121-122) reports that the two are somewhat similar, though Cantonese laa3 has a much wider range of uses than Mandarin le. Conversely, Yip and Matthews (2000, 131) and Matthews and Yip (2011, 402) claim that laa3 is less commonly used. As explained in Chapter 1, it is highly problematic to explain the meaning of a particle in one language by reference to a particle in another language, for several important reasons. It is extremely unlikely, then, that Mandarin le is equivalent to laa3. At least, rigorous semantic analysis of le (as well as laa3) should be carried out before such a statement is made. As le is reported to be used for more grammatical functions in Mandarin, it may be that this function has been confused with its meaning in Cantonese.

Even if we accept that le is indeed equivalent to laa3, Kwok’s description of le itself could be clearer. Meanings given by Kwok for le include ‘inchoative’,
‘progress in story’, ‘completed action as of the present’, ‘consequent clause to indicate situation translatable as “then”’, and ‘isolated event in the past’. This seems somewhat contradictory, since for example, ‘progress in story’ implies that something is ongoing, whereas ‘completed action as of the present’ and ‘isolated event in the past’ suggest that something has ended. It is unclear whether laa3 refers to things in the past or in the present. Words like ‘inchoative’ are obscure even to many native speakers of English, and it is counterintuitive to describe a frequently used, everyday word via one that is less well known and understood.

Considering the Cantonese particle, Kwok (1984, 46-47) writes that laa3 ‘is related to the expression of aspect, and is used to indicate that an action (or a certain situation) is beginning, has begun, will begin, or has ended, and so on’. This appears to cover all the conceivable possibilities, and is therefore not very useful. She states that laa3 frequently co-occurs with time expressions like ‘now’, ‘at that time’, ‘soon’, and aspect markers like the perfective, experiential, or completive marker, and so on. This faces similar problems.

Yip and Matthews (2000, 131) state more specifically that laa3 occurs especially with the perfective aspect marker zo2 and with other particles expressing completion. In fact, the Cantonese zo2 is usually viewed as corresponding to Mandarin le, with both marking the perfective aspect (Matthews and Yip 2011, 226, Goddard 2005, 111-112). Yip and Matthews’ description may therefore have something to do with comparisons of laa3 and le. However, Matthews and Yip (2011, 225) state that although the Cantonese system of aspect corresponds approximately to that in Mandarin, aspect is an area where the two languages differ significantly. Furthermore, such interpretations of laa3 in relation to time and aspect may conflict with Matthews and Yip’s (2011, 402) shorter explanation of ‘to emphasise a point of current relevance’. Matthews and Yip also report that laa3 is used for ‘advice’ and place it in the category of ‘assertive’ particles (Matthews and Yip 2011, 391, 401-403).

In contrast to descriptions of laa3 and the perfective aspect, Yiu (2001) states that it is actually the Cantonese particle lei4 that conveys the perfective aspect. She states that lei4 ‘expresses past time, assertion, emphasis, etc.’ as well as ‘completion of a situation’ and ‘change from a previous state’ (Yiu 2001,
Confusingly, these descriptions have also been used for laa3 – as mentioned, Kwok (1984) reported that laa3, or rather, Mandarin le, could be used to refer to completed actions and events in the past, and Matthews and Yip (2011) reported that laa3 could be used for emphasis; additionally, Gibbons (1980, 770) reported that laa3 indicates ‘assertion’, and Yiu herself stated that laa3 can indicate change of state (Yiu 2001). Thus this provides yet another example of a widespread problem in the literature, of confusingly similar descriptions being used for multiple particles (see also Chapter 7). According to Yiu (2001), the particles laa3 and zyu6 are both associated with the imperfective and inchoative aspects.

It may be relevant that Cantonese is an isolating language with a lack of inflection (as is Mandarin), and Cantonese speakers simply need to use aspect markers regularly. Given that laa3 also occurs frequently, it may not be surprising that they often co-occur. What would be more revealing is whether laa3 co-occurs with aspect markers significantly more often than other utterance particles co-occur with aspect markers – this appears to be unknown and would need further investigation. From the corpus examples reproduced below, it does not appear that aspect markers occur especially frequently. In any case, since many English speakers and other readers may also be unfamiliar with Cantonese aspect marking, it is not the best way to explain particles.

Yiu (2001) states that the ‘core meaning’ of laa3 is a change of state. She claims that laa3 is used to mark the anticipation of the beginning of a new situation, or to emphasise a change of state, or both. The continuation of the new state is also expressed. She relies on constructed examples. Yiu’s description has some overlap with Kwok’s (1984, 46-47) statement that laa3 can relate to something possibly beginning soon, but this is probably due to Kwok’s description being too broad. Yiu’s description also has significant overlap with other particles which are not interchangeable with laa3. It has overlap with Yiu’s own description of lei4, mentioned above. The task of emphasising, as Yiu assigns to laa3, is especially common in the area of Cantonese utterance

---

69 Yiu (2001) analysed laa3 as a ‘toneless’ particle which can supposedly be pronounced with a number of different tones. She treats laa3 in the mid-level tone as the ‘basic form’ from which the other ‘variants’ such as laa1 in the high level tone are derived. For more on the sub-syllabic morphemes hypothesis, see Chapter 8.
particles, and overlaps with numerous particles. ‘Emphasis’ is not an informative description of any particle, particularly when used to describe multiple particles. Furthermore, Yiu (2001, 138) herself states that the particle lo3 also expresses a change of state. Little explanation is given for the difference between laa3 and lo3, except reports by others that lo3 gives: additional emphasis. Readers cannot know the differences between different particles.

Use of laa3 in utterances which convey a ‘change of state’ or ‘newness’ are indeed common, but this is not a rule. For example, we can construct utterances like ‘I haven’t gone [to Japan, Australia, etc.] in 10 years laa3’, or ‘I haven’t smoked for 20 years laa3’, where laa3 is still able to be used naturally in Cantonese. Laa3 is also compatible with hou2 noi6 ‘very long time’. Surely, at some point (10 or 20 years in these examples), the description ‘change of state’ is no longer applicable. If ‘change of state’ refers to the change from visiting a place to not visiting a place, or from smoking to not smoking, this could be attributed to the utterance even without the presence of laa3. Some real examples from the corpus are given further below which will show ‘change of state’ is not an ideal definition of laa3.

As can be seen, some descriptions of laa3 such as ‘emphasis’ contribute little, and references to ideas of time and aspect are varied. In addition, all these descriptions are just as frequently applied to other Cantonese particles, such as lei4, lo3, and zyu6.

6.2 NSM semantic analysis of laa3

6.2.1 The NSM explication for laa3

The proposed explication for laa3 is given below, labelled [6A]. Examples of laa3 are discussed further below, and it becomes clear that this explication is valid in a range of situations where laa3 is found to be used, including informing, teaching/instructing, answering questions, and even in jokes and teasing.
[6A] Final explication for laa3:

I want you to think now: ‘I know it’s like this’

ngo5 soeng2 lei5 ji4gaa1 lam2:
I want you now think
‘ngo5 zi1 hai6 gam2joeng2’
I know is this-way

The explication shows that when using laa3, the speaker can sound like they are telling the hearer something they should pay attention to (regardless of whether the hearer already knows). The explication may help explain why laa3 is often described as adding ‘emphasis’ (e.g. Matthews and Yip 2011, Yiu 2001). The kind of ‘emphasis’ sometimes conveyed by laa3 often indicates that the hearer should ‘note’ or ‘register’ something, that something is worth thought, or should not be overlooked. The English word emphasise usually implies that out of multiple things, one thing stands out (therefore one cannot ‘emphasise everything’) (Wierzbicka 1987, 338). The kind of ‘emphasis’ conveyed by laa3 seems to be different, as it does not necessarily mean that there are other alternatives or options which are relatively unimportant or not worth highlighting. The explication does not specify that a change has occurred or that advice is being given, but it is compatible with both. The thing that the speaker wants the hearer to know might not be the thing that has been explicitly stated, just as with laa1.

The first part of the explication in Cantonese seems like a blunt command, and therefore sounds strange. Saying either lam2 haa5 [think DEL] or the reduplicated lam2 lam2 would be an improvement, but only a marginal one. The explication is proposed to be as shown, since the English explication also sounds like a command.

6.2.2 Testing of explications and descriptions using the corpus

This section presents the semantic analysis of laa3, and discussion of more than 20 examples from the Hong Kong Cantonese Corpus. Other tested NSM explications, which were rejected, are also discussed. As in preceding chapters, the proposed explication is shown to represent the invariant meaning of laa3. The subheadings used in this section do not differentiate clear-cut or
comprehensive meanings of laa3, and many examples can fit into multiple categories.

**Laa3 used in informing**

Typical examples of laa3 can be found in situations of giving information. In the first example below, A is talking about her trip to Indonesia, in which she felt that her time was not well spent. In (6.1a), she is talking about the consequences of her late flight, and in (6.1b), she is talking about her long travelling time between two places.

(6.1)

A:

\[
\text{Gwo3 } \text{dou3 } \text{heoi3 } \text{ne1 } \text{seng4 } \text{sap6jat1 } \text{dim2 } \text{gei2}
\]

over arrive go PRT whole eleven o’clock some

**laa3 (6.1a)**

\[
\text{ji5ging1, } \text{zik1hai6 } \text{heoi3 } \text{dou3 } \text{ne1}
\]

PRT already meaning go arrive PRT

\[
\text{faan6 } \text{dou1 } \text{mou5dak1 } \text{sik6 } \text{ne1}
\]

rice/meal also/even not-can eat PRT

\[
\text{zau6 } \text{fan3 } \text{lo1}.
\]

then sleep PRT

‘[When we] arrived ne1 it was already eleven-something o’clock laa3, meaning when we arrived ne1 we couldn’t even eat dinner ne1 then went to sleep lo1.’

[Nine turns omitted]

A:

\[
\text{Gam2 } \text{zau6 } \text{jiu3 } \text{daap3 } \text{gau2 } \text{go3 } \text{zung1tau4}
\]

like-that then need ride nine CL hour

\[
\text{ce1 } \text{lo1}. \text{Gam2 } \text{zau6 } \text{mou5 } \text{zo2}
\]

car PRT like-that then not-have PFV

\[
\text{jat1 } \text{jat6 } \text{laa3 (6.1b) } \text{jau6}.
\]

one day PRT again

‘Then we needed to ride in a car for nine hours lo1. So then we lost one day laa3 again.’

Applying explication [6A] to example (6.1a), A’s use of laa3 means that she wants the hearer to think now that ‘I know it’s like this’, i.e. that it was already late when she arrived in Indonesia. Using laa3 ‘emphasises’ the information that is being given, conveying to the hearer that A wants her to ‘note’ or ‘register’
the information, or perhaps inviting the hearer to think about the situation from the perspective of the hearer by standing in her shoes and thinking ‘I know it’s like this’.

In (6.1b), A’s use of laa3 is also expressing ‘I want you to think now: “I know it’s like this”’. In other words, A wants the hearer to think now that she knows A spent a day travelling for nine hours and thus ‘lost’ that day. Just as with (6.1a), use of laa3 encourages the hearer to ‘note’ the information and perhaps to think about the situation from the perspective of the hearer and to know what she is saying or feeling. It is possible that the speaker wants the hearer to know that she lost a day, but perhaps more likely, it is possible that the speaker wants the hearer to understand that she was dissatisfied with her trip or felt like she wasted her time or was cheated of a nice holiday, etc.

(6.2)

A:

... gam2 mou5 laa4? Soeng6ci3 nei5 zau6 gam2
then not-have PRT last-time you then like-that
gong2 zo2 niidit zau6 mou5 gaa4 laa4?
say PFV these then not-have PRT PRT

‘That’s all laa4? Last time you only spoke about these and there’s nothing else gaa4-laa4?’

B:

Mou5 laa3, gong2 zo2 loeng5 geoi3 zaa3,
not-have PRT say PFV two sentences PRT
gani1zyu6 zihau6 dou1 mou5 wan2 keoi5.
then after also/even not-have find him

‘No laa3, [we] spoke for only a little (two sentences) zaa3, then after that I haven’t contacted him.’

In example (6.2), A and B are discussing their mutual friend, with whom A is no longer on good terms. Speaker B has spoken to this friend recently about A and their falling out, and is relaying to A what was said. B does not have much to report, but A continues to press him for more information. Laa3 means ‘I want you to think now: “I know it’s like this”’, and is used by B to express to A that nothing else was discussed with their friend, and that he has nothing else to say. His use of laa3 makes sense particularly because we know from the conversation that A seems not to believe him, and seems to wants more
information. He wants her to think now something like ‘I know nothing else was said’ and perhaps this signals to her that he is being open and honest. Conversely, descriptions like ‘newness’, ‘change of state’, and ‘advice’ (Gibbons 1980, 770-772, Yiu 2001, Matthews and Yip 2011, 391) do not describe examples like this well. From the way that A constructs her question, we can surmise that B’s answer is not surprising or new to her (this is different from the answer not being what she wants to hear), and B is not giving advice.

(6.3)

\[ Gam2 \ keoi5 \ gaak3lei4 \ go2di1 \ aa6 \ teng1 \ dou3 \]
then \ s/he \ next \ those \ PRT \ hear \ until
\[ jim3, \ mun6 \ dou3 \ sei2 \ laa3 \ ji5ging1. \]
bored/fed up/dislike \ bored \ until \ die \ PRT \ already

‘And the people beside him/her aa6 were bored/fed up of listening, bored to death laa3 already.’

In example (6.3), the speaker is talking about a company talk that he attended, which he thought was a waste of time. By using laa3, the speaker lets the hearer know that he wants her to think now: ‘I know it’s like this’. What he wants the hearer to know could be the objective facts that he is stating. Alternatively, it could be the general ‘feeling’ he is describing, of himself in that situation – perhaps he wants the hearer to understand how he was feeling. This is believable especially since he talks about this for a long time, and clearly has strong feelings about how bad the talk was, that he wants the hearer to understand. Whichever is the case, the explication makes sense. Descriptions like ‘advice’ (Matthews and Yip 2011, 391) are not applicable to examples like this.

(6.4)

B: ‘How much money does a CM get aa3?’

A: ‘CM gets around \$4000 less [than GM] laa1.’

B: ‘So that’s around \$15 000.’

A:
\[ Maan6cat1, \ m4 \ zii \ maan6luk6 \ ding6 \ maan6cat1 \]
17 000 \ not \ know \ 16 000 \ or \ 17 000
\[ ngo5 \ dou1 \ m4 \ zii \ gei2do1 \ cin2 \ laa3. \]
I \ even \ not \ know \ how-much \ money \ PRT
‘$17 000, [I] don’t know $16 000 or $17 000 I don’t even know how much money laa3.’

Example (6.4) shows a discussion about the salaries of two positions, namely, CMs (Certificate Masters) and GMs (Graduate Masters). *Laa3* is used when A tells B how much a CM supposedly earns. Particularly because B has asked about the salary of a CM, it makes sense that A wants B to think now ‘I know it’s like this’, i.e. that he knows she does not know the exact amount a CM gets but that it is around $16 000 – $17 000. Examples like (6.4) show the inadequacies of descriptions like ‘indicate that an action (or a certain situation) is beginning, has begun, will begin, or has ended, and so on’ (Kwok 1984, 46-47). A is not talking about an event. Descriptions such as ‘change of state’ and ‘newness’ are also not applicable (Gibbons 1980, 770-772, Yiu 2001). More examples of *laa3* being used in responses to questions are given below.

**Answering questions**

In example (6.5), *laa3* means that speaker A now wants speaker B to think ‘I know it’s like this’, i.e. to know that A is going to have a look at universities/information before making a decision about where to study. Using *laa3* highlights that A wants B to think that she knows this now. Perhaps this reinforces the sincerity of her answer, which may seem worthwhile because her answer could be seen as evasive, or at least not as straightforward as might have been expected. The *laa3* in (6.6) further below is also in response to a question.

(6.5)

B: ‘So if you study Accounting, then which university were you thinking of going to?’

A: *Gam2 ngo5 jiu3 tai2 maai4 sin1 laa3 jiu3.*

then I have-to look QPRT first PRT have-to

‘Then I have to see first [have a look first] *laa3.*’
(6.6)

A:
Nei5 zeoi3gan6 jau5 mou5 dak1 interview aa3
you recently have not-have can interview PRT
kei4sat6?
actually
‘Have you recently been able to get any interviews aa3, actually?’

B:
Zeo3gan6 aa4? Zeo3gan6 - zeo3gan6
recently PRT recently recently
dou1 mou5 laa3.
also not-have PRT
‘Recently aa4? Recently – recently I haven’t laa3.’

A:
Hai6 aa4? Ji2! Ngo5 go2 ci3 hou2ci5
is PRT hey I that time seems-like
teng1gong2 nei5 hou2ci5 waa6 heoi3 in jat1
hearsay you seems-like say go in one
fan6 m4 zi1 dim2joeng2...
CL not know how
‘Yes/really aa4? Oh! That time I thought I heard that you were going to interview for I-don’t-know-what...’

In (6.6), B is saying that she has not had any job interviews lately. She uses laa3 ‘I want you to think now: “I know it’s like this”’ because she wants A to think now that A knows B has not had any interviews recently. B’s utterance can give the feeling of her not wanting to say more, as if she is slightly embarrassed. The use of laa3 is logical, and would be compatible with ‘embarrassment’, since B wants A to understand the situation and, perhaps if she thinks ‘I know it’s like this’, to not question her more. This also sheds some light on A’s response, because A is surprised and chooses to explain that she thought B went to an interview. Perhaps A says this because she recognises that B does not want to talk about it, and A wants to explain why she brought up this topic – she mistakenly thought B went to an interview (also, see from (6.7) below that they were already discussing job-hunting). This example suggests that Gibbons’ (1980) claim that laa3 indicates ‘newness’ is incorrect, since nothing new has
happened. Descriptions like ‘advice’ (Matthews and Yip 2011, 391) are also not supported.

(6.7)

B:
Jessica, nei5 wan2-gung1 wan2 sing4 dim2 aa3?
Jessica you find-job find complete how PRT
‘Jessica, how’s your job-hunting aa3?’

A:
Aai1-jaa3! Dou1 m4 zi1 dim aa3. Kei4-sat6
aiya also/still not know how PRT actually
ngo5 dou1 gin3 zo2 hou2-d01 fan6 laa3.
I also see PFV lots CL PRT
daan6-hai6 dou1 mei6 dak1 aa3.
but also/all not-yet done PRT
‘Aiya! I don’t know aa3. I actually went to lots [of jobs/interviews] laa3, but I’m still not done [still haven’t gotten a job] aa3.’

Example (6.7) is from the same conversation as (6.6), although (6.7) occurs much earlier. A uses laa3 to express to B that she wants her to think now ‘I know it’s like this’, i.e. that she has been to lots of job interviews. As above, the meaning of laa3 is well-suited to responses to questions, as it is unsurprising that one would want the person asking the question to think now ‘I know it’s like this’. Unlike B in (6.6), A here does not seem to mind talking about her job search, but the explication is still valid. Again, this example does not show ‘advice’ being given.

(6.8)

J:
Nei5 seng4jat6 duk6syu1 bin3 zo2 mou5
you always study change PFV not-have
kapinaap6 dit gung1zok3 ging1jim6 hou2ci5
absorb/take-in CL work experience seems-like
hou2ci5 e6 bin3 dak1 go3 jan4 hou2ci5
seems-like eh change ADV CL person seems-like
hou2 taa13gw03 gaan2daan1 laa3.
very too simple PRT
‘If you always study it means you don’t get work experience and it seems like, seems like, eh, it makes your person seem very, too simple laa3.’
In example (6.8), speaker J is responding to A’s query about whether or not she will continue to study. J explains her feeling that constant studying without work experience is not good. From her use of laa3, we see that she is saying to A that she wants A to think now: ‘I know it’s like this’. The proposed explication is valid. In this particular scenario, it seems she wants A to understand what she is saying, or to think about it so she will know.

(6.9)

M: ‘Where is your hometown/country home?’

E: ‘My – my hometown/country home is this place Haifeng gaa3. [Do you] know where it is ne1?’

M: ‘Don’t know.’

E: ‘I don’t really know myself, I just know it’s in Guangdong Province lo1. Even I don’t know where exactly which location I also don’t really know. Ai I think it’s near Shantou or Shanwei or those kinds of places.’

M:

Hai6 me1?

is PRT
‘It is me1?’

E:

Hai6 laa3. Daan6hai6 ngo5 nam2 dou1
is PRT but I think still/even
soeng1gaak3 gei2 jyun5 gaa3 wo3...
separated few/quite far PRT PRT

‘It is laa3. But I think it’s still quite far apart gaa3-wo3...’

In example (6.9) above, E tries to explain where his hometown of Haifeng is, but M does not know where it is. The question preceding the laa3-suffixed utterance appears to be more of an acknowledgement of what is being said, rather than a genuine question, such as in examples (6.2), (6.5), and (6.6). Nevertheless, laa3 is used by E to say to M ‘I want you to think now: “I know it’s like this”’. It is possible that he expresses this to compensate for being unclear about where his hometown is, as he does not want M to think that he is deliberately being vague or evasive. It is also possible that he simply wanted her to feel more reassured, or that he wanted to reinforce and back up his own uncertain description. Prior descriptions of laa3 such as ‘change of state’ and
that an action or situation ‘is beginning, has begun, will begin, or has ended’ (Kwok 1984, 46-47, Yiu 2001) do not explain this instance of laa3 well.

**Teaching, instructing**

As with example (6.9) above, laa3 is often used in contexts where a speaker is teaching someone something, or giving instruction or advice, and this fits well with the proposed explication ‘I want you to think now: “I know it’s like this”’. These examples also have similarities with examples of laa3 above which ‘inform’.

(6.10)  
B:  
Keoi5 zuu6 waa6 saa1leot2 baa1 go2di1 m4 hou2  
s/he then say salad bar those not good  
sik6 lo1. So2ji5 nei5 haa6ci3 heoi3 Pizza Hut  
eat PRT so you next-time go Pizza Hut  
ne1 m4 hou2 siz6 go2di1 **l**aa3.  
PRT not good eat those PRT  
‘S/he said don’t eat those [foods] from the salad bar lo1. So next time you go to Pizza Hut ne1 don’t eat those **l**aa3.’

Example (6.10) shows a conversation where speaker B has been telling speaker A about her student, who is working at Pizza Hut. B asked her student the truth about the cleanliness at the Pizza Hut, and the student responded that it is quite clean, except for the salad bar. B relays this to A, and tells A not to eat from the salad bar. Speaker B in this example wants the hearer to think now that he knows it is like this – that he should not eat from the salad bar. This particular example supports Matthews and Yip’s (2011, 391, 401-403) description that laa3 is used to give advice. It also conveys a mild ‘warning’, unlike most other examples of laa3.

(6.11)  
Dou2bat1jiyu4 gong2 faan1 kek6cing4 aa1.  
why-not talk return plot PRT  
Nei5 dou1 lei4tai4maan6zoeng6 **l**aa3 hoi1ci2.  
you even digress/leave-topic-far-away PRT start  
‘Let’s talk about the plot again aa1. You’re going very far off-topic **l**aa3 starting to.’
Example (6.11) above shows a similar example, in that the speaker is instructing the hearer to do something. The speaker is ‘complaining’ that the hearer has digressed from their topic of a television show, and is now very far off-topic. The speaker wants the hearer to go back to talking about the plot of the television show. The speaker uses laa3 ‘I want you to think now: “I know it’s like this”’ to point out to the hearer that he is going off-topic. Though this example is like (6.10) in that the speaker tells the hearer to do something, this example cannot really be described as giving ‘advice’, because the speaker here is not trying to help the hearer. Therefore, although these examples seem similar, they cannot necessarily use the same descriptors from the literature. The proposed explication makes sense for both examples.

(6.12)

Bat1gw03 nei5 m4hou2 zing2 faa1, zing2
but you don’t make patterned/dirty make
zip3, jau5 jam6ho4 syun2wai2 laa3, jyu4gw02
fold have any damage PRT if
m4hai6 zau6 ham6baang6laang6 mou5, zik1hai6
not-is then all not-have meaning
gaa3 -, mou5, dim2 aa3? Mou5 go2 go3
val - not-have how PRT not-have that CL
maai6 faan1 ceottihei3 ge3 gaa3zik6 lo1...
sell return/back out-go LP value PRT

‘But don’t make it dirty [/write on it], fold it, or damage it in any way laa3, otherwise you’ll lose all, I mean val – it won’t have, how [do I say] aa3? It won’t have that value when you want to sell it lo1...’

In example (6.12), a speaker is explaining at length about collecting postage stamps. This is similar to the examples above in that the speaker is ‘instructing’ the hearer to do something. It is like (6.10) in that the speaker is trying to be helpful, with a mild sense of ‘warning’. The speaker in this example also conveys a definite air of ‘teaching’ and passing on information, unlike (6.11). Whereas the speaker in example (6.11) was being impatient or frustrated and complaining, the speaker in example (6.12) is being very patient and helpful. Laa3 means ‘I want you to think now: “I know it’s like this”’, and in (6.12) indicates that the speaker wants the hearer to think now that s/he knows and understands this information about stamp collecting.
Fittingly for contexts where a speaker is teaching someone something, another explication considered earlier for laa3 was ‘this thing was not like this before, I want you to know this now’\textsuperscript{70}. This kind of explication would have been compatible with, for instance, the previous descriptions ‘change of state’ and ‘newness’. However, examples such as (6.13) below support the rejection of this explication. (6.13) shows a speaker explaining that only one university teaches physiotherapy. Here, the laa3-suffixed utterance does not indicate that anything has happened or changed, so ‘this thing was not like this before, I want you to know this now’ is not applicable. Note descriptions such as ‘change of state’, ‘newness’, and ‘advice’ are also not valid (Gibbons 1980, 770-772, Matthews and Yip 2011, 391, 401-403, Yiu 2001).

(6.13)

A: ‘There is only Poly [Hong Kong Polytechnic University] ze1-maa3.’

B: ‘Really me1? Only Poly offers/teaches physiotherapy zaa3-me1?’

A:

\begin{verbatim}
Hai6 aa3. Kei4taa1 dou1 mou5 gaa3 laa3.
\end{verbatim}

is PRT other all not-have PRT PRT

‘Yes aa3. The others don’t have it gaa3-laa3.’

Instead, the current proposed explication is substitutable for laa3 in (6.13), and indicates that A wants B to think ‘I know it’s like this’, i.e. that she knows that other universities do not offer physiotherapy. If B thinks ‘I know it’s like this’ and therefore understands the information, this would be in line with the air of ‘finality’ that comes across, or the feeling that this is the end to this part of the discussion. This sense of finality is not present in all other examples, such as (6.1a) or (6.12).

Example (6.4) above also showed that the current proposed explication [6A] is better than rejected explications like ‘this thing was not like this before, I want you to know this now’. In example (6.4), nothing had changed, hence ‘this thing was not like this before’ is not always substitutable for laa3.

\textsuperscript{70} A variant was ‘I want you to know this now, it was not like this before’, but this is poorly constructed because the problem of the ‘sequential’ reading occurs. The ‘it’ in the second part of the explication was intended to refer to the same thing that ‘this’ from the first part refers to. However, this explication was open to the interpretation that the second part of the explication referred to the first part. This would have meant that the speaker did not want the hearer to know this before.
Does the hearer know?

The next few examples suit the idea that laa3 indicates that the hearer does not know something (or rather, is perceived by the speaker not to know something). This clearly fit instances where laa3 is used to answer questions, or to teach or inform someone of something. Following this thinking, the earliest proposed full explication of laa3 was as presented in [6B.1] below. It is long compared to the final particle explications proposed in this thesis, since the short nature of the explications was at first unexpected. The explication was later shortened to [6B.2], and then to [6B.3]. As can be seen, these early explications all hinged on the speaker’s apparent impression that the hearer does not know the thing being spoken about. This idea was eventually dismissed based on examples such as those given further below. Similar components which were also tested include ‘you aren’t thinking about this’, and variations of ‘someone can think not in this way/not like this’. These components could be paired with other components (including each other) to create longer explications, but they were all ultimately rejected.

[6B.1] First full explication proposed for laa3:

I think you don’t know this
I say this now because I want you to know this now
I think you can know what I think now

[6B.2] Second full explication proposed for laa3:

you don’t know this
I want you to know what I think now

[6B.3] Third full explication proposed for laa3:

you don’t know this

In example (6.14), E and M are talking about introducing E to some of M’s new colleagues. As M has mentioned some colleagues whom she describes as ‘grandmas’ [women old enough to be grandmas], E, who is 23 years old, asks whether the women she will introduce him to are young. M says they are, and so E approves. It is unlikely that E meant ‘you don’t know this’ or ‘someone can think not in this way/not like this’. E’s request to be introduced to someone closer to his own age is presumably not surprising to M since it is a normal and reasonable request, and in the previous turn E has already stated that that is
important to him. Instead, the current explication ‘I want you to think now: “I know it’s like this”’ makes sense. E seems to be expressing to M that he wants her to think now something like ‘I know that it will be acceptable if they are young’.

(6.14)

E: ‘But are they young though, that’s the most important.’

M:

\[E6, \text{ dou1 jau5 di1 hau6saang1 ge3.}\]

eh also have some young PRT

‘Eh, there are some that are young ge3.’

E:

\[Gam2 \text{ dak1 laa3...}\]

then/like-this can/okay PRT

‘Then it’s okay laa3…’

Example (6.15) is from the same conversation as example (6.2). Recall that A and B are discussing their mutual friend with whom A has had a falling out. A wants to know what B discussed with this friend. At this point in the conversation, speaker B has already told her some things, and now has no more to say. In this excerpt, B expresses in the first turn that he has said all there is, but A continues to press him for information. In the third turn, speaker B states that there is nothing else. His use of laa3 is unlikely to mean ‘you don’t know this’, because he has already indicated in the preceding turns that there was nothing more (including through his use of laa1, which indicates ‘you now know how I think about this, I can not-say more’). This example also helps to reject the tested explication ‘this thing was not like this before, I want you to know this now’, as well as previous descriptions of laa3 such as ‘newness’, ‘change of state’, and ‘advice’ (Gibbons 1980, 770-772, Yiu 2001, Matthews and Yip 2011, 391).

(6.15)

B:

\[Gam2 \text{ e6 zau6 gam2 laa1.}\]

then eh so/then like-this PRT

‘So eh it was like this laa1.’ [‘That’s all.’]
A: Gam2 dim2 aa3?
then how PRT
‘Then what aa3?’

B: Mou5 laa3 gan1zyu6...
not-have PRT then/following-this
‘Then there was nothing else laa3.’

Instead, the laa3 in example (6.15) can be substituted by the current proposed explication [6A]. Speaker B conveys to A that he now wants her to think something like ‘I know there was nothing else’. This can be substitutable in context. As mentioned with regards to example (6.2) and evidenced by the repeated questioning by A, it seems that A does not fully believe B, or at least still wants to know more. It makes sense that B wants her to think now something like ‘I know there was nothing else’, and perhaps signals to her that he is being open and honest. B’s utterance also gives a feeling of finality (it would be understandable that B feels overly questioned, as A is quite fixated on finding out about this mutual friend) and it seems like B is trying to end this topic of conversation or at least does not have more to say. This also makes sense with the current explication because, assuming that A thinks ‘I know it’s like this’, there is not more that needs to be said.

Though components featured in the early explications, like ‘you don’t know (something)’, were not always applicable where laa3 was used, this component still seemed very fitting with most examples. Furthermore, in the case of examples in which it would be wrong to say that laa3 meant ‘you don’t know (something)’, this component nevertheless gave the correct ‘feeling’. It seemed that the speaker always intended to give the air or impression of ‘telling’, ‘teaching’, ‘emphasising’ etc. to the hearer something that the hearer did not already know. This led to consideration of a series of explications which were to do with the speaker saying something as if the hearer did not know something. These tested explications included ‘it is like you don’t know this thing; because of this, I want to say this thing to you now’. Another possibility was ‘I want to say this thing to you now, because it is like you don’t know this thing’. This attitude or quality is conveyed by the speaker, regardless of whether
the hearer knows or not. These tested explications were supported by most examples of laa3 in the corpus, and were also more informative than some others considered. For some time, the favoured explication was ‘I want you to know this now; I say this to you like you don’t know it’.

The idea of the speaker talking ‘as if’ the hearer did not know something was expressed in NSM using LIKE, but unfortunately, this is not easily translatable into Cantonese without the use of dong3 ‘presuming/assuming’ in place of LIKE. Dong3 is not a semantic prime and does not have a simple meaning. A workaround would have required a long and more convoluted explication. Ultimately, these explications were not chosen.

**Forgetful speakers**

Other explications considered for laa3 include ‘it is good if you think about this’ and ‘I want you to think about this’. These potential explications were still related to ideas such as ‘you don’t know this’, but were found to be more fitting for most examples. Nevertheless, these were still odd when substituted in examples such as (6.16) and (6.17), which both show someone forgetting something.

(6.16)

\[
\begin{align*}
\text{Gam2} & \quad \text{zung6} & \quad \text{jau5} & \quad \text{heoi3} & \quad \text{zo2} & \quad \text{cin2cou2} & \quad \text{laa1}, \\
\text{then} & \quad \text{still} & \quad \text{have} & \quad \text{go} & \quad \text{PFV} & \quad \text{Asakusa} & \quad \text{PRT} \\
\text{heoi3} & \quad \text{zo2} & \quad \text{daai6baan2} & \quad \text{laa1}, & \quad \text{heoi3} & \quad \text{zo2} & - \\
\text{go} & \quad \text{PFV} & \quad \text{Osaka} & \quad \text{PRT} & \quad \text{go} & \quad \text{PFV} \\
\text{m4gei3dak1} & \quad \text{zo2} & \quad \text{laa3,} & \quad \text{hou2} & \quad \text{noi6} & \quad \text{lu3}. \\
\text{not-remember} & \quad \text{PFV} & \quad \text{PRT} & \quad \text{very} & \quad \text{long-time} & \quad \text{PRT}
\end{align*}
\]

‘And then [we] also went to Asakusa laa1, went to Osaka laa1, went to – I forgot laa3, [it was] a long time [ago] lu3.’

In example (6.16), the speaker talks of her last trip to Japan, listing some of the places she went to. It is unlikely that she used the particle laa3 to express ‘I want you to think about this [the fact that I’ve forgotten]’ or ‘it is good if you think about this [the fact that I’ve forgotten]’. That one has forgotten something may be something that one wants the hearer to know, but wanting the hearer to

---

71 As well, ‘I want you to know this now’ did not contribute much, but if omitted, the remaining explication would be open to unwanted ‘sequential’ readings.
THINK ABOUT this is less plausible, especially given the trivial nature of the thing she has forgotten. Instead, the current explication is substitutable in examples of forgetfulness – it is believable that the speaker here wants the hearer to think something like ‘I know she’s forgotten’.

(6.17)
A:

\[
\text{Go2zan6si4} \quad \text{ngo5dei6} \quad \text{heoi3} \quad \text{zo2} \quad \text{bin1dou6} \quad \text{sik6} \quad \text{aa3?}
\]

at-that-time we go PFV where eat PRT

\[
\text{Daa2bin1lou4} \quad \text{aa3} \quad \text{ho2?}
\]

hot pot PRT PRT

‘Where did we go to eat that time aa3? Hot pot aa3-ho2 [right]?’

J:

\[
\text{M4gei3dak1} \quad \text{laa3.}
\]

not-remember PRT

‘Can’t remember laa3.’

In example (6.17), A is recalling a dinner she and her friends, including J, once had. As in example (6.16), J says that she cannot remember something, and uses the particle laa3. The thing she has forgotten is again trivial, and it is unlikely that she means ‘it is good if you think about this [the fact that I can’t remember]’ or ‘I want you to think about this [the fact that I can’t remember]’. Conversely, the current explication ‘I want you to think now: “I know it’s like this”’ is substitutable. In this example, J wants A to think something now like ‘I know she can’t remember’.

It is not only examples where the speaker has forgotten something which show that explications like ‘it is good if you think about this’ and ‘I want you to think about this’ do not represent the meaning of laa3 satisfactorily. Reconsider examples (6.13) and (6.15). In example (6.13), B asks whether only one particular university offers physiotherapy, and A replies ‘Yes aa3. The others don’t have it gaa3-laa3.’ A’s short answer sounds very ‘final’ and ‘certain’, and does not seem to encourage B to really think about it. In example (6.15), A asks B what he spoke about with someone else, and he replies ‘Then there was nothing else laa3.’ B also does not seem to encourage A to really think about it. On the contrary, he sounds like he does not want A to dwell on this anymore, and wants to end this topic of conversation.
In cases such as forgetting something or wanting to draw an end to a topic of conversation, it is more likely that this is something the speaker wants the hearer to know, rather than think about. A better explication already mentioned would be 'I want you to know this (now)'. A drawback was that its meaning can be applied to a very large number of utterances, and it seemed to perhaps be missing a component. Variations such as 'because I want you to know what I think now' also did not add significant meanings. Nonetheless, at an earlier stage of analysis these were paired with various other tested components mentioned to create explications of one or two lines. For instance, 'I want you to know this' was long considered a potential component alongside components indicating change or events happening (see below). Other tested explications involving knowing included 'I know you want to know this now' and 'I think you want to know this now'. All of these explications and components were eventually dismissed.

Changes and happenings

Many of the previous descriptions of laa3 involve references to time and aspect and 'changes of state', so potential explications concerned with these ideas were also considered. A series of tested explications involved happen. This included the simple 'something happened (is happening, will happen)', and 'I want you to know: something happened'. Such explications/components could be paired with various other considered components, such as 'I want you to know this now'. A series of components intended to refer to a current state of affairs (but which needed to remain applicable to any possible state of affairs), like 'it is like this (now)' were also tested.

72 Again, some explications were rejected based on poor construction. For instance, 'I want you to know this', paired with variations of 'someone can think not in this way/not like this', resulted in flawed explications since the separate lines of the explication could be interpreted as 'sequential'. The second line of the explication was composed as a comment on whatever is being said in the utterance, but could easily be misunderstood as a comment on the first line of the explication. Therefore, it could be interpreted as something like 'I want you to know this ('this' being whatever was in the utterance), and someone can think I don't want you to know this'. Nor could this problem of the unintended 'sequential' reading be easily solved by reversing the lines of the explication. The explication might be interpreted as 'someone can think not in this way, and I want you to know that someone can think not in this way'. Furthermore, the use of someone in 'someone can think...' invited ambiguity. This person could be interpreted as any third person who could potentially think anything.
Explications using variations of ‘something happened’ were quickly able to be rejected. We have already seen examples above which do not support claims that laa3 is to do with time and change, such as (6.4) and (6.9). Some more are given below. Example (6.18) shows a compliment being given to the hearer. The speaker is commenting on the (unchanged) fact that her friend is the smartest of their friend group. Nothing has actually happened, and this does not support the previous descriptions that laa3 indicates change (Gibbons 1980, Kwok 1984, Yiu 2001).

(6.18)
\[
\text{Ngo5dei6 gam3 do1 go3 zi1zung1}
\]
we like-this much/many CL out-of/from
\[
\text{zi3 sing2 hai6 nei5 gaa3 laa3.}
\]
most smart is you PRT PRT

‘Out of all of us, you are the smartest gaa3-laa3.’

It was considered possible that the change is supposed to happen in the opinion of the hearer. Potential explications then included ‘you did not think about this thing before’ and ‘I want you to think not like you thought before’. These could be paired with various other components already mentioned, such as ‘I want you to know this now’ or ‘I want you to know what I think now’. These proposed explications seemed possible for many examples of laa3, such as example (6.19) below.

(6.19)
\[
\text{B: Gam2 mai6 m4 dim6.}
\]
so/then then not okay

‘Then that’s not good.’

A:
\[
\text{Mou5 ge2, gam2 ji4gaa1 gam3 naan4 wan2,}
\]
not-have PRT so/then now so hard find
\[
\text{di1 gaau3syu1. Keoi5 send zo2 hou2do1 fan6}
\]
CL teaching she send PFV very-many CL
\[
\text{seon3 heoi3 laa3.}
\]
letter go PRT

‘No ge2, they’re so hard to find nowadays, those teaching jobs. She has already sent many letters out [job applications] laa3.’
In this example, A and B are talking about their mutual friend who is going to teach at a primary school. B indicates that the job is not ideal, but A says it is okay because it is really hard to find teaching jobs and their friend has already sent many letters [job applications] (and she also goes on to tell B about other friends of hers who sent out hundreds of letters without getting any responses). Speaker A is not just stating the fact that their friend sent out many letters, but is using this to show sympathy, and support her stance that the job their friend has is not bad. This is a typical use of laa3 and fits many of the tested explications. In the context here of showing sympathy for a friend, A’s utterance could mean something like ‘I want you to know that she has already sent many letters out [she is already trying hard/out of options]’, and this could acceptably be paired with, say, ‘you did not think about this thing [that she is already trying hard/out of options] before’ or ‘I want you to think not like you thought before [I want you to stop thinking that her current job is not good]’.

Nonetheless, these explications/components were all ultimately rejected after consideration of more examples. Recall, for instance, situations where the speaker is unsure or has forgotten about something. Examples have been given above, such as (6.4), (6.16) and (6.17), where components like ‘I want you to think not like you thought before’ are not applicable. The current explication ‘I want you to think now: “I know it’s like this”’ is valid for example (6.19). It indicates that A wants B to think now ‘I know it’s like this [she has already sent out many letters]’. This also achieves the speaker’s apparent goal of telling her friend to show some sympathy.

(6.20)
A: ‘... You know I don’t contact him gaa3-laa1.’
B: ‘I know.’
A: ‘Yes laa1.’
B: ‘That’s also right ge2, you not contacting him. Not contacting him is also right ge2.’
A: *Syn3 laa1, aai6ji5deon6 lei4 ngo5 ji4 heoi3 laa3* let-it-go PRT Elton leave I and go PRT
In example (6.20) above, two speakers are talking about speaker A’s poor relationship with their mutual friend, ‘Elton’. In the preceding conversation, including the first few turns of this excerpt, both parties make it clear that they know A and Elton have no relationship anymore. We know this from, for example, their use of ‘you know/I know’, and A’s use of laa1, which indicates that B knows how A thinks. Both speakers also explicitly state that A does not contact Elton anymore. Therefore, in the laa3-suffixed turn at the end of this excerpt, it would not make sense that laa3 includes the meaning ‘you did not think about this thing before’.

**Incorporating components from laa1**

Since it has been claimed that laa1 and laa3 share some meaning components because they share phonological features (except for tone), it was considered that the additional laa3 component might be taken from the explication of laa1 (for more on the sub-syllabic morphemes hypothesis, see Chapter 8). Recall that the two components of the explication for laa1 are ‘you now know how I think about this’ and ‘I can not-say more’. Added onto other considered components for laa3, these two components seemed plausible for laa3 and were actually found to be acceptable in many cases. Furthermore, it was not the case that all the acceptable examples were situations in which laa3 could have been replaced by laa1. ‘You now know how I think about this’ seemed, for a few examples of laa3, to be slightly more fitting than ‘I can not-say more’. Overall, tested components and explications still seemed better without the added components from laa1. Examples from the corpus supported the rejection of the two laa1 components.

To begin with, reconsider the laa3 in example (6.9). Since it is obvious and expected that M does not know where E’s hometown is (E is not even sure himself), it is not plausible that laa3 includes ‘laa1-like’ meanings of either ‘you now know how I think about this’ or ‘I can not-say more’. It is also relevant that continuing on from the excerpt shown in (6.9), E continues to talk more about this place and how long it takes to travel there etc. This supports the conclusion
that E did not mean ‘I can not-say more’ or ‘you now know how I think about this’.

(6.21)

J: ‘Oh your boss is such a nice person ge2?’

A: ‘S/he is quite a nice person ge2.’

J: 

Gam2 mai6 hou2 lo1. Aai5, ngo5

so/then then good PRT ai [sigh] I

lou5baan2 aa6 m4 zi1 laa3.

boss then not know PRT

‘Then that’s good lo1. Ai, I don’t know about my boss laa3.’ [I.e. ‘I don’t think my boss is that nice’.]

A: ‘Have you ever seen your boss aa3?’

Consider also example (6.21), where J indicates that her boss may not be that nice of a person. Use of laa1 might point out that what is being said is fairly obvious, easy to understand, common-sense, etc. – that the hearer can reasonably know and understand what the speaker is trying to convey. But A’s response and the subsequent turns indicate both that A does not know anything about J’s boss, and that J is happy to tell A more. Given also that J’s laa3-suffixed turn is a slight change of topic towards J’s boss, whom A does not seem to know about, it is not very reasonable to say that J meant either ‘I can not-say more’ or ‘you now know how I think about this’. (If we pretend that speaker A actually does know something about J’s boss, ‘you now know how I think about this’ would be more fitting than ‘I can not-say more’.)

Now consider the current proposed explication ‘I want you to think now: “I know it’s like this”’. It does not require assumptions about what the hearer already knows. In example (6.21), the speaker points out the unchanged fact that her boss is not very good. She does not mean that her boss has become worse. This also supports the rejection of components such as ‘something happened’. Again, this example and the laa3-suffixed utterance do not indicate that anything has happened or changed, and the above example is not well explained by the description ‘change of state’.
Example (6.22) also supports the rejection of the laa1 components. This conversation was also featured in Chapter 4. Two people are talking about a Cantopop singer. Speaker B uses laa3 to say that the singer is only good to listen to. It is unlikely that he meant ‘I can not-say more’ or ‘you now know how I think about this’. Firstly, the meaning behind his laa3-suffixed utterance is fairly obscure and puzzling. The most common and straightforward appeal of a singer should be their singing/songs; that B seems to downplay this and is judging her appeal by her height is unusual. Therefore, it is understandable that speaker A would not know what B means when he says ‘listening a bit is good/enough laa3’. Secondly, we can reason that both speakers think the meaning of the laa3-suffixed utterance is unclear. We can see that speaker A does not understand because she directly questions B on what he means. We can also deduce that B knows his statement is unclear, because he tries to explain what he means. The two turns after laa3 actually overlap somewhat; B says ‘teng1 haa5...’ roughly at the same time as A is saying ‘keoi5’. It seems that B said ‘teng1 haa5...’ because he was starting to explain himself further, since he knew he was not clear enough. When A finishes her question, B then explains: ‘gam3 gou1’, [she’s] so tall’. Since this laa3-suffixed utterance is logically
strange and seems to be judged by both speakers as being unclear, this suggests that ‘you now know how I think about this’ and ‘I can not-say more’ do not belong in the explication for laa3 here. The current explication ‘I want you to think now: “I know it’s like this”’ is more acceptable, as it does not require that the hearer know something before/without being told.

As demonstrated here, the explication for laa3 does not include part of the explication for laa1. This seemed a worthwhile line of investigation, although this should not be considered rigorous testing of the similarities and differences between laa1 and laa3. Nonetheless, this may have implications for the sub-syllabic morphemes hypothesis (see Chapter 8).

**Jokes and teasing**

Finally, consider a context where laa3 can be used, which has not been seen above, and is not mentioned in the literature. Laa3 can be used in scenarios of joking and teasing, and explication [6A] is still applicable. Example (6.23) occurs later in the same conversation as (6.17). Recall that A is reminiscing about a hot pot dinner she shared with J and their friends, but J does not remember it. After A describes it in detail, J remembers it in the fourth turn of the excerpt below. Speaker A talks about the excessive amount of food that their friend bought, and J teases that A is so pitiful/poor for having too much food. J’s laa3-suffixed utterance is clearly a sarcastic joke, pointing out the ridiculousness of A’s ‘complaint’. Nonetheless, the explication of laa3 can be substituted here. J wants A to now think something like ‘I know I am very poor/pitiful (for having too much food)’.

(6.23)

A: ‘Yes aa3, Linda said why not eat hot pot ne1, it was like [the food] was free wo3 the way she was buying things.’

J: ‘Yes me1?’

A: ‘And then we ate leaving so many leftovers ge2.’

J: ‘Ah, yes aa3, yes aa3, yes aa3, yes aa3.’

A: ‘How can we eat all of it zek1?’
J: Ai1jaa3 zan1hai6 caam2 laa3 nei5. aiya truly poor/pitiful/sad PRT you
‘Aiya you are so poor/pitiful laa3.’

In example (6.24), E and M are talking about M’s job as a teacher. They joke about certain things she could say which, if said to her students, would be ridiculous, outrageous, or funny. This leads E to tease that M will have to look for a new job soon (if she says those things), and that she better start looking at newspapers [job advertisements]. In both instances of laa3, laa3 means ‘I want you to think now: “I know it’s like this”’, and this can be substituted in context.

(6.24)
E:
Gam2 aa6 hou2 faai3 zau6 wui5 jau6 jiu3 then/so then very fast then will again need
wan2 gwo3 fan6 gung1 laa3 (6.24a) nei5. find over/another CL job PRT you
‘Then you have to find another job again very quickly laa3.’

M:
Hai6 laa1. is/yes PRT
‘Yes laa1.’

E:
Wan2 ding6 bou3zi2 dou1 dak1 find already/certain newspaper also/even can
laa3. (6.24b) PRT
‘You can (/should) even start looking at newspapers [job advertisements] laa3.’

Both (6.24a) and (6.24b) give a sense of ‘finality’, that the hearer should accept that she should look for another job / start looking at advertisements. This is in line with the proposal that laa3 means ‘I want you to think now: “I know it’s like this”’, since if M really thinks ‘I know it’s like this’, then she would do what E is saying. E could even be interpreted as being (jokingly) ‘pushy’ in telling M to do these things.
6.3 Concluding remarks

This chapter has explored the meaning of laa3 using NSM and naturally-occurring examples of laa3 from the corpus. Most prior descriptions of laa3 associate it with ideas related to time and aspect, although the descriptions are found to be quite confusing and vague. Other descriptions rely problematically on familiarity with Mandarin, or use descriptors such as ‘emphasis’ and ‘relevance’, which do not tell us much about what the particle means. Furthermore, descriptions of laa3 overlap significantly with descriptions of other Cantonese particles, even in the work of the same authors. Finally, examination of examples of laa3 from the corpus shows that the descriptions are not always valid. This chapter has proposed an NSM explication of laa3 which appears to represent its invariant meaning and which can explain what it means in a variety of contexts. Through testing by substitution, the proposed explication is demonstrated to be superior to various other NSM explications and components which were considered at earlier stages.
Chapter 7:
The semantics of particle zaa3

Zaa3 is the 17th most frequently used utterance particle in the Hong Kong Cantonese Corpus, occurring 173 times. It is the 144th most frequently used word overall. Zaa3 readily combines with gaa3 and wo3, but not with laa1 or laa3. The clusters gaa3-zaa3, zaa3-wo3, and gaa3-zaa3-wo3 are all possible and indeed common. As in the preceding chapters, this chapter first provides an overview of the existing literature on zaa3, which in this case focuses mainly on the relationship between zaa3 and English only, and between zaa3 and another Cantonese utterance particle ze1. Then, an NSM analysis is presented, with real examples of zaa3 from the corpus provided throughout. This is the last chapter of Part Two.

Unlike in the preceding chapters, the analysis here is presented in a way which attempts to steer readers through the investigative process via which an acceptable NSM explication was found. The process of NSM semantic analysis usually requires consideration of many intricate, interrelated but diverse possibilities, and there is seldom time or space to go over these processes in detail. This chapter, however, provides a linear narrative as much as possible. It aims to guide the reader from a native speaker intuitive understanding of zaa3 through to an NSM explication substantiated by native speakers’ real, recorded speech.

7.1 Previous descriptions of zaa3

The earliest studies of particles did not reveal much about the meaning of zaa3. Gibbons (1980) made no mention of such a particle, and according to Yau (1965, 112-116, 1980, 50), zaa3 as an independent particle did not convey any of his 12 tested connotation concepts. As was the case with laa3 (see Chapter 6), this would suggest that zaa3 has no meaning, but this most likely

---

73 The particle zaa3 appears to have had several other written forms in earlier studies of particles, including dzaa3 (Yau 1965), t/ya (Yau 1980), and lbza (Kwok 1984). Presumably, these different forms are due to differences in romanisation only, and this paper treats these as the same particle expressed here as zaa3.
reflects a flaw in Yau’s study. From these early studies we learn only that zaa3 can be categorised as an ‘S-type’ particle which is ‘obviously affirming’ (Yau 1965, 66, 1980, 46). Kwok similarly categorised zaa3 as one which may be suffixed to statements without altering their grammatical status as statements (Kwok 1984, 41), and added the insight that it is semantically similar to the English word only (Kwok 1984, 50-51). Since then, there has been a general consensus that zaa3 means only. Wakefield (2008) concluded that the English equivalent of zaa3 was a combination of ‘only’ or ‘an equivalent’ such as ‘just’, combined with stress intonation on the English word put into focus by zaa3.

However, as discussed in Chapter 1, using single English words and translations such as only to explain Cantonese utterance particles generally does not work well, and this is not an adequate or accurate description for several reasons. First, English words like only and just are themselves fairly complex, despite being, like Cantonese utterance particles, fairly short and common words. From previous studies of these English words (e.g. Aijmer 2002, 153-174, Wierzbicka 2003, 346-347, 350-354), it is unlikely that only and just are semantically equivalent even to each other, let alone to zaa3 (Wakefield’s claim that only and just are equivalents indicates a loose use of the word ‘equivalent’).

Second, even if zaa3, only and just were semantically equivalent, using a complex English word as a definition is of no help except to English speakers; an explication in NSM would be much more easily understood and translatable. Mentioning stress or intonation in one’s definition, as Wakefield does, also only helps English speakers with a native-like grasp of intonation and nuances in speech, and assumes that all English speakers use intonation in exactly the same way.

Third, as has been noted throughout this thesis, it often happens that proposed translations of the utterance particles turn out, on closer inspection, to be translations of other parts of the utterance. We can see an example of this in Kwok’s treatment of zaa3. In addition to describing the basic particle, Kwok (1984, 52) describes the cluster ˦ge ˧dza [ge3zaa3], using examples such as that reproduced below as (7.1). As is evident, the ‘only’ in her free translation ‘I only

---

74 This particle occurs five times on its own in Yau’s ‘Recording A’, which ran for approximately 15 hours, with 104 participants/informants (Yau 1980, 39).
translate’ seems to come from *dziŋ hai* [*zing6hai6*] ‘only’, leaving it unclear what *ge3-zaa3* actually contributes to the sentence. Surely, the Cantonese utterance ‘I only do translate’ gives the free translation ‘I only translate’, even without any particles.

(7.1)

\[
\begin{align*}
\text{yno} & \quad \text{dziŋ hai} & \quad \text{dzou} & \quad \text{fa:n jik} & \quad \text{ge dza} \\
\text{I only do translate} & \quad \text{ 동시에} & \quad \text{ge dza} \\
\end{align*}
\]

‘I only translate.’ (Kwok’s gloss and translation)

Kwok adds some explanatory text which states what the implication of this utterance could be in certain contexts, but without a reliable definition as a starting point, this is only helpful for explaining those limited examples which are provided. Needless to say, it is impractical and impossible for explanatory text to accompany every possible utterance, and unfortunately the reader is left to use their own linguistic intuition. NSM avoids this problem, since NSM explications are substitutable in context. Kwok herself explains that the expression *zing6hai6* ‘only’ is occasionally used to reinforce the meaning of *zaa3* (Kwok 1984, 51). It is possible that she gets the idea of *only* from words like *zing6hai6* rather than *zaa3*. And if not, does this mean that Cantonese speakers essentially say ‘only’ twice in the same utterance (i.e. *zing6hai6*... *zaa3*)?

Fourth, when we investigate whether all examples of the particle *zaa3* can be explained using English *only*, it turns out that it cannot be used in all examples. For example, (7.7) below would have to be interpreted as something like ‘you only don’t have benefits’, which makes little sense. Example (7.18b) where the speaker says ‘there is not much to eat inside’ would be read as something like ‘there is only not much to eat inside’, which again makes little sense. A common feature of these examples is that *zaa3* is used in close proximity to the word *mou5* ‘don’t have’. Perhaps it is due to some sort of incompatibility between *mou5* and *only* that causes *only* to be a particularly poor translation in these cases.

---

75 And further, even if we were to accept that *zaa3* means ‘only’, we would still face the conundrum of what *ge3* has added. What would be the difference between *ge3-zaa3* and *zaa3*?
Finally, many other Cantonese utterance particles – which cannot all be used interchangeably – are also described as meaning *only*. Kwok (1984, 53) describes particles such as ldze [ze1] as meaning *only*, saying that they have similar meanings and overlap in function. Wakefield (2012b, 1) also claimed that the English near-equivalent of *ze1* is expressed by the word ‘only’, plus a rise-fall-rise pitch contour – this is very similar to his description of *zaa3*, mentioned above. Fung (2000, 30) describes no less than seven Cantonese *z*-particles as all meaning *only*: *ze1*, *zek1*, *ze4*, *zaa3*, *zaa4*, *zaa5*, and *zaak1*. Even if these particles are indeed semantically similar, describing them all in terms of *only* is clearly not helpful. It ignores the fact that they are mostly non-interchangeable. Fung herself admits that *z*-particles have been associated with ‘a bewildering variety of meanings such as delimiting, reporting, refuting, downplaying, persuading, interrogating, asserting, and the marking of affection, coquettishness, impatience, pride, contempt, jealousy, etc.’ (Fung 2000, 30).

Kwok states that *zaa3* frequently seems to have ‘a negative value in the sense that its presence indicates that what is being stated is not more, or bigger, or longer, or better or more desirable, and so on’ (Kwok 1984, 51-53). As a result, it easily conveys the idea of insufficiency and ‘is sometimes associated with an attitude of disdain, of scorn and of disapproval’. To her credit, Kwok stops short of saying that *zaa3* always indicates something negative. She adds that disapproval or disdain are not essential components in the particle’s meaning. For example, someone who has passed their driving test could produce (7.2), with an attitude of pride (Kwok 1984, 53).

(7.2)

```
hok dzo sap gei go dzun τau dza
```

‘I had only ten odd hours of driving lessons.’ (Kwok’s gloss and translation)

Kwok believes the difference between *zaa3* and *ze1* to be that *zaa3* suggests ‘not enough’, and that *ze1* carries the meaning of ‘not too excessive’ or ‘not too much’ (Kwok 1984, 53). She uses the example of ‘*saam1 man1 zaa3*/ze1’, ‘three dollars

---

76 Two *z*-particles described by Fung (2000), namely *ze4* and *zaak1*, seem to me to not be in use in contemporary Hong Kong Cantonese. Other studies which look at a broad range of Cantonese utterance particles, including Kwok (1984), Matthews and Yip (2011), Sybesma and Li (2007), Yau (1965), and Yau (1980), also do not mention these two particles.
zaa3' to make her point, stating that zaa3 suggests three dollars is not enough, while ze1 suggests that three dollars is not too much. Fung (2000, 59-60) follows Kwok’s thinking with a very similar example, noting that zaa3 and ze1 are not always interchangeable because they convey different expectations of the speakers. Use of zaa3 supposedly means that the speaker assumed the price to be higher, but that the price turned out to be lower than expected.

I disagree, because I believe that zaa3 can indicate either ‘not enough’ or ‘not too much’. One can imagine a scenario where a shopkeeper advertises something as ‘three dollars zaa3’, meaning that the goods for sale are really cheap: ‘not too much’. Equally, one can imagine a scenario where the shopkeeper says ‘three dollars zaa3’, upset that profits are too low: ‘not enough’. If Kwok and Fung are wrong about the speaker’s expectations, then Fung’s (2000, 63-64) argument based on this – that zaa3 does not carry a downplaying function – is also called into question. Note also that even when zaa3 and ze1 do signal that expectations have been too high or too low, it may not necessarily be the expectation of the speaker – it could be the expectations of somebody else.

According to Fung, senses conveyed by zaa3 are less varied than ze1 and are confined to the delimitative and contrastive senses, with most tokens of zaa3 found in her corpus performing the delimiting function (Fung 2000, 58-59). The delimiting sense apparently frequently co-occurs with semantically similar adverbs such as zi-hai ‘only’ (Fung 2000, 34-35); while this may be true, it could again make it easy to mistakenly attribute the meaning of ‘only’ to the particle. Zaa3 can also connect two clauses and thereby induce a contrastive reading by juxtaposing contradictory premises side by side (Fung 2000, 61-62). This seems to be an observation of the kinds of sentences that zaa3 can be used in, more than a semantic description. More importantly, the contrastive reading remains even if the particle is omitted. Therefore, one can hardly say that the juxtaposition is due to the particle; only that the two are compatible. One of Fung’s examples is reproduced below.
Fung (2000, 71-72) summarises her discussion by saying that the core semantic feature of Z- is ‘[restrictive]’, with a list of six ‘functional primes’ for zaa3: -propositional, +h-assumption ['assumption of a higher value on the evaluation scale'], +/-connective, -concessive, and -exhortative, -epistemic. Each functional prime does not appear to be explained further. This summary may be somewhat effective in internally comparing and contrasting particles discussed in Fung’s work, but could seem too technical and obscure for the average native speaker or language learner. If descriptions can be understood by native speakers, they have the advantage of also being able to be verified by native speakers. NSM can explain meanings in simple, clear, and translatable natural language which can be more easily understood and tested.

Sybesma and Li accept without question Fung’s (2000) analysis that z-particles have the ‘core meanings’ of ‘delimitation’ and ‘contrast’, and state that zaa3 conveys this meaning and nothing else (Sybesma and Li 2007, 1754). Their short description does not contribute any new ideas about the meaning of zaa3. They also claim that ‘all sources and informants agree that zaa3 conveys “only” in the neutral sense of “not more than that” or “and not something else as well”’ (Sybesma and Li 2007, 1754).

This brings us to another descriptor found with zaa3: ‘neutral’. Following Fung (2000) and Sybesma and Li (2007), Wakefield (2012b, 1) also accepts that zaa3 has ‘the neutral meaning of “only”, and the implication that this is less than the speaker expected is pragmatically understood’. Setting aside the problems of using ‘translations’ such as only, their repeated description of zaa3 examples as ‘neutral’ and as describing merely factual things is puzzling. Firstly, one would presume that if an utterance were truly neutral, that it would have no utterance particles attached at all. Secondly, Sybesma and Li’s own

---

77 Her abbreviation ‘FP’ stands for ‘final particles’.
analyses suggest that their provided sentences are not truly neutral. For instance, they provide example (7.4) below and stress that ‘2000 people’ is *factually* a small number (Sybesma and Li 2007, 1754-1756). This is problematic as the significance of a given number is relative. From the perspective of someone who thinks 2000 students is a lot, this sentence would be odd.

(7.4)

\[
\text{ngo5-dei6} \quad \text{hok6hau6} \quad \text{jau5} \quad \text{loeng5-cin1} \quad \text{jan4} \quad \text{zaa3} \\
\text{1P} \quad \text{school} \quad \text{have} \quad \text{2000} \quad \text{people} \quad \text{ZAA3}
\]

‘our school only has 2000 people’ (Sybesma & Li’s gloss and translation)

Sybesma and Li claim that the same sentence with *ze1* in place of *zaa3* would not be neutral and would mean that 2000 is being thought of as a big number. They believe the distinction between *zaa3* and *ze1* is due to a difference in how a number or event is perceived. Setting aside the fact that ‘2000 people *ze1*’ could actually imply that 2000 people is being considered few, why this asymmetry whereby one particle is neutral and implies that 2000 is factually small, while the other particle is non-neutral and implies that 2000 is a big number? It seems more logical to instead say that *zaa3* is not neutral, but is in this example adding some meaning which presents 2000 as a small number.

We can see a similar difficulty with another of their examples, ‘he is only 50 years old’, provided below as (7.5). Sybesma and Li (2007, 1755-1756) claim that the sentence below is a ‘neutral, factual statement… used when the age of 50 is deemed relatively young’. This description is perplexing as, presumably, there could be many people who find the age of 50 to be relatively old. Their own description implies that the sentence is not truly neutral. Surely the more neutral sentence would be simply ‘he is 50 years old’, and the addition of ‘only’ and/or *zaa3* is what conveys the additional, non-neutral idea that 50 is young. As with their first example, *zaa3* is clearly conveying some additional meaning compared with the same statement without any particles.

(7.5)

\[
\text{keoi5} \quad \text{ng5-sap6} \quad \text{seoi3} \quad \text{zaa3} \\
\text{3s} \quad \text{50} \quad \text{year} \quad \text{zaa3}
\]

‘he is only 50 years old’ (Sybesma & Li’s gloss and translation)
Again, Sybesma and Li contrast this with a sentence which is identical save for the utterance particle being replaced by ze1. They claim that this ze1-suffixed statement also means ‘he is only 50 years old’, but would express that 50 is not old enough. In the context of retirement, therefore, the zaaz3-suffixed statement would imply ‘wow, so young!’ whereas the ze1-suffixed statement would imply ‘too early!’ As before, this chapter argues that both interpretations are available for the zaaz3-suffixed statement. Additionally, it is confusing that Sybesma and Li translate both the zaaz3-suffixed sentence and the ze1-suffixed sentence equally into English as ‘he is only 50 years old’. Could not this English sentence imply, equally, both ‘wow, so young’, as well as ‘too early’? Their translations have not shed any light on why the two particles zaaz3 and ze1 seem to add different colouring to an otherwise identical sentence.

According to Wakefield (2012b, 1-2), the meaning of only, and therefore zaaz3, can be paraphrased in NSM as ‘it is not more than this’. Unfortunately, the comparative construction is not universal and does not adhere to the rules of NSM. The English word ‘than’, though seemingly basic, is not an NSM semantic prime, and cannot be translated into all languages. Wakefield’s NSM explication is actually more helpful to us in a different way – in understanding the relationship Wakefield sees between zaaz3 and ze1. According to Wakefield (2012b, 2), the explication for ze1 is as follows:

**Ze1 (explication proposed by Wakefield):**

a. it is not more than this  
b. you think this much/many is bad  
c. I want you to know this much/many is not bad

Ignoring the non-NSM syntax and whether the proposed meaning of ze1 is accurate or not\(^7\), the relevant point to note about Wakefield’s explications is that he believes the meaning of zaaz3/only is supposedly entailed in the meaning of ze1. That is, his explication for zaaz3/only doubles as line (a) of his explication for ze1. This appears to follow the ideas of Fung (2000) and Sybesma and Li (2007) who suggested that zaaz3 is somehow ‘neutral’ and that ze1 adds to that. As mentioned above, the argument that zaaz3 is ‘neutral’ is

---

\(^7\) According to Wakefield’s explication, ze1 should be fairly positive – it is pointing out that something which the other person thinks is bad is actually not bad. But ze1 can be used for negative views too.
puzzling, and there appears to be no evidence to support Wakefield’s claim that 
zaa3 is semantically simpler than ze1.

Strangely, Wakefield makes a contrary point in a different paper, in which 
he states that zaa3 and ze1 do not mean *only* (Wakefield 2011b, 250-251):

A very interesting contrast between the particles zaa3/ze1 and English’s
‘only’ or Cantonese’s zing6hai6 ‘only’, is that zaa3/ze1 can be suffixed to a
proposition such as ‘It’ll be fast’ to get the meaning ‘It won’t take a long
time’. It is not possible to get this meaning with the adverb ‘only’ in either
English or Cantonese, i.e. ‘It’ll only be fast’ does not mean ‘It won’t take
long’. This does not appear to be a difference in the pragmatics of English
speakers vs. Cantonese speakers; if so, then the Cantonese adverb
zing6hai6 ‘only’ should be able to produce the same meaning as zaa3.
This is a clear indication that not all SFPs [sentence-final
particles/utterance particles] have exact counterparts in English.

Since this passage makes an interesting point, it is unclear why Wakefield
(2012b) overlooks it. He does say that the semantics of ze1 is more complicated
than what is shown in the explication provided, partly because ze1 does not
always include the meaning of *only* (Wakefield 2012b, 2). But if ze1 does not
always include the meaning of *only*, then how can ‘it is not more than this’ – i.e.
the supposed meaning of zaa3/only – be included in the explication of ze1?

Wakefield (2012b) explains that he assumed ze1 to be polysemous,
although he does not reveal what the other meanings of ze1 might be. As
explained, it is generally safer to initially treat all words as having a single
meaning, until or unless there is firm evidence to the contrary. Assuming in
advance, before or without semantic analysis, that words are polysemous makes
it possible to settle on incomplete or inaccurate explications, as it is all too easy
to disregard and dismiss complications as issues for a different problem.
Whether or not zaa3 means *only*, and whether or not the meaning of zaa3 is
included in ze1 cannot be known unless the meanings of both particles are first
reliably and accurately identified.

It is also worrying that Wakefield (2012b) does not show any examples of
zaa3 or ze1 from real, natural conversation. NSM semantic analysis works best
when there are many naturally-occurring examples. Despite the limited number
of semantic primes, the primes can be combined in diverse combinations to construct a potentially large number of possible explications, and simply guessing from intuition alone rarely, if ever, yields a correct explication. The semantic analysis presented in this chapter hopes to demonstrate exactly this.

As a final idea in the literature concerning zaa3, consider briefly Matthews and Yip’s (2011, 391) claim that zaa3 is a contraction of ze3 and aa3. Although the formula zaa3 = ze3 + aa3 fits neatly with other supposed particle formulas like gaa3 = ge3 + aa3, mentioned in Chapter 5, ze3 does not appear to be a particle in use, and the evidence for the contraction seems rather weak. Particle contractions, including the supposed zaa3 contraction, are revisited in Chapter 8. The next section, 7.2, will guide readers from an initial native speaker intuitive understanding of zaa3 (which turns out to be inaccurate) through to an NSM explication substantiated by native speakers’ real, recorded speech.

7.2 NSM semantic analysis of zaa3: testing possible key primes and components using the corpus

We know from examples of zaa3, the literature, and by using native speaker intuition, that zaa3 often seems to have some sort of ‘minimising’ or ‘downplaying’ effect. It is much more difficult to establish an actual invariant meaning that can be stated in a clear and cross-translatable way. This section attempts a linear narrative which shows how the final explication was arrived at. It demonstrates one way in which NSM semantic analysis can be conducted, as well as justifying the final explication proposed for zaa3. The final explication is able to incorporate and explain ideas about zaa3 which have been mentioned in the literature review above. The final explication [7E] is arrived at in section 7.2.10, but is first copied here (without explanatory discussion) to give readers an idea of the direction this analysis takes.

[7E] Final explication for zaa3:

- it is like this, (it is) not more
- someone can feel something because of this
- hai6 gam2joeng2, m4 hai6 do1di1
- is like-this/this-way not is more
It will be shown that the corpus examples are invaluable, and that native speaker intuitive understanding of utterance particles can be very unreliable. Roughly 20 examples from the corpus are presented in the remainder of this chapter. Many potential primes and components are rejected before a final explication is proposed. Possible key primes for the explication, and then full components, will be presented one by one, and each considered against examples of zaa3 from the corpus. At first, it seemed the explication of zaa3 would have to include several descriptors in order to cover different situations in which zaa3 can be used. This resulted in an explication which went along the lines of ‘someone can think like this about it: it is very little~few/small’. Eventually, as more real-life examples of zaa3 from the corpus were considered, some components (e.g. to do with small size) had to be omitted and other components (e.g. to do with time) were included. Additionally, it was felt necessary to include a line in the explication to explain the feeling or reaction someone might have to a zaa3-suffixed utterance.

Any explication must have an equivalent in Cantonese NSM, and it became significant that the Cantonese prime THIS explicitly requires THING in a statement like ‘this is very little~few’. Cantonese speakers are required to say li1 joeng6 je5 hai6… [THIS CL THING IS…]. *Li1 hai6… [THIS IS…] is unacceptable. Of the potential explications considered for zaa3, something like ‘this (thing) is very little~few’ was one of the most plausible, and this translatability issue caused problems because the referent need not be a physical thing. Omitting THING was made possible if THIS was also omitted. The preferred solution at the early stage was to use hai6 how2 siu2. This is literally [IS VERY LITTLE~FEW], i.e. with a zero-subject, and corresponds to English ‘it is very little~few’. This

---

79 The examples presented in this chapter feature more particle clusters than in previous chapters, partly because zaa3 is the least common of the five particles studied, seeming to occur in clusters more often. Several of the examples in this chapter reappear in Chapter 9 when clusters involving zaa3 are discussed.

80 This how2 ‘very’ could truly mean VERY or may be interpreted as a ‘dummy’.
seemed to provide a viable frame for using LITTLE~FEW siu2 in the explication, but complications are discussed further below.

7.2.1 Possible key prime: LITTLE~FEW siu2 (or NOT MUCH~MANY m4 do1)

Since zaa3 appears to have a ‘minimising’ effect, LITTLE~FEW\(^{81}\) was one of the primes which first sprung to mind when searching for an explication. A component like ‘it is very little~few’ seems to work well with many examples. For instance, (7.6) shows zaa3 being used in a sentence about having only said VERY FEW things to someone.

(7.6)

A:
\(\text{Soeng6ci3 nei5 zau6 gam2 gong2 zo2} \) 
last-time you then/just like-this say PFV
\(\text{niidi1 zau6 mou5 gaa4 laa4?} \)
these then not-have PRT PRT
‘Last time you only talked about these and then there was nothing else gaa4-laa4?’

B:
\(\text{Mou5 laa3, gong2 zo2 loeng5 geoi3 zaa3,} \)
not-have PRT say PFV two sentence PRT
\(\text{ganizyu6 ziihau6 dou1 mou5 wan2 keoi5.} \)
following-this after also/all not-have find him
‘No laa3, said two sentences [didn’t speak for long] zaa3, and then after that I didn’t find [speak to] him.

Example (7.7) shows someone talking about how if you buy postage stamps and their value does not rise later, then there are very few (or, not many) advantages or benefits to having them. In this example we can use ‘it is very little~few’ to explain what the speaker means – that there would be few benefits to you in this scenario. Note that using the English word only to explain zaa3 would not have worked well here – it would not make sense to say something like ‘you only don’t have benefits’.

---

\(^{81}\) Recall that LITTLE~FEW is a single prime, as the two English exponents mean the same thing, although English makes the distinction between count and mass nouns. In Cantonese, no such distinction is made and the exponent for this prime is siu2.
Because at the beginning of the year they take/collect all your money in one go, you book/reserve everything for the whole year so/like that then at that time PRT rise/increase then of course you advantage/benefit PRT if not rise/increase that time PRT you then you then really not-have what you then really not-have what not-have what advantage/benefit gaa3 zaa3 nei5 zan1hai6. PRT PRT you really

'Because at the beginning of the year they take/collect all your money in one go, you book/reserve everything for the whole year gaa3-laak3, so then, later on aa3, if it [the value] rises then of course it's good for you laa1, if it doesn't rise nei, you then, you then really don't have much/any, you then really don't have much/any, don't have much/any advantage/benefits gaa3-zaa3 you really.'

LITTLE-FEW also seemed like an appropriate key prime for an explication for example (7.8). A and B are discussing B's student who has a part-time job at Pizza Hut. Both instances of zaa3 are used to talk about a low wage/salary, which can be interpreted as 'very) little money'.

(7.8)

B:

Tung4maai4 bat1gw03 jan4gung1 dou1 hou2 and but wage/salary also/even/still very
ngan1 zaa3 (7.8a) wo3. Zou6 jat1 go3 skinny/little PRT PRT do one CL
jyut6 nei1 dou1 hai6 dak1 month PRT also/even/still is only-have
And but the salary is still very low zaa3-w03. For one month’s work ne1 it’s still only three-thousand-and-something dollars zaa3-w03.’

Notwithstanding its initial appeal, it became clear fairly quickly that ‘it is very little~few’, at least on its own, was not an acceptable explication for zaa3. There are many examples for which this component would not work well. For example, (7.11) and (7.12a, b) are to do with time. Examples (7.14) and (7.15) are both to do with there being one of something. Example (7.22) is about stamps losing their value and becoming worthless if they are not in good condition. While some options such as ‘very little time’ make sense, and ‘very few’ might be acceptable to some people as including ‘one’, these phrases would not be ideal NSM. Furthermore, (7.19) and (7.20) further below both show zaa3 being used, not about a physical thing, nor even about quantity.

One of the earliest considered explications for zaa3 was ‘it is not much~many’. This is essentially the opposite of using LITTLE~FEW. This component turned out to be subject to the exact same problems. The next possibility considered was that of SMALL or NOT BIG.

7.2.2 Possible key prime: SMALL sai3 (or NOT BIG m4 daai6)

Native speaker intuition placed something like ‘it is small’ (or perhaps ‘not big’) as among the most plausible and appealing meaning components for zaa3. Surprisingly, however, it was found that the vast majority of real-life examples from the Hong Kong Cantonese Corpus were not to do with things being small in size. This mismatch between intuition and real data was so undetectable that for some time, analysis continued to include the prime SMALL as an apparently necessary part of an explication of zaa3. There did not appear to be any reason to exclude such a meaning component. It was the near-imperceptible evidence of absence which eventually revealed its being out-of-place. A thorough search through the Hong Kong Cantonese Corpus revealed 137 instances of zaa3 but, astonishingly, only one example, (7.9) presented below, was used in the context of something being small in size. This indicates only 0.7% of zaa3 examples would use the prime SMALL as a descriptor. This was surprising and clearly demonstrates that native speaker intuition can be inaccurate.
A very small number of other naturally-occurring examples of zaa3 from the corpus were related in some way to size, but ultimately would have used LITTLE~FEW in an explication. For instance, the following example could be interpreted as something (the building/area) being SMALL in size, but should more accurately be described as a comment on how FEW rooms there are.

Even previous work such as Kwok (1984, 50-53) states more than once that zaa3 can be used to convey the sense that something is not bigger or not big enough. This supports the notion that size seems intuitively relevant to zaa3. Yet a thorough look at all of Kwok’s 16 examples of zaa3 shows that none of them are to do with things being small in size. Of Fung’s (2000) 20 examples of zaa3, none are in the context of describing something as small. Similarly, Sybesma and Li (2007, 1754-1755) do not provide any examples of zaa3 being used to comment on something’s small physical size.

So where does that leave us? Should SMALL still be included as one of the options in an explication of zaa3, or is its occurrence so rare that the one example found in the corpus can be written off as an anomaly? On one hand,
given the extremely low percentage of SMALL’s relevance, that example could be disregarded as just an exception. On the other hand, there do not appear to be any examples of zaa3 which could be described as NOT SMALL, and it is possible to construct an imaginary but natural-sounding example of zaa3 being used for something’s small size, which would support the native speaker intuition of myself and others such as Kwok. Ultimately, the real, naturally-occurring data should take precedence, and the data available essentially shows evidence of its absence. Therefore, a meaning component including SMALL will be omitted from the proposed explication of zaa3. Another option is to consider the phrase ‘like something small’, but for now we can say that other primes such as LITTLE~FEW discussed above, as well as some others to be discussed below, appear to be more applicable to zaa3.

7.2.3 Possible key prime: A SHORT TIME jat1 zan6

If a component focussed on the prime SMALL is eliminated, the only prime we have reason to include up to this point is LITTLE~FEW. However, as more examples of zaa3 were studied, an increasing number of examples were found to be better described by the NSM prime A SHORT TIME, rather than by LITTLE~FEW. Some time-related examples can be argued to use LITTLE~FEW, as in ‘very little time’ or hou2 siu2 si4 gaan3, but this was not the most ideal in all cases. For example, A SHORT TIME describes (7.11) below better than LITTLE~FEW. In this example, A and B are talking about postage stamps, and B asks A about old ones, specifically with the British queen on. A explains that he has not been collecting stamps for that long and does not have any with the Queen on. A actually says in this instance ‘a very short time’ (rather than something like ‘very little time’).

(7.11)

B:

Gam2 hai6 jing1, e6, jing1neoi5wong4 ziicin4 ne1?
Then/so is British eh British-female-royal before PRT
Kei4sat6 nei5 gei2si4 hoici2 cou5 jau4piu3 gaan3?
Actually you when start collect postage-stamp PRT
‘Then eh, before the Queen ne1? When did you start collecting postage stamps gaan3?’
A: *Kei4sat6 ng05 cou5 jau4piu3 ge3 si4gaan3*
Actually I collect postage-stamp that time
*hou2 dyun2 gaa3 zaa3.*
very short PRT PRT

‘Actually I’ve been stamp-collecting for a very short time *gaa3-zaa3.*’

Example (7.12) below is from a conversation about collecting postage stamps. The speaker is saying that some time ago, before he realised that $1.80 stamps would eventually be worth a lot more, he thought they would become useless, and so he tried to quickly use all the $1.80 stamps that he had. In this excerpt he uses the particle *zaa3* twice. Example (7.12a) is clearly better explicated using **A SHORT TIME** than **LITTLE~FEW**. Example (7.12b), though, could indicate that one or two months is **A SHORT TIME**, or that one or two months is very **FEW**.

(7.12) ...
...faai3faai3ceoi3ceoi3 maang5gam3 zoeng1 di1
quickly ferociously/furiously take/put CL
go3baat3 nim4 nim4 saai3, waa3, nam2 haa5, ei3
one-eight stick stick all wow think DEL hey
lau4 hai2dou6 mou5 jung6 gaa3 maa3, gei3
keep/stay here/there not-have use PRT PRT send
gei3 gei3. Hou2 dyun2 si4gaan3 gaa3 zaa3 (7.12a)
send send very short time PRT PRT
hai2 ng05 gei3 saai3, zik1hai6 jau4 keoi5
at I send all/complete meaning from it/they
hooi1ci2 gaa1 jau4piu3, zi3dou3 ng05 gei3
start add/increase postage-stamp until I send
saai3 di1 go3baat3 ge3 jau4piu3, zo3
all/complete CL one-eight those postage-stamp again
zi3dou3 go3baat3 hooi1ci2 sing1zik6 ge3 si4hau6
until one-eight start increase-in-value that time
ne1, zi2hai6 go3 leng4 loeng5 go3 jyut6
PRT only CL/one zero two CL month
gaan1 ge3 si6 lai4 gaa3 zaa3. (7.12b)
space/between that event come PRT PRT
Go2di1 go3baat3 jau4piu3 ne1, waa3, sing1
Those one-eight postage-stamps PRT wow rise/increase
dou3 zan1hai6 din1.
until really crazy
‘... quickly, ferociously take the $1.80 [stamps] stick them all, wow, think
about it, hey keeping them here/there is useless gaa3-maa3, send send
send. A very short time gaa3-zaa3, from when I sent them all, I mean
from the time they increased the postage stamps, until I’d sent off all the
$1.80 postage stamps, and then until the time the $1.80 started increasing
in value ne1, only took one or two months’ time gaa3-zaa3. Those $1.80
postage stamps ne1, wow, rose crazily.’

Similarly, (7.13) could be interpreted as meaning that two days is a SHORT TIME,
or that two days is very FEW. A and J are talking about A’s recent holiday. On
the first of five days, their flight arrived at almost midnight. On the third day,
they travelled for nine hours from one area to another area. On the last of the
five days, their flight left immediately after breakfast. Therefore, A is saying the
holiday was not good, as out of five days, they only had two days for activities.
This example was also shown in Chapter 5.

(7.13)
A:
Gei1bun2soeng6 ne1 waan2 ne1 hai6 dak1 loeng5
basically PRT play PRT is only two
jat6 gam3 do1 zaa3.
day like-this many PRT
‘Basically ne1 play [/do things] ne1 was only for two days zaa3.’

Notice that example (7.13) clearly conveys a sense of ‘not enough’. Examples
above, like (7.8), also convey this sense. However, consider examples like (7.11)
or (7.12), which would instead be interpreted as ‘not too much’. Hence, the
supposed ‘not enough’/‘not too much’ distinction that Kwok (1984, 53) and
Fung (2000, 59-60) endorse (see section 7.1) appears to be inaccurate – zaa3
can express both. Moreover, other examples shown in this chapter show that
some instances of zaa3 cannot be easily categorised into either ‘not enough’ or
‘not too much’.

While a SHORT TIME and LITTLE~FEW are different primes, the two are
logically related and are both to do with a ‘small’ amount of something. There

82 Unlike in English, the Cantonese exponent for LITTLE~FEW, siu2, cannot directly be used to
say e.g. *siu2 jat6 [few days] or *siu2 jyut6 [few months]. Instead, a Cantonese speaker can say
that e.g. two days/months ‘is very little~few’ hai6 hou2 siu2.
may be examples where both A SHORT TIME and LITTLE~FEW are acceptable. What is more important is that there are some examples where A SHORT TIME is acceptable, but not LITTLE~FEW, and vice versa. Therefore, both are necessary. Use of A SHORT TIME to refer to time ensures that the explication makes sense in all languages (as opposed to ‘little time’, which some languages might be unable to express). In English it is possible to say ‘small amount of time’, but this is not NSM. In any case, the Cantonese exponent sai3 cannot be used to describe lengths or amounts of time.

7.2.4 Possible key prime: ONE jat1

Besides examples of zaaz3 to do with A SHORT TIME, there were another group of examples which could not be explicated satisfactorily using LITTLE~FEW – where speakers talk of there being ONE of something. For example, in (7.14), the speakers are talking about postage stamps, and zaaz3 accompanies an utterance about there only being one set of stamps released. In (7.15), the speakers are talking about how A’s friend wants to study physiotherapy, but it is only taught at one university (example (7.15) was also shown in Chapter 6). Note that (7.14) and (7.15) happen to show questions ending in zaaz3-me1. These examples may challenge Kwok’s (1984, 41) report that zaaz3 is suffixed to statements without altering their grammatical status as statements (although of course these examples are strongly affected by me1, which is an interrogative particle).

(7.14)

A:
Ng5jyut6, luk6jyut6 gei2 tou3 jau4piu3
cfive-month six-month few/some set/collection postal-ticket
cceot1.

tout
(In) May, June, (there will be) a few sets of postage stamps coming out.’

B:
Gei2 tou3 aa4? Keoi5 m4hai6 waa6 dak1
few/some set/collection PRT he/she/they not-is say only
jat1 tou3 zaaz3 me1?
one set/collection PRT PRT
‘A few sets aa4? Didn’t he/she/they say (there will be) only one set zaaz3-
me1?’
A: 
*M4hai6...*
not-is
‘No...’

(7.15)
B: 
*Bin1 gaan1 U aa3 nam2 zyu6 gaan2?*
which CL U [university] PRT think CONT choose
‘Which university is he thinking about *aa3*?’

A: 
*Poly ze1 maa3 dak1.*
Poly PRT PRT only-have
‘There is only Poly [Hong Kong Polytechnic University] *ze1-maa3.*’

B: 
*Hai6 me1? Dak1 Poly jau5dak1 duk6*
is PRT only-have Poly can/have read/study
*Physiotherapy zaa3 me1?*
Physiotherapy PRT PRT
‘Really *me1*? Only Poly offers/teaches physiotherapy *zaa3-me1*?’

A: 
*Hai6 aa3. Kei4taa1 dou1 mou5 gaa3 laa3.*
is PRT other all not-have PRT PRT
‘Yes *aa3*. The others don’t have it *gaa3-laa3.*’

Could *LITTLE~FEW*, or *VERY LITTLE~FEW*, be used to refer to one item only? It might be possible to think of *ONE* as *VERY FEW*, such as if one exclaims ‘So few! Only one’. But though there is clearly a relationship or similarity between the two, *VERY LITTLE~FEW* does not mean *ONE*. Moreover, NSM primes are meant to be basic to the extent that they cannot be further decomposed using reductive paraphrase, and saying that *VERY LITTLE~FEW* could mean *ONE* would weaken the accuracy of the NSM. Consider examples (7.16) and (7.17) below.

(7.16)
B: ‘But I want the most to go to Australia, New Zealand that area, because I still haven’t been.’

A: ‘Really *aa4*? You haven’t even been to Australia *me1*?’

B: ‘Mm, not yet *aa3*.’
A: ‘Oh, I thought you just haven’t, haven’t been to New Zealand tim1. But I also want to go to New Zealand.’

B: M6, mai6 heoi3 can1 dou1 hai6 lang3 mm so-it-is go whenever also/all is connect/hang maai4 nau5sai1laan4 jat1cai4 gaa3 laa1?
as-well New-Zealand together PRT PRT

‘Mm, [isn’t it] whenever you go it’s always lumped together with New Zealand gaa3-laa1?

A: M4hai6 aa3. Soeng6ci3 ngo5 zing6hai6 heoi3 ou3zau1 not-is PRT last-time I only-is go Australia zaa3. Mou5 heoi3 nau5sai1laan4 aa3. PRT not-have go New-Zealand PRT

‘No aa3. Last time I only went to Australia zaa3. Didn’t go to New Zealand aa3.’

(7.17)

Ngo5 loeng5 go3 B saam1 gaa3. Jat1 go3 I two CL B three PRT one CL

B sei3 zaa3. (7.17a) Zau6 hai6 go3 zung1man2 B four CL then/just is CL Chinese
gam3 seoi1 gaa3 zaa3. (7.17b) this/so bad PRT PRT

‘I got two B3s [the better subdivision of B] gaa3. One B4 [the worse subdivision of B] zaa3. Just the Chinese that was so bad gaa3-zaa3.

Looking first at the zaa3-suffixed statement in (7.16), soeng6ci3 ngo5 zing6hai6 heoi3 ou3zau1 zaa3 ‘last time I only went to Australia zaa3,’ it would not be accurate to say that the speaker went to VERY FEW places. It would be more accurate to say that she went to ONE place, namely Australia. (It could be that she visited several places within Australia, but from the context we see that here she means one place, i.e. Australia, and not the suggested two places in the conversation, i.e. Australia and New Zealand.) Similarly, the speaker in (7.17) is saying that there was only ONE subject in which he got an exam mark of B4 (which is a lower grade than B3). Following the same reasoning, examples (7.14) and (7.15) should also use ONE.
Recall that in the earlier discussion, it was explained that although zaa3 frequently seems to be used in a negative way, this is not a requirement or invariant meaning of zaa3. Kwok (1984, 53) also pointed out that zaa3 frequently seems to have a negative value, easily conveying the idea of insufficiency, although this was not an essential part of the particle’s meaning. Example (7.17) shows that this is true - the speaker is actually boasting about his good results. This is part of a fairly lengthy discussion between two people about their examination results. In (7.17a), the speaker is proud to have gotten only one B4 grade, and his friend is also surprised and impressed at his good grades. Example (7.17b), though it is expressed as a slight complaint, really serves to reinforce that his other marks were good.

Even though LITTLE~FEW and ONE are separate primes, it seems obvious that they are related to small amounts or to being ‘minimal’ in some way. This is like LITTLE~FEW and A SHORT TIME (and even SMALL), where even though they are separate concepts, they seem logically similar. As above, since VERY LITTLE~FEW and ONE are not equivalent and are both necessary, the NSM explication would need to include both.

7.2.5 An early partial explication for zaa3

Based on the concepts and primes found necessary so far, the partial explication labelled [7A] below was created for zaa3. At this point, it appears that the explication needs to make mention of LITTLE~FEW, A SHORT TIME, and ONE. The prime VERY is used with LITTLE~FEW and A SHORT TIME to ‘intensify’, which also helps give the feeling that it is worth noting and sometimes unexpected. Originally, one line featured ‘it is very small’, but this is omitted below due to the reasons discussed above.

[7A] Early partial explication proposed for zaa3:

someone can think like this (about it):
’it is/there are very little~few
/ it is a very short time
/ there is one thing/place/someone’

If adopted, explication [7A] could have been regarded as a case of polysemy, whereby three similar meanings exist for zaa3. In effect, each of the separate
possibilities (indicated by the three alternative final lines) could have been assigned as a separate explication. Against this, however, was the strong impression that native speakers would not endorse such an interpretation. Polysemy is usually ‘noticeable’ to native speakers if they think about it.

Sections 7.2.6 to 7.2.9 below consider other possible components for \textit{zaa3}, which could be used in addition to the partial explication [7A] above. These are motivated by the intuition that \textit{zaa3} seems to do more than just point out that something is \textit{LITTLE~FEW/A SHORT TIME/ONE}. Some time before the partial explication above was reached, it had already become clear that an extra component was necessary, possibly to do with a subjective feeling or reaction. Some of the considered components are discussed below.

**7.2.6 Possible components involving ‘not what you think’ and ‘true’**

An early considered meaning component was ‘it is not what you think’. The initial idea was that this could be followed by a meaning component which conveyed that something was less than what one thought. For example, in (7.16), speaker B thinks that holidays to Australia always include New Zealand, but speaker A corrects her faulty assumptions – speaker A could be saying that it is ‘not what you think’. Also, in (7.17), the speaker could be saying something like ‘I got better grades than you think’.

This possible component was rejected when natural examples were found to show \textit{zaa3} being used even when speakers were thinking the same thing. For instance, example (7.18) shows two friends agreeing on something. In (7.18b), speaker A seeks speaker B’s confirmation that there is not much to eat at Disney World. Speaker B confirms that there are only hamburgers and fries, and speaker A agrees. This shows that a component like ‘it is not what you think’ would not be fitting because the two speakers are thinking the same thing (i.e. that there is not much to eat at Disney World). Note that (7.18b) is not easily translatable into English using the word \textit{only}.

(7.18)

\begin{verbatim}
A: Daan6hai6 dou1 gei2 gwai3 wo3. Dou1 jiu3
but still quite expensive PRT still need
\end{verbatim}
seng4 maan6sei3 man1, sei3 ng5 jat6. Zing6hai6
whole fourteen-thousand dollar, four five day only
heoi3 Orlando go2dou6 waan2 zaa3 (7.18a) w03.
go Orlando that-place play PRT PRT
Mou5 je5 zou6.
not-have thing(s) do

‘But it’s still quite expensive w03. Still need $14 000, [for] four or five
days. Only going to play at that Orlando place zaa3-w03. Nothing to do.’

B:  
Gam2 mai6 sik6 tung4 waan2 mai6 aam1. Daan6hai6
so so eat and play so correct but
ngo5 –  
I
‘So eat and play that’s right. But I –’

A:
Hai6 leoi5bin6 sik6 lo1, leoi5bin6 mou5 mattje5
is inside eat PRT inside not-have what-thing
sik6 gaa3 zaa3 (7.18b) w03. Hai6 m4hai6 aa3?
eat PRT PRT PRT is not-is PRT
Nei5 dou1 heoi3 gwo3 laa1.
you also go EXP PRT

‘Eat inside lo1, there’s not much to eat inside gaa3-zaa3-w03. Right aa3?
You’ve also been laa1.’

B:  
Hai6 sik6 go2di1 mattje5 hon3bou2baau1,
is eat those what-thing hamburger
‘Just eat those what hamburger;’

A:  
Mai6 hai6 lo1.
so is PRT
‘Yeah lo1.’

B:  
Syu4tiu2 gaa3 zaa3 (7.18c)
french-fries PRT PRT
‘French fries gaa3-zaa3.’

Some previously considered meaning components for zaa3 indicated that
something might be unexpected or different from the norm, which led to
explication variants such as, roughly, ‘this is not like what you are thinking’. Such a meaning component would have been plausible for (7.18a, b), but does not work for (7.18c). In other words, speaker A in (7.18b) might have meant ‘there’s not much to eat inside, contrary to what you think’. However, speaker B in (7.18c) could hardly have meant to say to speaker A that there are only French fries to eat and that this is not what speaker A was thinking, since speaker A is the one who brought up this point in the first place about there not being much to eat. The zaaz in (7.18c) is more likely to be helping speaker B express that hamburgers and fries make for very few options, or perhaps that they are ‘low quality’ foods.

Such meaning components would also not have worked well with zaaz as used in questions. It is possible that the person asking the question wants to cast doubt on something, but this was considered unlikely in most cases. For instance, in (7.14), speaker A says there will be a few sets of postage stamps coming out, and speaker B says ‘Didn’t he/she/they say (there will be) only one set zaaz met?’ This could perhaps be interpreted as speaker B using a ‘polite’ way to challenge A, but in the context of the rest of the conversation this seems unlikely – B is constantly asking A questions about postage stamps and seems to want to learn from A, who is explaining a lot. Similarly, in (7.15), B’s question to A seems genuine.

Examples (7.14) and (7.15) of zaaz being used in questions also helped to eliminate several other considered meaning components, such as ‘this is true’, ‘other things are not true’, and even ‘don’t think something else/other things’. These meaning components were considered because zaaz often seems to be used in utterances where the speaker is trying to be convincing in some way. But questions, as in examples (7.14) and (7.15), show that this is not an invariant meaning of zaaz. That something is true seemed not to be the main motivation in using zaaz in an utterance.

7.2.7 Possible component involving ‘it can’t be... like this’

At an earlier point in the analysis (when the first part of the explication still used the options LITTLE–FEW and SMALL), an extra component considered was ‘it can’t be like this’, resulting in explication [7B.1] below. This last component was
intended to convey something like ‘unexpectedness’ and that something was 'surprisingly' LITTLE–FEW etc., such that it was noteworthy in some way. This was rejected though, because this is not always the case, and because ‘it can’t be’ implies disbelief. For example, in (7.18) above, speaker A uses zaa3 twice (7.18a and b) to convince speaker B that there is not much to do or eat at Disney World. ‘It can’t be like this’ would have indicated that she was surprised, but she clearly was not, as she is the one trying to convince B. In example (7.18c), it is actually B that uses zaa3, agreeing with speaker A. Again, it seems implausible that zaa3 includes the meaning ‘it can’t be like this’, because B herself has been to Disney World and is agreeing. It could be argued that the ‘someone’ in the explication could be a third person, but this feels like ‘stretching’ the explication to make it work.

[7B.1] Early proposed explication for zaa3:

someone can think like this (about it):
  ‘it is very little–few/small
  it can’t be (little–few/small) like this’

Another possible argument is that ‘it can’t be like this’ means something like ‘this is not acceptable/possible’, but this seems to be moving further from the meaning of zaa3, and can quickly be rejected after considering other examples mentioned above, such as example (7.16) ‘I only went to Australia zaa3’. Example (7.14) also shows that this early explication should be rejected. In this example, speaker B has been told there will be several sets of postage stamps coming out, and he is asking something like ‘Didn’t they say there will only be one set of postage stamps?’ Ignoring for now the problem with ONE and LITTLE–FEW, the component ‘it can’t be like this’ would not make sense since speaker B is asking a question, and the way that the question is asked shows that speaker B expected there to only be one set. He cannot be expressing ‘it can’t be [little–few/one] like this’. The problem is not due to the fact it is a question, but rather due to the expectations of the speaker. (Example (7.15) features a very similar question ‘Only the Polytechnic offers physiotherapy?’ but with expectations reversed. That is, the speaker expected there to be more universities which offer physiotherapy courses, but finds out that there is only
one. In that case, it is plausible that she was expressing something like ‘there can’t only be one place’.)

Another problem with this proposed component is that when the first part of the explication is updated to [7A] as per the earlier discussion, the explication seems slightly illogical. With SMALL omitted and A SHORT TIME and ONE inserted, the explication would look like that below. For example, this explication states that it is possible that someone think like this: ‘there is one thing, it can’t be [one thing] like this’. This would make little sense in a typical scenario where the speaker is telling another that there is one of something. Thus the extra component which seems to be missing cannot be ‘it can’t be... like this’.

[7B.2] Early proposed explication for zaa3:

someone can think like this (about it):
‘it is/there are very little~few
/ it is a very short time /
/ there is one thing/place/someone
it can’t be like this’

7.2.8 Possible component involving ‘this someone can feel something because of this’

The next component which was considered was ‘this someone can feel something because of this’. This seemed better than the previously considered ‘additional’ components discussed in the last few sections. Note that this component is incompatible with the idea that zaa3 is ‘neutral’, as Sybesma and Li (2007) and Wakefield (2012) said. The new explication is presented below, labelled [7C].

[7C] Early proposed explication for zaa3:

someone can think like this (about it):
‘it is/there are very little~few
/ it is a very short time /
/ there is one thing/place/someone
this someone can feel something because of this

We can revisit examples (7.14) and (7.18) to check the suitability of [7C], since these examples posed difficulties for many of the previously considered components. Example (7.14) did not work well with ‘it can’t be like this’, because the zaa3 in this example was being used in a question. With explication
[7C], however, no such problem arises. The line ‘this someone can feel something because of this’ is deliberately vague as to exactly what the person felt. In this example we can speculate that speaker B felt something like surprise, because he thought that there would only be one set of stamps, but is now being told there will be a few. It is also possible that he felt any other emotion, for example, confusion or disbelief. Perhaps he felt good or happy because there were more than he expected, or anger at having been given the wrong information previously. In any case, the explication would still work. Example (7.15) is the same. The last line of this explication ‘this someone can feel something because of this’ still works even if the line preceding it is ‘there is one thing’.

In example (7.18a), speaker A could be thinking: ‘there is only one place’ (or even ‘there are very few things to do/play’), with the added component pointing out that she ‘can feel something because of this’. In this case, speaker A would likely be feeling something bad or negative, as she is commenting that it is a very expensive trip for such a limited itinerary. This works well with the explication. Examples (7.18b, c) are similar to (7.18a) – each speaker can feel something bad or negative because of the limited food options.

Similarly to example (7.18), speaker A in (7.13) is telling J about her holiday. If the final line of the explication is ‘this someone can feel something because of this’, it is plausible that speaker A feels, for example, angry at not spending enough time abroad. Explication [7C] is much better than explication [7B.2] for this example of zaa3. For A to express ‘it can’t be like this’ would have implied she was in disbelief, which is unlikely given that the speaker is the one informing the other.83

Example (7.8) also works better with the line ‘this someone can feel something because of this’, rather than with the line ‘it can’t be... like this’. Since in this example speaker B is informing speaker A of a fact (that their friend earns very little), it would be odd for her to say ‘it can’t be like this’, but it

83 If ‘it’ is meant to refer to good holidays or some such, it does not flow well because the preceding ‘it’ in ‘it is a very short time’ refers to ‘two days’ (/‘it is very few days’). If the ‘someone’ in the top line of the explication were a third person who was surprised, the explication is not straightforward enough.
it makes good sense for her to mean ‘this someone can feel something because of this’ – perhaps this could be a feeling of surprise or sympathy.

In example (7.7), we could explain what the speaker means by using proposed explication [7C]. The speaker is saying that someone can think about the situation like this: that there would be very few benefits if one bought stamps and their value did not increase, and that this someone would feel something because of this. In this case, the feeling would probably be something negative, such as being disadvantaged or having wasted money.

Because of the fact that zaα3 is often used to point out how little~few etc. something is (with a ‘minimising’ or ‘downplaying’ effect), the added line ‘this someone can feel something because of this’ often refers to bad or negative feelings. Nonetheless, as we have seen, this is not always the case. For example, we saw with (7.14) that it was possible the speaker felt happy, and in (7.17) that the speaker felt proud. In summary, then, the proposed component ‘this someone can feel something because of this’ seems very promising.

Still, the first section of the explication (with its three similar-but-different ‘minimising’ components) was not very satisfactory. At this point, a new option came into consideration: ‘it is not more’.

7.2.9 Possible component involving ‘it is not more’

It was observed that when zaα3 is used with utterances about non-physical things, this posed problems for early explications, because in such cases none of the ‘minimising’ components (‘very little~few’, ‘very short time’, ‘one (thing/place/someone)’) seemed to fit well. The prime THING was deliberately omitted in all considered explications and components, but the explication still caused problems. Examples (7.19), (7.20), and (7.21) are cases in point. If one says, as in example (7.19), ‘it’s hearsay zaα3’, the only possible interpretation might have been ‘there are very few facts’ or ‘there is very little evidence’, but this is quite a stretch, as the topic of conversation is not the amount of facts or evidence. The same is true of (7.20), where the speaker says something like ‘everybody seems to be guessing’. Both are from a conversation about postage stamps. Example (7.21) could possibly be interpreted as meaning that there is only one person, Lily, who said something, but this is again a stretch, and is not
the most natural interpretation of the sentence. The most natural interpretation would be something like ‘Lily told me, so I don’t know how reliable this information is’.

(7.19)

\[ \begin{align*}
\text{Batigwo3} & \quad \text{hou2ci5} \quad \text{waa6,} \quad e6, \quad \text{keoi5} \quad \text{wu}i5 \quad \text{sau1} \\
\text{but} & \quad \text{very-similar} \quad \text{say} \quad \text{eh} \quad \text{they} \quad \text{will} \quad (\text{re})\text{collect} \\
\text{faan1} & \quad \text{di1} \quad \text{neoi5wong4} \quad \text{tau4} \quad \text{jau4} \quad - \quad \text{zik1hai6} \quad \text{jau5} \\
\text{back} \quad \text{CL} \quad \text{female-royal} \quad \text{head} \quad \text{postal} \quad \text{meaning} \quad \text{have} \\
\text{neoi5wong4} & \quad \text{tau4} \quad \text{ge3} \quad \text{jau4piu3}, \quad \text{daan6hai6} \quad \text{ne1}, \quad \text{zau6} \\
\text{female-royal} \quad \text{head} \quad \text{LP} \quad \text{postage} \quad \text{stamp} \quad \text{but} \quad \text{PRT} \quad \text{then} \\
\text{m4hai6} & \quad \text{sap6zuk1} \quad \text{gaa3} \quad \text{gam2} \quad \text{sau1} \quad \text{faantlei4}, \\
\text{not-is} \quad \text{full} \quad \text{price} \quad \text{like}-\text{that} \quad (\text{re})\text{collect} \quad \text{back} \quad \text{here} \\
\text{teng1gong2} & \quad \text{zaa3}. \quad \text{Jau4guk2} \quad \text{hou2ci5} \quad \text{keoi5} \quad \text{waa6} \quad \text{hear-say} \quad \text{PRT} \quad \text{post-office} \quad \text{very-similar} \quad \text{they} \quad \text{say} \\
\text{m4} & \quad \text{zi1} \quad \text{ho2nang4} \quad \text{baat3} \quad \text{zit3} \quad \text{aa3...} \\
\text{not} \quad \text{know} \quad \text{maybe} \quad \text{eight} \quad \text{break/discount} \quad \text{PRT} \end{align*} \]

‘But I think they/some people said, eh, they will take back the queen’s head postal – I mean the postage stamps with the queen’s head, but ne1, not just taking them back at the full price, from what I heard zaa3. I think the post office said they I don’t know maybe (take them back at) at 80% [of the original] price aa3...’

(7.20)

\[ \begin{align*}
\text{zik1hai6} & \quad \text{mattje5} \quad \text{zok3} \quad - \quad \text{zik1hai6} \quad \text{go2di1} \quad \text{gau3si1} \quad \text{aa3}, \\
\text{meaning} \quad \text{what} \quad \text{make} \quad \text{meaning} \quad \text{those} \quad \text{idea} \quad \text{PRT} \\
\text{zik1hai6} & \quad \text{bin1903} \quad \text{hai6} \quad \text{cit3gai3} \quad \text{aa3}, \quad \text{ji5} \quad \text{mattje5} \quad \text{wai4} \\
\text{meaning} \quad \text{who} \quad \text{is} \quad \text{design} \quad \text{PRT} \quad \text{by} \quad \text{what} \quad \text{be} \\
\text{zyu2ta1i4} & \quad \text{aa3}, \quad \text{dim2joeng2} \quad \text{joeng2}, \quad \text{ham6baang6laang6} \\
\text{theme} \quad \text{PRT} \quad \text{how} \quad \text{like} \quad \text{all} \\
\text{cyun4bou6} & \quad \text{mou5} \quad \text{jan4} \quad \text{gong2} \quad \text{gw03}, \quad \text{so2ji5} \\
\text{all} \quad \text{not-have} \quad \text{people} \quad \text{say} \quad \text{EXP} \quad \text{that’s} \quad \text{why} \\
\text{go3g03} & \quad \text{hou2ci5} \quad \text{mung4caa4caa4} \quad \text{aa3}, \quad \text{hou2ci5} \quad \text{CL CL (everyone)} \quad \text{seems like} \quad \text{confused} \quad \text{PRT} \quad \text{seems like} \\
\text{dyun3gu2} & \quad \text{gam2} \quad \text{gaa3} \quad \text{zaa3}. \\
\text{guess} \quad \text{like that} \quad \text{PRT} \quad \text{PRT} \end{align*} \]

‘I mean what will make – I mean those ideas aa3, I mean who is doing the design aa3, using what as the theme aa3, how it will be like, all of these nobody said anything about, so everyone seems confused aa3, just like guessing gaa3-zaa3.'
Prompted by these problems, another possibility, ‘it is not more’, came under consideration. Such a component could be acceptable even when the referent is not a physical thing. Two full explications which were considered are shown below, labelled [7D.1] and [7D.2]. Each explication uses the new component ‘it is not more’, with [7D.1] using it in addition to the three previously-considered ‘minimising’ options. In [7D.2], the new component is used in place of those three options.

[7D.1] Early proposed explication for zaa3:

someone can think like this (about it):
  ‘it is/there are very little~few
  / it is a very short time
  / there is one thing/place/someone
  it is not more’
this someone can feel something because of this

[7D.2] Early proposed explication for zaa3:

someone can think like this (about it):
  ‘it is not more’
this someone can feel something because of this

Weighing these two alternative explications against one another, the new line ‘it is not more’ works better on its own, i.e. as in [7D.2] (examples like (7.19), (7.20), and (7.21) still support the rejection of the other three lines). The final
line ‘this someone can feel something because of this’ appears still to be valid; for instance, in examples (7.19), (7.20), and (7.21), someone could feel doubtful or unsure.

Another example of zaaz is provided below in (7.22), which also supports explication [7D.2] over [7D.1]. (This example was also shown in Chapter 6, and is part of the conversation about postage stamps.) Zaaz here would not have been described well using ‘it is/there are very little~few’, ‘it is a very short time’, nor ‘there is one thing/place/someone’. It can be explained by ‘it is not more’ quite well though.

(7.22)
A:  
_Batigwo3 nei5 m4hou2 zing2 faa1, zing2 zip3, but you don't make patterned/dirty make fold_  
jau5 jam6ho4 syun2wai2 laa3, jyu4gw02 m4hai6 zau6 have any damage PRT if not-is then  
ham6baang6laang6 mou5, zik1hai6 gaa3-, mou5, dim2 all not-have meaning val- not-have how  
aa3? Mou5 go2 go3 maai6 faan1 ceot1heoi3 ge3 PRT not-have that CL sell return/back out-go LP  
gaa3zik6 lo1, bin3 zo2 nei5 zan1hai6 wing5jyun5 zau6 value PRT change PFV you really forever so/then  
zi6gei2 zaa1 zyu6 lai4 tai2 gaa3 zaaz3. self hold CONT come/to watch/look at PRT PRT  
‘But don’t make it dirty [/write on it], fold it, or damage it in any way laa3, otherwise you’ll lose all, I mean val – it won’t have, how [do I say] aa3? It won’t have that value when you want to sell it lo1, it will mean you’ll forever just have it to hold yourself and to look at gaa3-zaaz3.’

The important question then arises of whether the component ‘it is not more’ is applicable to all the examples of zaaz previously considered. An examination of each example showed that this was possible (and the line ‘this someone can feel something because of this’ is also justified). For instance, example (7.6) can be interpreted as ‘I said not more than two sentences to them’, and (7.7) can be interpreted as ‘there are not more advantages/benefits’. Example (7.18) also makes sense using this new explication, which appears to explain it even better than explication [7C]. For instance, in (7.18c), when speaker B says that there
are only burgers and fries to eat, she seems to have meant that there are not more food choices.

Examples (7.12) and (7.13) are to do with short amounts of time, and they also work well with ‘it is not more’. Example (7.12a) can be interpreted as someone thinking ‘it was a very short time, not more’, and (7.12b) can be interpreted as someone thinking ‘it was only one or two months, not more’. The last line of the explication points out that this someone can perhaps feel the urgency, or feel surprised or impressed at how short a period of time it was. Example (7.13) can be interpreted as ‘there were not more than two days to play’, and that someone can feel something bad about this.

‘It is not more’ also works for questions like in (7.14), where it can be interpreted as ‘there will only be one set of stamps, not more?’. The question in (7.15) can be interpreted as ‘only the Polytechnic offers physiotherapy, not more [universities]?’ Where previously the explication would have included ‘there is one thing/place’ for these examples, the newer line ‘it is not more’ also works and even seems more suited to questions. ‘It is not more’ implies there might have been an expectation of something more. Therefore, the speaker seems to have more reason to be asking a question. The other examples that would have used ONE are also compatible with ‘it is not more’. For instance, in (7.16), the speaker is saying that she went to one country, Australia, and not to more places. In (7.17) the speaker is saying that he got one B4 grade and not more. In all these examples, the final line of the explication ‘this someone can feel something because of this’ works in the same way as before and is unchanged.

Importantly, to be a viable replacement of the earlier explication, the new explication also needed to exclude any interpretations which were deliberately excluded by the previous explications. In particular, the new explication should reflect the earlier findings that zaa3 does not really apply to things being SMALL in the literal sense. Interestingly, MORE (and its Cantonese equivalent do1di1) do seem to meet this requirement – MORE is generally used when there is more in number/amount (including more than one) or more time, but is less often used when something is larger. That is, saying that something is more (do1di1) in size seems strange.
The component ‘it is not more’ can give the feeling that something is lacking or is reduced, just as *zaa3* often does. *Zaa3* is often used when speaking negatively about something, as in the majority of the examples in the corpus and in this chapter, but this is not always the case. Often, the negative feeling comes from context and would be present even if *zaa3* were not used. Recall, though, that some examples such as (7.2) and (7.17) show *zaa3* being used as boasts and proud statements. Therefore, the explication of *zaa3* should reflect the tendency for *zaa3* to be used in situations where something is bad, but should not explicitly state that something is definitely BAD. This component seems to fit this requirement well.

7.2.10 The final NSM explication for *zaa3*

One final issue is that the component ‘it is not more’ sometimes implies not more *than* something. Recall that the comparative construction is not universal and does not adhere to the rules of NSM, so there is no solution available using the comparative. At this point, the idea appeared of placing ‘it is like this’ in front of the new component ‘it is not more’. This helps, as the explication can be read as being like ‘it is not more than this’. As well, the final explication removes the line ‘someone can think like this’. This reduces unnecessary components and makes this explication more similar to the proposed explications of other Cantonese utterance particles. This leads us to the final proposed explication for *zaa3*, which is shown below, labelled [7E]. The Cantonese version is also provided.

[7E] Final explication for *zaa3*:

\[
\text{it is like this, (it is) not more} \\
\text{someone can feel something because of this}
\]

\[
\text{hai6 gam2joeng2, m4 hai6 do1di1} \\
\text{is like-this/this-way not is more} \\
\text{jau5jan4 ho2ji5 jan1wai6 gam2 gok3dak1} \\
\text{someone can because like-this/this-way feel} \\
\text{jau5 je5} \\
\text{there-is something}
\]
As can be seen, explication [7E] is a modified version of [7D.2]. Just like [7D.2], it is able to explain examples of zaa3 which are to do with hearsay, such as (7.19), (7.20), and (7.21), as well as all the previous examples looked at in this chapter which are to do with LITTLE~FEW, A SHORT TIME, and ONE, while excluding the possibility of a meaning to do with SMALL. The explication is compatible with the tendency for zaa3 to be used to describe something negatively, while not making it part of the particle’s invariant meaning, since zaa3 can occasionally be used in a positive way, such as in example (7.17). All these considerations mean that it is unnecessary to seek any additional explications for zaa3, i.e. to posit polysemy.

The Cantonese NSM exponents identified in Chapter 2 are fully capable of rendering this explication, making the explication understandable by native speakers of Cantonese. Note that the Cantonese expression gok3dak1 jau5 je5 [FEEL THERE-IS SOMETHING] does mean ‘feel something’, just like in English. Though as an independent word jau5 can mean ‘have’, and thus jau5 je5 can be interpreted as ‘have something’, in this context jau5 je5 is acting as an indefinite substantive. (Jau5 di1 je5 [THERE IS SOME SOMETHING] might also have been possible, but emphasises that there is definitely something. This variation felt stronger than the intended meaning and the English version.) The tests which have been carried out throughout this chapter, where the explication is substituted for zaa3 and applied to the utterance and context of each example, should be equally valid if conducted in Cantonese.

Explication [7E] is similar in some important respects to Wakefield’s (2012b, 1-2) explication of zaa3 and to Wierzbicka’s (2003, 346-347) explication of English only. Compared to both of these, explication [7E] has the added ‘feeling’ component (i.e. ‘someone can feel something because of this’). Explication [7E] is also an improvement on Wakefield’s in terms of the well-formedness of the NSM, and is additionally provided in Cantonese. This chapter has provided detailed argumentation and corpus-based evidence for the explication, including methodically ruling out other potential primes and components (Wakefield (2012) provided almost no reasoning for his claims – the short paper in question focussed on the particle ze1).
7.3 Concluding remarks

This chapter has guided the reader through the investigative process by which the final proposed NSM explication [7E] was found for zaa3. This chapter considered many key primes and components which had seemed possible at earlier stages, and systematically justified the rejection or acceptance of those primes and components using evidence from a corpus of real, recorded Cantonese. The final explication was found to be substitutable in numerous examples of zaa3, thereby justifying the explication and also showing what zaa3 meant in each example. Considering the wide variety of examples considered, it appears that the proposed explication is the invariant meaning of zaa3. The explication is clearly expressible both in English NSM and in Cantonese NSM.

Needless to say, the exact trajectory that led the present researcher to her conclusions is not the only possible trajectory; other researchers might have reached the same final conclusion using a different route. The important thing is not the pathway that happens to lead a particular researcher to a particular final proposal, but rather how the final proposal stands up against the evidence: whether it predicts the particle’s attested range of use and provides intuitively plausible interpretations of uses across a range of contexts.

This is the final chapter in Part Two of the thesis. Five individual Cantonese utterance particles have been examined in turn. In each case, it has proved possible to propose a specific, clear, and cross-translatable NSM paraphrase that arguably captures the invariant meaning expressed by the particle. The explications show that previous claims that the particles have no meaning should be reconsidered.

Part Three of this thesis looks at particle combinations. Chapter 8 considers proposals according to which monosyllabic particles, such as those considered so far, are actually decomposable into smaller morphemic units, either as contractions of other monosyllabic particles, or as sub-syllabic morphemic units. This will include a discussion of the alleged zaa3 = ze3 + aa3 contraction. Chapter 9 considers polysyllabic particle clusters, using the findings of Part Two.
Part Three:
Particle combinations
Chapter 8:
The contractions hypothesis and the sub-syllabic morphemes hypothesis: a semantic critique

This chapter and the next together comprise Part Three of this thesis. Both chapters are concerned with questions about ‘combinations’ of particles, but their foci are quite different. As noted in Chapter 1, one of the distinctive properties of Cantonese particles is that they can occur in sequences of two or more, forming what are termed here ‘particle clusters’, such as gaa3-zaa3 or gaa3-laa3-wo3. This phenomenon is the topic of Chapter 9. Before moving to that discussion, however, the present chapter reviews several proposals in the literature, according to which individual particles like gaa3, zaa3, and wo3 can be analysed as combinations – I divide these into two main sub-types, which I refer to as ‘particle contractions’ and ‘sub-syllabic morphemes’. The first describes monosyllabic particles claimed to be made up of two particles, such as the claim that gaa3 = ge3 + aa3. The second refers to morphemes which are not particles themselves and which cannot occur independently, but are said to ‘combine’ to form individual particles. For example, gaa3 can be analysed into three separate morphemic components, none of which occur on the surface: g-initial + aa rhyme + mid-level tone 3.

Though quite different in their details, hypotheses of both these kinds propose that the monosyllabic particles heard in ordinary speech can be broken down into constituent units. They also share the fact that they depend on argumentation. The purpose of this chapter is to assess the plausibility of a selection of proposals for both kinds. It will be shown below that the literature in relation to both areas is opaque and contradictory. As with studies of individual particles, most prior work on particle combinations looks at large groups and ‘families’ of particles. Consideration of select particles and morphemes here shows that the arguments for particles being combinations of smaller meaningful units are not very convincing.

---

84 Note that in the existing literature, some authors refer to particle contractions as particle clusters, and vice versa.
For the sake of simplicity, this thesis looks at each kind of combination separately. It may be worth noting, though, that the different kinds of (supposed) particle combinations are not mutually exclusive. It is possible, for instance, to believe that some clusters consist of particles which are combinations themselves. For example, Kwok (1984, 8-11) proposed that laa3 and aa4 combine to form the contraction laa4. Laa4 in turn can be preceded by ge3 to create the cluster ge3-laa4. In other words, ge3-laa4 = ge3 + laa3 + aa4. Gibbons (1980, 764-765) proposed the following, in which four syllables are said to be comprised of five particles:

When the second particle follows there is no change in form if it begins with a consonant, e.g. la bo. However, if the second particle begins with a vowel, there is deletion of the vowel and tonal element of the first particle... and the two particles are joined, e.g. la + ak = lak. More particles can be added, up to a total of five, e.g. tim-ge-lak-wá.

The outline of this chapter is as follows. Section 8.1 overviews the idea that certain particles seem to share similarities. Section 8.2 then looks at examples of the contractions hypothesis, focusing on proposals about gaa3 and zaa3. Section 8.3 explores the more radical sub-syllabic morphemes hypothesis, looking at a range of initials, rhymes, codas, and tones, particularly those that are claimed to make up the ‘wo3 family’ (wo3, wo4, wo5, and bo3) and some members of the l- initial family. Section 8.4 looks briefly at possible historical explanations. Finally, in section 8.5, a more moderate and plausible explanation for semantic similarities among particles is proposed – that Cantonese utterance particles may have some qualities of or similarities with phonesthesmes and sound symbolism.

8.1 Overview of approaches to similarity between particles

In previous literature on Cantonese particles, phonologically similar particles are often grouped together and described with reference to one another. For example, the particles wo3, wo4 and wo5 are sometimes said to be variations of each other, and the particles zaa3, ze1, and zek1 seem to be semantically alike.

**Note:** Others claim that the -k coda is not possible in clusters in a non-final position (see section 8.3.7).
For ease of comparison, the discussion in this section focusses on *wo3* as an example.

Some scholars, such as Yau (1965), have attempted to assess the similarities between large numbers of utterance particles by applying a common set of test criteria, such as the ‘S-Q test’ described in previous chapters. Since Yau included *wo3*, *wo4*, and *wo5* in his test, we can observe whether the test shows similarities and differences among the three. According to the ‘nominal interpretation’, based on the opinion of the majority of respondents, *wo3* is a Q-type particle, while *wo4* and *wo5* are S-type (Yau 1965, 53-57). According to the ‘numerical interpretation’, which relied on each particle’s mean and standard deviation, *wo3* belongs in the ‘mixed’ S-Q category, in contrast with *wo4* and *wo5*, which remain in the S-type category (Yau 1965, 58-66). In other words, Yau’s S-Q test suggests that the particle *wo3* is fairly different from *wo4* and *wo5*, although the latter two particles may be similar to each other.

Yau (1980, 37) acknowledged that a few of the particles or clusters he identified are probably free or contextual variants, but left these as unrelated items so the ‘C-tests’ could determine whether they were variants or not, based on their connotative properties. This could have been a good idea, but as mentioned throughout this thesis, there are problems with the C-test. Yau himself (1980, 49) admits that the tests did not capture some ‘peculiar’ functions of certain particles, such as *wo4*’s function of expressing reported speech.

A different approach was used by Gibbons (1980), though he maintained that each particle be subjected to the same categories and criteria. He tabulated each particle into the categories of assertion, mand, or question, and gave each a ‘strength’ from 1 to 3. He also considered things like ‘commitment to truth of proposition’ and ‘newness’. As with Yau’s tests, each particle could be easily compared with the others along a scale, but some categories and criteria would be less relevant for certain particles. Overall, the results of Yau (1980) and Gibbons (1980) are best described as inconclusive.

Luke (1990, 307) opined that while some connections and similarities might exist between particles which share the same segmental composition, ‘the degree and nature of their relatedness is unknown, and extremely difficult to pin
down’. He believed, for example, that wo3, wo4 and wo5 have rather different uses (albeit without having conducted in-depth analyses of wo4 or wo5). He stated that wo3 conveys an element of ‘unexpectedness’ which is less relevant to the other two particles, whereas he felt that wo5 has a ‘quotative’ or ‘hearsay’ use which is less prominent in the other two.

In summary, the scholars reviewed so far have discussed perceptions that there are semantic affiliations between various formally similar particles, but without firm conclusions and without proposing any concrete morphological analyses. Other scholars have looked at this from different angles. We turn now to specific proposals, according to which individual, monosyllabic particles can be broken down into smaller meaningful units.

8.2 The particle contractions hypothesis

As explained, the term ‘particle contraction’ is used in this thesis to describe a claim that a monosyllabic particle is actually made up of other particles combined. The idea seems to have been around for some time and is accepted widely, seemingly without question. Kwok (1984, 8) summarises the ‘contractions hypothesis’ as follows:

The process of derivation is one of addition, when two or more particles are added together. Various phonological changes may occur to prevent the ‘surface’ forms from being recognised as combinations of the basic particles. The semantic value of the ‘derived’ particle cluster is seen to be the sum of the meanings of its component parts. This treatment simplifies description, and allows us to see the relationships among the particles more easily.

This sounds attractive, but the details are not so straightforward. For one thing, a closer look reveals that Cantonese linguists do not agree on which particles are contractions, and if they are contractions, which particles they are contractions of. In this area, seemingly minor details can lead to quite different conclusions. Since contractions by definition lead to sounds being omitted or modified, it is unsurprising that there is potential ambiguity as to the identity of the omitted or modified sounds.
For instance, Kwok (1984) treats zaas as a ‘basic’ particle, while Matthews and Yip (2011) treat it as a contraction of ze3 + aa3, even though ze3 is not a particle that occurs in isolation (see section 8.2.2). Another puzzle is that of ze1. As mentioned in Chapter 7, Wakefield (2012b) believes that the meaning of zaas is entailed in and part of the meaning of ze1, i.e. ze1 is a contraction of zaas + something else. This is similar to but not the same as Law’s (1990, 132-135) claim that ze1 is derived from an inherently toneless zaas, combined with e, and a tonal ‘weaker’. These beliefs are shared by neither Kwok (1984) nor Matthews and Yip (2011), who do not regard ze1 as a contraction at all. It also contradicts Fung’s (2000, 59-61) report that ze1 is not interchangeable for zaas and vice versa. (Departing from all of these scholars, Sybesma and Li (2007) propose that ze1 = z- + tone 1 (+ default vowel), and that zaas = z- + aa (+ default tone). This will be dealt with in section 8.3.) All this is rather confusing, with no firm outcome as to whether zaas or ze1 are contractions at all, and, if they are, what particles/morphemes they might be contractions of. Furthermore, these hypotheses cannot be definitively tested, since the putative components such as z- and ze3 never occur in isolation.

Similar problems affect other proposed contractions. Kwok (1984, 48) states that the particles ge3 and laak3 combine to give the cluster ge3-laak3, and that a vowel change in the first syllable often causes this to be pronounced gaa3-laak3. Others believe that the change from ge3 to gaa3 is not just a sound change, but rather the addition of the particle aa3. For instance, Matthews and Yip (2011, 391-392) state that gaa3 = ge3 + aa3 and that the contraction is obligatory since ge3-aa3 does not occur. Kwok (1984, 46), however, provides examples that show that ge3-aa3 is possible as a disyllabic particle, though she does not believe that this precludes its possible contraction into gaa3. (Again, Sybesma and Li (2007) propose a more extreme combination g3 + aa (+ default tone); see section 8.3.)

Even within individual works, inconsistencies arise. For instance, Matthews and Yip (2011, 391) list certain related particles. Most of their
proposed contractions are given below (for simplicity, the list below excludes combinations resulting in polysyllabic clusters and those with a \(-k\) coda)\(^{86}\).

\[
\begin{align*}
Gaa3 &= ge3 + aa3 \\
Gaa2 &= ge3 + aa1 \\
Gaa4 &= ge3 + aa4 \\
Zaa3 &= ze3 + aa3 \\
Zaa2 &= ze3 + aa1 \\
Zaa4 &= ze3 + aa4 \\
Laa4 &= laa3 + aa4 \\
Laa5 &= laa3 + aa5
\end{align*}
\]

An obvious puzzle is that the three ‘laa family’ particles (laa1, laa3, and laa4) do not follow the same pattern as that proposed for the g- and z-families. Rather, laa3 is proposed to be a ‘basic’ particle which helps form laa4, as shown above. But why is laa3 not claimed to be le3 + aa3, just like gaa3 = ge3 + aa3 and zaa3 = ze3 + aa3? Le3 is not consistently recognised as a particle, and many native Cantonese speakers do not use it, but it is not unheard of in the literature; for example, Fung (2000) describes it in her study. Even if le3 is not recognised by Matthews and Yip (2011) as a particle, this still should not affect a potential le3 + aa3 contraction, because ze3 is not generally recognised as a particle either, though it is still proposed as a building block for zaa3, zaa2 and zaa4. Laa1 is another frequently used particle, but is not suggested by Matthews and Yip to be a contraction either.\(^{87}\)

As mentioned, Kwok (1984, 8-11) does not list any hypothetical ze3, stating instead that zaa3 is a ‘basic’ (as opposed to ‘derived’) particle. She thus uses zaa3 to produce the particles zaa4 and zaa5 – a different analysis from that of Matthews and Yip. Kwok’s zaa5 also takes the place of Matthews and Yip’s zaa2, which she does not mention. Zaa5 is not the only tone 5 contraction listed by Kwok, as she also adds gaa5 and laa5. The additional particles she includes, which were not reported by Matthews and Yip, are below (again excluding those resulting in polysyllabic clusters and those with a \(-k\) coda).

\[
\begin{align*}
Zaa4 &= zaa3 + aa4 \\
Zaa5 &= zaa3 + aa5 \\
Gaa5 &= ge3 + aa5 \\
Laa5 &= laa3 + aa5
\end{align*}
\]

\(^{86}\) It seems fairly straightforward to posit that two tone 3 particles would create another tone 3 particle. As for tones 3 and 1 creating tone 2 particles, this seems at first less logical, but this thesis makes the speculation that this is because of the rising tone. Tone 3 is a ‘mid level’ tone, 33 in IPA. Tone 1 is a ‘high level’ tone, 55 in IPA (or ‘high fall’, 53). Tone 2 is a so-called ‘high rise’ tone, 35 or 25 in IPA, so this can plausibly explain why some believe tone 3 followed by tone 1 can sound like tone 2. The tone 4 particles being contractions of tone 3 and tone 4 particles appears to be a simple omission or modification of tone 3.

\(^{87}\) Laa2 does not occur as a particle in speech, so its omission is understandable and we cannot expect it to correspond with gaa2 and zaa2.
As can be seen, there is disagreement among scholars about which particles are in use and which are not, let alone which particles are contractions of which other particles.

8.2.1 The proposed *gaa3* contraction

As described, the particle *gaa3* is claimed in the literature to be made up of the particles *ge3 + aa3* (Kwok 1984, 8-9, Fung 2000, 168-171, Matthews and Yip 2011, 391-392). Kwok does not explain why she believes this. Matthews and Yip (2011, 391-392) state that the *ge3 + aa3* contraction is obligatory, but Kwok (1984, 46) gives examples of *ge3-aa3* as a disyllabic particle. As a native speaker, I also believe that *ge3-aa3* is possible. Even if it were true that the particles *ge3* and *aa3* never occur together, this would not be a sufficient justification of the claim that *gaa3* as a contraction of the two. Plenty of other Cantonese utterance particles never occur together, without being claimed to be contractions.

If we take Matthews and Yip's (2011, 391-392) own description, we see that *ge3* is defined as ‘affirmative: “this is the case”’ and *aa3* as a ‘softening statement or question’. Regarding the contraction, they claim that ‘the assertive particle *ge3* combined with *aa3* results in *gaa3*, which has the effect of seeking confirmation of a statement’. This would seem to imply that *gaa3* is used only in questions. This contradicts the definition of *ge3* as ‘affirmative: “this is the case”’, which implies that *gaa3* is used in statements. Descriptions such as ‘assertive’ for *ge3* and ‘softening’ for *aa3* are disconnected and do not support the theory that the meaning of *gaa3 = ge3 + aa3*.\(^88\)

Matthews and Yip also display a classic case of explaining an utterance rather than the particle's meaning. The supporting example given by Matthews and Yip (2011, 392) is shown in (8.1) below. The question on the right, 'hai6 m4 hai6 zan1 gaa3?', ‘is it true?’ is a standard ‘A-not-A question’. As Matthews and Yip (2011, 360-363) themselves explain, this is the most neutral form of yes/no

---

\(^{88}\) Matthews and Yip’s (2011, 394-397) subsequent section on particle clusters omits clusters consisting of *gaa3*. This is strange, considering how frequently *gaa3* occurs in speech and combines with other particles. All of the other particles analysed in this thesis are able to occur with *gaa3* (see Chapter 9). Other possible clusters include but are not limited to *gaa3-met, gaa3-zet, gaa3-bo3, and gaa3-aa1-maa3*. These not only come immediately to a native speaker’s mind, but are present in the corpus.
question in Cantonese, and is like asking ‘is A the case or not?’ As can be seen, this particular example uses ‘is-not-is’ before ‘true’ to ask the question ‘is it true?’ If gaa3 were not included at the end, the sentence would sound odd, but would nevertheless be a question, and would still have ‘the effect of seeking confirmation of a statement’. We must therefore question whether this effect is due to the utterance particle gaa3.

(8.1)

\[
\text{Haih jân ge} \rightarrow \text{Haih-mhaih jân ga?}
\]

is true SFP\(^{89}\) is-not-is true SFP

‘It’s true.’ (statement) ‘Is it true?’ (question)

(Matthews and Yip’s gloss and translation)

Recall from Chapter 5 that gaa3 regularly occurs with statements too. But example (8.1) gives the impression that gaa3 is a question particle, essentially equal to ge3, which is its counterpart in statements. In fact, in the statement on the left, the particle ge3 could naturally be substituted for gaa3. This would result in another declarative statement ‘hai6 zan1 gaa3’, which would also be translated as ‘it’s true’. Actually, a number of other particles or clusters could be substituted here.

More confusingly for the contractions hypothesis, it is unclear how the differences on each side of example (8.1) are to be attributed to the supposed addition of aa3 ‘softening statement or question’. This example does not show evidence that gaa3 is derived from ge3 + aa3. It is possible that ge3 and gaa3 are truly related, but this example does not demonstrate how; it could be argued that the two sentences are similar largely because of zan1 ‘true’, and the most drastic difference would appear to be the transformation into an A-not-A question. Ultimately, example (8.1) fails to explain the meaning of gaa3, or the relationship between ge3 and gaa3.

Kwok’s (1984) description has parallels with Matthews and Yip’s. According to Kwok (1984, 42-43), ge3 shows that a ‘sentence is a factual statement expressing what the speaker regards as true’. It is ‘used to strengthen the force of the assertion, and is like prefacing the sentence with “It is a fact...”’. This is compatible with Matthews and Yip’s (2011, 391-392) explanation of ge3.

\(^{89}\) Their abbreviation ‘SFP’ stands for ‘sentence-final particle’.
Regarding *aa3*, Kwok (1984, 45, 71) acknowledges that *aa3* is found in declarative, interrogative, and imperative structures, but states that *aa3* does not carry much semantic content, and ‘does not add a great deal’ to the meaning of the sentence. She mentions that the main function of *aa3* is to make a sentence sound less abrupt, which corresponds with Matthews and Yip’s (2011, 391-392) description of ‘softening’. These descriptions face the same problems as those of Matthews and Yip. If *gaa3* = *ge3* + *aa3*, but *aa3* has no meaning, what is the difference between *ge3* and *gaa3*? This kind of analysis can hardly convince the reader that added particles are contributing additional meanings.

Fung (2000, 168-171) admits that there are situations where *gaa3* can occur naturally, but where *ge3* would sound odd. This would seem to indicate that *gaa3* is not made up of *ge3* + *aa3*, since if *ge3* were not applicable in a situation, there is no apparent reason why the addition of *aa3* would alter things. If *ge3* is included in the meaning of *gaa3*, this should restrict when and where *gaa3* can occur. Fung overlooks this and does not question the contractions hypothesis. From all these studies, it does not appear plausible that *gaa3* = *ge3* + *aa3*.

As a final argument, it would seem logical to expect that highly frequent particles should have meanings which are relatively simple or broad, to reflect the fact that they can be used in more situations. In contrast, less frequently used particles can be expected to have meanings that are relatively narrow or complex. In this instance, the claim is that the meaning of the particle *ge3*, which occurs 183 times and is the 16th most frequently used utterance particle in the Hong Kong Cantonese Corpus, is part of the meaning of the supposed ‘contracted’ particle *gaa3*, which occurs 1832 times and is the 2nd most frequently used utterance particle. This would cause a mismatch whereby the simpler, broader particle supposedly contains the meaning of something more narrow and less applicable. This mismatch does not appear to have been previously noted in the literature.

**8.2.2 Contraction analyses for *zaa3***

As mentioned, *zaa3* is claimed by Matthews and Yip (2011) to be a contraction of *ze3* + *aa3*, and thus to neatly parallel their *gaa3* = *ge3* + *aa3* proposal. The
glaring puzzle of this proposal, however, is that ze3 is does not occur in isolation, if it is a particle at all. Matthews and Yip (2011, 391) list ze3 as meaning ‘just, only’, but as explained in Chapter 7, ‘just, only’ is a very common definition of particles with a z-initial, so this description of ze3 does not help identify or differentiate it from other z-initial particles which are in use and are not all interchangeable. As a native speaker I do not believe ze3 is in use in Cantonese. Other studies, such as Yau (1965, 1980) and Gibbons (1980), make no mention of it either. Kwok (1984, 8-11) does not identify any particle ze3, nor does she state that zaaz = ze3 + aa3, despite listing out other combinations. Even Fung does not mention any particle ze3, despite devoting one chapter of her thesis to a ‘family’ of particles with z as the onset (Fung 2000, 30-73), and describing two other particles, ze4 and zaakt, which other scholars do not mention.

It is possible that ze3 may sometimes be heard in speech as a result of a vowel change when particles combine to form clusters (see Chapter 9). For example, some speakers might pronounce gaa3-zaa3-wo3 as ge3-ze3-wo3. Thus it might be proposed that ze3 is a variation of zaaz (as le3 may be laa3). However, this would still not be satisfactory, because if we treat ze3 and zaaz as the same particle, then the formula zaaz = ze3 + aa3 makes no sense. It would seem unnecessarily complicated to posit that ze3 is a particle which only ever occurs together with aa3 in the form zaaz, just for the sake of a formula zaaz = ze3 + aa3 which would fit neatly with the proposed particle contractions of the g-family. Problematically, it would also mean that ze3 and all of the formulae for zaaz, zaaz and zaaz are impossible to refute or confirm, given that we cannot be sure what ze3 means on its own.

Even if we accept that ze3 exists and matches with Matthews and Yip’s description of ‘just, only’, there are still problems. According to Matthews and Yip, aa3 is ‘softening statement or question’, and so the particle zaaz is supposedly a combination of ze3 ‘just, only’ and aa3 ‘softening statement or question’. No other works in the literature have described zaaz as ‘softening’, however, which casts doubt on the proposed analysis. In short, the argument that zaaz = ze3 + aa3 is rather weak.

Surely, it would be better to study zaaz as it is, than to explain it in terms of other, hypothetical particles. This was done in Chapter 7, where it was
concluded that zaaz means: ‘it is like this, (it is) not more; someone can feel something because of this’. The next section considers the contractions hypothesis from an NSM perspective.

8.2.3 The contractions hypothesis from an NSM perspective

It would be valuable in the future to explicate more particles using NSM methods, which could shed more light on the contractions hypothesis. For example, having explications of ge3 and aa3 in hand would help to determine whether gaa3 = ge3 + aa390. However, a speculation can be made that particle contractions such as ge3 + aa3 behave differently from the particle clusters to be discussed in Chapter 9. With regards to the NSM explications, it seems highly improbable that the explications which make up the hypothetical particle contractions can be combined as straightforwardly as those found in particle clusters such as gaa3-laa1 or gaa3-wo3. In the attested particle clusters, it will be shown in Chapter 9 that the individual particles each apply to the utterance to which they are appended, but this is less plausible for the supposed ‘contractions’.

As mentioned above, gaa3 occurs 10 times more frequently in speech than ge3 (1832 versus 183 occurrences in the corpus). The analysis presented in this thesis supports the idea that gaa3’s high frequency has to do with it having a very simple meaning which is applicable in a broad range of contexts, and that ge3’s relatively lower frequency is to do with it having a more specific meaning, which Cantonese speakers are generally less likely to want to express. If relatively uncommon particles like ge3 were part of the meaning of relatively common ‘contracted’ particles like gaa3, this would cause a paradox where a specific or exclusive meaning is part of a particle whose meaning is very broad and simple. Gaa3 is not the only supposed ‘contraction’ which would pose this problem. The more likely possibilities are either that particle contractions do not behave like particle clusters semantically, or that they are not ‘contractions’ at all.

90 A tentative and provisional proposal for an NSM component of aa3 is ‘I want to say this to you’.
Even if we say that particle contractions occur, NSM would still be an invaluable tool. One major problem with the contractions theory is that it leaves abundant opportunity for circular definitions. Circularity is one of the most widespread pitfalls of semantics. It is not helpful to explain, for example, zaa3 in terms of ze3/ze1, and vice versa. The NSM method used in this thesis avoids circularity and offers precision. If NSM semantic analysis were conducted on e.g. ze1, this would shed light on the similarities and differences between zaa3 and ze1, without defining one in terms of the other. It could be that explications for particles share similar components.

This section has discussed only a small selection of possible examples, but the main problems with the contractions hypothesis have been laid out. The next section investigates the sub-syllabic morphemes hypothesis.

8.3 The sub-syllabic morphemes hypothesis

Another response to the perception that similar-sounding particles often have similar meanings or functions is the notion that particles can be broken down into abstract, sub-syllabic units, which cannot be pronounced in isolation. This line of analysis opens up the attractive possibility that ‘families’ of phonologically similar particles are semantically related by shared morphemes. If it could be made to work, such an analysis would reveal a high degree of systematicity between the forms and meanings of Cantonese utterance particles.

Applying this line of analysis to wo3, wo4, and wo5, for example, it could be claimed that the meaning of wo (without a tone) is shared by each member, with the different tones adding distinct layers of meaning. This is quite a radical proposal, and as such not something that is fully accepted by all scholars in the area. Both Luke (1990) and Matthews and Yip (2011), for example, recognise that there are semantic similarities between wo3, wo4, and wo5, without suggesting that they can be broken down into sub-syllabic morphemes.

Literature in this area includes that of Law (1990), Fung (2000), Sybesma and Li (2007), and Li (2006). Fung (2000) proposed that particles with the same initial have the same core semantic feature. She describes z- initial, l- initial, and g- initial families. More radical proposals were made by Law (1990), Sybesma and Li (2007), and Li (2006). Law argued that the Cantonese particle
system can be simplified by proposing particles that have inherent tones, particles that are inherently toneless, and particles that consist only of tones. Interactions of these different types of particles then produce the surface forms of the particles which are heard in speech. Law argued that some particles are phonetically reduced forms of particle (segmental) clusters. Following this, Sybesma and Li (2007) and Li (2006), proposed that particles can be dissected into further sub-syllabic semantic units91: not just initials, but also rhymes, tones and codas. These kinds of sub-syllabic morphemes are not posited for other Cantonese words, but is claimed to be the case in the area of utterance particles only.

It is not feasible in this chapter to consider every possible sub-syllabic morpheme which, according to these radical analyses, could make up the utterance particles. Nonetheless, cases which relate to the particles investigated in Part Two are considered in the sub-sections that follow. Overview of a range of initials, rhymes, tones, and codas show that the problems discussed are not restricted to certain anomalous morphemes, but are an extensive and substantial issue. By definition, the would-be minimal units all connect with each other in use, so the following subsections are not fully independent of one another.

8.3.1 The w- initial

Sybesma and Li (2007, 1763-1764) believe that wo3, as well as wo4 and wo5, actually have no inherent onset, and that this is realised with w-. The same treatment is given in Li (2006). Why they believe this is not explained, but this zero (ø) initial and coda are presumed to have no meaning (Sybesma and Li 2007, 1740). Following this thinking, these three particles are made up only of ‘o + tone’. As mentioned in Chapter 4, some studies treat the infrequent bo3 as a variation of wo3. Sybesma and Li regard them as distinct but with some overlapping elements, though admitting that the status of the b- family is not clear (Sybesma and Li 2007, 1756). Given that the initial is the only difference

91 Sybesma and Li refer to these as ‘minimal meaningful units’. Since the resulting units are not words themselves, they are referred to as ‘sub-syllabic’ here to differentiate them from the components in particle contractions and clusters.
between \textit{wo3} and \textit{bo3}, the analysis that \textit{w-} has no meaning and that \textit{b-} has an unclear meaning provides no explanation of their difference.

If \textit{w-} has no meaning, but is nevertheless used in particles, why do we not find it used with \textit{aa} or \textit{e} rhyme particles? This cannot be due to semantic incompatibility (since \textit{w-} has no meaning), nor is it likely to be due to problems with pronunciation, since \textit{waa} in various tones exists in other word classes. Nor is it the case that the vowel \textit{o} cannot be pronounced without an onset – \textit{o} in various tones is available in other word classes. In short, it is unclear why a meaningless \textit{w-} would be necessary with the \textit{o} rhyme, but not with other rhymes/particles.

8.3.2 The ‘mid level’ \textit{tone 3}

Throughout Sybesma and Li’s (2007) paper, the ‘mid level’ tone indicated by the number 3 is assumed to be the neutral tone for all particles based on the work of Law (1990), and is assumed not to contribute any meaning. The same stance is taken by Li (2006). Because of its tone, the particle \textit{wo3} is assumed to be the ‘base particle’ of its family. The same would be true of \textit{zaa3}, \textit{gaa3}, and \textit{laa3}\textsuperscript{92}. These scholars thus claim that all particles with a tone other than \textit{tone 3} can be analysed as the \textit{tone 3} particle plus something else, e.g. \textit{wo4} and \textit{wo5} would have the meaning of \textit{wo3} plus something else, and the same effect would occur in other families.

Other analyses seem to disagree about tones contributing stable meanings. For example, Fung (2000, 119-124) proposes that the difference between \textit{lo1} and \textit{lo4} is parallel to the difference between \textit{laa1} and \textit{laa3}, being that of knowledge assumption. Since this relationship is in one case across \textit{tones 1} and 4 and in another case across \textit{tones 1} and 3, this is at odds with the idea of tones contributing stable meanings.\textsuperscript{93}

\textsuperscript{92} As noted in Chapter 6, \textit{laa3} was analysed by Yiu (2001) as a ‘toneless’ particle which can be pronounced with a number of different tones. She treats \textit{laa3} in the mid-level tone as the ‘basic form’ from which the other ‘variants’ such as \textit{laa1} are derived. She gives no reason for this treatment.

\textsuperscript{93} Furthermore, Fung (2000, 119-124) states that \textit{lo4} and \textit{lo1} share many more similarities with each other than either of them do with \textit{lo3}, which is unexpected if \textit{tone 3} is taken to be the default, most neutral tone.
The hypothesis that tone 3 is the neutral base upon which other particles are built is also problematic in relation to particles which have no tone 3 counterpart. Specifically, Sybesma and Li mention the particles haa2, he2, ho2, and waa2, which have no counterpart in tone 3. These particles would seem, then, to have been derived from hypothetical tone 3 particles which are not in use (see Sybesma and Li (2007, 1771-1772)). Again, using hypothetical particles to create other particles is unnecessarily complicated and unverifiable.

It is possible that the desire to propose a ‘neutral’ tone may be partly due to a mystery of what Cantonese tone 3 might contribute semantically, or partly due to influence from Mandarin. However, the Cantonese tone system is generally treated as an independent system very different from that of Mandarin, being considerably more complex (Matthews and Yip 2011, 27-28). The Mandarin tone system includes four distinctive tones, and an additional ‘neutral’ tone which occurs at various heights, determined by the preceding syllable. The Mandarin neutral tone is also assigned to many grammatical words, including all six of the sentence-final particles Li and Thompson (1981, 238). In comparison, Cantonese has six clearly distinctive tones94, i.e. there is no phonologically ‘neutral’ tone. Every particle, just like other Cantonese syllables, bears one of the six tones, and the mid-level tone 3, labelled as ‘neutral’ by Sybesma and Li (2007) and Li (2006), is a distinct tone.95 There seems to be no reason to believe that Cantonese tone 3 is neutral.

It appears that the idea that there are underlying ‘toneless’ particles in Cantonese had its origins in analyses of Mandarin; specifically, Chao’s (1968, 795-814) treatment of Mandarin particles, which included two non-segmental tonal particles96. The idea put forth for Cantonese by those such as Law (1990),

---

94 As mentioned in Chapter 1, checked syllables were traditionally counted contrastively, giving three additional tones, i.e. nine distinct Cantonese tones altogether. However, the three additional tones which occur before unreleased consonants are now often seen as abbreviated counterparts of the three level tones which occur in other contexts. It is often assumed that the ‘high level’ and ‘high falling’ tones are not distinctive. All this results in six distinctive tones (Matthews and Yip 2011, 27-37).

95 As pointed out by Matthews and Yip (2011, 29), it is a fact that many grammatical words in Cantonese, such as the possessive/linking particle ge3 and the experiential aspect marker guo3, have the mid-level tone but equally, many others have high tones, such as the adverbial dak1 and the progressive aspect marker gan2.

96 Chao (1968, 795-814) gives a list of 28 particles in Mandarin. Of these, 26 are unstressed and in the neutral tone, taking their pitch from the preceding syllable, like any other neutral syllable in Mandarin. The other two are reported to be non-segmental tonal particles, affecting the intonational endings of phrases and sentences. One is a rising ending and the other a falling
is that a ‘toneless’ particle like \textit{aa} (i.e. not differentiating between \textit{aa1}, \textit{aa3}, \textit{aa4}, or \textit{aa5}, as recognised by other Cantonese scholars), is ‘semantically empty’, acting only as a carrier for intonation. This would complement Law’s claim that tonal particles exist in Cantonese. A similar analysis by Li (2006) was explicitly taken from Mandarin. Li studied Mandarin, Cantonese, and Wenzhou particles, applying an analysis of the pitch variation of Mandarin \textit{a} to explain tonal variation in Cantonese particles. The Mandarin particle \textit{a} may appear with a high pitch or a low pitch, and these are claimed to perform semantic functions independent of the particle. Li claimed that Cantonese \textit{aa}-particles are very similar to the Mandarin \textit{a}, and that Cantonese examples can be translated into a corresponding \textit{a}-suffixed sentence in Mandarin. The two particles are therefore proposed to have the same function (Li 2006, 43-46, 98). There is clearly a strong influence from Mandarin, as well as an apparent lack of understanding of the meanings of Cantonese \textit{aa1}, \textit{aa3}, \textit{aa4}, and \textit{aa5}. Many ideas about tone that originated from Mandarin and are applied to Cantonese may be problematic.

8.3.3 The \textit{o} rhyme

The \textit{o} rhyme is said to have the meaning of ‘noteworthiness’. For the three particles \textit{wo3}, \textit{wo4} and \textit{wo5}, Sybesma and Li (2007, 1763-1764) conclude that ‘noteworthiness’ is their ‘core meaning’. As discussed in Chapter 4, noteworthiness is rather subjective and vague notion. Quite possibly, speakers always believe what they are saying to be noteworthy. But even if we accept the label ‘noteworthiness’, there are other questionable issues for Sybesma and Li’s sub-syllabic morpheme analysis. For example, since according to them, neither \textit{w}- nor tone 3 have any apparent meaning (see 8.3.1 and 8.3.2), the only meaningful unit of the particle \textit{wo3} must be the \textit{o} rhyme. The particle \textit{wo3} expresses only ‘noteworthiness’, and nothing else.

Furthermore, there are inconsistencies with the \textit{o} rhyme, evident if we look at the \textit{o} rhyme as used with other initials. Consider Sybesma and Li’s description of the particle \textit{lo3}. They suggest that the \textit{l}- initial indicates ‘realisation of state’. The particle \textit{lo3} should therefore have the meaning of \textit{l}-

---

ending. He argued that these are better treated as particles than as sentence intonation, since they do not affect the intonational pattern of the whole construction, but only the voiced part of the last syllable.
‘realisation of state’ + o ‘noteworthiness’ (since tone 3 is claimed to contribute no meaning). Despite this, Sybesma and Li do not mention ‘noteworthiness’ in their description of lo3. In fact, they describe lo3 only as a stronger or more intense form of laa3 (Sybesma and Li 2007, 1753). This is unexpected and inconsistent. They do not argue that the o rhyme is a stronger or more intense form of the aa rhyme97, nor do they argue that the aa rhyme indicates any weak version of ‘noteworthiness’.98 If lo3 does not include the supposed meaning of o, the claim that each particle can be further broken down is not convincing. Not only does ‘noteworthiness’ not feature in Sybesma and Li’s description of lo3, it is not mentioned either in their descriptions of lok3, lo4, lo1, or ho2, all of which feature an o rhyme (Sybesma and Li 2007, 1753, 1756-1757). Nonetheless, Sybesma and Li still conclude later in their paper that lo1, lo3, and lo4 are ‘l (‘realisation of state’) + o (‘noteworthiness’) + [tone]’. Lok3 is claimed to be the same, with the addition of the -k coda acting as ‘emotion intensifier’ (Sybesma and Li 2007, 1764). Without explanation, they write ‘we may safely say that “noteworthiness” is part of the semantics of these [particles]’. The reader is left confused, if not baffled.

To make matters worse for their analysis, in other literature on Cantonese utterance particles, particles such as lo3, lok3, lo4, lo1, and ho2 have not generally been reported to be similar in meaning to wo3, or to include any meaning of ‘noteworthiness’. On the contrary, Luke (1990, 193-194) states explicitly that lo1 can be regarded as marking ‘unnoteworthiness’: lo1 is ‘a means of constructing natural, reasonable, and unnoteworthy descriptions’. Kwok (1984, 57-59) reports that the particles lo1 and lo4 point out what is obvious; Luke (1990, 193-194) interprets ‘obviousness’ as ‘unnoteworthiness’. ‘Noteworthiness’ is also generally not used by Fung (2000) to describe any of the o rhyme particles she looks at. Judging by the existing literature, Sybesma and Li’s claim that the o rhyme means ‘noteworthiness’, and that this meaning

97 Sybesma and Li differentiate between aa rhyme particles with no onset and aa rhyme particles with an onset. This is strange because the simple ‘addition’ that they advocate should not be affected by the absence of an onset.

98 In a different study, Kwok (2006) places wo3 and ho3 into a group with aa3 and laa3 as particles which express the semantic feature of ‘informativeness’. Kwok does not claim that the particles can be dissected into sub-syllabic morphemes, but her grouping does indicate flaws in the sub-syllabic morphemes hypothesis.
is present in all Cantonese utterance particles with o rhymes, is not very convincing.

8.3.4 The 'low level’/‘falling’ tone 4

According to Sybesma and Li, tone 4 (variously described as a ‘low level’, ‘falling’, or ‘low fall’ tone) does not show the most consistency across particles, although all tone 4 particles are able to be characterised as ‘the tone 3 counterpart, plus something else’ (Sybesma and Li 2007, 1769-1770). Since Sybesma and Li claim that tone 3 has no meaning, this merely states that tone 4 indicates something. The question is: what is this meaning?

Sybesma and Li (2007, 1770) find that the difference between wo3 and wo4 is that wo4 additionally expresses the speaker’s sudden awareness. ‘Sudden awareness’ may be interpreted as ‘unexpectedness’. On one hand, this is supported by Kwok’s (1984, 68-69) report that wo4 indicates unexpectedness and surprise. On the other hand, it contradicts Luke’s observation that wo4 and wo5 do not have the element of ‘unexpectedness’, which he says is more relevant to wo3 (Luke 1990, 307). Matthews (2007) seems to support both views, stating that wo3 covers surprise and unexpectedness, and that wo4 has a greater force than wo3.

If we accept Sybesma and Li’s proposal that the meaning of wo4 is that of wo3 plus sudden awareness, other inconsistencies arise. Seeing as the only phonological difference between wo3 and wo4 is the tone, one would assume that the speaker’s sudden awareness is expressed by tone 4. However, the speaker’s sudden awareness is not apparent in any of the seven other tone 4 particles in Sybesma and Li’s list99. Instead, Sybesma and Li (2007, 1770) state that the function of tone 4 is either adding strength to the counterpart with tone 3, or turning it into a question particle (with one particle aa4 being ambiguous between the two). To begin with, these two effects are contradictory. As for wo4, the reader is left wondering where the supposed meaning of ‘sudden awareness’ comes from, and why wo4 is not effectively wo3 + tone 4. Sybesma and Li suggest that the reason some tone 4 particles do not have a strengthening

---

99 The seven other tone 4 particles in Sybesma and Li’s (2007, 1769-1770) list are e4, aa4, laa4, zaa4, gaa4, le4, and lo4.
effect may be that in those contexts, strength is not a relevant notion. It is a mystery to the reader, then, why those particles would occur in *tone 4* at all, and why they would not use the 'neutral' *tone 3* instead.

The idea of a 'strengthening' effect appears to be influenced by Law (1990), who generalised that a low tone is usually associated with a stronger force. Law states that the low tone, without specification of register, is a 'strengthener', having the effect of reinforcing the idea conveyed by a particle, expressing a lack of interest, or a sense of coldness, boredom, or bluntness (Law 1990, 94). Applied to *wo4*, this is odd, since *wo3* supposedly conveys 'notworthiness', and a strengthened form of it should not express a lack of interest or boredom. Sybesma and Li nevertheless conclude that *tone 4* can be isolated as a meaningful unit which strengthens.

Several other accounts of *tone 4* are also available. Matthews (2007, 15) states that *tone 4* is associated with assertiveness, and since it is shorter, it is more abrupt than the other tones. This at least is a clearer and more straightforward definition than that offered by Sybesma and Li (2007). Matthews and Yip (2011, 392, 407) say that *tone 4* is associated with surprise or disapproval, with *wo4* specifically indicating discovery. The various accounts of *tone 4* do not seem to match.

Overall, we have seen many confusing and inconsistent accounts of the function of *tone 4* in the literature. These include a strengthening effect, a questioning effect, unexpectedness, no unexpectedness, surprise, boredom, assertiveness, lack of interest, discovery, disapproval, and abruptness. It is difficult to believe that any consistent meaning can really be attributed to *tone 4*. There needs to be consistent and reliable analyses of all *tone 4* particles before it can be believed that *tone 4* contributes a stable meaning. So far, none of the sub-syllabic morphemes considered have yielded convincing analyses.

---

100 Matthews (2007, 14-15) advances several reasons for believing that *wo4* is a variant of *wo3* derived by addition of a tonal particle. This includes that there appears to be a system of low falling variants, including *aa3/aa4*, and *lo1/lo4*. This depends on the plausibility of the *aa3/aa4* and *lo1/lo4* analyses, which cannot be explored in detail here. Also, note that *lo1* is in *tone 1*, unlike *aa3* and *wo3* which are in *tone 3*. Matthews also states that *wo4* occurs less frequently than either *wo3* or *wo5*, which would be consistent with it being a derived rather than a basic form. However, in the Hong Kong Cantonese Corpus, *wo4* occurs nine times, which is slightly more than *wo5*, which occurs six times (*wo3* occurs in excess of 900 times).
8.3.5 The ‘low rise’ tone 5

The ‘low rise’ tone 5 is said to convey evidentiality, shared between wo5, aa5, and le5 (Sybesma and Li 2007, 1772-1773, Li 2006, 110-112). Wo5 is to do with the source of the information, and aa5 and le5 are to do with the reliability of the information. Which phonological segment differentiates between these two categories is not specified, and is thus another weakness in the sub-syllabic morphemes hypothesis. The remaining tone 5 particles gaa5, laa5, and zaa5 are handled by Li as being equivalent to ge-aa5, le-aa5, and ze-aa5 respectively. The plausibility of this depends partly on the plausibility of the ‘simplex particles’ ge, le, and ze, which are discussed in section 8.3.8.

It follows that the particle wo5 is said to be ‘noteworthiness’ + ‘evidentiality’. Sybesma and Li (2007, 1764, 1772) state that wo5 indicates noteworthiness, and that the speaker is not completely sure and is presenting hearsay information. Similarly, Matthews (2007) and Matthews and Yip (2011, 391, 406) state that wo5 is a ‘hearsay’ evidential particle used to indicate reported information, and Luke (1990, 307) suggests that wo5 has a ‘quotative’ or ‘hearsay’ use less prominent in wo3 and wo4. Kwok (1984, 67-68) reports that wo5 is used to report what someone else has said. This is also compatible with W-m Leung’s (2009, 1-3, 2010a, 94) analysis that wo5 expresses hearsay information. I agree that the particle wo5 seems to convey evidentiality. One example from Matthews (2007, 10-11) is reproduced below. As Matthews states, wo5 is the only indication in this utterance of the second-hand nature of the information, and such a reading would not be necessitated by wo3 in the same utterance.

(8.2)

Keoi5dei6 jau6 bun2 uk1 wo5

they again move house PRT

‘They’re moving house again (they say).’ (Matthews’ gloss and translation)

The relevant question here, however, is not whether evidentiality is conveyed by wo5, but whether evidentiality is conveyed by tone 5. On this point, scholars do not agree. Matthews (2007) treats all three particles wo3, wo4, and wo5 as belonging to the realm of evidentiality, as does Matthews and Yip (2011, 404-407). Matthews (2007) even considers gwaa3 to be an evidential. He believes
that \( w03, w04, \) and \( w05 \) are ‘tonal variants of an evidential morpheme which have developed specialised functions within the semantic field of evidentiality, broadly defined to include mirativity’ (Matthews 2007, 15-16). W-m Leung (2013) analyses \( w03 \) and \( w05 \) together as both being evidential particles, and \( w05 \)'s expression of hearsay information is said to also be a function of \( w03 \). These accounts would actually suggest that some part of a ‘toneless’ \( wo \), rather than \textit{tone 5}, conveys evidentiality.

If evidentiality is not conveyed exclusively by particles which are in \textit{tone 5}, the questions remain as to which phonological feature conveys evidentiality, and what meaning or function \textit{tone 5} conveys. There may be no answers to either of these questions, because it is most likely that there are no sub-syllabic morphemes. There is no apparent phonological feature which consistently and exclusively conveys evidentiality, and \textit{tone 5} may not have any independent meaning. This does not contradict the understanding that the particle \( w05 \) seems to convey evidentiality.

8.3.6 The \textit{l-} initial \textit{laa1} and \textit{lo1}

The \textit{l-} initial family of particles is rather large\textsuperscript{101}. Fung (2000) and Sybesma and Li (2007) believe that \textit{l-} indicates the realisation of state-of-affairs. Puzzlingly, this is despite confessing that not all of the \textit{l-} initial particles encode all the semantic features of \textit{l-} (Fung 2000, 83), that the \textit{l-} family's shared meaning is not very obvious, and that \textit{le-} particles do not indicate this meaning at all (Sybesma and Li 2007, 1754).

The two particles \textit{laa1} and \textit{lo1} are relatively common in speech and have been studied in relative detail in the literature. Fung (2000) claims that \textit{laa1} and \textit{lo1} both mark the realisation of a physical state or the realisation of an epistemic state. \textit{Laa1} is additionally used to give commands, advice, and suggestions. Sybesma and Li (2007) claim that \textit{laa1} is \textit{l-} ‘realisation of state’ + \textit{a} ‘smooth-alert’ + \textit{tone 1} ‘forward-looking’, and \textit{lo1} is \textit{l-} ‘realisation of state’ + \textit{o} ‘noteworthiness’ + \textit{tone 1} ‘forward-looking’. These descriptions make it seem like \textit{laa1} and \textit{lo1} are very similar particles. However, Luke's (1990) much more

\textsuperscript{101} Fung (2000) describes a family of 12 \textit{l-} particles which includes \textit{lei4}, \textit{lai4}, \textit{lai3}, \textit{lai4}, \textit{laak3}, \textit{lo1}, \textit{lo3}, \textit{lo4}, \textit{lok3}, \textit{le3}, \textit{le4}, and \textit{le5}. Sybesma and Li (2007) propose the same, except for the addition of the particle \textit{le1} and the omission of \textit{lei4}.
detailed analysis of laa1 and lo1 found that they can be quite different. In fact, their functions are ‘diametrically opposed to each other’ (Luke 1990, 188). Laa1 is found to have a forward-looking quality, whereas lo1 is found to have a backward-looking quality (this explicitly contradicts Sybesma and Li’s claim about tone 1), and the two have distinct sequential implications. He explains that this accounts for a large array of differences in the distributional and co-occurrence behaviour of the two particles. His analysis can naturally explain speakers’ management of continuation and extension, including for example the difference in likely responses or actions by both the hearer and the speaker after laa1- or lo1-suffixed turns (Luke 1990, 176-191). Based on this, it would seem that laa1 and lo1 are more dissimilar than would be predicted by the sub-syllabic morphemes hypothesis.

8.3.7 The -k coda and ‘intensifiers’

The -k coda

The -k coda, which is unreleased or realised as a glottal stop, is generally believed to give a stronger force to a particle (Yiu 2001, Fung 2000, Li 2006, Sybesma and Li 2007). Studies which give descriptions of multiple particles, where a comparison of particles with and without the -k coda can be made, usually also support this view, even if the -k coda is not specifically said to be an intensifying morpheme. For example, Kwok (1984, 46-50) reports that laak3 conveys more certainty than laa3, that lok3 reinforces the meaning carried by lo3, and so on.

One puzzle that arises is that if -k is simply ‘intensifying’, why is it not able to attach to all utterance particles? For instance, why is -k never found with the ‘wo’ particles of any tone, i.e. *wok3, *wok4, or *wok5? This is unlikely to be due to semantic conflict, since it is not the case that the ‘base’ particle wo3, meaning ‘noteworthiness’, is either already intense or inherently moderate. It cannot be due to phonological incompatibility either, since Cantonese allows -k to follow o. Both of these are demonstrated by the particle lo3, which shares the o rhyme; lo3 does have an ‘intensified’ version lok3, as documented in the literature and in the Hong Kong Cantonese Corpus. Similarly puzzling is why laak3 exists as an intensified version of laa3, while *laakt apparently does not
occur as an intensified version of laa1. Again, this is unlikely to be due to phonological conflict with the rhyme, or with the ‘high level’ tone 1, since for example zek1 is possible as an intensified version of ze1. Why does -k attach to only a limited number of utterance particles? Sybesma and Li (2007, 1754) acknowledge that this claim regarding -k is ‘vague and multi-interpretable’.

Law notes that particles with -k codas do not precede other particles in clusters, and proposes that this is due to coda deletion if the particle is not last (Law 1990, 182-187). This leads to the loss of –k’s (supposed) meaning, and the difference between e.g. laa and laak is not recoverable when this occurs. This essentially states that whenever laa occurs without a -k coda, it means laa – this is no surprise and one cannot help but feel that this analysis unnecessarily overcomplicates the matter. It would also mean that in many cases, one cannot know which particle was being used in the cluster. For example, she states that when the particle me is added to laa/laak, after coda deletion, the two clusters have the same form, laa-me. How would one ever be sure whether the first particle was laa or laak, and if we cannot know, how do we figure out the combined meaning of the components involved? Law gives some examples of coda deletion in other word classes, such as ceo(t)1-min6 ‘outside’ and zi(k)1-haak1 ‘immediately’, where the consonants in parentheses are deleted codas. However, the meaning in these cases stays the same, and they are equally acceptable with or without codas. It would be difficult to believe that the -t or -k codas in ceott ‘out’ or zik1 ‘immediately’ have independent meanings, so this seems not to help the argument.

Nonetheless, despite the weaknesses of current analyses, it seems possible to the author that -k may indeed carry some specifiable meaning. Consider the fact that Cantonese utterance particles vary in duration to a greater extent than other Cantonese words. The vowel of a particle may be pronounced short, or prolonged for up to a second or so, with this drawing out being a characteristic of Cantonese which distinguishes it from other varieties of Chinese including Mandarin (Matthews and Yip 2011, 390). A speculation is

There does seem to be an intuitive connection between e.g. laa3 and laak3, or gaa3 and gaak3, although they are not always interchangeable. Ze1 and zek1 seem to be interchangeable in even less situations, and it seems to the author, at first glance, that there may be more of a meaning difference between ze1 and zek1 than between laa3 and laak3 or between gaa3 and gaak3.
made here that the -k coda serves the purpose of shortening the syllable, just as unreleased consonants in final position in other areas of Cantonese result in abbreviated syllables (though the only consonant available at the end of an utterance particle is -k). If they are truncated variants, it might make sense that they convey some meaning related to immediacy. This is only a speculation, which of course needs much more investigation to verify and develop. Future studies would be more valuable if they avoid unclear and problematic descriptions such as ‘intensifying’ (also see below).

As a final comment about -k, it would be very interesting to explore whether this -k coda is the same as that which can attach to the common Cantonese interjection(s) ait-jaa(k)3, which expresses sudden bad feelings. This possible connection appears not to have been made in the literature so far. The second syllable of ait-jaa3 can also be particularly long, but from native speaker intuition, it seems mostly to be short when the -k coda is used. If ait-jaak3 is a truncated form of ait-jaa3, we can speculate, as above, that possible meaning differences between the two might be related to time or immediacy. Future research could shed more light on this area.

Problems with ‘intensifiers’

There is a problem with describing multiple morphemes as having an ‘intensifying’ effect. Not only is this ‘vague and multi-interpretable’, as Sybesma and Li admit, but another problem arises in that this must result in multiple levels of supposed intensity. Fung (2000, 107-110) and Sybesma and Li (2007, 1748-1749, 1753) state that laa3, the ‘core’ meaning of the l- family, means ‘it is now the case that’, and that lo3 is the ‘more intense’ form of laa3\(^{103}\). Kwok (1984, 48-49) also reports that compared to laa3, lo3 has an intensifying function. Additionally, Fung (2000) and Sybesma and Li (2007) share the view that adding -k to lo3 and laa3 produce the more intense lok3 and laak3 respectively. According to them, then, laa3 can be intensified as laak3 as well as lo3, which can in turn be intensified as lok3. Fung (2000, 109-110) even states that laa3 is stronger than laat. Furthermore, tone 4 is also said to ‘add

\(^{103}\) According to Sybesma and Li’s (2007) system of sub-syllabic morphemes, the difference between laa3 and lo3 should be between aa ‘smooth-alert’ and o ‘noteworthiness’, not intensity. This was discussed in section 8.3.3.
strength to the counterpart with tone 3’. This is supposedly the case with $e_4$, $aa_4$, $le_4$, and even $lo_4$ (Sybesma and Li 2007, 1770). This means that $lo_4$ is stronger than $lo_3$. It is a relief that $lok_4$ seemingly does not occur and is not proposed by Sybesma and Li as an intense variation of $lo_4$, although there is no reason for its absence, since it should theoretically be possible. As it is, it is extremely unclear where the intensity expressed by $laak_3$, $lo_3$, $lok_3$, and $lo_4$ fit in with each other. It is only known that these are all stronger or more intense than $laa_3$. (Recall from Chapter 6 that $laa_3$ is itself also said to be used for ‘emphasis’.) Even if the level of intensity of each of these particles could be clarified, it needs to be seriously questioned whether this chaotic hierarchy of intensity is plausible or helpful.

If $laa_3$, $laak_3$, $lo_3$, $lok_3$, and $lo_4$ are all the same save for varying degrees of intensity, should they not be able to be used in the same context? But Fung (2000, 104-110, 124-125) reports that some cases of $lo_3$ cannot be substituted with $laa_3$, the phrase-final $lok_3$ is not interchangeable with $laak_3$, and some cases of $laak_3$ sound odd substituted with $laa_3$. This further reduces the plausibility of this multi-level scale of intensity.

Still other particles seem to ‘intensify’ differently, according to other authors. For example, Kwok (1984, 99-100) states that $ge_3$ and $ge_2$ can help the speaker to express views concerning the degree of certainty or reservation in relation to what is being discussed, with $ge_3$ indicating more certainty than $ge_2$. This would seem to be yet another way to add intensity, following the -k coda and tone 4 (and the unclear relationship between $laa_3$ and $lo_3$). Kwok does not promote the dissection of monosyllabic particles into smaller morphemes, so she does not claim that the ‘high rise’ tone 2 always indicates more certainty. It could be that some particles have a more ‘intense’ counterpart, but that this is not due to any phonological segment in particular, i.e. they do not behave as predicted by the sub-syllabic morphemes hypothesis.

It is possible that the numerous cases of describing one particle as more ‘intense’ than another particle is a somewhat lazy or uncertain description. Describing particles by relying on reference to other particles is not ideal. It leads to circularity, and it is not helpful to define one word using another word which is itself not well defined. Confusion arises over the real similarities and
differences between particles. NSM and reductive paraphrase can avoid these problems.

8.3.8 Sub-syllabic morphemes as ‘simplex particles’

In the most radical form of the sub-syllabic morphemes hypothesis, Li (2006) proposes that the minimal units of initials, rhymes, codas and tones can be seen as independent particles themselves. There would be 11 such ‘simplex particles’ in Cantonese: ge3, le, ze, me, ne, aa, o, k, 1, 4, and 5 (Li 2006, 114-115). The first seven are all inherently toneless, and the last four are not autonomous in phonology. Le and ze cannot be used in isolation; they occur only in non-final position in particle clusters, e.g. le-aa5 (laa5), ze-aa5 (zaa5). Following this, what are traditionally thought of as single, basic particles are actually particle clusters, following certain combinatorial rules. The sub-syllabic morphemes hypothesis as overviewed above is already highly problematic and unconvincing, and this is even more extreme. It is not clear what the advantages are of proposing hypothetical particles and then proposing rules about which of those hypothetical particles can and cannot combine.

It is unknown why Li (2006, 116) states that neither the published sources nor her informants report any co-occurrence of aa and o. She predicts that gaa, laa, and zaa do not occur preceding wo. This is strange since gaa3-wo3, zaa3-wo3, and laa3-wo3, not to mention gaa3-laa3-wo3 and gaa3-zaa3-wo3, are all very common in speech. These are recognised in most of the literature on Cantonese utterance particles and found easily in the Hong Kong Cantonese Corpus (see Chapter 9). One cannot help but think that her sources have been misleading. Following Li’s own claims, if aa marks ‘discourse relevance’ and o marks ‘noteworthiness’, there would be no apparent reason for aa and o not to co-occur – so her belief that they do not co-occur is actually inconsistent with her own hypothesis.

To make matters worse, some particles (or particle ‘clusters’, according to Li) cannot be created using the 11 simplex particles identified. For example, the particle gwaa3 cannot be constructed from the 11 particles. Is there any plausible reason why particles such as aa1-maa3 could not be pronounced aa3-maa1 or some other variation? Following the thinking that any of the simplex
particles can combine, it should also be possible for e.g. the addition of \(-k\) the ‘emotion intensifier’, although \(laak3-w03, laa3-wok3, zaa3-wok3\) etc. never occur, and that is unexplained. Why do some ‘combinations’ not occur? This idea of simplex particles does not seem to help our understanding of particles any more than it raises additional, unnecessary questions.

It appears that Mandarin has again had an influence here. One of Li’s goals was apparently to reduce the number of Cantonese utterance particles, so as to make them more similar to the number of particles found in Mandarin.\(^{104}\) The idea of reducing the particle inventory, like many ideas associated with the sub-syllabic morphemes hypothesis, seems to have originated with Law (1990), who was one of the first to declare in favour of reducing the size of the Cantonese particle inventory. However, why this should be necessary or desirable is unclear. It seems to have an opposite effect of complicating the particle system. Can we not accept that this is simply an area where the two languages differ?

### 8.3.9 Further observations

Numerous other particles and supposed combinations suffer from general problems within the sub-syllabic morphemes hypothesis. For example, \(aa3-laa3\) is considered a particle cluster by Kwok (2006) and is found in her corpus\(^{105}\). \(Aa1-maa3\) is considered a particle by others such as Li (2006), Matthews and Yip (2011), and Wakefield (2011b). Regardless of whether they are considered ‘basic’ or ‘clusters’, this does not make much sense within the sub-syllabic morphemes hypothesis, under which it would mean either that certain morphemes (e.g. the \(aa\) rhyme, which can occur more than once) are repeated and combined with each other, or that these morphemes are inserted for phonological reasons only, implying that they have no independent meaning.

\(^{104}\) Li looked at five final particles in Mandarin: \(ne1, ba, ma1, ma2,\) and \(a\). She did not attempt to break down the Mandarin particles into sub-syllabic morphemes (although she argues that the two \(ma\) particles are actually one). She explains that Mandarin final particles cannot be dissected into smaller meaningful units, i.e. onsets, rhymes, tones and coda.

\(^{105}\) The Hong Kong Cantonese Adult Language Corpus, based on spontaneous speech recorded from Hong Kong radio.
Moreover, some semantically similar particles are not phonologically similar. For example, Kwok (1984, 59-60) reports that the particle nei is similar to the particle lo4. As can be seen, the two particles do not share an initial, rhyme, or tone. Under the sub-syllabic morphemes hypothesis, it should be the case that these two particles’ meanings are significantly different, since they share no sounds in common.

A great deal of the problem with proposals about sub-syllabic morphemes is that many of the researchers in question have not conducted independent empirical research, but rely heavily on pre-existing literature, which is itself unreliable. Law’s (1990, 7) descriptions and examples, for instance, are taken largely from Kwok (1984), and this causes problems from the outset, as it has been shown throughout this thesis that Kwok’s descriptions are generally not comprehensive, only accounting for certain aspects of each particle. For example, Law accepts from Kwok that laa1 conveys a sense of indefiniteness, tentativeness, and uncertainty (Law 1990, 104-107). This proposition would have been easily disproven by looking at examples of laa1 in use in ordinary conversation, or indeed at other accounts of the particle found in the literature. (Multiple examples from the Hong Kong Cantonese Corpus, featured in Chapter 3, show the particle being used in contexts of ‘certainty’, co-occurring with gang2hai6 ‘of course’ and its synonyms.) If Law’s description of laa1 is inaccurate, this means that her analysis of the differences between laa1, laa3, laa4, and other particles are also likely to be inaccurate.106 Similarly, Li’s (2006) account relies on pre-existing literature, rather than on original research into usage.107 But any analysis which claims that a meaning of a particle combination is the combined meaning of its parts presupposes a rigorous and independent description of the meanings of those parts.

Finally, it should be noted that much of the literature on hypothetical sub-syllabic morphemes seems motivated not by any special interest in their semantics, but rather by the goal of examining the syntactic functions of the

106 Law (1990, 104-107) states that the difference between laa1 and laa3 is that laa1 makes the utterance sound less tentative and less certain, but we know that this is not necessarily the case. This calls into question Law’s claim that the high tone is a ‘weakener’, which is supposedly what makes the utterance sound more tentative. If laa1 is not truly ‘weaker’, can laa3 still be said to be a ‘neutral’ variant?

107 See also discussion of Sybesma and Li’s (2007) treatment of zaaz3 versus ze1 in section 7.1.
particles and the structural positions in which they occur. In other words, the motivation is theory-driven. For example, Law (1990) aimed to show that some particles are syntactic function words occupying specific syntactic positions, and to show the distributional properties of the particles. Li (2006) and Sybesma and Li (2007) investigate how Cantonese final particles relate to sentence structure, assigning each sub-syllabic morpheme a position in abstract syntactic d-structures. Each attempts to reduce the particle inventory, which is not in line with explaining many intricate meanings.

8.4 Other possible reasons for similar particles

There may be historical reasons for some particles seeming similar to others. For example, according to W-m Leung (2009, 1-3, 2010a, 94), wo3 has been in use as an utterance particle indicating ‘hearsay’ since the late 19th century. The functions of wo3 changed over time and by the second half of the 20th century, its functions expanded to ‘realisation’, ‘reminder’, ‘showing contrast’, ‘unexpectedness’ and ‘noteworthiness’. She further states that wo5 was not found in the earlier literature, but has now taken on wo3’s earlier function of expressing hearsay information.

Matthews (2007) states that the relationship between wo3 and wo5 can be described in a cross-linguistically recognised way. He posits that wo3 conveys unexpectedness and noteworthiness, belonging in the category of mirativity. This follows DeLancey (1997), who introduces the category of mirativity as marking a proposition as one which is new to the speaker, regardless of whether the information source is first- or second-hand. Matthews refers to DeLancey and Aikhenvald (2004), who discuss cases from several languages which show that languages often use the same or related forms to express mirative and

---

108 Most work in the area of sub-syllabic morphemes readily accepts the hypothesis without much justification, starting with the basic assumption that there is a correlation between sound structure and interpretation. An exception is Fung (2000), who acknowledged that Law’s study was not comprehensive with regards to semantics.

109 Possibly also relevant is that tone 3 and tone 5, although contrastively distinctive phonetically, are being merged in some varieties of Cantonese spoken outside but near Hong Kong, such as in Guangdong Province (W-m Leung 2009, 8-10). Some words originally in tone 5 are merging into tone 3, e.g. ci5 ‘to resemble’, and vice versa, e.g. si3 ‘to try’. In Hong Kong Cantonese, tones 3 and 5 are overall still very distinctly separated, but this may nevertheless cause problems for the hypothesis that particles’ tones hold independent, stable meaning. (W-m Leung (2010a, 94, 2009, 4) and Matthews (2007) report that wo3 and wo5 can sometimes be used interchangeably, although this seems to be due to genuine semantic overlap.)
hearsay functions. Matthews reasons that if wo3 is a mirative, the close relationship between the mirative wo3 and the hearsay evidential wo5 makes sense. In relation to Cantonese, he adds that this is supported by the fact that in metalinguistic commentaries native speakers tend to merge wo3 and wo5, and that there are hearsay contexts in which either wo3 or wo5 may appear (Matthews 2007, 11-13). Matthews’ view is that wo3 is the ‘basic particle’ from which wo4 and wo5 are derived.

Other Cantonese scholars believe that wo5 is a combination of the verb waa6 ‘say’ and the particle aa3, an idea which appears to have originated from Chao (1947, 121). W-m Leung (2009, 3) points to cross-linguistic studies that have shown that the verb ‘say’ tends to develop into a hearsay marker of a complementiser. This is also possible, though it does not necessarily prove that the wo family are all to do with hearsay (or saying things, any more than other particles). It is unclear whether wo3 and wo4 are also historically related to waa6 ‘say’ and aa3/other particles, and whether this is why the three particles seem to be related. It is also possible that wo5 alone originated from waa6 ‘say’, and wo3 and wo4 did not. Obviously, this is different from Matthews’ opinion described above. Without further investigation, both theories seem plausible enough. In any case, this does not seem very relevant to the meanings of the particles as they are used today. From informal discussions with native Cantonese speakers, it seems that the average native speaker is unaware of any historical connections between, say, the particles and waa6 ‘say’, so they are not influencing present-day speakers’ usage consciously.

A similar observation can be made about certain proposed particle contractions, such as the claim that gaa3 = ge3 + aa3. Even if there were some historical evidence that gaa3 was first used as a contraction of the two particles ge3 + aa3, this does not seem to be in the minds of present-day speakers of Cantonese. Based on the author’s intuition and informal discussions with other native speakers, it seems unlikely that speakers are consciously expressing ge3 + aa3 when they say gaa3 (though native speakers agree that ge3 and aa3 exist as independent particles). If gaa3 really originated as a contraction of ge3 + aa3 (and the author is not aware of any historical evidence for this), it probably happened a long time ago, and the same is true for other supposed contractions.
There does not appear to be any evidence that *zaa3*, for instance, originated as *ze3 + aa3*.

### 8.5 An alternative approach: phonesthemes and sound symbolism

A more moderate and promising approach, which does not appear to have been proposed for Cantonese utterance particles thus far, is that some of them share phonesthemes or phonesthemic qualities. Phonesthemes are recurrent pairings of sound and meaning, such that some phonemes or groups of phonemes have recognisable semantic associations without being clearly contrastive morphemes. Though pursuing this idea in any detail is beyond the scope of this thesis, the following brief review may help to lay the groundwork.

Phonesthemes in English can occur in different parts of the word, as initials, rhymes, or codas.\(^{110}\) Some well-known English phonesthemes are syllabic onsets, such as *sl-*, *gl-*, or *tw-*. Bergen (2004, 290) explains that English word onsets such as *gl-* and *sn-* are relatively infrequent, except among words to do with vision and light or to do with the mouth and nose respectively.\(^{111}\) Phonesthemes in English can also be rhymes, such as the *-ash* in *bash*, *crash*, or *smash*, or syllabic nuclei as in *drip* and *drop*, *clip* and *clop*, *ding* and *dong*, or *tick* and *tock*. Generally speaking, English phonesthemes consist of multiple segments, although this is not always the case.\(^{112}\) A selection of English examples is given in Table 13, adapted from Kwon and Round (2015) and Bergen (2004). Of course, these sounds do not have the same apparent role in all English words in which they appear.

\(^{110}\) Bergen (2004, 293-294) claims that some phonesthemes can even be whole syllables, which may possibly affect Cantonese utterance particles, if e.g. *wo* has phonesthemic qualities.

\(^{111}\) Phonesthemes are better attested in the lexicon of a language than would be predicted, all other things being equal. For example, 39% of word types and 60% of word tokens starting with *gl-* have meanings related to light or vision, compared to 0% and 1% respectively for *sn-* words. For *sn-* words, 28% of word types and 19% of word tokens are to do with the nose or mouth, compared to 4% and 1% respectively for *gl-* words (Bergen 2004, 293).

\(^{112}\) As noted by Kwon and Round (2015, 18), there is numerically more evidence for *t-* being a phonestheme to do with ‘pointed’ things (e.g. *tip*, *toe*, *tine*, *tack*, *tusk*, *taper*) than for many other English phonesthemes. The connection is less obvious, however, because so many other words also begin with initial *t-*.
Table 13: Select phonemes of English

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Meaning</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>cl-</td>
<td>'denoting sound'</td>
<td>cluck, click, clap, clack, clash, clutter, clang, clank, clamber, clamour, clam, clump, clip</td>
</tr>
<tr>
<td></td>
<td>'grasp'</td>
<td>clutch, claw, cling, close, clasp</td>
</tr>
<tr>
<td>gr-</td>
<td>'grasp'</td>
<td>grasp, grip, grab, grapple, grope</td>
</tr>
<tr>
<td>gl-</td>
<td>'visual'</td>
<td>glow, glance, glare, gleam, glim, glimmer, glimpse, glint, glisten, glister, glitter, gloaming, glower, gloom, gloss, glassy, glaze</td>
</tr>
<tr>
<td>sl-</td>
<td>'slimy/slushy matter'</td>
<td>slime, slush, slop, slough, slobber, sludge, slosh, sloppy, sluggard, sleet</td>
</tr>
<tr>
<td></td>
<td>'a falling blow'</td>
<td>slay, slaughter, slit, sling, slash, slap, slam, slog</td>
</tr>
<tr>
<td></td>
<td>'falling or sliding movement'</td>
<td>slide, slither, slip, slouch, slump, sled, slick, slack, slope, sleek, slant</td>
</tr>
<tr>
<td>sn-</td>
<td>'nasal/oral'</td>
<td>snaffle, snap, snarl, sneer, sneeze, snicker, sniff, snuffle, snore, snack, snout, snort, snigger, snivelling, snorkel, snuff</td>
</tr>
<tr>
<td>-irl/url</td>
<td>'circular'</td>
<td>twirl, curl, furl, burl, knurl, whirl, hurl, swirl, purl</td>
</tr>
</tbody>
</table>

The vast majority of phonemes in English decompose into a phoneme and a residue which is not itself a meaningful unit (Kwon and Round 2015, 4, Bergen 2004, 292-294). For example, in slobber the phoneme sl- leaves a residue -obber. Likewise, gl- can leave -int, -ow, etc. These residues appear to have no recurrent meaning in the rest of the lexicon. Notably, scholars refrain from proposing hypothetical meanings to each residue or trying to break them down further. This is reminiscent of the proposal put forward by Fung (2000) regarding Cantonese utterance particles, who attributed meanings to the three initials z-, l-, and g-, but without proposing that all the ‘residual’ rhymes or tones have independent meaning. In this respect, phonemes and Fung’s proposals differ from the more extreme claims of the sub-syllabic morphemes hypothesis, according to which all phonological segments of the particle have specifiable meanings. If we can compare the particle combinations considered in this chapter with phonemes, we can argue that even if certain shared features exist within...
some groups of similar particles, it is not necessary to require that all particles fully decompose and that all the ‘residues’ have meaning.

Bergen (2004, 290) explains that phonesthemes have been documented in such diverse languages as English, Indonesian and other Austronesian languages, Japanese, Ojibwa, and Swedish. According to Yip (2000), a similar phenomenon is found in Mandarin Chinese. Table 14 below is reproduced from Yip (2000, 178), who lists several groups of words in Mandarin with phonesthetic similarities (these are accompanied by Chinese characters in the original text). He also lists close to 30 words beginning with \( m- \), which he claims have meanings associated with things being covered or unclear, and close to 10 \( m- \) words with meanings associated with boundlessness (Yip 2000, 178-179).

Table 14: Xi- phonetheme of Mandarin

\[
\begin{array}{ll}
\hline
\text{xì} & \text{‘thin’} \\
\text{xiăo} & \text{‘small’} \\
\text{xì} & \text{‘sparse’} \\
\text{xiăn} & \text{‘rare’} \\
\text{xì} & \text{‘tall and slender’} \\
\text{xiá} & \text{‘narrow’} \\
\text{xiăn} & \text{‘fine’; ‘minute’} \\
\text{xiàn} & \text{‘limit’} \\
\hline
\end{array}
\]

Likewise, Yip (2000, 217) reports that in the lexical field of words for ‘honesty and hypocrisy’, the positive terms seem to favour retroflex consonants such as \( zh/ch/sh/r \), whereas the negative terms favour palatal consonants such as \( j/q/x/y \), as shown in Table 15. As with the English examples above, this does not mean that these sounds have the same apparent role in other areas of the lexicon, or even in all words in this field.

Table 15: Phonesthemes in Mandarin words for ‘honesty and hypocrisy’

\[
\begin{array}{ll}
\hline
\text{positive terms (favour zh/ch/sh/r)} & \text{negative terms (favour j/q/x/y)} \\
\text{chéngshí} & \text{jiăoxiá} \quad \text{‘honest’; ‘dishonest’; ‘sly’} \\
\text{zhēnzhì} & \text{jiănxié} \quad \text{‘sincere’; ‘crafty’; ‘evil’} \\
\text{zhíshuài} & \text{xūjiă} \quad \text{‘straightforward’; ‘above board’} \\
\text{zhīshuài} & \text{yīnxīăn} \quad \text{‘hypocritical’; ‘false’} \\
\text{rèchén} & \text{xūjiă} \quad \text{‘enthusiastic’; ‘earnest’} \\
\text{rèchén} & \text{yīnxīăn} \quad \text{‘insidious’; ‘treacherous’} \\
\hline
\end{array}
\]

This phenomenon is not limited to the onset in Mandarin, as Yip (2000, 218) argued that for terms indicating outward shape and size, ‘positive’ terms
apparently favour \( a/e/o/u \), while ‘negative’ terms favour \( i/ia \). His list of these terms is reproduced in Table 16.

**Table 16: Phonesthemes in Mandarin words for ‘shape and size’**

<table>
<thead>
<tr>
<th>‘positive’ terms (favour ( a/e/o/u ))</th>
<th>‘negative’ terms (favour ( i/ia ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( dà ) ‘big’</td>
<td>( xiǎo ) ‘small’</td>
</tr>
<tr>
<td>( gāo ) ‘high’</td>
<td>( dī ) ‘low’</td>
</tr>
<tr>
<td>( kuān ) ‘wide’</td>
<td>( xiá ) ‘narrow’</td>
</tr>
<tr>
<td>( shēn ) ‘deep’</td>
<td>( qiǎn ) ‘shallow’</td>
</tr>
<tr>
<td>( hòu ) ‘thick’</td>
<td>( biān ) ‘thin’</td>
</tr>
<tr>
<td>( cū ) ‘thick all around’</td>
<td>( xi ) ‘thin all around’</td>
</tr>
</tbody>
</table>

In view of these reports, it seems possible that phonesthemes may exist in Cantonese and that they may affect Cantonese utterance particles. Of course, much more research would need to be carried out before we could be sure whether any Cantonese utterance particles or groups of particles had phonesthemic qualities, but the basic idea is surely prima facie more plausible than the sub-syllabic morphemes hypothesis. With regards to NSM, it might turn out to be the case that particles which share certain phonesthemes contain similar NSM components.

As with English, we obviously cannot argue that this extends to all words in a language. Just as English words such as *glove*, *glue*, *glottis* and *gluten* contain *gl-* without being related to vision, any potential phonesthemes in Cantonese would also be limited to a certain set of words. Equally, if it so turned out that certain particles or particle groups have phonesthemic qualities, there would be no reason to assume a priori that the same applies to all particles (or that perception of them is at the same degrees). For example, the author personally perceives some intuitive similarities among particles of the \( z- \) initial family and among particles of the \( wo \) family, but particles with an \( l- \) initial or \( aa \) rhyme seem to have no obvious connection with other particles with similar formal makeup.

Finally, I would not want to give the impression that phonesthemes are fully understood, even in English. On the contrary, they seem to remain a relatively peculiar and poorly understood feature of language.
8.6 Concluding remarks

This chapter has assessed the plausibility of breaking down monosyllabic particles into smaller meaningful units. It has reviewed various proposed dissections of monosyllabic particles at different levels, with some claims being more extreme than others. Two main types of hypothetical ‘combination’ were identified and evaluated, namely contractions and sub-syllabic morphemes. It has been shown that the specifics of both are very unclear, and neither hypothesis is very convincing or helpful. The overview of two particle ‘contractions’ and a range of initials, rhymes, tones, and codas show that the problems discussed are not restricted to certain anomalous particles or morphemes, but are extensive and substantial. Generally speaking, the situation described in this chapter is the same as has been demonstrated throughout this thesis regarding individual particles – that the existing literature on Cantonese utterance particles is often inaccurate, vague, and contradictory. The semantics of the supposed meaningful units within monosyllabic particles, and the way they combine, have not been well thought through, and numerous questions are left unanswered.

Studies in this area would surely be more rewarding if more precise and accurate definitions can be established for individual particles. Without understanding the meanings of individual particles well, it is difficult to draw accurate conclusions about potential shared features between particles. We can speculate that if some particles, such as wo5, were to be explicated in NSM using the same method as in Part Two, that its explication might have some similarities or even shared components with that of wo3 (and possibly wo4). More particles first need to be explicated, and this is beyond the scope of this thesis.

We have to agree with Luke (1990, 307) that while some connections and similarities exist between particles which share overlapping segmental composition, the degree and nature of their relatedness is unknown. We suggest an alternative approach for why some particles share semantic similarities, in that some particles might involve phonemes. This cannot be investigated in detail, but could be an interesting area for further research.
Chapter 9 looks at a different kind of particle combination, namely particle ‘clusters’ which result in transparent polysyllabic combinations. It will be shown that in the case of clusters, it does appear to be true that the meaning of the combination is the combined meaning of its parts.
Chapter 9:  
The semantics of particle clusters

Returning to NSM explications of particles, this chapter explores the phenomenon of particle clusters. Recall that the term particle ‘cluster’ as used in this thesis refers to polysyllabic combinations where two or more utterance particles can be heard being used one after the other in immediate succession. Examples include gaa3-zaa3 (two particles) and gaa3-laa3-wo3 (three particles). Particle clusters have often been claimed to have the combined meaning or function of the separate particles of which they are made up (Yau 1965, Gibbons 1980, Kwok 1984, 8-15, Yip and Matthews 2000, 131-132, Matthews and Yip 2011, 396-397). As Wakefield (2011b, 13-14) notes, all studies on Cantonese utterance particles appear to agree on this point, presumably based only on native speaker intuition and assumption. In the absence of accurate semantic analyses of individual particles, this claim has hitherto been unable to be tested in any rigorous fashion. The NSM analyses in this thesis provide a new and very concrete way of testing the idea that the meaning of particle clusters is equal to the meaning of the individual particles combined.

The clusters considered in this chapter are those which feature the particles individually analysed in Part Two, i.e. laa1, wo3, gaa3, laa3, and zaa3. These five particles are obviously only a small sub-set of the ‘basic’ utterance particles, which is generally considered to be 30 or more. They are used because the meanings of the particles that make up the clusters need to themselves be accurately and systematically identified, before methodical analysis of the clusters can be conducted. The NSM explications proposed in Part Two are used to test whether they can be combined to explain the meanings of the resulting particle clusters. This exercise bears on the second main research question, outlined in section 1.4: to what extent particle combinations in Cantonese can be explained semantically, and whether the potential meaning of polysyllabic particle clusters is the same as the combined meanings of the particles of which they are made up. A preliminary investigation into the
particles gaa\(^3\), laa\(^1\), and wo\(^3\) using NSM was outlined in HHL Leung (2013, 18-24); the analysis has since been updated and expanded.

The five particles analysed in Part Two can together produce eight particle clusters which occur in speech.\(^\footnote{13}\) Therefore, eight NSM explications are assessed in this chapter, i.e. one for each possible cluster. As with the explications proposed in Part Two, they are tested by substitution into examples found in the Hong Kong Cantonese Corpus. A selection of examples of each, from varied contexts, accompanies each explication. This chapter starts by reviewing the limited existing literature on particle clusters, in section 9.1. In section 9.2, it then introduces the occurring and non-occurring clusters based on the five particles chosen. The bulk of this chapter, sections 9.3–9.5, considers each of the eight possible clusters in turn, using the corpus data to show that their combined NSM explications can explain their meanings. Finally, some possible reasons for non-occurring clusters are discussed in section 9.6.

9.1 Previous descriptions of the meanings of particle clusters

One of the rare investigations which touched on the meanings of particle clusters is that of Yau (1965). Yau’s S-Q test suggested that the last particle of any particle cluster was the dominating factor as far as the S-Q function was concerned (Yau 1965, 68). Nevertheless, Yau does acknowledge exceptions to this rule, and it has already been noted that Yau’s S-Q test is flawed. Moreover, Gibbons (1980, 774), the other to have touched on the subject, gave a different basis for determining the function of the combination (contraction or cluster). Gibbons stated that single particles of the ‘Q class’, i.e. those occurring in questions, appear to be dominant when combined with particles of other or no class(es). In other words, a combination of particles which contains a question particle will also be used in questions. Other features Gibbons calculated, such as ‘commitment to the truth’ and ‘newness’ were also supposedly carried by the

\(^{13}\) As mentioned in Chapter 8, the different kinds of (supposed) particle combinations are not mutually exclusive. Due to the problems discussed in Chapter 8 and for the sake of simplicity, this chapter will treat particle clusters as clusters only. That is, gaa\(^3\)-wo\(^3\) is treated as gaa\(^3\) + wo\(^3\) only, rather than, for example, ge\(^3\) + aa\(^3\) + w + o + tone 3, or any variation of such.
question particle, although regarding ‘strength’, combinations took on the ‘strength’ of the strongest component particle. It is of note that all of the combinations he considered were classified as having the highest strength. The usefulness of such a category is weakened if every possible combination belongs in the same class.

Regarding Yau’s C-test, one of the relevant ideas was that if a certain emotive connotation was ‘significantly’ felt in an utterance particle, then that emotive connotation would be felt again in any particle cluster which includes that utterance particle (Yau 1965, 120). Yau does not appear to have ever verified this, mentioning it only briefly. We can pick one particle cluster from Yau’s own study to test his hypothesis. His data shows that the cluster gaa3-laa1 was tested to only have the connotation ‘coaxing’. However, gaa3 and laa1 contained other connotations ‘conceited’, ‘persuading’, ‘reminding’, and ‘politely urging’. Another point to note is that the potential connotations can be contradictory. In any case, his hypothesis about the connotations of particle clusters relies on the correct analysis of the independent particles, and Part Two has shown that Yau’s connotation concepts were highly unreliable and often incorrect, even for independent particles.

A note on vowel changes

It has been observed in the literature that when utterance particles combine in clusters, vowel changes can occur. Yau (1965, 29) observed that the particles can occur in weak forms, where their vowels would shift to the corresponding central position and also become shorter. Matthews and Yip (2011, 395) state that the cluster tim1-ge3-laa3-wo3 can be pronounced tim1-ge3-lo3-wo3 or tim1-go3-lo3-wo3, i.e. with the vowel of the last particle spreading to one or both of the preceding particles. Sybesma and Li (2007, 1775) state that utterance particles that occur in non-final position in particle clusters generally feature schwas instead of e or aa.114 As Yau explains, however, the weak forms

114 Sybesma and Li (2007, 1776) view such sound changes as confirmation of their belief that the particles are just the onset with a default vowel to enhance pronounceability. Firstly, this would suggest that different vowels (rhymes) do not have separate meanings, and this seems to go against the major claims of their own paper. Secondly, as the points made throughout Chapter 8 should make clear, there is no solid evidence that specific phonological segments have independent semantic content, so even if sound changes occur, this does not mean that semantic changes occur.
are regarded as alternatives of the norm and are not contrastive to the norm. Following convention, particles in clusters are written in their usual citation forms, even if vowel changes can occur in speech.

9.2 Clusters made up of laa1, wo3, gaa3, laa3, and zaa3

9.2.1 Possible clusters and the mechanics of their combination

The particles laa1, wo3, gaa3, laa3, and zaa3 can occur in the following clusters and sequences:

- Gaa3-laa1
- Gaa3-wo3
- Gaa3-laa3
- Laa3-wo3
- Gaa3-laa3-wo3
- Gaa3-zaa3
- Zaa3-wo3
- Gaa3-zaa3-wo3

Several observations can be made about these clusters. These can be summarised as follows:

- In a cluster that includes gaa3, gaa3 always occurs first.
- In a cluster that includes wo3, wo3 always occurs last.
- Laa3 and zaa3 can both occur with gaa3, with wo3, or with both gaa3 and wo3.
- Laa1 can only occur with gaa3.
- The particles laa3 and zaa3 never occur together – they therefore appear to occupy the same position, i.e. first position of two if gaa3 is not present but wo3 is present; second and last position if gaa3 is present but wo3 is not present; or second and middle position if gaa3 and wo3 are both present.
- The positioning of the particles within clusters is rigid. The same particle combinations cannot occur in any other order. For example, *laa1-gaa3* is unacceptable and does not occur.
As can be seen, it is possible to describe and to a certain extent even predict what orders the particles will occur in when they combine. Based on this small selection of particles, we do not know for sure whether laa1 occurs in the second position or the third position. Laa1 only occurs with gaa3 out of the particles available here, so we know only that laa1 cannot be in the first position. Given that in clusters involving wo3, wo3 always occurs last, we speculate that the case would be the same with *laa1-wo3, and that *wo3-laa1 is even less plausible due to positioning of wo3. This would mean that laa1 appears to occupy the same position as laa3 and zaa3.

The possible particle clusters are shown in the table below, with each horizontal line of the table showing a possible particle cluster. As can be seen, gaa3 always occurs first, wo3 always occurs last, and laa1, laa3, and zaa3 appear to share the same position in the second slot.

**Table 17:** Possible particle clusters comprised of laa1, wo3, gaa3, laa3, and zaa3

<table>
<thead>
<tr>
<th>Gaa3</th>
<th>Laa1/laa3/zaa3</th>
<th>Wo3</th>
</tr>
</thead>
<tbody>
<tr>
<td>gaa3</td>
<td>laa1</td>
<td></td>
</tr>
<tr>
<td>gaa3</td>
<td>laa3</td>
<td>wo3</td>
</tr>
<tr>
<td>gaa3</td>
<td>laa3</td>
<td></td>
</tr>
<tr>
<td>gaa3</td>
<td>laa3</td>
<td>wo3</td>
</tr>
<tr>
<td>gaa3</td>
<td>zaa3</td>
<td></td>
</tr>
<tr>
<td>gaa3</td>
<td>zaa3</td>
<td>wo3</td>
</tr>
</tbody>
</table>

Although constraints on particle ordering can be described, the reason for this ordering is at present unknown. Researchers in the area of sub-syllabic morphemes have investigated the ordering of particles within clusters from a syntactic point of view, but without conclusive or comprehensive findings; for example, Li (2006, 116) and Sybesma and Li (2007, 1777) admit that they cannot explain why the co-occurrence of ze and le, or z and l respectively are not commonly accepted. Additionally, some of their claims are puzzling, for example Li (2006, 116) states that aa and o never co-occur, which is shown to be false by the clusters above, substantiated by the Hong Kong Cantonese Corpus.
In any case, since their analyses are based on claims about sub-syllabic morphemes which have been shown in Chapter 8 to be highly problematic, their analyses of ordering must be viewed with similar scepticism.

Could the explanation be phonological? It happens to be that in these clusters, the place of articulation of the onsets or initials of each particle moves from the back of the mouth to the front. Roughly speaking, we can say that the onset of gaa3 begins in the back of the mouth with a velar sound, then the alveolar initials of laa1, laa3, and zaa3, and finally moving towards the front of the mouth, at least if we treat the onset of wo3 as a bilabial. Considering particles not analysed in this thesis, we note that particles with a bilabial m-initial occur at the end of a cluster, in the same position as w- particles. Nonetheless, this is far from a satisfactory explanation for the ordering of the particles within clusters. Even if this ‘rule’ were always followed, it does not explain, crucially, why it is always followed. Since no logical reason for combining particles in this way is apparent, this thesis treats this phonological ‘rule’ as a mere coincidence. Furthermore, consideration of a larger number of particle clusters would show (a limited number of) deviations from this ‘rule’. For instance, laa3-gwa3 is possible and occurs in the corpus, although it is relatively less common. The puzzle of ordering is revisited in section 9.6.

In this chapter, the explications that make up the meaning of a particle cluster are placed one after the other in the same order as in speech, and refer to the utterance it is attached to, just like the individual particles in Part Two. Examining examples in the Hong Kong Cantonese Corpus can then show whether the combined explications can be substituted in place of those clusters and plausibly explain what the speaker was expressing at that moment.

The order of the analyses of particle clusters is as shown in Table 17. Firstly, gaa3-laaz and gaa3-wo3 are analysed in section 9.3. Next, three clusters involving laa3 are examined: gaa3-laaz, laa3-wo3, and gaa3-laaz-wo3. Finally, three clusters involving zaa3 are presented: gaa3-zaa3, zaa3-wo3, and gaa3-zaa3-wo3. This grouping is for expository convenience.
9.2.2 Non-occurring clusters

Some particles never combine to form clusters. The particles *laa1, laa3, zaa3, and wo3* are unable to be combined as follows, regardless of order or sequence (the particle *gaa3* is not in this list because it is able to combine with every other particle analysed in Part Two). Any longer clusters (e.g. of three particles) which include any of the unacceptable pairings below are also unacceptable.

*Laa1-wo3*  
*Laa1-laa3 / *laa3-laa1*  
*Laa1-zaa3 / *zaa3-laa1*  
*Laa1-wo3 / *wo3-laa1*  
*Laa3-zaa3 / *zaa3-laa3*

Note that besides *laa1-wo3* (assuming, as above, that *laa1* remains in the middle position and that *wo3* remains in the last), all the other non-occurring clusters are comprised of particles sharing the same position in Table 17, i.e. the middle column. It may be that the particles *laa1, laa3, and zaa3* do not occur together because they each must be in the same position of a cluster. More observations regarding non-occurring clusters, including *laa1-wo3*, are discussed in section 9.6.

9.3 Clusters of two: *gaa3-laa1* and *gaa3-wo3*

We start by looking at two clusters, consisting of two particles each. The first is *gaa3-laa1*. *Laa1* does not combine with any of the other particles analysed in Part Two. The second cluster in this section is *gaa3-wo3*. *Wo3* does occur in several other clusters involving *laa3* and *wo3*, which will be discussed further below. In each of the following sections, the combined explication for the particle cluster is proposed, followed by examples of the cluster in use. It is not feasible to show as many examples for each particle cluster as were given for each individual particle in Part Two, partly for reasons of space and partly because some clusters may be less common. Nonetheless, as much as possible, a range of different examples are given to show various interpretations of the particle cluster/explication.
9.3.1 Gaa3-laa1

If the particle cluster gaa3-laa1 has the same meaning as gaa3 + laa1, then the NSM explication for gaa3-laa1 can be expected to consist of the explications for gaa3 and laa1 combined. By placing the NSM explications for gaa3 (from Chapter 5) and for laa1 (from Chapter 3) one after the other, the resulting explication is as shown in [9A]. Combination of the explications for gaa3 and laa1 together happens to be possible because the proposed explications are semantically compatible with each other. The explications do seem to be able to ‘combine’.

[9A] Proposed explication for gaa3-laa1:

\[
\text{it is good if you know this} \\
\text{you now know how I think about this} \\
\text{l can not-say more}
\]

\[
jyu4gwo2 \text{ lei5 z}1\text{dou3 zau6 hou2} \\
\text{if you know then good}
\]

\[
\text{lei5 ji4gaa1 z}1\text{dou3 ngo5 dim2 lam2} \\
\text{you now know I how think}
\]

\[
\text{li1 jat1 joeng6 je5} \\
\text{this one CL thing}
\]

\[
\text{ngo5 ho2ji5 m4 zoi3 gong2} \\
\text{l can not more say}
\]

This proposed explication can indeed be substituted where gaa3-laa1 occurs, and can explain what gaa3-laa1 means in each case. This is revealed by looking at examples from the Hong Kong Cantonese Corpus. Note that not all examples can be described as ‘coaxing’ (Yau 1965, 120).

(9.1)

A:

\[
\text{Ciisin3. Daap3 fei1gei1 heoi3 gwo3 go2bin1} \\
crazy ride airplane go over that-side
\]

\[
fan3gaau3 \text{ aa4?} \\
sleep PRT
\]

‘Crazy. Fly all the way over there to sleep aa4?’
B: 
Gam2 fong3gaa3 hai6 relax. Hai6 gam2
then holiday is relax is like-that

gaa3 laa1.
PRT PRT

‘Holidays are for relaxing. It’s like that [that’s the way it is] *gaa3-laa1.*’

Example (9.1) above is about going on holiday to the USA. Speaker B has suggested that speaker A sleep while in the US, and speaker A seems shocked at this. As with the explications in Part Two, we can gain an insider’s perspective into what the speaker is saying/thinking. Speaker B uses *gaa3-laa1* to indicate firstly ‘it is good if you know this’ – that it is good if speaker A knows that she should relax while on holiday. This might be because it supports B’s argument, or because B is reminding her friend to relax. *Gaa3-laa1* also indicates ‘you now know how I think about this’ – speaker A should know how speaker B thinks of holidays and that they are for relaxing. Finally, *gaa3-laa1* expresses ‘I can not-say more’ – speaker B does not need to say more, because relaxing on holiday is to be expected. In this way, speaker B has defended the idea of sleeping in the US while on holiday. Speaker A understands and moves on to the next topic. This use of *gaa3-laa1* is friendly and casual, like a cool reminder, piece of advice, or comment.

If speaker B’s utterance had used only *gaa3*, this could have implied that speaker A would not understand his statement ‘holidays are for relaxing, that’s the way it is’. The utterance would have sounded like speaker A was not expected to know that holidays are for relaxing. This is because *gaa3* ‘it is good if you know this’ on its own can sound like an instruction or teaching, and it is *laa1* that conveys that the hearer is expected to know and understand what is being said.

If the utterance had used only *laa1*, this utterance would probably have come across less like a piece of friendly advice or friendly reminder. Speaker B might even have sounded annoyed or impatient, because *laa1* indicates only that the speaker assumes the hearer knows what the speaker thinks, and that the speaker does not need to say more. Use of *laa1* by itself does not indicate that it is good if the hearer knows this.
A similar analysis can be made of example (9.2), which is from a much later part of the same conversation. Speaker B uses *gaa3-laa1* in an utterance saying that holidays are great for eating, playing, and shopping. Using *gaa3-laa1*, speaker B indicates that it is good if speaker A knows this, that A now knows how she thinks of this, and that she can not-say more.

(9.2)

B:

\[
\begin{align*}
\text{Gam2} & \quad \text{fong3} & \quad \text{gaa3} & \quad \text{zeoi3} & \quad \text{hou2} & \quad \text{heoi3} & \quad \text{sik6} & \quad \text{aa3}, \\
\text{so/then} & \quad \text{holiday} & \quad \text{best} & \quad \text{go} & \quad \text{eat} & \quad \text{PRT} \\
\text{waan2} & \quad \text{aa3}, & \quad \text{maai5} & \quad \text{je5} & \quad \text{aa3}, & \quad \text{zau6} & \quad \text{zeoi3} & \quad \text{hou2} \\
\text{play} & \quad \text{PRT} & \quad \text{buy} & \quad \text{thing} & \quad \text{PRT} & \quad \text{then} & \quad \text{best}
\end{align*}
\]

**gaa3 laa1.**

PRT PRT

‘Well the best thing to do on holiday is to eat aa3, play aa3, shop aa3, that’s the best *gaa3-laa1*.’

A:

\[
\begin{align*}
\text{Hai6} & \quad \text{lo1...} \\
\text{yes} & \quad \text{PRT}
\end{align*}
\]

‘Yes lo1...’

Example (9.3) is from a conversation about postage stamp collecting, and the cluster here refers to something slightly different. Speaker A has been talking about the quick and large increase in the value of old $1.80 stamps (old stamps with the Queen’s image are no longer issued in Hong Kong). Speaker A is recalling his past actions of using them up quickly, without realising their worth. The *gaa3-laa1* here is meant in a comforting way. The thing which from speaker B’s perspective is ‘good to know’ is that $1.80 stamps are no longer so popular and no longer increasing in price. Speaker B believes that A knows what she thinks about this now – that there is no point lamenting a missed opportunity – and can not-say more about it, since it is over.

(9.3)

A:

\[
\begin{align*}
\text{So2ji5} & \quad \text{ngo5} & \quad \text{ji4} & \quad \text{gaa1} & \quad \text{nam2} & \quad \text{hei2} & \quad \text{di1} \\
\text{that’s-why} & \quad \text{I} & \quad \text{now} & \quad \text{think} & \quad \text{of} & \quad \text{CL} \\
\text{g03baat3} & \quad \text{jau4piu3}, & \quad \text{ngo5} & \quad \text{maang5gam3} \\
\text{one-dollar-eighty} & \quad \text{stamp} & \quad \text{I} & \quad \text{ferociously/persistently}
\end{align*}
\]
That’s why when I think of those $1.80 stamps now, I ferociously/persistently sent, ferociously/persistently sent [kept sending in large amounts] all those at that time.

B:
Samitung3 laa3. Daan6hai6 ji4gaa1 jau6
heart-pain PRT but now instead/again
mou5 ni1 jat1 joeng6 gaa3 laa1,
not-have this CL thing PRT PRT
hai6 m4hai6 aa3?
is not-is PRT

‘Heartache laa3. But there aren’t these things now gaa3-laa1, right aa3?’

The cluster gaa3-laa1 is used slightly differently in example (9.4) below. This time, it is more ‘blunt’, not intended to be a friendly, casual comment or reminder as in examples (9.1) or (9.2), nor comforting as in example (9.3). Here, two people are having an enthusiastic debate about football clubs and players. When speaker B says that he is going to support Germany in the World Cup, speaker A says that the German team is boring to watch. Speaker B agrees that they are indeed boring, but produces the following example of gaa3-laa1. This conversation was also shown in Chapter 4.

(9.4)
B:
Jeng4 aa6 dak1 gaa3 laa1 zeoi3 gan2jiu3.
win then/so can PRT PRT most important
Nei5 pung2 deoi6 bo1 ci3 ci3 syu1
you support CL ball CL CL lose
hou2ci5 jing1gaak3laan4 gam2 dim2 gaau2 aa3?
like England then/so how make PRT

‘It’s okay as long as they win gaa3-laa1, that’s the most important. If you support a team that loses every time like England, then what aa3?’

Speaker B indicates that his friend’s arguments are irrelevant and that she should know that winning is the most important thing. This is presented as an easily understood fact that does not require more explanation. He poses a
rhetorical question to make his point, which he believes to be very obvious. Example (9.4) is not like (9.1) or (9.2) because it is presented almost like ‘teaching’ the hearer. It might be considered rude in other situations depending on the relationship between the speaker and the addressee. Speaker B’s utterance is presented as very obvious and is intended to put down speaker A’s views (though they are clearly good friends enjoying some banter).

(9.5)

A: 
Gam2  ho2  m4  ho2ji5  dak1haan4  lo2  
so/then can not can available/free take  
dii  din6sam1  faan1  uk1kei2  jung6  aa3?  
CL  batteries  return  home  use  PRT  
‘Then can you take some batteries back home to use when you are free aa3?’

J: 
M4  dak1  gaa3. (9.5a)  Nei5  dou1  zii  gaan1 
not can  PRT  you also/even know  CL  
gung1si1  gei2  zai3dou6faa3  gaa3  laa1. (9.5b) 
company  how  systematic  PRT  PRT  
Faan4hai6  ling1  dii  je5  nei5  dou1  
whenever  take  CL  thing  you  even/also/all  
jiu3  dang1gei3  lok3.  
need  register  PRT  
‘[I] can’t gaa3. You also know how systematic [/structured] the company is gaa3-laa1. Whenever you take things you have to register lok3.’

Example (9.5) is about someone’s work at a company that produces batteries, from a conversation also shown in Chapters 3 and 4. Here, the friend is asking if she can take the batteries home, presumably for free. The first gaa3, labelled (9.5a), indicates ‘it is good if you know this’. Gaa3-laa1, labelled (9.5b), indicates ‘it is good if you know this’, and also ‘you now know how I think about this, I can not-say more’. Gaa3 adds value in answering the question and providing relevant information, and laa1 ‘reminds’ the hearer that the company is very systematic. Addition of laa1 here makes sense because the utterance refers to something that the hearer supposedly knows, as indicated by ‘nei5 dou1 zii’ ‘you also know’. Compare this with m4 dak1 gaa3 ‘can’t gaa3’ in (9.5a), where speaker B cannot really expect A to know how she thinks of it already,
since A is asking a question. Thus it makes sense that (9.5a) uses only gaa3, and (9.5b) uses the cluster gaa3-laa1.

(9.6)

G:

Wai3, gam2 gang2hai6 zou6 zou6 haa5 hok6 -
hey but/then of-course do do DEL learn
jat1bin1 zou6 jat1bin1 hok6 gaa3 laa1. (9.6a)
one-side do one-side learn PRT PRT

‘Hey, but of course you work a bit then learn – work and learn at the same
time gaa3-laa1.’

B:

Ngo5 zii.
I know
‘I know.’

G:

Gam2 go3go3 ceot1 dou3 lai4 dou1 hai6
so/then everyone out arrive come all is
gam2 gaa3 laa1. (9.6b)
this-way PRT PRT

‘Everyone is like that when they come out [to work] gaa3-laa1.’

In example (9.6) above, speaker B is talking about how he cannot always tell what people in business mean, because they do not necessarily say what they are really thinking. Speaker G is comforting him, telling him that this is normal and happens to everybody. The two occurrences of gaa3-laa1 in example (9.6) indicate ‘it is good if you know this, you now know how I think about this, I can not-say more’. Speaker G is pointing out to B that he can of course learn these things as he works, and that all new workers are the same. This is presented as ‘clear’ or ‘obvious’ information, and B should be able to understand this and to know that G is essentially comforting him or telling him not to worry. The information is so ‘obvious’ and ‘understandable’ that G also indicates she does not have to say more. As explained in Chapter 3, laa1 frequently attaches to utterances with gang2hai6 ‘of course’, and it seems that addition of gaa3 does not affect this. This is not surprising since there is nothing in the meaning of gaa3 which clashes with ‘of course’ or its synonyms. The following section looks at the particle cluster gaa3-wo3.
9.3.2 Gaa3-wo3

Using the same reasoning as with gaa3-laa1 above, we can propose a combined explication, as shown in [9B] below, for the common cluster, gaa3-wo3, and assess this against real examples of gaa3-wo3 from the Hong Kong Cantonese Corpus. Gaa3-wo3 generally points out to the hearer something that the speaker perceives them to not know or not be thinking about, but which would be good for the hearer to know.

[9B] Proposed explication for gaa3-wo3:

it is good if you know this
you aren’t thinking about this at this moment
it is good if you think about it

\[
\begin{align*}
jyu4gw & \text{wo2 lei5 zi1dou3 zau6 hou2} \\
\text{if you know then good} & \\
lei5 & \text{li1 jat1haak1 m4 hai6 lam2 gan2} \\
\text{you this moment not is think PROG} & \\
\text{li1 jat1 joeng6 je5} & \\
\text{this one CL thing} & \\
jyu4gw & \text{wo2 lei5 lam2 haa5 zau6 hou2} \\
\text{if you think DEL then good} & \\
\end{align*}
\]

As with gaa3 and laa1, the combination of the explications for gaa3 and wo3 was found to be possible because the proposed explications are semantically compatible. Promisingly, this proposed composite explication can be substituted where gaa3-wo3 occurs, and plausibly explains what gaa3-wo3 means in various corpus examples.

(9.7)

A: ‘Australia’s very hot? When aa3?’

B: ‘If you go in summertime, it means you are going when it’s winter in Hong Kong.’

A: ‘You mean in December?’

B:

\[
\begin{align*}
\text{Hai6 aa3, hoeng1gong2 sap6jii6jyut6 go2si4 heoi3,} & \\
is/yes PRT Hong-Kong December that-time go & \\
\end{align*}
\]
We can first consider the three examples of *gaa3-w03* in (9.7), where two people are talking about going on holiday to Australia. Part of this excerpt was also discussed in Chapter 5. In (9.7a), speaker A clearly does not know when it becomes hot in Australia. Speaker B states that Australia gets very hot in December and this utterance is suffixed with *gaa3-w03*. In this example, the speaker may be expressing ‘it is good if you know this’ for several reasons. Speaker B is giving speaker A some useful general knowledge, particularly
because both speakers have mentioned in this conversation that they want to visit Australia, and also because speaker A has explicitly asked about it. Speaker B also expresses ‘you are not thinking about this at this moment’, because it is clear that A does not know Australia gets very hot in summer, and ‘it is good if you think about it’, because it is a useful fact or even a warning to consider before going. Speaker A’s response shows that this information is news to her and that she considers it, even wanting to know more.

After (9.7a), speaker B goes on to state that it is very sunny in Australia, providing (9.7b). As above, she feels it is good for speaker A to know this because it is useful general knowledge and because B knows that A wants to go to Australia. Speaker B indicates that speaker A is not thinking about this at this moment, firstly because B is introducing a new fact, and secondly because A has stated that she does not know this about Australia. Again, she expresses ‘it is good if you think about it’ because it is a useful fact to consider. It is evident from A’s response that this makes sense to her and that she considers how serious it sounds.

Speaker A provides the next case of gaa3-w03 in (9.7c), asking about the flies in Australia. As the proposed explication of gaa3-w03 shows, speaker A is indicating that it is good if speaker B knows there are many flies in Australia, that speaker B was not thinking about this at this moment, and that it is good if B thinks about it now. It is good if B knows about the flies in Australia because A clearly wants B to tell her about them, and A acknowledges and draws attention to the fact that B is not thinking about the flies because it is another slight change in topic. It is good if B thinks about it, because A is asking a question and clearly wants to know.

(9.8)

A:
Wai3 ci4 di1 heoi3 m4 heoi3 leoi5hang4 aa3?
hey late bit go not go holiday PRT

Nei5 lou5gung1 jau5mou5 peng4 gei1piu3 aa3?
you husband have-not-have cheap flight-ticket PRT

‘Hey, [are you] going on holiday later aa3? Does your husband have cheap flight tickets aa3?’
Example (9.8) above shows another instance of gaa3-wo3. Speaker A’s question implies that she thinks cheap flights are available, and B’s answer rejects A’s apparent assumptions (relevant background we know from the rest of the conversation is that B’s husband works in the travel industry, thus A mentions him). Speaker B expresses ‘it is good if you know this’ because the attached information is general knowledge and because her statement is the answer to A’s question. Speaker B also highlights that A is not thinking about the fact that the cost of flights varies according to season. She indicates that it is good if she thinks about it, perhaps because this is ‘obvious’ information, or perhaps because she thinks that if A thought about it more, she would realise the answer to her own question.

Note that if example (9.8) had only used wo3, the statement that cheap flight tickets are only available in the off-season would not sound as much like speaker A should know it or that B is ‘teaching’ A something useful or obvious. If this example had only used gaa3, it would have sounded more like B was teaching A something that she might not be expected to know.

(9.9)

B:
Keoi5 battinau1 dou1 cat1 hou6
He [Beckham] always also/still seven number
gaa3 wo3. (9.9a)
PRT PRT
‘He [Beckham] was always number seven gaa3-wo3.’

A:
Bik1haam4 sap6 hou6 gaa3 wo3 (9.9b) gau6nin2.
Beckham ten number PRT PRT last-year
‘Beckham was number ten gaa3-wo3 last year.’
Examples (9.9a) and (9.9b) show gaa3-wo3 being used in a conversation with a different tone or mood compared to examples above. The excerpt is from the same conversation as example (9.4), where two friends are having a debate about football. Here, they are arguing about what number shirt Beckham wears. In each turn, the speaker uses gaa3-wo3 to indicate firstly that it is good if the other person knows Beckham wears a certain number shirt, i.e., number seven or number ten. This corresponds with the first line of the proposed explication [9B], ‘it is good if you know this’. At the same time, both speakers register a disagreement and that the other thinks Beckham wears a different number. This corresponds with the second line of the explication, ‘you aren’t thinking about this at this moment’. Both try to back up their arguments to persuade the other of their belief. They each want the other to think about the fact that Beckham ‘was always number seven’ or ‘was number ten last year’. This corresponds with the last line of the explication, ‘it is good if you think about it’.

(9.10)

J:
Gam2 nei5 gok3dak1 hou2 m4 hou2waan2 aa3?
then you feel good/fun not fun PRT
‘Do you feel that it’s fun aa3?’

A:
Hou2 m4 hou2 waan2?
good not good fun
‘Is it fun?’

J:
Hai6 aa3.
is/yes PRT
‘Yes aa3.’

A:
M4 ho2ji5 jung6 go3 waan2 zi6
not can use CL fun/play word
lai4 jing4jung4 gaa3 wo3.
come/to describe PRT PRT
‘Can’t use the word “fun” to describe it gaa3-wo3.’
J:

Gam2 aa6 hai6, gam2 nei5 jau5mou5
like-this then is then you have-not-have
hing3ceoi3 zou6 niidit je5 aai1?
interest do these thing PRT

‘That’s true, then do you have any interest in doing these things aai1?’

In example (9.10), A and J are discussing A’s work as a researcher. J asks A whether she finds the work ‘fun’, and A responds that ‘fun’ is not the right word. The two turns before the gaa3-wo3 suffixed utterance also show A’s bemusement or confusion, as A repeats J’s question ‘hou2 m4 hou2 waan2’ ‘is it fun?’. A uses gaa3-wo3 to indicate that it is good if J knows ‘fun’ is the wrong descriptor. This should be something that J can be expected to work out or understand if she thinks about it, since they are talking about work. A’s use of gaa3-wo3 also indicates to J that she is not thinking about this at the moment, and that it is good if J thinks about it. It appears that after A says this gaa3-wo3 suffixed utterance that J understands where the ‘problem’ has occurred. J responds with ‘that’s true’, and then rephrases her question in a way that does not use the word ‘fun’.

(9.11)

A:

Tung4maai4 tai2 haa5 dii jan4 gong2 dak1
and see DEL CL people speak ADV
hou2 m4 hou2. Di1 gaau3sau6 hou2 m4 hou2.
good not good CL professor good not good

‘And [I’ll] see whether the people speak well. [See if] the professors are
good or not’

B:

Waa3 daan6 nei5 m4 jap6heoi3
wow but you not go-in
nei5 m4 zi1 gaa3 wo3.
you not know PRT PRT

‘Wow but you can’t know that without going in [to study] gaa3-wo3.’

In example (9.11), speaker B asks A how he will choose which university to go to. He says he will see which one has the best professors, but B is surprised, and points out that he cannot know these things until he is already studying there.
She uses *gaa3-wo3* to convey to A that it is good if he knows he cannot know whether the professors are good before studying there, and since this is fairly ‘obvious’, she points out that he is not thinking about this at this moment, and that it is good if he thinks about it.

Example (9.12) shows two instances of *gaa3-wo3*. Speaker A is asking B about whale-watching tours, and at this point in the conversation has just mistakenly brought up feeding the whales. More turns preceding this excerpt are shown in example (9.19).

(9.12)

A:

\[O3 \ zik1hai6 \ wai3 \ keoi5 \ sik6 \ je5 \ go2di1\]

oh meaning feed them eat thing those

\[laa3 \ w03\] . (9.19b)  

PRT  PRT

‘Oh you mean like feeding them [the whales] *laa3-wo3*.’

B:

\[M4hai6 \ aa3. \ Keoi5 \ ni1go3 \ nei5 \ hai2...\]

not-is PRT it/they this you at

\[keoi5 \ go2 \ zek3 \ m4hai6 \ gaa3 \ keoi5 \ go2 \ zek3\]

it/they that CL not-is PRT it/they that CL

\[hai6 \ tai2 \ tin1jin4 \ go2di1 \ gaa3 \ w03, \ (9.12a)\]

is watch natural those PRT PRT

\[m4hai6 \ jan4dei6 \ joeng5 \ gaa3 \ w03. \ (9.12b)\]

not-is people raise/rear PRT PRT

‘No *aa3*. This one is... that one is not *gaa3* that one is to watch the natural ones [whales] *gaa3-wo3*, they’re not raised by people *gaa3-wo3*.’

In the examples of *gaa3-wo3* labelled (9.12a) and (9.12b), *gaa3-wo3* expresses ‘it is good if you know this’ because it is good if A knows that the whales are wild and not raised by humans (and therefore you do not feed them). *Gaa3-wo3* also expresses ‘you aren’t thinking about this at this moment’ as a reference to A’s belief that you can feed the whales – A clearly is not thinking that the whales are not fed by humans. Finally, *gaa3-wo3* expresses ‘it is good if you think about it’, to stress the significance of the wildness of the whales and to help point at the silliness of A’s thinking that tourists can feed them.
9.4 Clusters involving laa3

This section looks at clusters which involve laa3 and are comprised of the particles analysed in Part Two, i.e. gaa3-laa3, laa3-wo3, and gaa3-laa3-wo3.

9.4.1 Gaa3-laa3

In the same way as previously, a composite explication can be proposed for the cluster gaa3-laa3 by placing the individual explications for gaa3 and laa3 one after the other. Because the two explications are compatible, they can be combined. Moreover, the proposed composite explication [9C] is able to explain examples of gaa3-laa3 from the corpus.

[9C] Proposed explication for gaa3-laa3:

it is good if you know this
I want you to think now: ‘I know it’s like this’

\[\text{ju}u4\text{gwo2 lei5 zi1dou3 zu}u6 hou2\]
if you know then good
\[\text{ngo5 soeng2 lei5 ji}4\text{gaa1 lam2:}\]
I want you now think
\[\text{‘ngo5 zi1 hai6 gam2joeng2’}\]
I know is this-way

In example (9.13), speakers A and B are talking about A’s friend, who wants to study physiotherapy. This example was also shown in Chapters 6 and 7. Speaker B is seeking confirmation that only one university in Hong Kong offers physiotherapy, and A confirms this using a gaa3-laa3 suffixed utterance. Gaa3-laa3 indicates firstly that it is good if B knows that the other universities do not offer physiotherapy, since B has asked explicitly. Gaa3-laa3 then indicates that, above and beyond it simply being good if B knows this, A wants B to ‘register’ or ‘note’ the knowledge. A wants B to think ‘I know it’s like this [that the other universities do not offer physiotherapy]’. B understands and moves on to talk about something else. This use of gaa3-laa3 is relatively without emotion, compared to some of the other examples shown further below.

(9.13)

B: ‘Really met? Only Poly [Hong Kong Polytechnic University] offers/teaches physiotherapy zaaz me1?’
A:  
_Hai6  aa3. Kei4taa1  dou1  mou5  gaa3  laa3._

is  PRT  other  all  not-have  PRT  PRT

‘Yes aa3. The others [universities] don’t have it _gaa3-laa3._’

B:  ‘What about [your] other classmates _ne1_ what are they thinking of studying?’

In example (9.14) below, speaker M is talking about the lack of men in her workplace that she might be able to date. She uses _gaa3-laa3_ to lament and express (jokingly) to speaker E that it is good if he knows she has no hope, and that she wants E to think ‘I know it’s like this’, perhaps to reinforce the finality of her dire situation.

(9.14)

M:  ‘Me _ne1_, out of those new teachers _ne1_, there’s one there’s only one teacher who is male _laa1_.’

E:  ‘Yes _aa4_?’

M:  ‘But that male teacher _ne1_, has to be about forty-something years old the kind that’s all married with children _ne1_.’

E:  ‘Oh, you even know that he’s got children.’

M:  _Hou2  caam2  aa3. Ngo5  ji5ging1_  very  poor/suffering  PRT  I  already  
mou5  saai3  heimong6  gaa3  laa3.  
not-have  all  hope  PRT  PRT

‘Very poor _aa3_. I already have no hope [/I’ve already lost all hope] _gaa3-laa3._’

Example (9.15) shows a speaker commending the hearer’s school and fellow students. The speaker uses _gaa3-laa3_ as part of a compliment, indicating to the hearer that it is good if he knows that his schoolmates are much better than at other schools, and that she wants him to think now ‘I know it’s like this’, i.e. to accept what she is saying. This use of _gaa3-laa3_ in a compliment is different to the ‘complaint’/‘grievance’ in (9.14) and the objective statement in (9.13).

(9.15)

_Kei4sat6  nei5dei6  gaan1  hok6haau6  go2di1  jan4_  actually  you:PL  CL  school  those  people
beih2hei2 kei4taa1 hok6hau6 go2di1
compared-to other school those
hou2 hou2do1 do1 gaa3 laa3.
good very much PRT PRT

‘Actually the people in your school are much better than those in other schools *gaa3-laa3.*’

Example (9.16) shows two people discussing satisfactory A-level grades. In (9.16a), in response to speaker B’s question about his confidence leading up to his A-level exams, speaker A states that he has no confidence and will be satisfied with a B grade. Using *gaa3-laa3*, he expresses ‘it is good if you know this, I want you to think now: “I know it’s like this”’. This is probably because it is new information to B and the answer to her question. A wants her to understand how he feels and accept that a B grade will be good enough for him. This is somewhat like example (9.14), in which the speaker is talking about a ‘sad’ situation that s/he wants the hearer to understand and accept. B does understand and accept this and agrees that B is good enough (the example of *gaa3-laa3-wo3* is discussed in section 9.4.3). Example (9.16b) is the same as (9.16a), as speaker A states that Bs and Cs are acceptable to him.

(9.16)

B: ‘Do you have confidence about your A-level exams aa3?’

A: ‘No [don’t have].’

B: ‘No confidence?’

A:

*Ng05 jau5 dii B zau6 syun3*
I have some B then let-it-go/enough

*gaa3 laa3.*(9.16a)
PRT PRT

‘If I have some Bs it’s enough/I’ll be satisfied *gaa3-laa3.*’

B:

*Jau5 dii B dou1 gau3*
have some B already/even enough

*gaa3 laa3 wo3.*(9.24)
PRT PRT PRT

‘Getting some Bs is [good] enough *gaa3-laa3-wo3.*’
A: 
B B C C ji5ging1 ho2ji5 gaa3 laa3. (9.16b)  
B B C C already can PRT PRT  
‘B, B, C, C, [Bs and Cs] can already [is already okay] gaa3-laa3.’

B: 
Hai6 aa3. Ji5ging1 ho2ji5 gaa3 laa3 (9.16c) 
yes PRT already can PRT PRT  
zan1thai6.  
truly  
‘Yes aa3. That will do gaa3-laa3 really.’

In example (9.16c), it is speaker B who uses gaa3-laa3. Her utterance sounds accepting and comforting. We know from the explication of gaa3-laa3 that she is agreeing by saying to A ‘it is good if you know this [that it will really be okay/Enough]’ and ‘I want you to think now: “I know it’s like this [that it will really be enough]”’. Her nice tone is somewhat similar to the use of gaa3-laa3 in example (9.15), although this utterance is not a compliment.

(9.17)  
Zik1hai6 bat1gw03 jiu3 hai6 daam6gwai3  
meaning but need in/at off-season/low-season  
ge3 si4hau6 ne1. Zau6 m4 m4 –, jat1ding6  
LP time PRT so not not definitely  
m4 wui5 cat1 baat3jyut6 gaa3 laa3.  
not will seven eight-month PRT PRT  
‘I mean but it needs to be at the time of the off-season ne1. So not not –, definitely not in July or August gaa3-laa3.’

Finally, gaa3-laa3 is used in a conversation about going on holiday, shown in example (9.17). The speaker states that she will definitely not go on holiday in the peak season of July or August. Gaa3-laa3 indicates to the hearer ‘it is good if you know this’, as this is a good and relevant point to the conversation, and ‘I want you to think now: “I know it’s like this”’, as the speaker wants the hearer to really know and ‘register’ this fact.

9.4.2 Laa3-wo3

The composite explication for laa3-wo3 is presented in [9D]. This combined explication makes good sense and can explain the seemingly highly varied
functions of the cluster. Matthews and Yip (2011, 396) state that laa3-wo3 indicates ‘current relevance + noteworthiness’. This corresponds well with explication [9D], where the first line (i.e. the explication for laa3) emphasises the utterance’s relevance, especially as the line includes the prime NOW. The last two lines (i.e. the explication for wo3) can convey ‘noteworthiness’ as the speaker feels that it is good for the hearer to think about something.\textsuperscript{115}

[9D] Proposed explication for laa3-wo3:

I want you to think now: ‘I know it’s like this’
you aren’t thinking about this at this moment
it is good if you think about it

ngo5 soeng2 lei5 ji4gaa1 lam2:
I want you now think
’nngo5 zi1 hai6 gam2joeng2’
I know is this-way

lei5 li1 jat1haak1 m4 hai6 lam2 gan2
you this moment not is think PROG

li1 jat1 joeng6 je5
this one CL thing

jyu4gwo2 lei5 lam2 haa5 zau6 hou2
if you think DEL then good

Example (9.18) below features speaker A’s shaky defense about her relationship with her and B’s mutual friend. It is from the same conversation as one shown in Chapter 5. B is incredulous that A is talking about her past relationship with their mutual friend and stresses that the good times of their relationship were a long time ago (in fact, three years ago). B clearly thinks this is too long ago to be relevant to the present time – as B shows in the first turn of this excerpt, A’s relationship with their friend is not good.

\textsuperscript{115} Laa3-wo3 is sometimes considered to be one single particle. For example, Kwok (1984) lists it as a ‘basic’ particle (as opposed to ‘derived’). In Kwok’s list, it is one of two disyllabic particles which are considered ‘basic’; the other ‘basic’ particles are all monosyllabic. Following the simple and straightforward distinction used throughout this thesis, I treat laa3-wo3 as a cluster comprised of the two particles laa3 and wo3, because it is disyllabic and audibly made up of those two particles combined.
(9.18)

B:
Nei5 loeng5 go3 gwaan1hai6 gam3 m4 hou2 gam2.
you two CL relationship so not good like-that
‘You two have such a bad relationship like that.’

A:
Ngo5dei6 ji5cin4 di1 gwaan1hai6 hou2 hou2 gaa3.
we before CL relationship very good PRT
‘We used to have a very good relationship gaa3.’

B:
Ji5cin4 gwaan1hai6? Waa3! Hou2 noi6
before relationship wow very long-time
laa3 w03. Gei2do1 nin4 cin4?
PRT PRT how many year before
‘Before? Wow! [That was] a very long time ago laa3-w03. How many years ago?’

A: ‘Three years.’

B: ‘Three years ago laa1!’

A: ‘Three years ago our relationship was very good gaa3.’

The laa3-w03 suffixed utterance in (9.18) shows B trying to ‘remind’, point out, or ‘emphasise’ to A how long ago her relationship with this person was. It could be interpreted as a mild teasing or joking, but it appears that there is an element of truth here. B’s amazement or surprise at A bringing up their past is also demonstrated by her use of waa3 ‘wow’. Laa3-w03 helps B to express firstly that B wants A to think now ‘I know it’s like this’ – that it was a very long time ago. A needs to know, to understand and accept the fact that the good times in their relationship were a very long time ago. B also expresses ‘you aren’t thinking about this at this moment’, because it seems that A does not realise the significant length of time that has passed, or perhaps is not thinking about how their relationship is not good in the present. Finally, B expresses ‘it is good if you think about it’ to A. This is a way for B to express to A that she needs to realise and accept these facts. This seems to be a way for B to tell his friend to move on and not dwell on their past relationship which is, in B’s eyes, no longer relevant. The use of laa3-w03 helps B simultaneously express his incredulity.
and his appeal that she accept the facts. This and similar examples in the corpus show that Kwok (2006, 58) is incorrect in saying laa3-wo3 is mostly used with objective information where neither the speaker nor the hearer is the focus of attention in the utterance.

Two examples of laa3-wo3 are shown in example (9.19) below, which are very different from the one used in example (9.18). This example shows a conversation about whale-watching in Australia. It directly precedes the excerpt shown in (9.12) on p.326. Speaker B would like to go on a whale-watching tour and is explaining what that is, while A, who clearly does not know anything about whale-watching, asks B questions about it. While (9.18) above showed laa3-wo3 used in an utterance where the speaker had lots of certainty and wanted to ‘emphasise’ what he was saying, speaker A in (9.19) below is uncertain about whale-watching and uses laa3-wo3 to check what she thinks with B. (9.19a) and (9.19b) function somewhat like tag questions.

(9.19)

B:

Zik1hai6 keoi5 wui5 jau5 - jau5 zek3
meaning they will have have CL
syun4 ceot1hoi2. Gam2joeng2 zau6.
boat go-out-to-sea like-that then
‘I mean they’ll have – have a boat to take you out to sea. Like that.

A:

Zik1hai6 dou3si4 zi6gei2 join go2di1
meaning at-that-time self join those
local tour laa3 w03, (9.19a) ji3si1 hai6.
local tour PRT PRT meaning is
‘You mean when you get there you’ll join those local tours by yourself laa3-w03, you mean.’

B:

E6, jyu4gw02 hai6 zi6gei2 heoi3 zau6 hai6 lo1...
eh if is self go then is PRT
‘Eh, if you go by yourself then yes lo1...’

[Three turns omitted]
B:

... Hai6 tai2 hoi2 aa3, tai2 saan1wu4
is look sea PRT look coral
go2di1 je5 lo1.
those thing PRT

‘Look at the sea aa3, look at the coral and things lo1.’

A:

O3 zik1hai6 wai3 keoi5 sik6 je5 go2di1
oh meaning feed them eat thing those

laa3 wo3. (9.19b)
PRT PRT

‘Oh you mean like feeding them [the whales] laa3-wo3.’

B:

M4hai6 aa3. Keoi5 ni1go3 nei5 hai2...
not-is PRT it/they this you at
keoi5 go2 zek3 m4hai6 gaa3 keoi5 go2 zek3
it/they that CL not-is PRT it/they that CL
hai6 tai2 tin1jin4 go2di1 gaa3 wo3, (9.12a)
is watch natural those PRT PRT
m4hai6 jan4dei6 joeng5 gaa3 wo3. (9.12b)
not-is people raise/rear PRT PRT

‘No aa3. This one is... that one is not gaa3 that one is to watch the natural ones [whales] gaa3-wo3, they’re not raised by people gaa3-wo3.’

In (9.19a), speaker A is making an uncertain assumption that people join a local whale-watching tour after they get to Australia. She is unsure about this, and so seeks confirmation from speaker B. Her utterance with laa3-wo3 functions like a tag question, as if she is saying something like ‘you join the local tours when you get there, right?’ although the same effect may be noticeable without the particles present. Using the proposed explication [9D], we can understand her utterance to mean firstly ‘I want you to think now: “I know it’s like this”’. This is plausible as A wants B to know the answer to her question, and to think about whether you join local tours when you arrive. She also expresses ‘you aren’t thinking about this at this moment’, because she has changed the topic slightly to how to join such tours, and ‘it is good if you think about it’, because if B thinks about it she can ‘check’ A’s assumption and respond with the relevant information. The use of laa3-wo3 successfully expresses to B that A wants her
statement to be checked and corrected if necessary, and B responds appropriately in the next turn.

Example (9.19b) is similar to (9.19a). Speaker A ‘asks’ B whether people feed the whales. She says this like a statement and uses laa3-wo3 to check this or get confirmation, again a little like a tag question. A is expressing ‘I want you to think now: “I know it’s like this”’ because she wants B to know and be able to tell her. She also expresses ‘you aren’t thinking about this at this moment’ since B has not mentioned feeding the whales in the previous turn. She indicates ‘it is good if you think about it’. If she is correct that you feed the whales, this serves as a way to ‘remind’ B of this missed point, or perhaps to elicit more information about it from B. She successfully ‘checks’ whether people feed the whales – she has completely misunderstood, and B corrects her in the next turn. The two examples of laa3-wo3 in (9.19) show that Kwok (2006, 58) is incorrect in saying that ‘laa3-wo3 occurs when the speaker presupposes the hearer may not have the knowledge on the proposition conveyed by the utterance’ – in these examples, the speaker is clearly hoping that the hearer will have the knowledge, such that she hearer can enlighten the speaker.

(9.20)
B: Nei5dei6 haa6nin2 saam1jyut6 zau6 haau2 laa3 wo3 you(PL) next-year March then test PRT PRT jing1man2.

English
‘You are sitting your English exams next March laa3-wo3.’

A: Hai6 aa3.
is PRT
‘Yes aa3.’

Example (9.20) shows another different use of laa3-wo3. Here, speaker B brings up the topic of an English exam that speaker A is to sit. Presumably, A already knows that he is to sit the exam, and this is supported by his response in the next turn. In this context, it is less likely that B is ‘asking’/‘checking’ about the exam as with the whale-watching examples in (9.19) (although this interpretation is possible). It is more likely that B is saying this as a kind of
‘reminder’ or to bring it up so that the conversation can move on to this new topic. Either way, the proposed explication still explains its use. Speaker B expresses ‘I want you to think now: “I know it’s like this”’ because A should know that he is to sit the exam. Moreover, if he knows this, he can give extra information about it to B. The speaker expresses ‘you aren’t thinking about this at this moment’ since it is a change in topic. She also expresses ‘it is good if you think about it’, as a reminder or even warning to him that his exams are coming up, or as a way to indicate that B would like A to tell her more about it.

(9.21)
B:
‘And but the salary is very low zaa3-w03 (9.32a). For one months’ work ne1 you still only get three-thousand-and-something dollars zaa3-w03 (9.32b).’

A:
‘Oh.’

B:
Ng05 go2zan6si4 Form Five heoi3 zou6 I at-that-time form five go do syu2kei4gung1 saam1cin1leng4 man1. summer-job three-thousand-something dollar Ji4gaa1 jau6 hai6 saam1cin1leng4 man1. now also/still is three-thousand-something dollar Daan6 ji5ging1 hai6 gei2do1 nin4 cin4 aa3? but already is how-many year before PRT Ng5 nin4 cin4 laa3 w03. five year before PRT PRT ‘When I had a summer job in Form Five [I got] three-thousand-and-something dollars. Now it’s also/still three-thousand-and-something dollars. But that was already how many years ago aa3? Five years ago laa3-w03.’

Example (9.21) shows another example of laa3-w03. A student of speaker B’s has gotten a summer job and speaker B is telling speaker A about it. The speaker is commenting that the pay for a summer job has not changed in five years. The use of laa3-w03 here has the effect of ‘pointing out’ or ‘emphasising’ that the speaker is referring to something that happened five years ago. This example has some similarities with the laa3-w03 in (9.18), although the speaker
is pointing out the ridiculousness of the situation rather than the ridiculousness of the hearer. The first turn includes two instances of zaa3-wo3 which are featured in example (9.32) on p.347.

Laa3-wo3 in (9.21) above indicates that B wants A to know that the pay has not changed in five years. B wants A to think now: ‘I know it’s like this’. B points out that A is not thinking about this at this moment – B believes this because of A’s unenthusiastic response to B’s original statement that the job only pays three-thousand-and-something dollars. The speaker indicates that it is good if speaker A thinks about it, so that he can grasp how low the pay is.

(9.22)

Gan1zyu6 dai6ng5 jat6 ne1 jau6 zou2 gei1
and-then fifth day PRT again/also early flight
zau2 wo3. Jat1 hei2san1 ne1 sik6 jyun4
leave PRT as-soon-as wake-up PRT eat finish
zou2caan1 zau6 zou2 laa3 wo3.
breakfast then leave PRT PRT
‘And then on the fifth day ne1 [we] left on an early flight wo3. As soon as [we] woke up ne1 finished breakfast and then had to leave laa3-wo3.’

Example (9.22) above shows a speaker describing her trip to Indonesia. She is saying that the trip was very bad because she did not get much time to really enjoy Indonesia, since they were always travelling. This conversation was also shown in Chapter 6. The speaker uses laa3-wo3 to emphasise that she had to leave early on the last day of her trip and did not get to spend much time there. Laa3-wo3 indicates firstly that the speaker wants the hearer to now think ‘I know it’s like this’, i.e. to know what happened and understand the situation. The speaker also expresses ‘you aren’t thinking about this at this moment’ as it is an added point to a longer conversation about how little time she had free in Indonesia. Laa3-wo3 also indicates ‘it is good if you think about it’ because the speaker is trying to be very ‘persuasive’ or ‘convincing’ about how poorly her trip went.

9.4.3 Gaa3-aa3-wo3

This section looks at the final cluster involving laa3. It is the first cluster examined in this chapter that is made up of three particles. The proposed
explication for the cluster gaa3-laa3-wo3 is presented in [9E] below. Examples of gaa3-laa3-wo3 from the corpus are given further below.

[9E] Proposed explication for gaa3-laa3-wo3:

\[
\text{it is good if you know this} \\
\text{I want you to think now: 'I know it's like this'} \\
\text{you aren't thinking about this at this moment} \\
\text{it is good if you think about it}
\]

\[
jyu4gwo2 \ \text{lei5} \ \text{zi1dou3} \ \text{zau6} \ \text{hou2} \\
\text{if you know then good}
\]

\[
go5 \ \text{soeng2} \ \text{lei5} \ \text{ji4gaa1} \ \text{lam2:} \\
\text{I want you now think}
\]

\[
'ngo5 \ \text{zi1} \ \text{hai6} \ \text{gam2joeng2'} \\
\text{I know is this-way}
\]

\[
\text{lei5} \ \text{li1} \ \text{jat1haak1} \ \text{m4} \ \text{hai6} \ \text{lam2} \ \text{gan2} \\
\text{you this moment not is think PROG}
\]

\[
\text{li1} \ \text{jat1} \ \text{joeng6} \ \text{je5} \\
\text{this one CL thing}
\]

\[
jyu4gwo2 \ \text{lei5} \ \text{lam2} \ \text{haa5} \ \text{zau6} \ \text{hou2} \\
\text{if you think DEL then good}
\]

Example (9.23) shows some friends talking about where to go for speaker A’s birthday. Speaker A says she does not want to go to karaoke because she does not want them to spend too much money. Her two friends want to go to karaoke and point out that not going to karaoke will be even more expensive. Speaker E’s gaa3-laa3-wo3 indicates that it is good if A knows that the karaoke would include dinner, and is therefore not too expensive. E wants A to understand and to think now ‘I know it’s like this’. E also expresses ‘you aren’t thinking about this at this moment’ because A does not seem to be taking into account the inclusion of dinner, and ‘it is good if you think about it’, so that she will think about it. Using gaa3-laa3-wo3 in this way, the speaker points out that karaoke is not too expensive because it includes dinner, and prompts A to understand and think about it. Thus we can see that the explications for gaa3, laa3, and wo3 combine together and all apply to the utterance. The example of zaa3-wo3 is discussed in section 9.5.2.
A: 
Fai3si6 jiu3 nei5dei6 po3fai3 laa1.
not-bother need you (PL) spend-too-much-money PRT

‘I don’t want to bother [make] you guys to spend too much money [by going to karaoke] laa1.’

M: 
Daan6hai6 ngo5 geng1 m4 kaa1laa1OK
but I scared not karaoke
zung6 gwai3 wo3.
even-more expensive PRT

‘But I’m scared [worried] that not going to karaoke will be even more expensive wo3.’

E: 
Zung6 po3fai3 wo3 ngo5 geng1.
even-more spend-too-much-money PRT I scared
Kaa1laa1OK sik6 maai4 faan6
ekaraoke eat with/include rice/meal

gaa3 laa3 wo3 go2 zek3.
PRT PRT PRT that CL

‘I’m scared that it will be even more expensive wo3. Karaoke would include eating dinner gaa3-laa3-wo3.’

M: ‘Yes lo1.’

E: ‘Big Echo advertises a ten dollar meal za3-wo3 (9.33).’

M: ‘I mean I mean cheaply everybody contribute a hundred – a hundred or so dollars then that’s acceptable/satisfactory gaa3-laa1-maa3.’

Example (9.24) is from the same conversation as (9.16) above. Speaker B comforts speaker A that getting a B grade is good enough, saying it is good if A knows this, and that B wants A to think now ‘I know it’s like this’. She also expresses ‘you aren’t thinking about this at this moment, it is good if you think about it’. She is trying to be comforting and to point out to the hearer that B will be good enough. Recall from (9.16) that she is trying to do the same in the next utterance, i.e. example (9.16c). It makes sense that her earlier utterance uses gaa3-laa3-wo3 while her second uses gaa3-laa3 only, since it becomes clear between her two turns that speaker A understands, and therefore B does not need to express ‘you aren’t thinking about this at this moment’.
(9.24)
B: ‘Do you have confidence about your A-level exams aa3?’
A: ‘No [don’t have].’
B: ‘No confidence?’
A: ‘If I have some Bs it’s enough/I’ll be satisfied gaa3-laa3 (9.16a).’
B:
Jau5 di1 B dou1 gau3
have some B already/even enough
gaa3 laa3 wo3.
PRT PRT PRT

‘Getting some Bs is [good] enough gaa3-laa3-wo3.’

Later in the conversation, in example (9.25), the two are discussing past exams. Speaker A tells speaker B that he got a B3 (the better subdivision of B) grade for English. B is impressed by this, and states that getting a B (of any subdivision) in English is already very impressive. Speaker B’s comment seems to be motivated by her genuine admiration of his grade, and/or was intended as a compliment. B’s gaa3-laa3-wo3 expresses that it is good if A knows getting a B in English is very impressive, and that she wants him to think now ‘I know it’s like this’. It also conveys ‘you aren’t thinking about this at this moment, it is good if you think about it’ – immediately before this excerpt, he is boasting about having received A grades in other subjects, so it is possible that B wanted to reassure him that a B in English is very impressive too. Alternatively, it is also possible that she simply wanted him to take in her compliment; this would be very different from example (9.23) above, but is still explained by the proposed explication.

(9.25)
A: ‘[I got] B3 [the better subdivision of B] aa3.’
B:
Jing1man2 haau2 B dou1 hou2
English test B also/already/still very
saiilei6 haau5 gaa3 laa3 wo3.
impressive DEL PRT PRT PRT
‘Getting a B in English is very impressive \textit{gaa3-laa3-w03}.’

A: ‘Yes aa3.’

In example (9.26), the speakers are fiddling with something (the thing is not identified). Speaker E uses the cluster \textit{gaa3-laa3-w03} three times within a short time. Since he is essentially repeating himself and is referring to the same thing, each occurrence of \textit{gaa3-laa3-w03} appears to be expressing the same meaning as the others in this excerpt. Speaker E seems to be expressing ‘it is good if you know this, I want you to think now: “I know it’s like this”’ because he is informing M of something which she is not aware of. We know that she is not aware of it because of her response ‘huh?’. At the same time, he expresses ‘you aren’t thinking about this at this moment, it is good if you think about it’ because he wants to draw her attention to the thing. This is successful as she pays attention to it after this excerpt.

(9.26)

E:

\begin{verbatim}
Ji2 ho2ji5 \underline{gaa3 laa3 w03. (9.26a)}
\end{verbatim}

\begin{verbatim}
oh/eh can PRT PRT PRT
\end{verbatim}

\begin{verbatim}
Nei5 ho2ji5 mang1 \underline{gaa3 laa3 w03. (9.26b)}
\end{verbatim}

\begin{verbatim}
you can pull PRT PRT PRT
\end{verbatim}

‘Oh can \textit{gaa3-laa3-w03}. You can pull \textit{gaa3-laa3-w03}.’

M:

\begin{verbatim}
Haa2?
\end{verbatim}

huh/what

‘Huh?’

E:

\begin{verbatim}
Nei5 ho2ji5 mang1 go2 gau6 je5
\end{verbatim}

\begin{verbatim}
you can pull that CL thing
\end{verbatim}

\begin{verbatim}
\underline{gaa3 laa3 w03. (9.26c)}
\end{verbatim}

\begin{verbatim}
PRT PRT PRT
\end{verbatim}

‘You can pull that thing \textit{gaa3-laa3-w03}.’

This section has overviewed all the possible clusters using laa3 created from the set of particles analysed in Part Two. The next section looks at the three possible clusters involving zaa3.
9.5 Clusters involving zaaz3

Clusters involving zaaz3 make up the three remaining clusters which could be produced using the particles from Part Two. Zaaz3 occurs in the same position as laaz3 in clusters, so the subsections here correspond to those in the previous section. The first cluster examined is gaaaz-zaaz3, followed by zaaz3-woz3, and then gaaaz-zaaz3-woz3.

9.5.1 Gaaaz-zaaz3

The composite explication for gaaaz-zaaz3 is shown in [9F]. Observe that the explications of gaaaz and zaaz3 are semantically compatible. Corpus examples of gaaaz-zaaz3 are shown below (some of them will be familiar from Chapter 7).

[9F] Proposed explication for gaaaz-zaaz3:

\[
\begin{align*}
\text{it is good if you know this} \\
\text{it is like this, (it is) not more} \\
\text{someone can feel something because of this}
\end{align*}
\]

\[
\begin{align*}
\text{jyu4gwo2} & \text{ lei5 zi1dou3 zau6 hou2} \\
\text{if you know then good} \\
\text{hai6 gam2joeng2,} & \text{ m4 hai6 do1di1} \\
\text{is like-this/this-way not is more} \\
\text{jau5jan4 ho2ji5 jan1wai6 gam2 gok3dak1} \\
\text{someone can because like-this/this-way feel} \\
\text{jau5 je5} \\
\text{there-is something}
\end{align*}
\]

In example (9.27), two speakers are discussing their friend who has been offered a job, but at the salary of the rank below him. Recall from Chapter 7 that zaaz3 can be used when talking about non-physical things. The same is true of gaaaz-zaaz3.

(9.27)

\[
\begin{align*}
\text{Gam2joeng2 keoi5 tung4 aa3 Lily jat1cai4 heoi3} \\
\text{this-way s/he with prefix Lily together go} \\
\text{in ge3, gan1zyu6 hau6mei1 - bat1gwo3} \\
\text{in [interview] PRT following-this later but}
\end{align*}
\]
Lily say give me listen PRT PRT
Keoi5 waa6 - zik1hai6 aa3 Lily jau6 hai6
She say meaning prefix Lily again/even is
m4 z1 teng1 - ho2nang4 teng1 aa3 Jenny
not know listen maybe listen prefix Jenny
gong2 gam2joeng2 laa...
say this-way PRT

‘So s/he went with Lily to the interview ge3, and then after that – but [this is what] Lily told me gaa3-zaa3. She said – I mean Lily also heard I don’t know – maybe heard it from Jenny like that laa1...’

Gaa3-zaa3 here means that the speaker thinks it is good if the hearer knows that she only heard this information from Lily. The speaker does not want to seem ‘responsible’ as the source of this information. She communicates to the hearer that Lily told her, and there is nothing more to it than that. The speaker also indicates that someone can feel something because of this – perhaps the hearer can feel doubtful or guarded, or careful about how much to trust this information.

Example (9.28) is similarly about how reliable some information is, and again shows that gaa3-zaa3 can be used even when not referring to physical objects. It shows a speaker discussing postage stamps. This is the same example as (7.20).

(9.28)
Zik1hai6 matije5 zok3 - zik1hai6 go2di1 gau3si1 aa3,
meaning what make meaning those idea PRT
zik1hai6 bin1go3 hai6 cit3ga13 aa3, ji5 matije5 wai4
meaning who is design PRT by what be
zyu2tai4 aa3, dim2joeng2 joeng2, ham6baang6laang6
theme PRT how like all
cyun4bou6 mou5 jan4 gong2 gw03, so2ji5
all not-have people say EXP that’s why
go3g03 hou2ci5 mung4caa4caa4 aa3, hou2ci5
CL CL [everyone] seems like confused PRT seems like
dyun3gu2 gam2 gaa3 zaa3.
guess like that PRT PRT
‘I mean what will make – I mean those ideas aa3, I mean who is doing the design aa3, using what as the theme aa3, how it will be like, all of these nobody said anything about, so everyone seems confused aa3, just like guessing gaa3-zaa3.’

In example (9.29) below, speakers A and B are discussing that to get into university, only the results of the best six subjects from secondary school are counted. In (9.29a), speaker A is saying to B that it is good if she knows that only the marks from her best six subjects are taken into consideration. This could be because it is a fact and speaker A wants to correct B’s mistaken thoughts. She conveys that it is only her best six subjects and not more. Someone – speaker B – can feel good about this, because speaker B was slightly worried about her marks in mathematics. This example is different to examples (9.27) and (9.28), as it refers to things of low quantities.

(9.29)

A:

Keoi5 zing6hai6 gai3 nei5 zei3 gaa1
they/it only calculate you most good
go2 luk6 fo1 gaa3 zaa3. (9.29a)

‘They only calculate [take into consideration] your best six subjects gaa3-zaa3.’

B:

Hai6 me1?
is PRT

‘Yes me1? [Really?]’

A:

Hai6 aa3. Jaa6sei3 gaa3 zaa3 (9.29b) so2ji5
is PRT twenty-four PRT PRT that’s-why
nei5 ge3 total go3 fan1 hai6 dak1. Keoi5 m4
you LP total CL mark is only they/it not
wui5 gai3 maai4 nei5 sou3hok6 go2di1 gaa3.
will calculate also you mathematics those PRT

‘Yes aa3. Twenty-four gaa3-zaa3, your total mark can only be. They won’t include your mathematics and such gaa3.’

In (9.29b), speaker A is saying to B that it is good if she knows that her maximum mark can only be 24. This could be good simply for clarification, but
also supports speaker A’s earlier remark that only the best six subjects are taken into consideration. Speaker A says that the maximum mark is 24, not more. Speaker B can again feel something good as a result, because she is worried about her marks in mathematics.

(9.30)

B:  
Gam2 hai6 jing1, e6, jing1neoi5wong4 ziicin4 ne1?  
Then/so is British eh British-female-royal before PRT  
Kei4sat6 nei5 gei2si4 hoii ci2 cou5 jau4piu3 gaa3?  
Actually you when start collect postage-stamp PRT  
‘Then eh, before the Queen ne1? When did you start collecting postage stamps gaa3?’  

A:  
Kei4sat6 ngo5 cou5 jau4piu3 ge3 si4gaan3  
Actually I collect postage-stamp that time  
hou2 dyun2 gaa3 zaa3.  
very short PRT PRT  
‘Actually I’ve been stamp-collecting for a very short time gaa3-zaa3.’

Example (9.30), a copy of (7.11), is different from the other examples of gaa3-zaa3 above, as it refers to a short period of time. Speaker B asks speaker A about old stamps, specifically those with an image of the British queen. Speaker A explains that he has not been collecting stamps for that long and does not have any with the Queen’s image on. Speaker A uses gaa3-zaa3 to indicate that it is good if speaker B knows that A has been stamp-collecting for a very short time. This is good because it directly answers A’s question. Speaker A also uses gaa3-zaa3 to express that s/he has been stamp-collecting for a short time, not more. Speaker B can feel something because of this – perhaps feel good that s/he knows the answer, or good that s/he can work out that A probably does not have any stamps with the Queen’s image on, or perhaps B can even feel bad/disappointed that A does not have any.

(9.31)

Ngo5 loeng5 go3 B saam1 gaa3.  
I two CL B three PRT one CL  
B sei3 zaa3.  
Zau6 hai6 go3 zung1man2  
B four CL then/just is CL Chinese
Example (9.31) shows a speaker talking about his exam results. The speaker has been asked about his exam results and in this part of the conversation is partly boasting about his good scores in English (in which he got B3s – the better subdivision of B) and partly a complaint/whinge about his poor scores in Chinese. He is saying to the hearer that it is good if she knows that it was just his Chinese that was so bad. He indicates that it was just his Chinese and not more. This achieves his part boast and part complaint about his English and Chinese results. Someone can feel something because of this – it is clear that he feels bad because of his poor marks in Chinese, but it could be the case that he also wants the hearer to feel impressed by his good marks in other subjects. We know from the rest of the conversation that his friend is indeed impressed and surprised at his good grades.

9.5.2 Zaa3-wo3

The explication proposed for zaa3-wo3, based on the explication of zaa3 and the explication of wo3, is presented in [9G]. Examples from the corpus show its use in natural, spoken Cantonese. Zaa3-wo3 appears to be relatively less common in conversation.

[9G] Proposed explication for zaa3-wo3:

\[
\text{it is like this, (it is) not more} \\
\text{someone can feel something because of this} \\
\text{you aren’t thinking about this at this moment} \\
\text{it is good if you think about it} \\
\text{hai6 gam2joeng2, m4 hai6 do1di1} \\
\text{is like-this/this-way not is more} \\
\text{jau5jan4 ho2ji5 jan1wai6 gam2 gok3dak1} \\
\text{someone can because like-this/this-way feel} \\
\text{jau5 je5} \\
\text{there-is something}
\]
The first example of zaa3-w03 is labelled (9.32). It is the first part of example (9.21). A student of speaker B’s has gotten a summer job and speaker B is telling speaker A about it. We can see that speaker B is commenting on the low pay. In (9.32a) she is saying that the pay is ‘skinny’ (i.e. little), not more. The someone who can feel something because of this may be herself, as she is clearly troubled or at the very least interested by this. It could also be that she believes her student would feel something because of this, or that she believes the hearer may be interested. She expresses to the hearer that he is not thinking about this at this moment, probably because she has only just brought it up. She also indicates that it is good if the speaker thinks about it. We know from the rest of the conversation that she wants to keep talking about this.

(9.32)

B:
Tung4maai4 bat1gwo3 jan4gung1 dou1 hou2
and but salary also very
ngan1 zaa3 w03. (9.32a) Zou6 jat1 go3
‘skinny’/little PRT PRT do one CL
jyut6 ne1 dou1 hau6 dak1
month PRT still/even is only
saam1cin1leng4 man1 zaa3 w03. (9.32b)
three-thousand-something dollar PRT PRT
‘And but the salary is very low zaa3-w03. For one months’ work ne1 you
still only get three-thousand-and-something dollars zaa3-w03.’

In (9.32b), the speaker comments that the pay is only three-thousand-and-
something dollars, and not more. Someone can feel something because of this – it is clear that speaker B feels alarmed and disturbed by it, and she seems also to want the hearer to feel alarmed and disturbed by it. She seems to want some kind of reaction or support from the hearer. She indicates that the hearer is not thinking about this at this moment, probably because he has not already
reacted, but possibly because this is new information to him. She then indicates that it is good if the hearer thinks about it. Again, she seems to want some kind of reaction from the hearer, so wants him to reflect on what she has said. Recall from example (9.21) that B then goes on to comment on how the pay has not risen since she had a summer job five years prior. Throughout (9.32) and (9.21) she uses wo3 to try to get a reaction out of the hearer.

(9.33)

E: ‘I’m scared that it will be even more expensive wo3. Karaoke would include eating dinner gaa3-laa3-wo3 (9.23).’

M: ‘Yes lo1.’

E:

Big Echo maai6 gwong2gou3 sap6 man1 jat1
Big Echo sell advertisement ten dollar one
caan1 zaa3 wo3.
meal PRT PRT

‘Big Echo advertises a ten dollar meal zaa3-wo3.’

M:

Zik1hai6 zik1hai6 peng4 peng4 dei2 mui5 jan4
meaning meaning cheap cheap ‘ish’ each person
gaap3 baak3 - baak3 leng4 man1
contribute/share hundred hundred or-so/around dollar
zau6 syun3 gaa3 laa1 maa3.
then let-it-go/finish PRT PRT PRT

‘I mean I mean cheaply everybody contribute a hundred – a hundred or so dollars then that’s acceptable/satisfactory gaa3-laa1-maa3.’

Example (9.33) is a continuation of (9.23), where friends talk about where to go for A’s birthday. Recall that speaker A has said she does not want to go to karaoke because she does not want her friends to spend too much, but her two friends say that not going to karaoke will be even more expensive. In (9.23), speaker E has indicated that the karaoke would include dinner, and is therefore not too expensive. Speaker E goes on to say that Big Echo, a karaoke venue, advertises a $10 meal, which is very cheap. His zaa3-wo3 indicates to his friends that the meal is $10 and not more. Probably, they can all feel glad about or comforted by this, or understand that going to karaoke will not be too expensive. He also points out that speaker A is not thinking about the fact that
they can eat dinner for $10 – this is apparent from her comment that it will be expensive. Finally, speaker E indicates that it is good if speaker A or their other friends think about this. Clearly, E would like to go to karaoke and wants his friend to reconsider.

(9.34)

\[
\text{Sap6baat3} \quad \text{dak1} \quad \text{hou2} \quad \text{siu2} \quad \text{zaa3} \quad \text{wo3}.
\]

eighteen have very few PRT PRT

‘Eighteen is very few zaaz-wo3.’

In example (9.34), two people are talking about points needed to get into certain study programmes. One of them points out that eighteen is very few. As explained by explication [9G], the speaker’s zaaz-wo3 indicates that eighteen is very few and not more. Because of this, someone may feel surprised or relieved or happy or sad about this. The speaker points out that the hearer is not thinking about this at this moment, as the hearer has not mentioned that this is very few. Finally, the speaker thinks it is good if the hearer thinks about it.

9.5.3 Gaa3-zaa3-wo3

This section describes the final cluster that can be made up of the particles analysed in Part Two, namely gaa3-zaa3-wo3. It is the second of two particle clusters comprised of three particles. The composite explication is proposed below in [9H]. Notice that all three of the original, separate explications are semantically compatible. Examples from the corpus show gaa3-zaa3-wo3 in use in ordinary Cantonese (some have been featured in earlier chapters).

[9H] Proposed explication for gaa3-zaa3-wo3:

- it is good if you know this
- it is like this, (it is) not more
- someone can feel something because of this
- you aren’t thinking about this at this moment
- it is good if you think about it

\[
\text{jyu4gwo2} \quad \text{lei5} \quad \text{zi1dou3} \quad \text{zau6} \quad \text{hou2}
\]

if you know then good

\[
\text{hai6} \quad \text{gam2joeng2}, \quad \text{m4} \quad \text{hai6} \quad \text{do1di1}
\]

is like-this/this-way not is more
Example (9.35) shows two speakers discussing Disney World in Orlando. This example of gaa3-zaa3-wo3 shows speaker A explaining that there is not much to eat inside the theme park.

(9.35)
A: ‘But it’s still quite expensive wo3. Still need $14000, four or five days. Only going to play at that place Orlando zaa3-wo3. Nothing to do.’

B: ‘So eat and play that’s right. But I –’

A: 
Hai6 leoi5bin6 sik6 lo1, leoi5bin6 mou5
is inside eat PRT inside not-have
mattje5 sik6 gaa3 zaa3 wo3.
what-thing eat PRT PRT PRT

Hai6 m4hai6 aa3? Nei5 dou1 heoi3 gwo3 laa1.
is not-is PRT you also go EXP PRT

‘Eat inside lo1, there’s not much to eat inside gaa3-zaa3-wo3. Right aa3? You’ve also been laa1.’

B: ‘Just eat those what hamburger,’

A: ‘Yeah lo1.’

B: ‘French fries gaa3-zaa3.’

In this example, speaker A is saying that it is good if speaker B knows that there is not much to eat inside, and that it is like this, not more. Someone can feel something because of this – this likely refers to herself, as she has already indicated that she does not find this holiday idea appealing. It could also refer to speaker B, as speaker A might feel that she will change her mind about
wanting to go upon realising that there is not more to eat. Speaker A also
indicates that speaker B is not thinking about this at this moment. This makes
sense because in the preceding turn, speaker B has said that you can eat and
play there, seemingly having overlooked that there is not much to eat there.
Finally, speaker A indicates that it is good if speaker B thinks about the fact that
there will not be much to eat there. This appears to be an attempt to change B’s
mind about going. It is also partly a question, as speaker A asks ‘is-not-is’
afterwards – so she may have wanted speaker B to think more about it to tell her
more about what it is like inside. All three of the particles *gaa3*, *zaa3*, and *wo3*,
and all three of their meanings, apply to the utterance. The example (9.35)
encompasses informing, warning, or reminding. Speaker A could be said to be
doing any of these things. One explication [9H] is versatile enough and allows
for all of these interpretations, without being restricted to any.

(9.36)

A:

*Keoi5 zou6. Ngo5 dou1 hai6 teng1 jan4 gong2*

he do I also/even is listen people say

*gaa3 zaa3 wo3. Nei5 zii ng05 mou5*

PRT PRT PRT you know I not-have

*wan2 keoi5 gaa3 laa1.*

find he PRT PRT

‘He’ll do it [take the job]. I heard it from other people *gaa3-zaa3-wo3.*
You know I don’t contact him *gaa3-laai.*’

B: ‘I know...’

Example (9.36) is from a later part of the conversation shown in (9.27). The
speakers return to the topic of their friend who has been offered a job at the
salary of the rank below him. Speaker B asks whether their friend will take the
job. Speaker A responds with an instance of *gaa3-zaa3-wo3* (and *gaa3-laai*).
*Gaa3-zaa3-wo3* can be explained by using explication [9H]. Speaker A says it is
good if B knows that she only heard this information from somebody else; she is
indicating that she is not the original source and that her information might not
be reliable. She explains that she heard it from other people and not more. She
did not witness this herself, and there is not more evidence for her statement.
Someone can feel something because of this – most likely, she expects speaker B
to feel that he should place less trust in what she is saying, or to feel that he knows more about the situation and how speaker A came by this information. Speaker A indicates that speaker B is not thinking about this at this moment, and that it is good if speaker B thinks about it, possibly because it is new information and speaker A wants to ‘warn’ him that she does not know for sure.

(9.37)

A:

\textit{Hou2do1 sei3 fan6 gyun2 dou1 A saai3.}

very-many four CL scroll/exam all A all

‘Many [people in my class] got As for all four papers.’

B:

\textit{Waa3 ng05 go2zan6si4 wui6haau2 dou1 hai6}

wow I at-that-time school-cert even/also is

\textit{lo2 C gaa3 zaa3 wo3. Zung6 zaang1 di1}

take C PRT PRT PRT even/still almost CL

\textit{lo2 D tim1.}

take/get D PRT

‘Wow, when I sat my exams [Certificate of Education Examination] I even [only] got \textit{C gaa3-zaa3-wo3}. I even almost got a D \textit{tim1}.’

In example (9.37), we see speaker B commenting that she got a C in her exam, and almost got a D. Her use of \textit{gaa3-zaa3-wo3} indicates that it is good if the hearer knows this. She states that she only got a C, and not a better grade. She can feel upset about this, or perhaps the hearer can feel sad or sympathetic too. The cluster also indicates that the hearer is not thinking about this at this moment. This could be because he has been talking about the grades of their other friends, or because B has never told him her grade before. She indicates that it is good if he thinks about it, perhaps as a way to show how impressive the others’ scores are.

(9.38)

\textit{Gam2 co1tau4 ne1 zau6 expect ne1, zik1hai6}

so/then at-the-beginning PRT then expect PRT meaning

\textit{ji5wai4 keoi5 gaan1 gung1si1 ne1 hou2 daai6 gaa3 wo3,}

thought it CL company PRT very big PRT PRT

\textit{gam2 ne1 daan6hai6 soeng5 dou3 heoi3 ne1, zau6}

so/then PRT but up arrive go PRT then
‘So at first I expected, I mean I thought the company was very big, so but when I arrived (up) there, then I realised it was very small.’

Example (9.38), a reproduction of (7.9), clearly shows an aspect of ‘unexpectedness’ which is not as prominent in the other examples of gaa3-zaa3-wo3. The speaker firstly indicates that it is good if the hearer knows she thought the company would be bigger, but that it turned out to be very small. She indicates that it was very small, and not more/bigger, and that she felt surprised by this. She expresses to the hearer that she isn’t thinking about this at this moment, and it is good if she thinks about it. This may be because she assumes the hearer also thought the company would be very big, or because she is recounting the story as if she is taking the hearer there and expects the hearer to also be surprised. It is also possible that she means that any person would be surprised and would think twice. Example (9.39), a copy of (7.10), is also very similar.

Example (9.39), a reproduction of (7.9), clearly shows an aspect of ‘unexpectedness’ which is not as prominent in the other examples of gaa3-zaa3-wo3. The speaker firstly indicates that it is good if the hearer knows she thought the company would be bigger, but that it turned out to be very small. She indicates that it was very small, and not more/bigger, and that she felt surprised by this. She expresses to the hearer that she isn’t thinking about this at this moment, and it is good if she thinks about it. This may be because she assumes the hearer also thought the company would be very big, or because she is recounting the story as if she is taking the hearer there and expects the hearer to also be surprised. It is also possible that she means that any person would be surprised and would think twice. Example (9.39), a copy of (7.10), is also very similar.
A:
Do1 z02 aa1 maa3 gau6nin2.
more PFV PRT PRT last-year
‘More were added aa1-maa3 last year.’

B:
O3 nei5 dou1 m4 wui5 gaan2 saai3
oh you even/still not will choose all
gaa3 laa1. Dim2 gaan2 saam1sap6saam1 fo1
PRT PRT how choose thirty-three subject
coot1lai4 duk6 aa3.
come-out read/study PRT
‘Oh but you don’t choose all [thirty-three] gaa3-laa1. How can you choose thirty-three subjects to study aa3.’

Example (9.40) above shows two people discussing selecting their university programmes. Speaker B uses gaa3-zaa3-wo3 to say to speaker A that when she went through the process, she only had 20 choices, as opposed to the 33 available to B. She says that it is good if B knows this, and that it used to be 20 and not more. She may feel surprised because of this, or may expect A to feel something because of this. She expresses ‘you aren’t thinking about this at this moment, it is good if you think about it’. This may be because she believed there to be 20, as shown in the preceding turns, and is concerned that B has made a mistake. This may also be because she believes 33 is too many, as shown in the following turns.

9.5.4 Summary

Based on the investigation into the semantics of particle clusters presented in sections 9.3–9.5, it appears to be true that the meaning of a particle cluster is that of the component particles combined. The composite meanings can be stated in English and Cantonese using NSM, and can explain examples of the clusters found in the corpus.

This finding also suggests that the explications proposed in Part Two are accurate, as the explications may have otherwise caused problems when combined. Furthermore, it reinforces the proposal that the explications for Cantonese particles are relatively short. This property makes them versatile,
being able to attach to a great many utterances, and being potentially compatible with various other particle explications.

9.6 Reasons for the non-occurrence of certain particle clusters

As described in section 9.2, some particle clusters never occur. Matthews and Yip (2011, 395-396) state that possible combinations are limited by pronounceability as well as semantic coherence. According to them, pronounceability affects clusters such as *laa3-le1, due to the clash of similar sounds (they imply that *laa3-le1 is semantically possible and otherwise acceptable). However, I believe that pronounceability is not a cause for non-occurring clusters, since this does not appear to affect other parts of speech. Tongue twisters in Cantonese and other languages, while acknowledged by speakers as difficult to pronounce, do not cause those utterances to be impossible to say. Reduplication in other Cantonese word classes is also common, suggesting that similar sounding words can co-occur acceptably. Moreover, if some words were forbidden due to phonological reasons, this might mean that utterances which end in certain words cannot be followed by certain particles. As far as I know, this does not happen, and native speakers are never prevented from using certain words or particles together for phonological reasons. Particles appear to be able to follow any word, limited only by semantics. This is also the only other possibility given by Matthews and Yip.

9.6.1 Semantic reasons for the non-occurrence of *laa1-w03

Laa1 and w03 do not combine, i.e. *laa1-w03 is unacceptable in Cantonese. Recall from section 9.2.1 that if we place laa1 in the middle column of Table 17 and w03 in the last column, then the non-occurring *laa1-w03 should be possible based on ordering alone. Why, then, does it not occur? As it turns out, semantic reasons can explain its non-occurrence, made clear by the NSM explications of laa1 and w03, which are reproduced below.
[3A] Proposed explication for laa1:

you now know how I think about this
I can not-say more

lei5  ji4gaa1  zi1dou3  ngo5  dim2  lam2
you    now    know    I    how    think

li1  jat1  joeng6  je5
this    one    CL    thing

ngo5  ho2ji5  m4  zoi3  gong2
I can  not    more    say

[4A] Proposed explication for wo3:

you aren’t thinking about this at this moment
it is good if you think about it

lei5  li1  jat1haak1  m4  hai6  lam2  gan2
you    this    moment    not    is    think    PROG

li1  jat1  joeng6  je5
this    one    CL    thing

jyu4gwo2  lei5  lam2  haa5  zau6  hou2
if      you    think    DEL    then    good

Importantly, the explications for laa1 and wo3 predict their non-occurrence in combination. Laa1 includes a component ‘you now know how I think about this’, while wo3 includes a component ‘you are not thinking about this at this moment’. These components are semantically incompatible, since a speaker should never want to express that the addressee knows how s/he thinks of something, while simultaneously expressing that the addressee is not thinking about it. The reason the two particles never combine to form a cluster is displayed clearly through the NSM explications. This again reinforces the accuracy and value of the NSM explications. *Gaa3-laa1-wo3, which is also technically possible if considering only the ordering of the particles, is impossible for the same semantic reason. We can speculate that non-occurrence of clusters formed with particles not analysed in this thesis may be explained with similar semantic reasons.
9.6.2 Semantic reasons for the non-occurrence of other clusters

While *laa1-w03 is impossible because laa1 and w03 express conflicting meanings, it may be that some other particles never occur together for the opposite reason, i.e. because they are too semantically similar. For instance, w03, w04, and w05 never occur together, as in *w03-w04, *w04-w05, etc. It could be that these particles would occupy the same position as each other in the ordering discussed in section 9.2, but this does not preclude possible semantic reasons as well. Since these particles are generally accepted to be semantically similar (see Chapter 8), and many scholars have claimed that they belong in the same ‘family’ or even share common meaningful morphemes, it seems possible that they never occur in clusters because their meanings are too similar. Speakers presumably do not want or need to express their meanings together in the same utterance. The same is presumably true for other particles which are generally understood to be semantically similar, such as zet and zaa3. This kind of ‘semantic incompatibility’ is a much more plausible reason for such particles never occurring together, compared to the problematic claims that these particles are made up of each other. (The sub-syllabic morphemes hypothesis would not handle particle clusters particularly well in any case, because of frequent reduplication of ‘meaningful’ morphemes, e.g. gaa3-laa3 has the ‘doubled’ features of aa and tone 3.)

9.6.3 Possible semantic reasons for rigid ordering in clusters

As observed in section 9.2, particles within clusters appear to follow rigid ordering, although neither phonological nor syntactic explanations for this are very satisfactory. Of course, semantic reasons are also possible. Moreover, any reasons do not have to be mutually exclusive – while *laa1-w03 would appear to be possible if assessed on ordering alone, other non-occurring clusters may be prevented by, say, both semantic and syntactic reasons. Notwithstanding this, the following focusses on possible semantic reasons. Since Cantonese utterance particles have fixed orders within clusters, the question arises of whether this may be due to how the meaning of each particle interacts with the others and/or with the utterance to which it is attached.
The explications for the individual particles presented in Part 2 were all created independently of the others, and each was constructed to refer to the utterances it attaches to. In other words, the explication for laa1 was intended to describe and explain what laa1 added to utterances it attached to. Generally speaking, parts of explications which are deictic and require context, such as the word ‘this’, refer to whatever might be in the utterance that the particle is attached to. The treatment given in the analyses above shows that it is possible that when clusters of particles are used, each particle ‘attaches’ to the utterance equally. However, assuming that phonological reasons do not rule out any clusters, this would not explain why the orders of the particles in each cluster are fixed.

Therefore, it seemed necessary to consider that when multiple particles occur together, some of the particles could be ‘attaching to’ or making a comment on the other particles in the cluster. An utterance particle might be attaching to a clause which already ends with an utterance particle. For instance, in the cluster zaa3-wo3, might it be the case that instead of the explications of zaa3 and wo3 each referring to the utterance, the explication of wo3 actually refers to the explication of the preceding particle, i.e. zaa3? Or, that wo3 attaches to the whole zaa3-suffixed utterance, i.e. including zaa3?

After much experimentation, this potential explanation was found not to work out particularly well, at least not with the current proposed explications. When tested with the real examples of particle clusters from the corpus, the hypothetical ‘nested’ explications did not always match the apparent intended meanings of the speakers. More experimentation may reveal different ways to ‘combine’ or present the explications, which can better address the remaining uncertainties about particle ordering within clusters. This is left for future research.

9.7 Concluding remarks

This has been an exploration into the hitherto untested idea that particle clusters have the combined meanings of the particles of which they are made up. This chapter has shown that the short but meaningful explications proposed for five particles in Part Two combine to yield eight plausible composite
explications for the particle clusters they form. Particles that occur together in clusters have semantically compatible NSM explications. The meaning of the particle clusters can be stated in English and Cantonese using NSM, and can explain the meanings of the clusters found in the corpus. This also suggests that the explications proposed in Part Two are accurate, as the explications may have otherwise resulted in problematic combinations.

The unacceptability of some non-occurring particle clusters appears to be due to semantic incompatibility. Some particles, such as laa1 and wo3, have conflicting meanings, as clearly evident from their explications. Others, such as wo3 and wo5 or zaa3 and ze1, are suggested to have meanings which are too similar to each other and therefore redundant if used together. Some other non-occurring clusters, however, and some non-occurring orderings, seem to be unacceptable for as yet unknown reasons.

Based on the NSM investigation into the semantics of particle clusters presented in this chapter, it appears to be true that a particle cluster’s meaning is that of the component particles combined. The examination of particle clusters conducted here appears to be the most in-depth study of the semantics of particle clusters to date. Part Three has considered all kinds of Cantonese utterance particle combinations and addressed the second main research question.
Part Four:
Conclusions
Chapter 10:
Review and prospects for further research

This thesis has challenged previous perceptions that Cantonese utterance particles have no meanings, or that their meanings are too elusive to identify. In this area, much of the existing literature published in English has many shortcomings, at least from a semantic point of view. This thesis has addressed the gap in Cantonese linguistics by carrying out in-depth semantic analyses of five common particles, namely laa1, wo3, gaa3, laa3, and zaa3. The core, invariant meanings of each particle were uncovered and explicated in simple, cross-translatable paraphrases. Eight particle clusters consisting of two or three of these particles have also been examined and it has been shown that their meanings consist of the combined meanings of the individual particles. This thesis has also included a semantic critique of the contractions hypothesis and the sub-syllabic morphemes hypothesis, according to which certain individual particles have an internal morphemic structure. The semantic methodology underpinning this investigation has been the Natural Semantic Metalanguage approach. This thesis has explored the viability of a Cantonese-specific NSM and found it to be practicable and able to be used to explicate particles. The following summarises the main findings and contributions of this thesis and then makes some suggestions for future research.

10.1 Main findings and contributions of the present study

10.1.1 Individual particles: laa1, wo3, gaa3, laa3, and zaa3

Using the NSM methodology and the evidence base of the Hong Kong Cantonese Corpus, the present study has demonstrated that the five Cantonese

---

116 It is difficult to believe that any of the existing descriptions on their own would be adequate for readers to fully grasp the meanings or usages of the particles. Many previous definitions are imprecise and do not allow for particles’ full ranges of use, are given in obscure or vague terms, contradict other definitions of the same particle, and/or are explained using words in other languages. Studies frequently mistake the meaning of the utterance that precedes the particle as being the meaning of the particle itself. Definitions are generally not tested against or even drawn from real, naturally-occurring speech data, relying mainly on notoriously unreliable native speaker intuition.
utterance particles laa1, wo3, gaa3, laa3, and zaa3 each have a stable, invariant meaning. For each particle, a single NSM explication was proposed as its ‘core’ or invariant meaning in Part Two. The explications are presented again in English and Cantonese NSM below. Each explication states the particle’s meaning clearly and precisely and can be expected to encompass its wide range of use. Recall from Chapter 1 that their alleged lack of semantic content was even said to be one of their distinctive features. By determining and stating the particles’ meanings explicitly, this thesis has shown that this is a misconception. This has answered the primary research question of whether or not the particles have meaning.

**Laa1:**

you now know how I think about this
I can not-say more

`lei5 ji4 gaa1 zi1 dou3 ngo5 dim2 lam2`

you now know I how think

`li1 jat1 joeng6 je5`

this one CL thing

`ngo5 ho2 ji5 m4 zoi3 gong2`

I can not more say

**Wo3:**

you aren’t thinking about this at this moment
it is good if you think about it

`lei5 li1 jat1 haak1 m4 hai6 lam2 gan2`

you this moment not is think PROG

`li1 jat1 joeng6 je5`

this one CL thing

`jyu4 gwo2 lei5 lam2 haa5 zau6 hou2`

if you think DEL then good

**Gaa3 (in statements):**

it is good if you know this

`jyu4 gwo2 lei5 zi1 dou3 zau6 hou2`

if you know then good
The NSM explications are given in paraphrases constructed of simple, cross-translatable semantic primes. NSM is able to overcome common problems in defining words, such as circularity, obscurity, untranslatability, and inaccuracy. Using NSM, it has been possible to construct explications from the speaker’s point of view, to reconstruct meaning, and to test these meanings with examples. Explications given in NSM can be understood by native speakers and be translated across languages and cultures, even by non-linguists.

A detailed investigation into Cantonese NSM was presented in Chapter 2. It was the first comprehensive account of Cantonese NSM. By identifying Cantonese exponents for all 65 NSM semantic primes and using them in explications, it has been shown that Cantonese NSM is a practicable system of semantic description.

For each particle analysed in this thesis, a wide range of naturally-occurring examples from the Hong Kong Cantonese Corpus were examined. Though the corpus is relatively small, the high frequency of utterance particles in ordinary speech meant that there were still copious examples of each particle. Testing of the explications was carried out primarily by substituting the proposed explications into real examples of the particle in use from the corpus,
and determining whether this made sense in context and seemed a plausible representation of what the speaker wanted to express. Testing in this way proved to be invaluable in eliminating poor explications or components of explications. Numerous explications or components seemed plausible at first, but were subsequently shown to be incorrect by careful examination of the data. This process increased the accuracy and reliability of the final explications. It satisfies one of the research questions that the meanings proposed for the particles can be tested and substantiated by evidence from a corpus of naturally-occurring, spoken Cantonese.

The corpus data also clearly showed that many previous descriptions of the particles are inaccurate, or at least not comprehensive. Native speaker intuition is often fuzzy and incomplete, and this is a weakness in most previous studies of Cantonese utterance particles, which relied on intuition alone. Furthermore, a wide range of examples is necessary. Previous definitions of particles have not allowed for particles’ full ranges of use, restricting themselves to certain contexts and ignoring others. Use of a corpus is more reliable and helps to prevent ‘cherry-picking’ of contexts/examples. The selection of examples presented in this thesis aimed to show varied contexts and different uses of each particle, to provide a picture that was as complete as possible.

One of the keys to explicating Cantonese utterance particles was to use NSM primes which allowed reference to deictic and changing elements. This was found to be crucial as the meanings of the particles are context-bound. This is relevant to the research question about the nature of the particle explications. Independent of context, these particles do not refer to ‘actual things and events’, in the way that words traditionally thought of as shizi ‘full words’ would be expected to. But deictic elements in the explications can apply to the hugely variable ‘actual things and events’ being spoken about. This is why the same particle can be used in a myriad of diverse and opposing situations. Understanding this helps to pinpoint the invariant meaning, and show that this meaning remains stable across different contexts and situations. It also helps to explain some previous observations about Cantonese utterance particles, that they are used in a wide range of seemingly random ways, and have either no meaning or numerous different meanings. From the examples presented throughout this thesis, it is clear that in context, the proposed explications apply
consistently, even to utterances which may be expressing quite different attitudes, moods, assumptions or intentions.

It should be noted that the explications proposed in this thesis look very different from the average NSM explication, even those of particles in similar languages (see section 1.6.3): specifically, they are much shorter and simpler. This was actually unexpected at first. In proposing explications and components, the researcher first attempted to include many more components than were ultimately retained, due both to their apparent hazy meanings and to an expectation based on the average length of other NSM explications. Because deictic elements exist, the explications could be very short but still accomplish a multitude of varied tasks in conversation. In retrospect, the fact that Cantonese utterance particles’ meanings are short and simple makes a great deal of sense. Their short, compact meanings help explain their remarkably high frequency in ordinary conversation, and their versatility across many contexts. Had particles represented long and complicated meanings, they might not occur as often in speech. Furthermore, their short and simple meanings make sense when considering particle clusters. Short explications would be expected to combine more easily, probably influencing the particles’ frequent combination with one another in particle clusters.

10.1.2 Particle combinations

Chapter 9 presented the most detailed investigation into the semantics of particle clusters to date. The NSM explications of the individual particles analysed in Part Two were ‘combined’ to investigate one of the research questions and the untested assumption in the literature that particle clusters have the composite meanings of the particles they consist of. The five explications in Part Two were all arrived at separately, but remarkably, it turned out that the explications were semantically compatible where those particles combined in speech. This was facilitated by the earlier finding that each particle’s explication is very short, thus readily combining with each other to produce explications for eight possible particle clusters. Occurring particle clusters reflect the semantic compatibility of the component particles and the combinability of those particles’ NSM explications. The explications could combine in ways predicted (or necessitated) by the actual particle clusters in
speech. When these composite NSM explications were tested by substitution
into corpus examples, they seemed to explain the meanings of the clusters.
Significantly, the semantically incompatible explications of laa1 and wo3 even
reflected the unacceptable clustering of the two particles in speech. This was
again remarkable, because the individual particles had been explicated
separately, without the prerequisite that the final explications be incompatible.
Non-occurring clusters appear to be partially explained by semantic
incompatibility, including both having opposing meaning and having meanings
that are too similar. All this can be considered as confirmation of the hitherto
untested hypothesis that particle clusters have the combined meaning of the
particles of which they are made up. As well, it is an additional test of the
individual explications proposed in Part Two, since incorrect explications may
have resulted in problematic combinations and explications being semantically
incompatible where this did not reflect the reality of the particles in speech.

In answer to another research question of this thesis, Chapter 8 found it
unlikely that monosyllabic particles are made up of smaller meaningful units.
The sub-syllabic morphemes hypothesis is highly implausible, as the supposed
meanings of the sub-syllabic morphemes do not always combine to create the
claimed meanings of the resulting particles, and the semantic analysis is highly
imprecise in general. The contractions hypothesis is more prima facie plausible,
but still suffers in some key areas, such as a lack of consensus regarding which
particles are contractions of which other particles. Problematically, both
hypotheses utilise hypothetical particles which are not found in present-day
Hong Kong Cantonese. Another question raised is whether particles likely to
have narrow meanings can be part of particles likely to have broad meanings.
An alternative view presented here is that some (possibly not all) monosyllabic
particles are semantically related through phonesthemes. This is less extreme
and more plausible than claiming that every single initial, rhyme, coda, and tone
has a distinct meaning (see also 10.2.3).

10.1.3 The particles in use

Since the particles selected for this study are common in speech, significant
portions of conversations can now be better understood. To demonstrate,
consider the following excerpt, (10.1), which shows a speaker describing a
student of hers whom nobody likes. Part of this excerpt was shown in Chapter 1 as example (1.1). Seeing how frequently the particles occur, it is remarkable that someone who does not know what the particles mean would be missing important information with every utterance. Even knowing the meanings of only a handful of particles significantly improves understanding of what the speaker is expressing.

(10.1)

A:

```
jyun4-loi4  ne1  ziicin4  hai6  go3  gaa1zoeng2  ne1
```

apparently **PRT** before is **CL** parent(s) **PRT**

deoi3  keoi5  ne1  taa13gw03  gun2  dak1

towards s/he **PRT** too control/management **ADV**

sailei6  gw03tau4  **laa3**. Gam2  ne1  so2jau5  je5

strong too-much **PRT** then **PRT** all thing

keoi5  dou1  hung3zai3  saai3  keoi5  **wo3**.

he/she/they also/even control all s/he **PRT**

**Pei3jiyu4  baau1kut3  saai3  keoi5dei6  zou2**

for-example including all they morning

```
g5  maan5  saam1  caan1  dou1  hung3zai3
```

afternoon night three meal all control/management

```
saai3  **laa1**,  pei3jiyu4  waa6  zi1tau4zou2  sik6
```

all **PRT** for-example say morning eat

```
zou2caan1  zau6  jat1ding6  jiu3  hai6  sik6  bin1
```

breakfast then definitely have-to is eat which

```
gei2  joeng6  **laa1**. Dou3  dou3  ng5caan1  **laa3**,  e6,
```

few **CL** **PRT** arrive until lunch **PRT** eh

dou3  dou3  siu2sik1  ge3  si4hau6  ne1,  go3

arrive until recess LP time **PRT** **CL**

```
gaa1zoeng2  zau6  wui5  jiu3  ling1  maa14
```

parent then will have-to take/bring close

```
saai3  di1  je5sik6  lai4dou3  hok6hau6  sung3  bei2
```

all **CL** food come-to school deliver for

```
go3  zai2  sik6  **wo3**. Hou2  **laa3**,  m4daan1zi2
```

CL son eat **PRT** good/okay **PRT** not-only

```
ni1joeng6  je5  **wo3**. Gam2  go2go3  hok6saang1  ne1
```

this thing **PRT** then that student **PRT**
keoi5 soeng5tong4 ge3 cing4jing4 ni1go3 gaa1zoeng2
he go-to-class LP situation this parent
ne1 dou1 wui5 ne1 jap6lai4 tai2 keoi5
PRT also will PRT come-in watch s/he
gaa3 wo3. Sam6zi3fu4 keoi5 zau6 wui5 ne1
PRT PRT even s/he then will PRT
gwaan1zyu3 dou3 ne1 hok6saang1 ge3 hung1hei3
pay-attention to PRT student POSS air
lau4tung1 ge3 man6tai4 wo3...
flow LP problem PRT

‘... apparently ne1 before it was the parent(s) ne1 towards him ne1 too controlled him too strongly laa3. And ne1 he/she/they controlled everything for him wo3. For example, this includes all their control of all his three meals morning afternoon and night laa1, for example in the morning when he eats breakfast he must eat which things laa1. Then at lunch laa3, eh, at recess ne1, the parent will then have to [want to] bring all the food to school for the son to eat wo3. Okay laa3, not only this wo3. Then that student ne1 his situation when he goes to class, this parent ne1 will also ne1 come in to watch gaa3-wo3. S/he even will ne1 pay attention to ne1 this student’s air flow [classroom ventilation] problems wo3.’

[one turn omitted]

A:
Kei4taa1 tung4hok6 ne1, e6, kei4sat6 ne1
other student PRT eh actually PRT
dou1 m4 zung1ji3 keoi5 gaa3. Jan1wai6 ne1
also/all not like him/her PRT because PRT
keoi5 m4 do1 ceottseng1 laa1, ji4ce2 ne1
s/he not much speak PRT also/moreover PRT
jau6 seng4jat6 tau4sou3 jan4dei6 wo3...
also always complain people/others PRT

‘Other students ne1, eh, actually ne1, they don’t like him/her either gaa3. Because ne1 s/he doesn’t talk much laa1, and ne1 also always complains about other people wo3.’

In this text, whenever laa3 has been used, the speaker wants the addressee to know and ‘register’ the content of the utterance. Laa3 points out where the speaker wants the hearer to note and understand something. Wo3 indicates where the speaker is saying something that might be unexpected or new information, or where the speaker thinks the hearer would not be thinking
about that thing. Additionally, wo3 signals to the hearer that the speaker wants them to think about that thing. Laa1 helps signal that the thing being said is something that the hearer can understand and which the speaker does not have to say more about, such as a simple point in a list. Finally, gaa3 points out something that the speaker thinks would be good for the hearer to know. Zaa3 does not occur in this example, but would indicate that something is a certain way and not more, and that someone can feel something as a result.

10.2 Suggestions for further research and future directions

10.2.1 More particles

One obvious possibility for further research is to study more particles using the NSM method. Given the significance and pervasiveness of utterance particles in ordinary, everyday Cantonese, it has been surprising that their meanings are so poorly understood. The present study has selected a limited number of particles for in-depth semantic analysis – a necessary constraint to ensure each particle received ample attention. The NSM method and corpus data used have proven effective and worthwhile, so investigation of more particles with this approach would no doubt yield valuable insights into the semantics of those particles. We can hypothesise that the remaining particles each have identifiable meanings too.117 As the number of 'basic' utterance particles in Cantonese is generally said to be 30 or more, there is much scope for expansion.

Furthermore, most other potential research areas, such as particle combinations, would benefit from systematic semantic analyses of individual particles. Many scholars have been interested in comparing the particles with each other and in identifying similarities and differences between them. Use of this same, proven approach to analyse more particles would allow easier comparisons, and ensure a solid starting point for further analysis. As shown, the existing literature sometimes builds on descriptions which are imprecise to begin with, which cannot result in accurate analysis at the next stage.

117 And for example, semantic analysis could clarify the differences between some non-interchangeable particles, such as the z- initial particles, which have all been given the same description 'only'.
10.2.2 Questions

This thesis has not focussed on the use of utterance particles in questions. This is a reflection of the particles selected for study, which are primarily used in statements. In addition, the meaning of gaa3 as used in questions may be somewhat different from its meaning in statements, and so is still unknown (see Chapter 5). It may be that gaa3 is polysemous and/or it may be that the current proposed explication can be improved.

10.2.3 Phonesthemes and particle families

Some particles have noticeable semantic similarities with each other, and some of these similar particles share phonological features in common. It was put forward in Chapter 8 that phonesthemes or sound symbolism may play a role, but a thorough investigation of such is beyond the scope of this thesis. Phonesthemes themselves are not well understood, and they have never been proposed for Cantonese utterance particles before, so this could be an ample area for future research. At this stage, it seems a plausible explanation for why some similar sounding particles seem to also be semantically similar.

Phonesthemic particles may have ‘residues’ whose meanings are not recurrent, but there is no reason this should negate any potential findings of phonesthemes. It seems possible to the author that some phonological features, such as z- initial or wo (any tone) can have phonesthemic qualities (those particles do have intuitively perceptible semantic similarities), while other features such as the e or aa rhyme may have none, or weaker phonesthemic qualities. It is possible that some sounds are phonesthemic while others are not. The assumed consistency between all particles and blanket ‘all or nothing’ claims, such as in the more radical versions of the sub-syllabic morphemes hypothesis, are attractively ‘neat’ and orderly, but may be premature and potentially hindering our understanding of meaningful sound segments.

If phonesthemes are not involved in Cantonese utterance particles, then perhaps future research can give another explanation for the groups of similar particles. The sub-syllabic morphemes hypothesis has been shown to be highly implausible, and the contractions hypothesis, at least in its current state, is not very satisfactory, so perhaps other new hypotheses can be put forth. These
might include historical or cross-linguistic reasons, such as those which have already been put forth for the ‘wo family’. Again, I would speculate that if certain groups of particles can be explained in these ways, this does not necessarily mean that all groups of particles can be. In any case, our current knowledge in this area is lacking, and more research could shed light on this.

10.2.4 Particle clusters

Following from the idea tested in Chapter 9 that explications of particle clusters can be created by combining the explications of the component particles, investigation into particle clusters would benefit from analyses of more individual particles using NSM. More research can confirm whether the observed semantic compatibility of the relevant explications is true of all particle clusters. Even more of a mystery is why the particles within the clusters must follow a rigid order, and why particles belonging in the same positions appear not to co-occur. Reasons or better evidence of reasons for their ordering may arise with further research.

10.2.5 Data

As mentioned in Chapter 1, the examples presented in this thesis were taken from the Hong Kong Cantonese Corpus, which includes audio files. The examples primarily featured informal, spontaneous Cantonese speech in ordinary settings among family, friends and colleagues. It may be that more and/or different data could be useful in future studies, possibly with speakers with different personal characteristics (e.g. age, sex, socioeconomic status), different relationships with each other, or in different settings, etc. As has been suggested by Chan (2002) and Botha and Barnes (2013), factors such as age and gender may influence the choice and usage of utterance particles. It is also easily noticeable that utterance particles are used less in formal speech, and it is possible that certain particles are considered more/less acceptable in formal speech (although very formal speech blurs the lines between natural and unnatural speech).118 If there were sociolinguistic elements involved in the

---

118 One study that explores an unusual kind of data (in the area of Cantonese utterance particles) is that of E Leung and Gibbons (2011), who studied Cantonese utterance particles in Hong Kong courtroom interaction, specifically in the speech of court interpreters interpreting between
usage of the particles, this would not be expected to change the proposed explications of their invariant meaning, but could nonetheless be an interesting area of research.

Also with regards to data, it may be that visual records of speakers and conversations can provide further insights which have hitherto gone unnoticed. These might have to do with body language or facial expression, for example. As far as I am aware, no studies of Cantonese utterance particles have used videos of real, spontaneous speech. The mechanics of such data collection would need to be carefully thought through, although recording devices are ubiquitous in Hong Kong and some people may not find them unreasonably intrusive.

10.2.6 Testing of NSM explications with native speakers

Future studies could take descriptions of particles, such as the NSM explications proposed, and ask native Cantonese speakers (not necessarily linguists) their opinions of them. NSM explications are designed to be able to be understood by native speakers, and this would reinforce the accuracy of the explications (and the acceptability of Cantonese NSM). This may be somewhat complicated by the fact that NSM explications do not look or sound like ‘normal’, everyday text, since explications are restricted in certain ways to maintain their simplicity, cross-translatability, and neutrality. Nonetheless, the unfamiliar presentation may be able to be overcome with some introductory explanation. Following this, native speakers could judge whether the explications seem intuitively to be correct. Of course, the researcher would need to bear in mind that native speaker intuition of particles is fuzzy and unclear, and most people do not consciously think about how they use particles, even if they use them frequently.

Another exercise which would involve acquiring native Cantonese speakers’ opinions about explications is to create some kind of survey. Gladkova, Vanhatalo, and Goddard (2015) experimented with an online survey technique to test NSM explications. They tested explications of certain English interjections using a method whereby subjects rated individual components separately, a test design that works well for short explications. This is

English and Cantonese in five rape trials. They found that the particles can have strategic roles of supporting, evaluating and questioning statements.
encouraging, as interjections and particles are somewhat similar, and the 
explications proposed in this thesis are unusually short compared to the average 
NSM explication. However, the design of a survey for utterance particles would 
need to be thoroughly considered. One relevant difference between utterance 
particles and interjections is that interjections can occur in isolation, while 
utterance particles cannot. It may therefore be a good idea to present survey 
respondents with the same utterance with different particles attached, and ask 
them about those utterances. As shown, it is a common pitfall in prior research 
that the meaning of the utterance preceding a particle is mistaken for the 
meaning of the particle itself, so careful thought would be required to avoid this 
as much as possible. Ideally, the utterances would be very simple and would not 
indicate any emotion themselves. These might include utterances like ‘five X’, 
‘yes X’, or ‘John X’, where several different particles could take the place of ‘X’. 
The format of the survey also requires thought. For example, the questions 
might be closed ended, since even native speakers have trouble explaining what 
the particles mean. This would have the disadvantage, though, that possible 
answers would need to be specified and restricted by the researcher.

10.2.7 Language learning

Language learners often find it difficult to grasp and use utterance particles 
appropriately, and often struggle to understand the subtle nuances between 
them. Native speakers have difficulty explaining the particles, and dictionaries 
and grammar books are usually unhelpful in this respect. Since the methods 
used in this thesis have successfully determined the meanings of certain 
particles and particle clusters (and more particles may be studied in the future), 
this could be of use to language learners. This is made easier by the format of 
the identified meanings – they are given not in obscure or abstract terms but in 
simple reductive paraphrases using NSM. NSM is designed to be easily 
understood and even universally translatable, so it is an excellent tool for 
language learners. Utterance particles may be learned easily given there is a 
single, succinct explication for each particle. Moreover, they can be combined 
readily to give the meaning of particle clusters. All these factors mean that they 
would be very easy to learn and remember, and learners could simply ‘apply’ the 
explications to various situations and utterances.
10.2.8 Culture

The link between particles and culture is another area for potential research. Language and culture are intertwined, and every language reflects unique and culture-specific ways of thinking. Beyond knowing what a particle means, there may be other factors which affect how particles are used or why they are used so often. Why do speakers of a language want to express that specific meaning, quickly and constantly? Do they in some sense ‘think differently’ from speakers of other languages, who do not have such a word (and therefore do not constantly express that meaning)? Besemer and Wierzbicka (2003, 30) and Wierzbicka (2003, 341) posit that particles can reveal a lot about the style of interaction prevailing in a given society, and that few aspects of language reflect speech culture better. Cantonese utterance particles, as an unusual linguistic phenomenon and a distinctive hallmark of Cantonese, may be doing something culturally unique, and may allow us to make some observations about Hong Kong culture or a Cantonese way of thinking.

As the dominant local language, Cantonese is important for the building of Hong Kong’s culture and identity, and understanding cultural background is important to understanding any language. Snow (2004, 194) calls Cantonese both a ‘language of identity’ and a ‘language of the heart’. He believes that use of Cantonese readily creates a sense of in-group belonging. It is distinctive enough from surrounding languages that communication becomes difficult, and this helps foster a group culture and identity by encouraging Cantonese speakers to communicate with each other much more than with others who do not share the language (Snow 2004, 75). Numerous historical events have also played a part in forming a unique, distinctive sense of Hong Kong identity and group personality (and a feeling of separation from China and other countries is well-documented) (Snow 2004, 68-77, Mathews, Ma, and Lui 2008, Chou 2013, Kwong and Yu 2013). Snow provides the insight that perhaps more important than whether or not Cantonese speakers have a distinctive group personality is the fact that Cantonese speakers themselves widely believe that they do.

---

119 The Hong Kong identity is argued by many to be more than a metropolitan identity within a national identity (such as Shanghainese in China or New Yorkers in the United States), because it is in many respects opposed to Chinese identity (see e.g. Mathews, Ma, and Lui 2008).
Language features such as particles might refer to or enable this sense of belonging to a connected group.

Parallels may be drawn with Singapore English. It has been said that Singapore English particles such as *one*, *meh*, *wut* (*what*), and *la* play a major role in the integrity and cohesiveness of the Singapore English speech community, and provide valuable insights into Singapore culture (Besemeres and Wierzbicka 2003, Wong 2004, 2005). For example, Besemeres and Wierzbicka (2003) state that *lah* marks speech involving familiarity, informality, solidarity, and rapport between participants, and that its pervasive use indicates the importance for Singaporeans to treat other Singaporeans as insiders who can know what one wants to say. They suggest that the small size of Singapore, and its special character as a city-state, leads to Singaporeans’ need to treat each other not as strangers but as insiders. Hong Kong is also a relatively small city (particularly compared with mainland China), separated from mainland China historically and politically, which may similarly foster desire to treat Cantonese speakers as insiders. Even in the handful of particles analysed in this thesis, it seems plausible that Cantonese speakers want to express closeness or connection with each other, for example, by indicating what the hearer should think or know.

Wong (2004, 2005) argued that some meanings contained in Singapore English particles indicate that Singapore English speakers have little regard for personal autonomy and freely attempt to change an addressee’s way of thinking. For example, the Singapore English *one* reflects speakers’ tendency to speak definitively and exaggeratedly, and is a way of connecting with people. He compared this to an Anglo English speaker, who would be more likely to acknowledge another individual’s point of view. He believes that using linguistic devices to overtly influence an addressee’s way of thinking is common among speakers of Chinese languages, suggesting that this is characteristic of traditional Chinese culture, including that of Hong Kong. In the context of Hong Kong, this would not be surprising to many, and again, the particles analysed here do show that the speaker often wants to influence what the hearer knows or thinks about.

Just like with Singapore English particles, it is likely that study of Cantonese utterance particles, particularly because of their salience and
pervasiveness in everyday Cantonese, can reveal something about the culture or ways of thinking of Cantonese speakers. Botha and Barnes (2013) suggested that utterance particles in Macau Cantonese are a resource for communicating speaker identity. Macau Cantonese may be slightly different to Hong Kong Cantonese, but their data showed that the frequency of particle use was the greatest when the affective quality between interlocutors was high and the conversation topic was more personal. Wakefield (2011b, 251) has stated that it is ‘not unlikely’ that some Cantonese utterance particles express culturally-specific meanings and meet culturally-specific communicative needs. James (2000, 56) seems to be the only one to have mentioned that Cantonese utterance particles have the effect of ‘enhancing solidarity and collectivity’. There appear to be no serious studies on the links between Cantonese utterance particles and Hong Kong culture so far, and this could prove very interesting.

10.2.9 Similar particles in other languages

As mentioned in Chapter 1, several other Southeast Asian languages have utterance-final particles comparable to those of Cantonese. Some Singapore English particles have even been suggested to have their origins in Cantonese – Lim (2007) considered the etymology of the Singapore English particles and suggests that Cantonese provided lor, hor, leh, meh, and ma, inclusive of tone, while Kwan-Terry (1978) and Gupta (1992) have suggested that Cantonese provided la (lah). Vietnamese has 25 to 45 particles, depending on the regional dialect (Tran 2010, 335-343), and Thai has roughly 40 (Peyasantiwong 1981). Mandarin has fewer particles, with Li and Thompson (1981, 238) listing six, and Chao (1968, 795-814) listing 26 (and two additional ‘tonal particles’).

No comparisons between particles in different languages were attempted in this thesis, though this may be an interesting area for further research. Other particles analysed using NSM may be a good starting point for comparisons. As mentioned, NSM has been used in studying Singapore English la/lah, wut, meh, one, á, and lőr (Besemer and Wierzbicka 2003, Wong 2004, 2005, 2014), Malay la and pun (Goddard 1994b, 2011, 180-183, 2001), and Mandarin me (Chappell 1991). Besemer and Wierzbicka (2003, 27-29) have already compared the semantics of Singapore English la with Goddard’s (1994b) analysis of Malay la.
10.3 Concluding remarks

This thesis has demonstrated that the Cantonese utterance particles laa1, wo3, gaa3, laa3, and zaa3 have meanings. These meanings can be stated clearly and precisely using English and Cantonese NSM, and are substantiated by examples of real, naturally-occurring speech found in the Hong Kong Cantonese Corpus. These findings refute existing literature which has claimed that Cantonese utterance particles have no meaning, or that their meanings are unidentifiable or unstable. Furthermore, the explications of the five particles can be combined as dictated by their acceptable clustering in speech, and the combined explications appear to explain the meanings of the particle clusters.
References


Tran, Ngoc Dung. 2010. *Cam nang ngu phap tieng Viet (Vietnamese grammar handbook).* California, USA: Tinh hoa Viet Nam Libray.


