



Universiteit
Leiden
The Netherlands

The dissection and structural mapping of Cantonese sentence final particles

Rint, S.; Li, B.

Citation

Rint, S., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua*, 117(10), 1739-1783.
doi:10.1016/j.lingua.2006.10.003

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3762434>

Note: To cite this publication please use the final published version (if applicable).

The dissection and structural mapping of Cantonese sentence final particles

Rint Sybesma^{*}, Boya Li

Universiteit Leiden, Department of Chinese studies, P.O. Box 9515, 2300 RA Leiden, The Netherlands

Received 22 May 2006; received in revised form 26 October 2006; accepted 30 October 2006

Available online 2 February 2007

Abstract

This paper explores the possibilities of dissecting the particles that we find in sentence final position in Cantonese into sub-syllabic semantic units. After characterizing the 40-odd most common particles, we study groups of particles, which share one feature (e.g., initial *g-*) to see what semantic feature they have in common, as well as groups of minimal pairs (*gaa3-gaa2* and *ge3-ge2*) to find out whether there are any correlative systemic semantic differences. We end up proposing 12 minimal semantic units, 4 initials (*l, z, l/n, m*), 2 rhymes (*aa* and *o*), 3 tones (**1, 4, 5**), 1 coda (*k*) and 2 such elements incorporating a tone (*g3* and *aa4*). Aside from these, we assume that there is one default rhyme (*e*) and one default tone (**3**). In the final part of the paper, all the minimal semantic units are given a place in the structure of the sentence, notably in the CP domain.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Cantonese; CP; Left periphery; Sentence final particles

1. Introduction

1.1. The subject of investigation

Cantonese has a rich inventory of sentence final particles (henceforth “SFP”). Estimates vary as different factors may be taken into account, but only taking formal properties into consideration, one will have to recognize at least about 40.¹ The main question we are concerned with in this paper is how Cantonese SFPs relate to the structure of the sentence, especially in the light of recent developments with respect to the expansion of the C-domain of the sentence. The

^{*} Corresponding author.

E-mail address: r.p.e.sybesma@let.leidenuniv.nl (R. Sybesma).

¹ Leung (1992) takes semantic properties into consideration as well and gets to more than 90; for instance, his list includes several *laa1*’s: identical in form but different in meaning.

investigations will ideally lead to a better understanding of both the function of SFPS as well as that of the outer layers of the sentence, that is, the structure of CP.

As a first step, we investigate the internal structure of the SFPS, dissecting them into smaller meaningful units, in as far as that will turn out to be possible. The second step involves assigning each meaningful unit a position in the (functional) structure of the sentence.² The main motivation for this approach lies in the high degree of systematicity which appears to underlie form-and-meaning of the particles in Cantonese: SFPS with onset *z-*, for example, share basic semantic properties, SFPS ending in *-aa4* are question particles, SFPS with the same tone to some degree show informational similarity, etc. In taking this approach, our work can be seen as a radical extension of the groundbreaking work done by Sampo Law as reported in her 1990 dissertation (Law, 1990), as well as Roxana Fung's important contribution (Fung, 2000).

Since Cantonese syllables consist of an initial, a rhyme, a coda and a tone, optimal dissection of the SFPS into minimal meaningful units (henceforth “MMU”) will leave us with a list of meaningful initials, rhymes, codas and tones. On the basis of the SFPS discussed in Cheung (1972, Chapter 5.2), Yau (1980), Law (1990), Matthews and Yip (1994:340–341, Chapter 18), Fung (2000), Fāng (2003) and Law (2004:62) we observe that SFPS may have any of the following³

- (1) ... eight initials: $\emptyset$, b, g, h, l, m, n, z;
 ... three rhymes: aa, e, o;
 ... five tones: 1^{(55; (53))}, 2⁽³⁵⁾, 3⁽³³⁾, 4^(21; 11), 5⁽¹³⁾;
 ... two codas: $\emptyset$, -k.

However, this elementary list does not correspond to the inventory of MMUS that we will eventually end up with, for two reasons. First, aside from the zero ($\emptyset$) initial and coda, which, presumably, do not contribute meaning, some of the other elements in (1) will turn out not to be MMUS because they are default (or “dummy”) elements, with no semantic content, present out of phonological necessity only. Secondly, we will end up postulating MMUS, which combine two different elements from the list in (1), such as segment and a tone.⁴ Still, taking (1) as our point of departure and basing ourselves on the references just mentioned, we have assembled Chart 1 (see next page) of the 40-some most common SFPS.⁵

² We use the term “particle” or “SFP” to refer to independently occurring elements, like *laa1*, *aa4*, etc., generally clusters of minimal meaningful units; we use “semantic unit” or “minimal meaningful unit”, abbreviated as “MMU”, to refer to the meaningful sub-parts of the particles. A more general term to use instead of MMU would of course be “morpheme”. We use MMU to make sure that tones are included as possible bearers of meaning.

³ With respect to the tone, some particles seem to have 53, high falling, a tone that is no longer part of the tonal inventory of current day Cantonese as it has been assimilated into the 55 (Matthews and Yip, 1994:21). Interestingly, especially some of the particles we will not deal with here have 53, viz., *tim1* and *sin1*. We will briefly return to this point later on (footnote 40).

⁴ Suppose, we end up with an inventory of MMUS close to what we see in (1), then, if all combinations were possible, Cantonese would have 240 particles. However, not all combinations are possible, and in the end we should be able to explain why this is the case. Due to limitations of space, this part of the research will not be reported on in the current paper.

⁵ Particles not included in the table (some of which are mentioned in the accompanying notes in Appendix A) are kept out of the discussion in this paper, e.g., *lei4*, *sin1* and *tim1*. For *lei4* see Yiu (2001) and Sybesma (2004), plus references cited there; for *tim1* Cheung (1972), Law (2004) and Wong et al. (2005); for *sin1* see Lucas and Xie (1994), Mài (1993) and Tang (2006). Bisyllabic particles, such as *aa1maa3* (Law, 2004, 1990; Matthews and Yip, 1994); *baa2laa3* (Cheung, 1972); *hai2laa3* (Cheung, 1972); *tsai1maa3* (Cheung, 1972; Law, 1990; Matthews and Yip, 1994) (among others) are excluded from the discussion as well. Note further, that we got hold of Leung (1992) too late to process it for Chart 1 and for the description of the particles in section 2.

Chart 1

The 40-odd most common sentence final particles of Cantonese

	e	aa	o	-k
ø	1, ø, 3, 4, ø ^{note: 1}	1, ø, 3, 4, 5 ^{note: 2}	(w?) ø, ø, 3, 4, 5 ^{note: 3}	aak3 ^{note: 3}
b			ø, ø, 3, ø, ø ^{notes: 3, 4}	
g	ø, 2, 3, ø, ø ^{note: 5}	ø, 2, 3, 4, ø ^{note: 6}		aak3 ^{note: 6}
h	ø, 2, ø, ø, ø ^{note: 7}	ø, 2, ø, ø, ø ^{note: 7}	ø, 2, ø, ø, ø ^{note: 7}	
l	(1), ø, (3), 4, 5 ^{note: 8}	1, ø, 3, 4, ø ^{note: 9}	1, ø, 3, 4, ø ^{note: 10}	aak3, ok3 ^{notes: 9, 10}
m	1, ø, ø, ø, ø ^{note: 11}	ø, ø, 3, ø, ø ^{note: 11}		
n	1, ø, ø, ø, ø ^{note: 8}			
z	1, ø, ø, ø, ø ^{note: 12}	ø, ø, 3, 4, ø ^{note: 13}		ek1 ^{notes: 12, 13}
other		gwaa3, waa2 ^{note: 14}		

The chart is three-dimensional: initials (vertical), rhymes (horizontal) and tones (the numbers). For bibliographical and other relevant comments, see the notes in [Appendix A](#).

A question that arises is whether it is really possible to analyze all SFPS as being assembled on the basis of, let's say, only three handfuls of semantic units (whatever the eventual inventory is going to be). As we just mentioned, important work has already been done, notably by [Law \(1990\)](#) and [Fung \(2000\)](#). For instance, with respect to the initials, [Fung \(2000\)](#) proposes that the core meaning of (all elements with) the initials *z*-, *l*- and *g*- can be characterized as follows (see p. 4)⁶:

z: +restrictive (example (2a))

l: +realization of state (example (2b))

g: +situation given; +focus; +deictic (example (2c))

We may add

m: +question (example (2d))

- (2) a. ji⁴gaa¹ zau⁶ waa⁶ hou² zaa³ (Fung, 59)
 now then say good ZAA3
 'it's quite good at this moment only'
- b. ngo⁵ jiu³ heoi³ Mei⁶gwok³ laa³ (Fung, 78)
 1s need go America LAA3
 '[it is now the case that] I have to go to America'
- c. aa³-ji⁶-suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ ge³ (Fung, 158)
 2nd-uncle will continue send-back-come GE3
 '[it is the case that] Second Uncle will continue to send them to us'
- d. lei⁵ sik¹ gong² jing¹man² maa³? (Law, 22)
 2s know speak English MAA3
 'do you speak English?'

⁶ All examples are given in the transcription proposed by the Linguistic Society of Hong Kong (though with superscript tone marks), even if the original source used a different method. Tones have been added if they were not represented in the original. In most cases, glosses and translation are ours, even if the example was taken from the published literature. "M&Y" stands for "Matthews and Yip (1994)".

We will rephrase some of these characterizations below, but Fung's insights are basically correct.

With respect to the coda, from the examples and summaries in Fung (2000), we conclude that *-k* (phonetically often realized as a glottal stop; see below) works as an “emotion intensifier”. Leung (1992:116) says that *-k* makes an utterance stronger. We will present examples below.

As for the tones, a lot of insightful work has been done by Law. Matthews and Yip (1994:339) summarize work done by Law, alone and with others, as follows: “[with respect to] tonal variants: these forms differ systematically in function. Typically, the high-tone variants are more tentative, the low-tone ones assertive and the mid-tone ones neutral.” (See also, e.g., Law, 1990:4.) Here, are some examples (from Law, 1990:107); we will look at Law's proposal more closely below.

- (3) a. ngo⁵-dei⁶ jat¹cai⁴ heoi³ tai²-hei³ e¹
 1P together go see-movie E1
 ‘how about we go see a movie?’ [it's up to you, not going is also fine]
- b. ngo⁵-dei⁶ jat¹cai⁴ heoi³ tai²-hei³ e³
 1P together go see-movie E3
 ‘how about we go see a movie’ [would be fun]
- c. ngo⁵-dei⁶ jat¹cai⁴ heoi³ tai²-hei³ e⁴
 1P together go see-movie E4
 ‘let's go see a movie!’ [I won't take “no” for an answer]

Turning, finally, to the rhymes, no systematic work has been done, though Law made many insightful comments and proposed noteworthy derivations for several individual SFPS.

With respect to the structural position of the SFPS, Law has done admirable work here as well. However, due to the state of the theory of sentential structure that prevailed in the late eighties, she had to make do with a much simpler functional structure than is assumed nowadays. Also, when it comes to the structural positioning, it must be taken into account that the SFPS are active at different levels: the sentential, propositional, discourse, speech act and epistemic domains, as Fung (2000) points out.⁷

Despite the fact that a lot of work has already been done, many questions still need to be addressed, as we will specify below. In sorting out the details, we will take the results reported above as our point of departure.⁸

1.2. Tasks and methods

The tasks that we have set for ourselves are the following:

- (i) Determine whether it is possible to dissect each actual SFP into smaller meaningful units as suggested above and, in as far as it is possible, determine what the core meaning of each minimal unit is.

⁷ As an aside, SFPS are generally considered close to impossible to acquire for non-natives trying to learn Cantonese. In their text book, Yip and Matthews (2000) devote one whole unit to them, stating encouragingly (p. 130) that they are “one of the most challenging features of Cantonese for learners of the language”. Caysac (1952), in contrast, explicitly refrains from paying any attention to them in his course book, as he is convinced that one can only learn them “de la bouche même des Chinois” (p. 19).

⁸ For a more complete literature overview, see Leung (1992:31ff), Law (2004, Chapter 2) and Fung (2000, section 1.2). Further also see Chao (1947), Alleton (1981).

- (ii) With respect to actual SFPS, see how the meaning of the constituting parts contribute to the meaning of the whole.
- (iii) Explain the collocational restrictions (why do not we find all possible combinations?) (but see footnote 4).
- (iv) Assign each minimal unit a place in the structure of the sentence.

Tasks (ii)–(iv) can only be done after those in (i) have been performed successfully. We approach the tasks in (i) by using two different methods:

Method I. Look at sets of minimal pairs and see whether they reveal regular semantic differences.

Method II. Look at all SFPS that have one formal feature in common (e.g., same tone or same initial) and see whether their common formal feature correlates with a common ground in meaning.

In doing the actual dissecting, we take into account the possibility that not every element has a meaning. In other words, we allow for default or dummy elements. Particularly, in view of its neutrality (noted above, and see below), we assume that tone 3 is a default tone, which means that it is not associated with any particular semantic content. The tone is added for phonological reasons, as in Cantonese all syllables, including the SFPS, have a tone (Yip, 2002:272).⁹ For similar reasons, we may end up proposing default vowels, for instance, if an SFP really consists of only a consonant and has to be made pronounceable.

In section 2, we look at all the SFPS in Chart 1 in constantly changing sets, applying Method I and Method II. In presenting the sets, we give a short characterization of the meaning of each SFP, partly basing ourselves on the sources mentioned above. The characterization of the meaning of the SFP has to be short for practical reasons; as a result, not all subtleties involved in the expressive power of the SFPS can be done justice to. In most cases, example sentences are provided for illustration.¹⁰ Many SFPS can occur in declaratives as well as interrogative sentences; we will only mention this kind of distributional properties if we think it is relevant.

There are two more factors that one has to be aware of when working on SFPS. The first such factor is that SFPS can have two different effects when added to a sentence. First, it may change the type of the sentence (it may change declarative into interrogative, for instance) or it may otherwise have a fundamental effect on the meaning of the sentence. Secondly, it may simply be compatible with a certain type of sentence, and do no more than add expressive coloring: an SFP may be compatible with sentences that are independently marked as suggestions and make it a stronger or weaker suggestion. The second factor that one has to be aware of is that in the literature many meanings have been ascribed to SFPS without much basis. Very often, the meaning that is reportedly expressed by the SFP, is already expressed by another element in the sentence (such as an adverb); the sentence has it even without the SFP. Also, it is not always clear whether the meaning described is the core meaning or a secondary connotation.

⁹ Cheung (1972:171) suggests that SFPS form an exception to this rule; this seems true only for particles in non-final position in clusters, as we will see below. A reviewer points out that although the idea that tone 3 is a default tone is not implausible, we have to keep in mind that there is no evidence for it in other areas of Cantonese.

¹⁰ Ideally, one base sentence is used to illustrate the semantic effect of each particle. However, since not every particle is compatible with every type of sentence (some are incompatible with modals, others with perfective aspect, etc.), this turned out to be impossible. We will, however, try to use the same sentence as often as possible so as to get minimal pairs and make comparison useful.

Finally, in this paper, the ideas are developed step by step as we go. Generally, we start out from the strongest hypothesis possible, but as more language material is discussed, we end up having to tone some of them down.

2. Dissecting Cantonese SFPS

2.1. *The initials*

In this subsection, we look at the semantics possibly ascribable to the initials. The order in which we look at them is chosen for expository reasons. We start with the three that are relatively straightforward, *g*, *l*, *z*, if only because they have been studied extensively.¹¹

2.1.1. *The g-, l-, z-families*

2.1.1.1. *The g-family.* The *g*-family has six members (not counting *gwaa3*, for which see below): *ge2*, *gaa2*, *ge3*, *gaa3*, *gaak3*, *gaa4*. The core meaning of all members is relatively constant: they are all involved in asserting (or questioning the assertion of) the relevance to the current conversation of the statement they are attached to. From the short descriptions that follow, it will become clear that we can characterize intra-familial relations by saying that *ge3* is the base and that the meaning of all others can be described as that of *ge3* plus something else. According to Fung (2000:157), “[t]he major differences [between the various *g*-particles] lie in the different assumptions of the existence of the situation and the knowledge of the situation on the part of the hearer.” What follows is a short characterization of the meaning of each SFP.

ge3: though generally described as an assertion marker,¹² it is probably best characterized as an “actuality marker”, asserting that the statement to which it is added is highly relevant to the current conversation (Sybesma, 2004). The higher degree of assertion and commitment, as well as the flavor of emphasis or focus (or even “foregrounding”) that is occasionally perceived in the context of *ge3*, seem to us to be side-effects of this high relevancy function. The particle is paraphrasable as: ‘It is a relevant fact that ...’ Consider the sentences in (4):

- (4) a. *go²-di¹-syu¹, aa³-ji⁶-suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴*
 that-CL-book 2nd uncle will continue send-back-come
 ‘as to those books, Second Uncle will continue to send them to us’
 b. *go²-di¹-syu¹, aa³-ji⁶-suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ ge³*
 that-CL-book 2nd uncle will continue send-back-come GE
 ‘as to those books, Second Uncle will continue to send them to us—for sure,
 don’t worry about it’ (based on Fung, 2000:158 (31))

¹¹ One more preliminary comment is in order. In this section (section 2), we look at the particles in their “citation form”; all example sentences only contain one particle invariably in sentence final position. In section 3, we will look at particles in series, and we will find that the particles that are not in final position are pronounced differently one way or another (see also footnote 9).

¹² “*Ge3* is used for assertions of facts, often marking focus or emphasis” (Matthews and Yip, 1994:349). According to Fung (2000:157), *ge3* “marks a high level of commitment on the part of the speaker to the proposition conveyed by the utterance, asserting the certainty of the proposition without any doubts” or *ge3* “emphasizes that the situation is given as a fact” (Fung, 2000:149; see also Leung, 1992:87). Kwok (1984:42) (quoted in Lee and Yiu, 1998:10) phrases it as follows: *ge3* “shows that the sentence is a factual statement expressing what the speaker regards as true. It is used to strengthen the force of assertion, and is like prefacing the sentence with ‘It is a fact.’” Others preface the English translation of sentences with *ge3* with “the situation is that” (Law, 1990), “such is the case” (Cheung, 1972:186), “it is the case that”, etc. For discussion, see also Fäng (2003).

The only formal difference between these two sentences lies in the absence and presence of *ge3*. Semantically, the difference lies in the link with the conversational context: whereas the sentence in (4a) is a neutral statement, the sentence in (4b) addresses, reassuringly, some concern expressed in the preceding part of the conversation, thus making the sentence more relevant to the current conversation. Whereas the sentence in (4b) is not likely to be uttered in isolation, the sentence in (4a) cannot be uttered in the context in which (4b) is uttered felicitously.

gaa3: is essentially the same as *ge3*; it may be seen as softening *ge3* a bit in the sense that by using *gaa3* the speaker says ‘it is a relevant fact that . . . but I don’t mind that you don’t know or forgot’. Fāng (2003:133) suggests that what *gaa3* adds to *ge3* is an element of “reminding”. Our informants agree; in comparison to (4b), the following sentence is still asserting but less so than (4b).

- (5) go²-di¹-syu¹, aa³-ji⁶suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ gaa³ (cf. (4b))
 that-CL-book 2nd uncle will continue send-back-come GAA3
 ‘you know, as to those books, Second Uncle will surely continue
 to send them to us’

gaak3: Fung (2000:176) calls it the “emotion intensifier of *gaa3*”, but she does not make it very clear in what sense *gaak3* emotionally intensifies *gaa3*, though it is true, as she notes, that it often co-occurs with utterances containing “intensified adverbs” meaning things like ‘definitely’.¹³ Fāng (2003:60, 136) notes that *gaak3* is especially used to assert, with surprise or indignation, that a certain situation prevailed in the past despite current appearances to the contrary, as if you want to remind yourself of how things used to be.

- (6) nei⁵ wong⁵-jat⁶ san¹tai² hou²-hou² gaak³ (Fāng, 136)
 2s in.the.past health good-good GAAK3
 ‘and you used to be in such good shape!’

In a similar vein, our informants report that *gaak3* is used when you argue against a presupposition held by the hearer or by the speaker him/herself. Thus, when we add it to the Second Uncle sentence, we express that we disagree with the assumption, implicit or explicit, that he’s not going to help us with those books.

- (7) go²-di¹-syu¹, aa³-ji⁶suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ gaak³ (cf. (5))
 that-CL-book 2nd uncle will continue send-back-come GAAK3
 ‘as to those books, Second Uncle will surely continue to send them to us—
 contrary to what you think and should not think’

In as far as correcting a wrong presupposition is a strong form of reminding, Fung is right in saying that *gaak3* is emotionally more intense than its unchecked counterpart.

To review all three *g*-particles we looked at so far, let’s look at two more example sentences, and attach the three particles to them to see what the semantic effect is. The differences in meaning are represented in the translations.

¹³ For some of our informants, there appears to be a difference between *gaak3* and *gak3* with a long and a short vowel ([a] or even [ə], actually, see below section 3.1), respectively. We have not been able to get hard facts on this. In some cases, we may perhaps be dealing with an emotionally intensified *ge3*.

- (8) a. ngo⁵ dim² dou¹ wui² bong¹ nei⁵ ge³
 1s how all will help you GE3
 ‘I will surely help you under all circumstances!’ [neutral assertion of relevance]
- b. ngo⁵ dim² dou¹ wui² bong¹ nei⁵ gaa³
 1s how all will help you GAA3
 ‘I will surely help you under all circumstances!—as you should know’ [reminding]
- c. ngo⁵ dim² dou¹ wui² bong¹ nei⁵ gaak³ (Fāng, 60)
 1s how all will help you GAAK3
 ‘I will surely help you under all circumstances!—contrary to what you seem to think’ [correcting presupposition]
- (9) a. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² ge³ (cf. Fāng, 145)
 Guangdong-people eat mouse GE3
 ‘Cantonese people certainly eat mice’ [neutral assertion of relevance]
- b. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² gaa³
 Guangdong-people eat mouse GAA3
 ‘Cantonese people certainly eat mice—as is common knowledge’ [reminding]
- c. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² gaak³
 Guangdong-people eat mouse GAAK3
 ‘Cantonese people certainly eat mice—as is common knowledge but what you seem to be surprised about’ [correcting presupposition]

gaa4: Fung (2000:177) says that *gaa4* “turns a factual declarative into a question”. It double-checks the existence of the given situation or the assumption of the situation conveyed by the declarative. It seems paraphrasable as ‘is it really a relevant fact that . . .?’ for checking a fact or its relevance to the conversation, not for casting doubt. For a similar characterization, see Fāng (2003:72, 145).

- (10) a. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² gaa⁴? (cf. (9))
 Guangdong-people eat mouse GAA4
 ‘do Cantonese people eat mice then?’
- b. nei⁵ gin⁶-saam¹ gam³ gwai³ gaa⁴? (Fāng, 145)
 2s CL-shirt so expensive GAA4
 ‘is your shirt so expensive then?’
- c. go²-di¹-syu¹, aa³-ji⁶suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ gaa⁴ (cf. (7))
 that-CL-book 2nd uncle will continue send-back-come GAA4
 ‘as to those books, will Second Uncle really continue to send them to us?’

ge2: *ge2* occurs in two different types of sentences, interrogative and declarative (Fāng, 2003; Fung, 2000; Law, 1990). Fung says (p. 158) that, as an interrogative particle, it is shorter and has a sharper rise. This is only partly true (see below), although the non-interrogative counterpart can certainly be characterized as being “prolonged”.

Declarative *ge2* is said to convey “the speaker’s reservation or uncertainty about the situation” (Fung, 2000:161). Its use “suggests that the speaker’s commitment is of a lesser degree [than with *ge3*]” (Law, 1990:96). Law suggests that it is paraphrasable as ‘Such would be the case if . . .’ or ‘Such would be the case but . . .’.

Our informants basically agree with these characterizations, but they do not think that sentences with *ge2* express that the speaker is less committed to the contents of the utterance than would be the case with *ge3*. It is true that sentences with declarative *ge2* are often followed by a “but”-sentence (and if it is not overt, it is certainly implicit), but the “but”-sentence, we get the impression, is not meant to weaken the commitment, but to tone down the hearer’s apparent expectation. We can see what this means when we add *ge2* to some of the sentences we have come across above.

- (11) a. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² ge² ... (cf. (10a))
 Guangdong-people eat mouse GE2
 ‘Cantonese people certainly eat mice [but it’s not our favorite food]’
 b. go²-di¹-syu¹, aa³-ji⁶suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ ge2 ... (cf. (10c))
 that-CL-book 2nd uncle will continue send-back-come GE2
 ‘as to those books, Second Uncle will certainly continue to send them to us [though not in the way you seem to think/not so fast]’
 c. ngo⁵ dim² dou¹ wui² bong¹ nei⁵ ge² ... (cf. (8a))
 I_s how all will help you GE2
 ‘I will surely help you under all circumstances! [but you should not only count on me]’

So-called interrogative *ge2* primarily occurs in sentences, such as the following:

- (12) a. hau⁶-min⁶ jau⁵ tiu⁵ ho⁴ ge²? (Fung, 159 (33))
 back-side there.be CL river GE2
 ‘how come there is a river in the back?’
 b. keoi⁵ m⁴ lei⁴ ge²? (Fāng, 89)
 3s NEG come GE2
 ‘why hasn’t he come?’
 c. ci⁴ dou³ ge²? (Fāng, 91)
 late arrive GE2
 ‘how come you are/he is (etc.) so late?’

Questions with *ge2* are interpreted as “why”-questions, despite the absence of any element with that meaning. Note that ‘why’ can be made explicit:

- (13) keoi⁵ dim²-gai² m⁴ lei⁴ ge²? (Fāng, 90)
 3s why NEG come GE2
 ‘why hasn’t he come?’

Interestingly, we observe that, while the rise of the *ge2*s in (12) is quite abrupt, this is not the case in (13). In the latter sentence, the rise and length is somewhere between those in (11) and those in (12). Fung (2000:160) claims that *ge2* cannot occur with any other *wh*-word than those meaning ‘why’ or ‘how come’ (see also Law, 1990). Fāng (2003), however, gives several examples, which counter this claim and our informants agree with Fāng, although there are restrictions (we have not been able to make out what they are).

It is possible that there are indeed two different *ge2*s, a declarative one, somewhat prolonged, and an interrogative one, with a sharp rise. It is also possible, however (and that is the line we

follow here, despite the fact that a lot more can be said about the question whether we have one *ge2* or two), that there is just one *ge2*, which is not an interrogative SFP. So-called interrogative *ge2* is then basically like *ge3* and declarative *ge2* except that it expresses a strong sense of surprise or unexpectedness. The externalization of the surprise subsequently gives rise to the apparent need for an explanation on the part of the speaker, whence the sense that a “why” is missing.¹⁴ The difference in length and rise may be related to the other difference we noted, viz., that the interrogative one occurs at the end of a sentence while the other one ends a clause and is explicitly, or implicitly, followed by a *but*-clause.

gaa2: from Fung (2000:171ff), we understand that *gaa2* shows doubt on the part of the speaker. It is paraphrasable as ‘I thought it was a relevant fact that ... (but somehow it doesn’t seem to be the case)’ or ‘It is a relevant fact that ... (isn’t it?)’. With *gaa2*, the speaker shows that he assumes the existence of a certain situation and is puzzled by the fact that this assumption seems incorrect. Sentences with *gaa2* are often followed by a sentence describing the situation as it really is, which is contrasted with the situation as the speaker assumed it to be. Note that the “*so why*”-sentence in the square brackets can also be rephrased using a *but*-clause, thus bringing *gaa2* in line with *ge2*.

- (14) a. gwong²-dung¹-jan⁴ sik⁶ lou⁵-syu² gaa² ... (cf. (10a))
 Guangdong-people eat mouse GAA2
 ‘[I thought that it was a fact that] Cantonese people eat mice [so why is it not on the menu?/so why aren’t there any mouse farms?]’
- b. go²-di¹-syu¹, aa³-ji⁶suk¹ wui⁵ luk⁶zuk⁶ gei³-faan¹-lei⁴ gaa² ... (cf. (10c))
 that-CL-book 2nd uncle will continue send-back-come GAA2
 ‘as to those books, [I thought it was a fact that] Second Uncle will continue to send them to us [so why are you throwing away our bookshelves?]’

Intermediate conclusion. On the basis of these short descriptions, we can formulate the following two generalizations. First, all members of the *g*-family are involved in asserting to some degree the relevance of the statement they are attached to the current conversation, or in the questioning thereof. The second descriptive conclusion is that the meaning that the different members of the *g*-family convey can be characterized as the meaning of *ge3* plus something else. What this something else is and whether it can be linked to the different formal properties of the various *g*-particles will be investigated below.

2.1.1.2. The *l*-family. The *l*-family has twice as many members as the *g*-family: *le1*, *le3*, *le4*, *le5*; *laa1*, *laa3*, *laa4*, *laak3*; *lo1*, *lo3*, *lo4*, *lok3*. Fung (2000) makes the claim that the central notion expressed by all these SFPs is [+realization of state]. This core meaning is exactly what is expressed by *laa3*; so let’s start our overview of the SFPs in the *l*-family and the meanings they convey with *laa3*.

laa3: expresses “realization of state” according to Fung (2000:93), that is “of a physical state in the real world [or] of an epistemic state” (Fung, 78–79), which means that *laa3* marks a change

¹⁴ A terminological note is in order here: the fact that a particle is compatible with an interrogative sentence, does not make it an interrogative particle; many particles, including *ge3*, can be found in questions. Only those particles, which can be argued to turn a declarative into an interrogative will be called “interrogative particles”.

of state, either objectively or subjectively.¹⁵ Much like Mandarin sentence-final *le*, *laa3* is best paraphrased as ‘it is *now* the case that’, implying that it was not the case earlier—at least, not as far as we were aware. According to our consultants, sentences with *laa3* are plain, neutral and factual.

- (15) a. ngo⁵ jiu³ heoi³ Mei⁶ gwok³ laa³ (Fung, 78)
 I_s need go America LAA3
 ‘[it is now the case that] I have to go to America’
- b. ai¹ jaa³, gam¹-jat⁶ jaa⁶-ng⁵ hou⁶ laa³! (Fung, 79)
 oh today 25 day LAA3
 ‘oh, today it’s already the 25th!’
- c. lok⁶ jyu⁵ laa³ (Fāng, 103)
 fall rain LAA
 ‘it’s raining now’

The sentence in (15c), for instance, implies that just a moment ago (or at least the last time we looked out of the window) it was not raining; a change of state has taken place. Without *laa3* the sentence would not have the change of state interpretation. *Laa3* can also be used in sentences that express completed events, in which case it may either be seen as “doubling” the completion semantics already expressed elsewhere, or as marking a change of state subjectively: according to the speaker, we are dealing with a new state of affairs. *Laa3* can also be used to express that a certain action is about to begin, in other words that a state is about to realize.

- (16) a. ngo⁵ sik⁶-jyun⁴ laa³
 I_s eat-finished LAA3
 ‘I’m done eating’
- b. coet¹ gaai¹ laa³! (Fāng, 105)
 go.out street LAA3
 ‘we are going out!’

Fāng (2003:159) reports that *laa3* is also used to change a declarative into an adhortative/directive, but in all examples, which supposedly illustrate this usage, other elements in the sentence are more likely to bring in the adhortative/directive sense than *laa3*. *Laa3* probably adds some sense of urgency due to its meaning ‘it is the case *now*—in contrast to earlier’; one of our informants reports that the sentences in (17) in some way “ask” for a change of state.

- (17) a. m⁵-hou² gong² gam³ do¹ laa³ (Fāng, 159)
 not-good talk so much LAA3
 ‘don’t talk so much!’
- b. ngo⁵-dei⁶ hou² zau² laa³ (Fāng, 159)
 I_P good leave LAA3
 ‘let’s leave!’
- c. ng⁵-dim² laa³, zau² laa³
 five-o’clock LAA3 leave LAA3
 ‘it’s five o’clock, go now’

¹⁵ See also Yiu (2001), where it is said that *laa3* marks inchoativity.

laak3: qua meaning, it is essentially the same as its unchecked counterpart, but it expresses more involvement of the speaker, i.e., it is less neutral than *laa3*, for example, in the sense that the new situation that has come about has consequences for the speaker. A clause with *laak3* is often followed by a follow-up clause.¹⁶

- (18) a. ngo⁵ sik⁶-jyun⁴ laak³ ... (cf. (16a))
 1s eat-finished LAAK3
 ‘I’m done now’ [so I’m off to work]
 b. lok⁶ jyu⁵ laak³ ... (cf. (15c))
 fall rain LAAK3
 ‘it’s started raining’ [so what are we going to do now?]

Thus, whereas the sentence in (15c) is just a factual statement, the sentence in (18b) is uttered (for example) when we had planned to go out. Fung (2000) claims that *laak3* is just like the other forms ending in *-k* in expressing a greater emotional involvement than their unchecked counterparts. Whether this is true here, is not entirely clear. In any case, with *laak3* the situation expressed in the sentence preceding it matters more to the speaker than is the case with sentences ending in *laa3*.

laa4: forms an interrogative to question or check whether a certain event did indeed take place or whether a certain state did indeed realize; it is used to seek factual confirmation, not to elicit new information (Fung, 2000:103; Fāng, 2003:70 a.o.p.). Our informants report that there is a short form and a long form, the former being more crude or rude than the latter. This point aside, *laa4* is truly the interrogative counterpart of *laa3*.

- (19) a. lei⁵ sik⁶-jyun⁴ laa⁴? (cf. (16a))
 2s eat-finished LAA4
 ‘it is indeed the case that you’re done eating?’
 b. nei⁵ bun² syu¹ wan²-faan¹ laa⁴?
 2s CL book look.for-back LAA4
 ‘it is indeed the case that you found your book back?’
 c. lok⁶ jyu⁵ laa⁴? (cf. (15c))
 fall rain LAA4
 ‘is it indeed raining now?’

laa1: would in comparison to *laa3* be more tentative, lacking forcefulness (Fung, 2000:96). Fung says that the difference lies in how much knowledge is assumed on the part of the hearer: *laa3* assumes no knowledge, *laa1* assumes the hearer should know—is not it obvious?¹⁷ Thus, using *laa1* in some of the sentences we have been using above (changing the subject in (20a)), we get the following results.

- (20) a. keoi⁵ sik⁶-jyun⁴ laa¹ (cf. (16a))
 3s eat-finished LAA1
 ‘[obviously] he’s done eating now’

¹⁶ See Fung (2000:104) and Leung (1992) for a similar view. Fāng (2003:108–109) has a different take on the checked particles, but our informants (all from Hong Kong, like Fung; Fāng describes the dialect of Guǎngzhōu) invariably agree with Fung.

¹⁷ Fāng (2003:56) has a slightly different intuition; our consultants agree with Fung; see footnote 16.

- b. ngo⁵ jiu³ heoi³ Mei⁶ gwok³ laa¹ (cf. (15a))
 1s need go America LAA1
 ‘[it is now the case that] I have to go to America [as you should know]’

Just like *laa3*, *laa1* is compatible with a directive or in any case a request/suggestion (not as strong as an imperative) (Fung, 2000; Fāng, 2003:158). The sentence in (21b) is much softer than its counterpart with *laa3* in (17c).

- (21) a. bong¹-haa⁵ sau² laa¹ (Fāng, 158)
 help-bit hand LAA1
 ‘help a bit please!’/‘let’s help’
 b. ng⁵-dim² laa³, zau² laa¹ (cf. (17c))
 five-o’clock LAA3 leave LAA1
 ‘it’s five o’clock, let’s go now’

le3: expresses deontic modality, according to Fung; in her words, it is used to “pose a suggestion of the occurrence of the state conveyed by the proposition”; it seeks the hearer’s agreement in a consultative tone (Fung, 2000:129; see also Fāng, 2003:73). As mentioned in the notes that accompany Chart 1, our native speaker consultants do not have this element in their speech, and we will not discuss it separately anymore, until section 3.

- (22) gam¹-maan⁵ ngo⁵-dei⁵ heoi³ sik⁶ hoi²-sin¹ le³ (Fāng, 73)
 tonight 1P go eat seafood LE3
 ‘shall we go and eat seafood tonight?’

le4: our consultants do, however, have *le4*. Fung (2000:130) and Fāng (2003:73) agree that *le4* is basically the same as *le3*, except that it is stronger mainly in that it assumes agreement from the hearer. For our consultants, sentences with *le4* are suggestions/propositions, which are not entirely neutral: there is a strong sentiment that it would be nice if the hearer were to agree. One informant says that suggestions with *le4* tend to be uttered “out of the blue”: they are not used to formulate the outcome of a discussion about immediate plans. The sentence in (23) will not mean “[in conclusion], let’s go and have seafood tonight then”.

- (23) gam¹-maan⁵ ngo⁵-dei⁵ heoi³ sik⁶ hoi²-sin¹ le⁴ (cf. (22))
 tonight 1P go eat seafood LE4
 ‘let’s go and eat seafood tonight, okay?’/‘shall we go and eat seafood tonight?’

le5: can be seen as an SFP of re-assertion: “It re-asserts a state-of-affairs that has been brought up before, but has not been properly acknowledged by the hearer” (Fung, 2000:128), paraphrasable as ‘I have told you’ (Law, 1990:137), with a slight touch of reproach (Fāng, 2003:41, 72). The speaker is sure, the hearer has apparently completely forgotten.

- (24) a. ngo⁵ zan¹ hai⁶ gin³-dou² keoi⁵ le⁵ (Fung, 129)
 1s really see 3s LE5
 ‘believe me, I really saw him’

- b. ngo⁵ m⁴ zung¹-ji³ sik⁶ min⁴ le⁵ (Fāng, 137)
 1s not like eat noodles LE5
 ‘I really don’t like noodles [as you should know]’

Fung and Law report that *le5* is also found as an interrogative SFP, in which the speaker wants confirmation or agreement from the hearer (Fung, 2000:126; Law, 1990:137; Fāng, 2003 agrees); we doubt that we are dealing with a separate particle. Meaning-wise, it is not different from non-interrogative *le5* in that it is asking for re-assertion. It is close to a rhetorical question: the speaker is sure about the contents of the utterance: ‘I told you, didn’t I?’ To illustrate, the following sentence, in (25), is used in a context in which we had planned to go and have seafood, when somebody proposes to go and see a film. The sentence wants to express that there is no time to see a film. The speaker, sure of himself, realizes that he is going against what seems to be the course of events and that is why he uses *le5*.

- (25) gam¹-maan⁵ ngo⁵-dei⁵ heoi³ sik⁶ hoi²-sin¹ le⁵ (cf. (23))
 tonight 1P go eat seafood LE5
 ‘but weren’t we going to have seafood tonight?’

The exact same semantics is preserved when we change the first-person subject into a second person subject:

- (26) gam¹-maan⁵ lei⁵-dei⁵ heoi³ sik⁶ hoi²-sin¹ le⁵ (cf. (25))
 tonight 2P go eat seafood LE5
 ‘but weren’t you guys going to have seafood tonight?’

It is interesting to see what happens when we contrast *le5* and *laa4*. Here is (26) with *le5* replaced by *laa4*.

- (27) gam¹-maan⁵ lei⁵-dei⁵ heoi³ sik⁶ hoi²-sin¹ laa⁴ (cf. (26))
 tonight 2P go eat seafood LAA4
 ‘[is it indeed now the case that] tonight you guys are going to have seafood?’

The latter is not only more neutral, but it very clearly incorporates the semantics associated with the [+realization of state], which Fung claims is expressed by all members of the *l*-family. We observe the same contrast in (28):

- (28) a. keoi⁵ m⁴ zung¹-ji³ nei⁵ le⁵? (Fāng, 41, 72)
 3s not like 2s LE5
 ‘so he doesn’t like you, right?’
 b. keoi⁵ m⁴ zung¹-ji³ nei⁵ laa⁴?
 3s not like 2s LAA4
 ‘he doesn’t like you anymore?’

Whereas the sentence with *laa4* in (28b) clearly conveys a change of state interpretation, it is not obvious that this also applies to (28a). We return to this point below.

le1: this particle has a variant form, *ne1*. Because none of the other *l*-particles have an *n*-variant, we assume that *le1* is a phonological variant of *ne1*, instead of the other way around. For *ne1*, see below.

(29) ngo⁵-dei⁶ hou² zau² lo³ (cf. (17b))
 1P good leave LO3
 'come on, let's leave!'

(30) a. ai¹jaa³, gam¹-jat⁶ jaa⁶-ng⁵ hou⁶ lo³! (cf. (15b))
 oh today 25 day LO3
 ‘oh, today is already the 25th!’ [it’s already too late!]
 b. lok⁶ jyu⁵ lo³ (cf. (15c))
 fall rain LO3
 ‘it’s raining now’ [not good!]

lo1: is very similar to *lo4*. It differs from it in that it is less factual and less blunt (Fung, 2000:119). One of our informants reports that with *lo1* the given explanation is “more evasive”, as if one does not want to admit the truth. For instance, when one is asked to explain certain aspects of one’s behavior (you’re buying suitcases, you have refused a wonderful job offer) you can say (31a), presenting a plain fact as a factual explanation of the situation, or utter (31b), which may imply something like ‘I don’t really want to go’ or ‘I know you don’t want me to go’.

- (31) a. ngo⁵ jiu³ heoi³ Mei⁶gwok³ lo⁴ (cf. (15a))
 1s need go America LO4
 'well, I have to go to America'
 b. ngo⁵ jiu³ heoi³ Mei⁶gwok³ lo¹
 1s need go America LO1
 'well, you know, I have to go to America'

¹⁸ We have not been able to get *lo4* clear for ourselves as not all our informants are comfortable with *lo4*, and some of those who do experience it as stronger than what suggested in the literature.

Fung (2000:119–120, 123) describes the difference between *lo1* and *lo4* in terms of knowledge assumption: *lo1* assumes knowledge on the part of the hearer, *lo4* does not (see the parallel with *laa1* versus *laa3*). Sentences which sound like suggestions with *lo1*, “become declaratives conveying the opinion of the speaker” with *lo4*; they are no longer a suggestion, but convey the opinion of the speaker.

Intermediate conclusion. One point we have to bring up is that, whereas with respect to the *g*-family, all members clearly had one meaning aspect in common (assertion of relevance), for the *l*-family it is less obvious that this is the case. The *laa*- and *lo*-members induce the change of state reading (in Fung’s words, they all share the feature [+realization of state]), but the *le*-members are found in suggestions/propositions and sentences that seek reassertion or confirmation, with no clear [+realization of state] semantics at all. (Note that all members of the *l*-family are compatible with suggestive sentences.) We may be dealing with two families instead of one; we return to this issue shortly.

Anticipating some of the discussions we will get into below, let’s make some observations regarding the other formal aspects of the SFPS we have seen so far (other than the initial, that is). First, in as far as *laa3* is the base particle of the *l*-family (or of one of the *l*-families), in both this and the *g*-family the particle that we identified as the “base particle” is a tone 3 particle. Secondly, both SFPS with *aa4* that we have seen are question particles. Next, Law (1990) seems to be right in claiming that SFPS with a high tone differ from their tone 3 counterparts in being a bit softer and less committed. Finally, from the examples we have seen we can also conclude that Fung is right about the *-k*, the coda consonant we have come across in some SFPS. Her claim (as we saw) is that *-k* makes an SFP “emotionally more intense”. Although in itself this characterization is vague and multi-interpretable, it is true, as we have seen, that checked SFPS show more involvement in what is going on the part of the speaker than is the case for the respective unchecked counterparts.

2.1.1.3. The *z*-family. Although the *z*-family may have other members for some speakers (see the notes to Chart 1; for *ze3*, see also section 3), we only look at the following four (anyway the most frequent ones): *zaa3*, *zaa4*, *ze1* and *zek1*. According to Fung (2000), all four have as their core meaning “delimitation”, with “contrast” as a derivative. We start with *zaa3*, because it seems to convey just this core meaning and nothing else.¹⁹

zaa3: 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’. The following sentences show its use in different contexts, scoping different parts of the sentence (Fāng, 2003:118, 133; Fung, 2000:58).²⁰

- (32) a. ji⁴gaa¹ zau⁶ waa⁶ hou² zaa³ (Fung, 59)
 now then say good ZAA3
 ‘it’s quite good at this moment only’
 b. ngo⁵ sik¹ da² gei¹ zaa³ (Fāng, 133)
 Is know play machine ZAA3
 ‘I can only play game machines’
 c. ngo⁵ heoi³ jau⁴-seoi² zaa³ (Fāng, 118)
 Is go swim ZAA3
 ‘I only go for a swim’

¹⁹ Li (2006a) also reports the existence of *zaa5* (as well as *laa5*, *gaa5*), most likely incorporating *aa5*. For discussion see Li (2006a).

²⁰ For *zaa3*, see Law (2004) and also Tang (1998:45ff).

- d. ngo⁵-dei⁶ hok⁶hau⁶ jau⁵ leung⁵-cin¹ jan⁴ zaa³ (Fāng, 134)
 1P school have 2000 people ZAA3
 ‘our school only has 2000 people’
- e. keoi⁵ ng⁵- sap⁶ sei³ zaa³ (cf. Fāng, 145)
 3s 50 year ZAA3
 ‘he is only 50 years old’

zaa4: turns a statement into a y/n-question, questioning and verifying, with some disbelief, the delimitation aspect of the semantics—‘really only that?’ (Fung, 2000:66; Fāng, 2003:145). The following examples illustrate this, and no further commentary is necessary.

- (33) a. gam¹-maan⁵ coeng³ ga¹-lai¹-OK zaa⁴? (Fāng, 145)
 tonight sing karaoke ZAA4
 ‘tonight we only sing karaoke?’
- b. lei⁵ heoi³ jau⁴-sei² zaa⁴? (cf. (32c), Fāng, 119)
 2s go swim ZAA4
 ‘you only go for a swim?’
- c. lei⁵-dei⁶ hok⁶hau⁶ jau⁵ leung⁵-cin¹ jan⁴ zaa⁴? (cf. (32d))
 2P school have 2000 people ZAA4
 ‘your school really only has 2000 people?’
- d. keoi⁵ ng⁵- sap⁶ sei³ zaa⁴? (cf. (32e))
 3s 50 year ZAA4
 ‘he is really only 50 years old?’

ze1: adds “down-playing” (in certain contexts even to the degree of disapproval or contempt) to the core meaning of delimitation (Fung, 2000:48; Fāng, 2003:55, 137). It expresses an implicit ‘it’s not a big deal’, ‘don’t make such a fuss’. We can see how this pens out when we compare *ze1* to *zaa3*.

- (34) a. lei⁵ heoi³ jau⁴-sei² ze¹ (cf. (32c))
 2s go swim ZE1
 ‘you’re only going for a swim’ [not a trip around the world]
- b. ngo⁵-dei⁶ hok⁶hau⁶ jau⁵ leung⁵-cin¹ jan⁴ ze¹ (cf. (32d))
 1P school have 2000 people ZE1
 ‘our school only has 2000 people’ [don’t think too much of it]
- c. keoi⁵ ng⁵- sap⁶ sei³ ze¹ (cf. (32e))
 3s 50 year ZE1
 ‘he is only 50 years old’ [not 150]

Whereas the sentence in (32c) is a neutral statement in the sense that it says that the person in question is only going for a swim and is not going to do anything else, the sentence in (34a) is uttered in a context in which someone is making a huge fuss one way or another while all she is going to do is going for a swim. Similarly, the sentence in (32d) is a neutral statement: 2000 people is not so big for our type of school. It may be followed by a sentence expressing the expectation that the government is not going to spend a lot of money on our school—why would it? 2000 people is *factually* a small number. The counterpart of the sentence with *ze1* in (34b) is used in a context in which 2000 is actually thought of as a big number, on the basis of which there seems to be some expectation that the government is going to pour huge amounts of money into our institution. The sentence in (34b) is used to say something like: ‘wake up!/be realistic’—2000 does not mean anything! Finally, we observe a similar difference when we compare (32e) with

(34c). The neutral, factual statement in (32e) is used when the age of 50 is deemed relatively young—the sentence can be followed by statements like “and he’s already the President of the United States” or “but he looks much older”, or you use it when you see that a certain person is retiring—wow, so young! In contrast, the sentence in (34c) expresses that 50 is not old enough. In the context of retirement, for instance, the sentence does not imply “wow, so young!” but “too early!” Or the person in question may be in very poor health (too young). Or the preparations for his 50th birthday are absolutely excessive—come on, it’s only his 50th birthday, not his 150th!

zek1: is in most respects the same as *ze1*, except that it “distinguishes itself from *ze1* by its high affective value” (Fung, 2000:57) or its “stronger emotional force” (Fung, 2000:50). Probably as a result, *zek1* is even less neutral than *ze1* and is often interpreted as adhortative and persuading. In one respect, *zek1* is more fundamentally different from *ze1*: it is used to report on things which the speaker thinks are only known to a very small number of people, which does not include the hearer, though it does include the speaker, much to her own content and pride (Fung, 2000:56; Fāng, 2003:31, 137, a.o.p.). The connection with a high emotive value is prominent in this use, though the link with the delimitation semantics seems to have been lost. Several sources mention that *zek1* is mainly used by female speakers (Chan, 2002; see also Law, 1990:196; Cheung, 1972:181–182).

Intermediate conclusion. What we observe in this family does not confirm all conclusions drawn in earlier sections. We do see confirmation of the conclusion that the base particle in a family is a tone 3 particle and that *aa4*-particles tend to be question particles. The claims with respect to coda *-k* are also confirmed. However, the idea that tone 1 *SFPs* are somehow “softer” than the base particles is not confirmed here.

2.1.2. b, h, m, n

b: as mentioned in the notes to [Chart 1](#), the status of the *b*-family is not clear. In any case, disregarding the sole mention of the form *be3* in [Matthews and Yip \(1994:340\)](#) (which none of our informants use), it has one member, *bo3*, which, according to a count reported in [Leung \(2005\)](#) is not very frequent.²¹ Some treat the SFPS *bo3* and *wo3* as free variants (or as two completely overlapping elements), but others do not. Notably, [Matthews and Yip \(1994\)](#) treat the two elements as different SFPS and ascribe different functions to them. The function ascribed to *bo3* is that of marking sentences as exclaimatives.

- (35) hou² ye⁵ bo³!
 good stuff BO3
 ‘well done!’

The work with our informants confirms that the situation is not entirely clear. We get the impression that Matthews and Yip are right in assuming that there are two sFpS, *bo3* and *wo3*: the forms generally overlap (*bo3* in (35) can be replaced with *wo3*), but there are contexts in which *wo3* cannot be replaced by *bo3* (particularly in ‘reminding’ sentences, such as (45b) in section 2.2.3, and for some of our informants also questions, but Fāng (2003) has a number of A-not-A examples with *bo3*).²² Pending more research, we essentially ignore *b*.

h: the three members of the *h*-family, *haa2*, *he2* and *ho2* are essentially adhortative-reminding SFPS. Sometimes they act as tags, if one wants someone's response to a proposition. Fāng (2003:147)

²¹ It occurs only 16 times in a 8-h spoken corpus, against 603 occurrences of *wo3*.

²² Leung (2005) argues that *wo3* in current Cantonese has two different historical sources, one of them *bo3*, the other *waa3*, now obsolete. The latter is an unlikely scenario, according to one of the reviewers, as it involves a sound change that is not attested anywhere else in the language.

claims that *haa2* is more “negotiative” than the other two; our informants report that when you use *haa2*, you do not necessarily expect an answer. With respect to *he2* and *ho2*, Fāng (2003:69) says that when one uses either of these two SFPs, one hopes that the other party agrees with you. This is confirmed by our consultants for *ho2*; they do not have *he2* as part of their speech.

- (36) a. m⁴ sik¹ zau⁶ man⁶ ngo⁵ haa²? (Fāng, 147)
 not know then ask me HAA2
 ‘how about if you don’t understand you ask me?’
 b. m⁴ sik¹ zau⁶ man⁶ ngo⁵ ho²? (cf. (36a))
 not know then ask me HO2
 ‘how about if you don’t understand you ask me, okay?’

We will not return to the *h* particles until section 2.5.3.

m: Cantonese has two SFPs that turn a declarative into a y/n-question, *maa3* and *me1*. There is reason to believe that only the latter is a particle like the other ones we are discussing here, and that the former consists of the negation marker *m* ‘not’ plus the SFP *aa3*, which means that sentences with *maa3* are A-not-A questions (Li, 2006a).²³

Most sources and all informants agree that questions with *maa3* are quite neutral y/n-questions (cf. Law, 1990:21), though Matthews and Yip (1994:310) note *maa3* down as “relatively formal”. *Me1* is used in confirmation questions (e.g., Fāng, 2003:144) but also in questions that express disbelief (Fāng, 2003:68; Law, 1990:18). All this is illustrated in (37).²⁴

- (37) a. ting¹ jat⁶ lei⁵ gin³-dou² keoi⁵ maa³? (Fāng, 67; see also (2c))
 tomorrow 2s see 3s MAA3
 ‘will you see him tomorrow?’
 b. nei⁵ m⁴ gei³-dak¹ me¹? (Law, 18)
 2s not remember ME1
 ‘you mean you don’t remember?’
 c. bak¹-fong¹-jan⁴ m⁴ sik⁶ se⁴ ge³ me¹? (Fāng, 144)
 northerners not eat snake GE3 ME1 (Fāng, 68)
 ‘so, northerners don’t eat snake?’
 ‘you mean, northerners don’t eat snake?!’
 d. lei⁵ heoi³ me¹? (Fāng, 22)
 2s go ME1
 ‘are you going?’

On the basis of the above, and ignoring footnote 23, we register *m-* as [+q].

²³ This view was already suggested in Chan (1949:264), who translates *maa3* as ‘or not’, analyzing it as *m* ‘NEG’ plus *aa* which, Chan says, “expresses smoothness or gentleness”. A reviewer suggests that interrogative *maa3* is a loan from Mandarin. S/he also mentions that (37a) is quite odd in Hong Kong Cantonese. The fact that it is good in Fāng’s Guǎngzhōu Cantonese confirms the loan theory, as Guǎngzhōu Cantonese is much more influenced by Mandarin than the variety spoken in Hong Kong.

²⁴ Fāng (2003) also mentions a non-interrogative use of *maa3*. It is used for “stating the obvious” (Fāng, 2003:63), often by way of explaining something. In some of these contexts it is interchangeable with the bisyllabic *a33-maa33*; according to a reviewer it is a variant of it.

(i) ngo⁵ go³ zai² zau⁶-hai⁶ zung¹-ji³ go²-go³ lei⁵-zai² maa³
 1s CL son just like that-CL girl MAA3
 ‘my son simply likes that girl [and there is nothing we can do about it]!’ (Fāng, 133)

It seems very close to Mandarin non-interrogative *ma*, which Li (2006a) analyses as a force modifier. We will leave it out of the discussion here.

n: there is only one SFP with initial *n*, *neI* (with phonological variant *leI*) and it occurs in interrogative and declarative sentences. With respect to the former, we see *neI* in sentences that are interrogative independent of *neI*, such as *wh*-questions or A-not-A sentences; in this type of sentence *neI* does not seem to do very much (Fāng, 2003:154; but see below). An A-not-A sentence with *neI* is given in (38a). *NeI* can also be argued to turn a sentence into an interrogative sentence, as illustrated in (38b,c). In such sentences, it is paraphrasable as ‘what if’ or ‘what about’. With respect to *neI* in declarative sentences, Law (1990:121) views its function as “drawing someone’s attention to something”; this use is illustrated in (38d), from Law (1990:122).

- (38) a. laai⁵-baai³-jat⁶ lei⁵-dei⁶ jau⁵ heoi⁵ gung-¹jyun⁴ mou⁵ ne¹? (Fāng, 155)
 Sunday 2P have go park not.have NEI
 ‘did you guys go to the park on Sunday?’
- b. keoi⁵ m⁴ lei⁴ ne¹? (Fāng, 110)
 3s not come NEI
 ‘what if he does not come?’
- c. John⁵ ne¹?
 John NEI
 ‘and (what about) John?’
- d. A: keoi⁵ hou²-noi⁶ mou⁵ daa²-din⁶waa² bei² ngo⁵ laa³
 3s very long not-have telephone to me LAA3
 B: wak⁶ce² keoi⁵ m⁴ dak¹-han⁴ ne¹
 maybe 3s not free time L/NEI
 A: ‘she hasn’t called me for a long time’ B: ‘she may not be free’

Following Li’s (2006a) analysis of *ne* in Mandarin, we distinguish two different *neI*s. One *neI* is an evaluative marker in the sense of Cinque (1999): by using it, the speaker marks the content of the utterance as extraordinary or unusual; it is used with declaratives, but the *neI* in A-not-A and *wh*-questions may at the core of the matter also be an evaluative marker (see Li, 2006a for discussion of Mandarin *ne*). The second *neI* is the one in (38b) and (38c). Following suggestions from the literature (notably Wu, 2005; see for details and more references, Li, 2006a), we take this *neI* to be a topic marker: it marks a topic but, unlike more common topic markers, it signals that the hearer is expected to provide the comment (rheme) rather than the speaker. The hearer orientation is a function of the high tone, as we will see. In any case, as an SFP, we take *neI* to be an evaluative marker.

At this point, let’s return to *le4* and *le5*. We noted above that they do not really belong to the *l*-family because it is not obvious that their meaning can be argued to incorporate any “change of state” semantics. We would like to submit the view that *neI*, *le4* and *le5* (as well as *le3* for those who have it) actually form one family of their own, which means that *le4* and *le5* are also evaluative markers. They suggest and re-assert, respectively, but, as we saw above, the contexts in which they are used fit the “evaluative” contexts, in that whatever is suggested or re-asserted can be contrasted to “usual” and “ordinary”: in the descriptions of *le4* and *le5* above, we used expressions, such as “out of the blue” and “going against the course of events”.²⁵

²⁵ If they form one family with *neI*, the question is, of course, why *le4* and *le5* do not have a variant with *n*. We have nothing non-speculative to say on this matter. Note that, as a reviewer notes, there is an ongoing sound change in Cantonese from initial *n*- to *l*-.

2.2. The rhymes

2.2.1. The *e*-family

As mentioned in the notes that accompany [Chart 1](#), *e* as a separate SFP is only mentioned in [Law \(1990\)](#) and [Yau \(1980\)](#).²⁶ None of the other sources mention it and our native speakers consultants do not have it in their speech. Here is how [Law \(1990:94, 107\)](#) describes them; example sentences have already been given in (3).

e1: suggestion, but tentative

e3: neutral suggestion

e4: suggestion, but the speaker expects agreement

Below is a list of most *Ce*-particles we reviewed above (excluding *he2*). They are followed by a very short indication of their meaning, excerpted from the descriptions provided above.

ge2: assertion of relevance but speaker is less sure than with *ge3*

ge3: assertion of relevance

me1: y/n²⁷

le4: evaluative: suggestion, seeking compliance from the hearer

le5: evaluative: seeking re-assertion or confirmation

n/le1: evaluative

ze1: ‘only’

Only *le4* is a suggestive SFP and as such may arguably be said to incorporate suggestive *e*. None of the other occurrences of the rhyme *e* seems to be related to the suggestive semantics, and for these cases we would like to propose that *e* is a default vowel, an element that is simply there to enhance pronounceability, devoid of any meaning. It is plausible that the *e* in *le4* is also a dummy; recall that most written sources and our native speaker consultants do not have *e* as a separate SFP. The suggestive semantics of *le4* must then be ascribed to something else. If this is the case, *e* in *Ce*-particles (that is, particles consisting of a consonant and *e*) is always a dummy.

This has the following consequence: the particles that we have called “base particles” above are only base particles in that they convey (“all and only”) the core meaning of the whole family. They are not base particles qua form, in that all other SFPs are formally assembled on the basis of them. Instead, the CVT (consonant–vowel–tone) base particles consist of a meaningful onset, a default rhyme and, presumably, a default tone. Leaving the tones undiscussed for now, this leads to the following proposal with respect to the internal structure of the *Ce*-particles (ignoring *he2*):

ge2: *g* (‘assertion’) + [default vowel] + [tone]

ge3: *g* (‘assertion’) + [default vowel] + [tone]

me1: *m* (‘+q’) + [default vowel] + [tone]

le4: *l* (‘evaluative’) + *e* (‘suggestion’)/[default vowel] + [tone]

²⁶ [Leung’s \(1992\)](#) list of over 90 particles does not include *e*.

²⁷ A reviewer remarks that describing the function of *me1* as y/n “is missing rather a lot of pragmatics”.

le5: *l* (‘evaluative’) + [default vowel] + [tone]

n/le1: *n* (‘evaluative’) + [default vowel] + [tone]

ze1: *z* (‘only’) + [default vowel] + [tone]

This is a preliminary proposal and all aspects will be discussed as we proceed.

2.2.2. The *aa*-family

2.2.2.1. *Aa-particles without onset.* The *aa*-family is much bigger than the *e*-family. It has the following members (excluding *maa3*, see above):

aa1, aa3, aa4, aa5, aak3

gaa2, gaa3, gaa4, gaak3, haa2, laa1, laa3, laa4, zaa3, zaa4

We look at the onset-less members first.

aa3: occurs with declaratives, imperatives, exclamatives, wh-questions and A-not-A questions. It is generally reported that it makes an utterance sound softer, less abrupt and more natural (e.g., Law, 1990:108, Matthews and Yip, 1994:340, Fāng, 2003:58). We think that *aa3* performs a very plain discourse function: it makes the utterance fit more smoothly into the conversational context. It has been said that the *aa*-particles alert the hearer by highlighting the relevance of the utterance to the discourse, but we see this as a side effect: by smoothing the entrance of the utterance into the discourse, *aa* makes the utterance more fit for the discourse, thus elevating the chance that it is taken seriously as a relevant contribution. In this sense, it may be said to have an alerting effect on the hearer. We will refer to *aa3*’s semantic contribution as “smooth-alert”

- (39) a. cin⁴-min⁶ jau⁵ hou²-do¹ jan⁴ aa³ (Fung, 169)
in front have very-many people AA3
‘there are lots of people in front’
- b. faai³-di¹ sik⁶ aa³ (Law, 108)
quick-bit eat AA3
‘eat a bit faster!’
- c. ni¹-di¹ ca⁴ zan¹-hai⁶ zeng³ aa³! (Fāng, 163)
this-CL tea really-be tasty AA3
‘this tea tastes really nice!’
- d. nei⁵ heoi³ bin¹-dou⁶ aa³? (Fāng, 152)
2s go where AA3
‘where are you going?’
- e. nei⁵ zi¹-m⁴-zi¹ aa³? (Fāng, 154)
2s know-NEG-know AA3
‘do you know?’

For all these sentences, it is true that their counterpart without *aa3* sounds more abrupt or harsher than they do now. Without *aa3*, the questions in (39d,e) could be used by the police during an interrogation. *Aa3* removes the sharp edges.

aa1: has the same function and distribution as *aa3*, but it is less neutral; it reveals more of the speaker’s emotion or attitude. It is said to make the utterance more “lively” (Law, 1990:109; Matthews and Yip, 1994:340); in any case, the speaker expects a reaction from the hearer, conversational interaction is expected, possibly because the statement preceding *aa1* is in

contrast with assumed information. This is clear from the sentences in (40), which are the same as those of (39), except that *aa3* has been replaced by *aa1*. The additional meaning is indicated. The expectation on the part of the speaker may be the reason why sentences with *aa1* are perceived as more lively than those with *aa3*.

- (40) a. cin⁴-min⁶ jau⁵ hou²-do¹ jan⁴ aa¹
 in front have very-many people aa1
 ‘there are lots of people in front [why did you say there were just a few?]
- b. faai³-di¹ sik⁶ aa1
 quick-bit eat aa1
 ‘Eat a bit faster!’ [nudging]
- c. ni¹-di¹ ca⁴ zan¹-hai⁶ zeng³ aa¹!
 this-cl tea really-be tasty aa1
 ‘this tea tastes really nice! [how come you don’t like it?]
- d. nei⁵ heoi³ bin¹-dou⁶ aa¹?
 2s go where aa1
 ‘where are you going?’ [challenging]
- e. nei⁵ zi¹-m⁴-zi¹ aa¹?
 2s know-NEG-know aa1
 ‘do you know?’ [challenging]

aa4: there are two *aa4*s, one of which is a question particle (which means that it turns a declarative into question), mainly used for checking facts, seeking confirmation (e.g., Law, 1990:19).

- (41) keoi⁵ jiu³ heoi³ Mei⁶gwok³ aa⁴?
 3s will go America aa4
 ‘is he indeed going to America?’

The other *aa4* is like the other *aa*-particles in that it occurs in all other types of sentences; it is like *aa3* qua meaning and function, except that it expresses certainty or even impatience on the part of the speaker—at least according to Law (1990:108–109): the other sources only have question particle *aa4* and our native speaker informants are divided.

aa5: not all our informants have *aa5*, but those who have it agree with Law (1990:18, 101) that it is used only in questions seeking confirmation. The difference between *aa4* and *aa5* is that the latter conveys a higher degree of confidence on the part of the speaker with respect to the nature of the expected answer.

- (42) a. keoi⁵ jiu³ heoi³ Mei⁶gwok³ aa5 (cf. (41))
 3s need go America aa5
 ‘he is going to America, isn’t he?’
- b. zou⁶ jat¹-tiu⁴ aa⁵? (Law, 19)
 make one-CL AA5
 ‘you’re making just one, right?’

aak3: Fāng (2003) reports that sentences with *aak3* are quite firm. According to Law (1990:196), *aak3* indicates that the information carried by the utterance is intended to contradict an assumption or an expectation held by the addressee; Fāng agrees.

- (43) ngo⁵ mou⁵ waa⁶ lei⁵ aak³ (Fāng, 133)
 1s not-have speak 2s AAK3
 ‘I didn’t criticize you!’

Conclusion. The *aa* particles make utterances fit into the context more naturally, more smoothly, as a result of which it alerts the hearer to pay attention. The base member of the onsetless *aa*-group is *aa3*, the other ones clearly mark “smooth-alert” plus something else.

2.2.2.2. *Aa-particles with onset.* When we look at the *aa*-particles with onset, we see that the base meaning of *aa3* does reveal itself. We will first look in some detail at the two *aa*-particles that form minimal pairs with *e*-particles: *ge3* versus *gaa3* and *ge2* versus *gaa2* (once again disregarding *he2* versus *haa2*). We start with *ge3* and *gaa3*. As we already discussed them as a contrastive pair, we only give a brief description here; see also (4b) and (5) above.

ge3: assertion of relevance marker, paraphrasable as: ‘It is a relevant fact that ...’

gaa3: essentially the same as *ge3*, just a bit “softer”: ‘it is a relevant fact that ... but I don’t mind that you don’t know or forgot’, “reminding” (Fāng, 2003:133).

Here is another set of facts, based on Fung (2000:169); cf. (39a), (40a).

- (44) a. cin⁴-min⁶ jau⁵ hou²-do¹ jan⁴ aa3
 front have very-many people AA3
 ‘there are lots of people in front’
 b. cin⁴-min⁶ jau⁵ hou²-do¹ jan⁴ ge3
 front have very-many people GE3
 ‘[it is a relevant fact that] there are/were lots of people in front’
 c. cin⁴-min⁶ jau⁵ hou²-do¹ jan⁴ gaa3
 front have very-many people GAA3
 ‘[it is a relevant fact that] there are lots of people in front [remember, be careful]’

Our informants report that the difference between (44a) and (44b) mainly lies in the assertion aspect: the former is quite neutral in that it reports on a situation, where the addition of *aa3* makes it a natural contribution to the discourse, thus alerting the hearer to pay attention, while the latter, with *ge3*, quite bluntly presents the situation as a relevant fact. *Gaa3* in (44c) seems to combine both features: it emphasizes that the fact reported is a relevant fact, but adds the aspect of “smooth-alert”, alerting the hearer. In other words, *gaa3* may indeed be analyzable as a member of the *g*-family, incorporating *aa3*.

We find a similar situation when we look at *gaa2* in comparison to *ge2*; see (11) and (14). *Gaa2* is clearly a soft version of *ge2*:

ge2: followed by a “but”-sentence (overt or implicit), but the “but”-sentence is not meant to weaken the commitment, but to tone down the hearer’s apparent expectation. ‘It is a relevant fact that ..., but ...’

gaa2: paraphrasable as ‘I thought it was a fact that... (but somehow it doesn’t seem to be the case)’ or ‘It is a fact that (isn’t it?)’

The other SFPS with *aa* have no minimal counterpart with *e* to compare them to: *gaa4*, *gaak3*, *laa1*, *laa3*, *laa4*, *zaa3*, *zaa4*. Can we nonetheless say that they incorporate the “smooth-alert”

2.2.2.3. Intermediate conclusion. With respect to the internal structure of the *aa*-particles, let us conclude the following, keeping in mind that we have not discussed the tonal aspect of the *SFPs* at all.

zaa3, zaa4: 𐀓 (‘restriction’) + 𐀓𐀓 (‘smooth-alert’) + [tone]

wo3: is supposed to be an “informative” SFP, indicating “noteworthiness” according to Matthews and Yip (1994:340) and Luke (1990), although it is also used for “reminding” according to Fāng (2003:67). *Wo3* is often, but not exclusively, found in adhortatives.

- (45) a. mei⁵-gam¹ sing¹-zo² wo³ (M&Y, 353–354)
US-dollar rise-ASP wo3
'look, the US dollar has gone up!'
b. lei⁵ siu²-sam¹ za¹-ce¹ wo³
2s careful drive-car wo3
'drive carefully!'

wo4: also conveys noteworthiness, though it is slightly stronger than *wo3* as here we are dealing with the speaker's sudden awareness or discovery of unknown or unexpected information, while with *wo3* this is not the case (Law, 1990:100; Matthews and Yip, 1994:354; Fāng, 2003:79).³⁰

³⁰ Sometimes, *wo4* can seem a question particle, but it is not really, we think: the speaker is so surprised that the hearer feels prompted to give an explanation.

- (46) a. jyun⁴-loi⁴ keoi⁵ hai⁶ ngo⁵ ji⁵-cin⁴-ge³ tung⁴-hok⁶ (M&Y, 354)
 after-all 3s be 1s before-MRKR classmate
 lei⁴-ge³ wo⁴
 “BE” wo⁴
 ‘it turned out she was my former classmate!’
- b. jat¹-go³ sin¹ dou¹ m⁴ sai² wo⁴ (Law, 99)
 one-CL cent all not need wo⁴
 ‘I didn’t even need to pay one cent!’

wo5: is similar, but here the speaker is less sure about the information s/he reports; with *wo5*, the information is presented as “hearsay” information (Law, 1990:100; Matthews and Yip, 1994:354; Fāng, 2003:66): ‘I only heard this but who knows?’.

- (47) a. ting¹-jat⁶ lok⁶-jyu⁵ wo⁵ (Fāng, 136)
 tomorrow rain wo⁵
 ‘I hear it will rain tomorrow’
- b. A: gam³ do¹ jan⁴ hai²-dou⁶ gau²-wai⁶-sang¹ ge²
 so many people here clean up GE2
 B: jau⁵ jan⁴ lei⁴ gim²-ca⁴ wo⁵ (Fāng, 66)
 there.be people come inspect wo⁵
 A: ‘why are so many people here cleaning up?’
 B: ‘it seems some inspector is coming’

We conclude from all this that the core meaning that these SFPS have in common is “noteworthiness”. We also conclude that *wo3* is the base particle of this group: while *wo3* only expresses “noteworthiness” and nothing else, the other SFPS express this notion plus something else: the speaker wasn’t aware earlier (*wo4*) or is not a hundred percent sure and presents the information as hearsay information (*wo5*).³¹

Turning to the *lo*-particles, when we look at the short descriptions below, which are based on the longer discussion presented above, we may safely say that “noteworthiness” is part of the semantics of these SFPS, in addition to the [+realization of state].

lo3: we find *lo3* in the same contexts as *laa3*; similar, just stronger in the sense of reporting new information

lok3: is basically the same as *lo3*, but with stronger emotion

lo4: is most often used in answers, explaining one’s behavior factually, neutrally

lo1: is very similar to *lo4*, just less factual and more evasive (giving an excuse rather than an explanation)

We conclude:

wo3, wo4, wo5: *o* (‘noteworthiness’) + [tone]

lo3, lo1, lo4: *l* (‘realization of state’) + *o* (‘noteworthiness’) + [tone]

lok3: *l* (‘realization of state’) + *o* (‘noteworthiness’) + *k* (‘emotion intensifier’) + [tone]

³¹ Stephen Matthews (p.c.) reports that the difference between *wo3* and *wo5* seems to be disappearing (a reviewer reports that in many young speakers tones 3 and 5 are merging more generally); Leung (2005) says that they are often used interchangeably. Our informants still feel a distinct difference (with *wo3* being more sure, no “heasay” factor).

2.3. Coda -k

With respect to the coda -k, we have been confirming Fung's (2000) claim that it signals greater emotional involvement on the part of the speaker than the unchecked counterparts. We will not discuss it more here.

2.4. Other particles

The two SFPS we have left undiscussed so far are *waa2* and *gwaa3*. Both are quite common SFPS, and the latter is mentioned in most sources.

waa2: is added to A-not-A and *wh*-questions and makes the interrogative nature more prominent (Fāng, 2003:37). Fāng's translations (into Mandarin) involve *nǐ shuō*, which may be translatable as 'did you say' or 'tell me'. Fāng (p. 75) also mentions that the particle is often used to ask about very current affairs, which is confirmed by Leung (1992:112) who says that the particle is used when one wants to verify whether one understood/heard the previous utterance correctly. Matthews and Yip (1994:318) characterize it as a marker of echo questions, "when the questioner wishes a particular word or phrase to be repeated". Here are some examples.

- (48) a. keoi⁵-dei⁶ bin¹-go³ saang¹-dak¹ gou¹ di¹ waa²? (Fāng, 37)
 3P who grow-EXTENT tall a.bit WAA2
 'which of them did you say is taller?'/ 'tell me: which of them is taller?'
 b. lei⁵ jiu³ gei²-do¹ waa²? (Leung, 1992:112)
 2s want how many WAA2
 'how many did you say you want?'
 c. keoi⁵-dei⁶ heoi³-m⁴-heoi³ waa²? (Fāng, 75)
 3P GO-NEG-go WAA2
 'are they going or not?'
 d. nei⁵ gong² mat¹-je⁵ waa²? (Fāng, 75)
 2P talk what WAA2
 'tell me: what are you talking about?'

It seems very much like an add-on, a tag. The SFP is quite obviously etymologically related to *waa3* 'say'. As a result, aside from the tonal issue, it is unlikely that it is further dissectable into smaller meaningful units.

gwaa3: Matthews and Yip (1994:353) state that *gwaa3* "indicates the speaker's uncertainty about the information in the sentence, like 'I suppose' in British or 'I guess' in American English", or, as they also mention, like Mandarin *ba* (Cheung, 1972:180). Matthews and Yip say that *gwaa3* is typically used in answers to questions and propositions, but Fāng (2003) and Cheung (1972) give several examples of *gwaa3* in questions, expressing that the speaker expects an affirmative answer. In fact, just like Mandarin *ba*, *gwaa3* can be attached to many different types of sentences.

- (49) a. lei⁵ m⁴-wui³ cheut³-heoi³ gwaa³? (Cheung, 1972:180)
 2s NEG-will out-go GWAA3
 'you're not going out, are you?'
 b. keoi⁵ seung⁵-lau² zou⁶ mat¹-je⁵ gwaa³ (Fāng, 53)
 3s go.upstairs do what GWAA3
 'he went upstairs to do something I suppose'

³² Note that for Law, not every tone **3** is a default tone; see below. The following tone **3** particles have no minimal counterpart that they can be contrasted with in tonal terms: *aak3*, *bo3*, *gaak3*, *gwaa3*, *laak3*, *lok3*; whether they are neutral is hard to assess, although the descriptions provided above give every reason to think that they are.

laa4: seeking confirmation of the realization of state-of-affairs

lo3: realization of state, noteworthiness

lo1: realization of state, noteworthiness, evasive, less factual than *lo4*

lo4: realization of state, noteworthiness, is most often used in answers, explaining one's behavior in a factual, neutral way

zaa3: restriction

zaa4: seeking confirmation of the restricted content

This overview confirms that, semantically, the tone 3 member is the most neutral member in each family.

Let us now concentrate on the other tones, and see what their meaning contribution is. The other tones are a high level tone (1) and what we may interpret as a low level tone (4); and two rising tones, one high (2), one low (5).

As noted above, Law (1990) has done important work in the tonal domain. Partly basing herself on work by others, especially Cheung (1986) (not available to us), she proposes that there are two non-segmental, purely tonal SFPS, a high one and a low one, “L” and “H”, respectively.³³ As Law observes, utterances to which L has been attached, are associated with a “stronger force”, while those with H generally convey a “weaker force”, whence she calls L a “strengtheners” and H a “weaker” (Law, 1990:94, and other places). In principle, the non-segmental SFPS can be attached to any utterance, whether it ends in an SFP or not, although, as Law acknowledges, cases of H or L attached to an utterance with no other SFP present are extremely rare (and it is unclear why). But here is an example:

- (50) a. keoi⁵ heoi³
 3s go
 ‘she is going’
 b. keoi⁵ heoi² (=3 + H) (Law, 126)
 3s go+weaker
 ‘she may go’ (more tentative)

In most cases, H and L are attached to utterances, which already have an SFP. According to Law, the SFPS that end up as tone 1 particles are combinations of toneless segmental SFPS with H, and most superficial tone 4 particles are combinations of toneless segmental SFPS with L. SFPS with a rising tone are combinations of segmental SFPS with a tone and H: 3 + H yields 2, 4 + H yields 5. Note that this implies that for Law, not all SFPS are inherently toneless (for the ones that are not, see footnote 38).

Others have proposed non-segmental tonal SFPS as well.³⁴ Chao (1968:812), for instance, postulates two such SFPS for Mandarin, one rising, one falling. Cheung (1972:170) suggests that Cantonese may have this kind of SFPS too, in any case a rising one, [35], and he explains [+q]-ge2 as a fusion between ge3 and this rising tonal SFP. Leung (1992) proposes six different tonal SFPS for Cantonese, some of which seem quite intonational in nature, a point to which we return.

³³ In fact she proposes three such particles, one L, two H, one of the H particles being used to form “echo questions”. See below and Law (1990:17, 86ff), Matthews and Yip (1994:318).

³⁴ We mention some proposals with respect to Chinese only, but in other areas we see similar proposals. Fabb (1992:5) postulates a “tone-only particle” in inland Ewe (spoken in Ghana) (which plays a role in syntax). For other African examples, see Yip (2002:272–273), and for tonal morphemes more generally, Yip (2002:106ff).

We adopt Law's H and L, though there are differences between our proposal and her's, some of which are a matter of fine-tuning, while some others are more substantial.

2.5.2. *High and low boundary tones: H and L*

2.5.2.1. *High and low boundary tones.* We would like to put Law's proposals with respect to H and L in a more general perspective, especially in reference to work on intonation by Pierrehumbert (1980), Pierrehumbert and Hirschberg (1990), Herman (2000) and Steedman (2000) as well as Chu (2002); for details and more complete references, see Kirsner et al. (1994) and Li (2006a). The notion relevant here is that of "boundary tone", which refers to the pitch at the boundary of a phrase or a sentence. In general, a high boundary tone can be interpreted as signaling that the phrase in question forms one interpretative unit with what follows, while this is not the case with a low boundary tone. The high boundary tone can be characterized as "forward looking". In case of a conversational setting, "forward looking" is interpreted as "hearer oriented": the speaker signals that it is the hearer's turn, that s/he wants input from the other interlocutor, be it confirmation or information. It is interpretable as that there is a sense of incompleteness, as that the speaker is not fully committed to what s/he just said. It is in this way that we have to interpret Law's (1990) claim that H is a weakener.

The low boundary tone does the exact opposite of what the high boundary tone does. The speaker is committed, does not show any need for confirmation or information from the side of the partner in the conversation, the speaker is, so to speak, not forward looking and not hearer oriented; we will call it "speaker oriented". This way we understand how L can be seen as a strengthener as Law does.

A question that is raised by this discussion concerns the status of these elements: are we indeed dealing with independent elements or is it just intonation? This issue is discussed by most of the literature we have been referring to. Yau (1980:51) suggests that there is "evidence supporting a tentative implication universal that there is a mutual compensation between S[entence] P[articles] and intonation patterns and that the more a language relies on the use of SP in expressing sentential connotations, the less significant will be the role played by intonation patterns, and vice versa." A similar suggestion is made in Cheung (1972).³⁵ Others have pointed out that the fact that Cantonese has three different register tones would provide its speakers with less space for intonational manoeuvring (see Fung, 2000:8ff, Matthews and Yip, 1994:27ff and Leung, 1992:23ff for discussion and references). Yip (2002, Chapter 9) makes it clear that although tone languages also have intonation, "tones bear a heavy functional load" (p. 272) as a result of which it could very well be case that the SFPS play a role which may be performed by pure intonation in other languages. This is especially true, Yip says, if the SFPS are toneless of themselves and in that way can be seen as "toneless carriers for intonation" (loc. cit). This fits well with Cheung's (1986) view, presented in Law (1990:83), that the non-segmental tonal SFPS are highly localized intonation. In other words, if Law's L and H can be associated with the low and high boundary tones, then we see that the SFPS in Cantonese, as a system, rather than replacing intonation, simply incorporate it.

Chao (1968:812) makes the following insightful comment regarding the two tonal SFPS he postulates for Mandarin: "I used to treat these as part of Chinese sentence intonation, but later

³⁵ Interestingly, as far as I know, Schubiger (1965) was the first to point at the interrelationship between particles and intonation in a comparative study on German and English, two non-tonal languages. It is good to realize that SFPS do not constitute an exclusive property of tone languages. For phonetic studies of such particles in Dutch, see Kirsner et al. (1994) and Kirsner and Van Heuven (1996).

found it better to treat them as particles, since they do not affect the intonational pattern of the whole construction, but only the voiced part of the last syllable.” We can view them as having an intonational role, but because they are realized on one segment only, we best treat them as separate non-segmental tonal SFPS. They partly support, partly represent intonation.

In view of all this, we treat them as separate MMUS.³⁶

2.5.2.2. Tone 1 particles. How do the SFPS with tone 1 and tone 4 fit in this framework? Let us look at the tone 1 particles first. In the spirit of the above, they should be analyzable as a base particle, combined with the high boundary tone. Let us check whether this is indeed the case. Here are the particles, contrasted with their respective tone 3 counterpart:

aa3: smooth-alert

aa1: smooth-alert; conversationally more lively than *aa3*

laa3: realization of state

laa1: realization of state, less committed

lo3: realization of state, noteworthiness

lo1: realization of state, noteworthiness, evasive

The picture is pretty consistent and well in line with the idea that the high boundary tone reflects a lower degree of commitment on the part of the speaker in the sense that the speaker is “forward looking” and expects something (confirmation, information) from the hearer.

Tone 1 particles that do not have an obvious base counterpart which they can be compared to are *me1*, *ne1*, *ze1*, *zek1*. Considering their meaning and function as reflected in the short description below, we conclude that for *me1* and *ne1* the tone can be seen as a reflex of the high boundary tone, but that this is much less obvious for *ze1* and even less so for *zek1*.

me1: confirmation seeking

n/le1: evaluative

ze1: adds the notion of “down-playing” to the core “delimitation” as expressed by the more neutral *zaa3*

zek1: like *ze1*, with higher affective value

In all, we accept the high boundary tone as a separate unit in our inventory of meaningful constituting parts of Cantonese SFPS.

2.5.2.3. Tone 4 particles. The picture presented by the tone 4 particles is less consistent than the tone 1 picture. Consider the following minimal pairs.

e3: neutral suggestion (see (2))

e4: suggestion, but speaker expects the hearer to comply/agree (see (2))

aa3: neutral smooth-alert

³⁶ It may be the case that not all non-segmental particles are the same: some may be more intonational than others. A criteria could be the extent to which a tonal particle can be attached to non-particles. The sharp rise we find at the end of non-particle sentences, representing a question (e.g. Matthews and Yip, 1994:318), may be more intonational in nature than the stylized/localized H and L. We leave the matter undiscussed for now.

aa4: (i) smooth-alert question particle, seeking confirmation

(ii) smooth-alert, with more certainty

laa3: realization of state

laa4: forms interrogative to question or check the existence of the event

zaa3: delimitative

zaa4: turns statement into y/n-question: only so little/few?

gaa3: quite neutral assertion of relevance

gaa4: seeking confirmation of the asserted content

le3: suggestion

le4: suggestion, seeking compliance from the hearer

lo3: realization of state, noteworthiness

lo4: realization of state, noteworthiness, is most often used in answers, explaining certain aspects of one's behavior in a factual, neutral way

wo3: noteworthiness

wo4: noteworthiness, expressing the speaker's sudden awareness

On the basis of these minimal pairs we can make the following observations. First, as before, in all cases the tone 4 particles are characterizable as the tone 3 counterpart, plus something else. Secondly, all *Caa4* are question SFPS, so the “something else” may be [+q]. On the other hand, (C)*e/o4* are not [+q] and *aa4* itself is ambiguous. In the non-q cases, we can easily observe the influence of the “speaker orientation” of the low boundary tone: sentences with the non-q tone 4 particles are more committed, express more force one way or another, than the same sentence with the respective tone 3 counterpart.

How about the (C)*aa4* question SFPS? First of all, it is clear that these SFPS are question particles in the sense that they turn a declarative into a question. We have seen several examples showing this. Secondly, taking one step back, we observe that the function of tone 4 is either adding strength to the counterpart with tone 3 (as is the case with *e4*, *aa4*, *le4*, *lo4*) or turning it into a question SFP (*aa4*, *laa4*, *zaa4*, *gaa4*). Interestingly, only one of these tone 4 elements is ambiguous, namely *aa4*, and the question is why none of the others is ambiguous. In other words, why do not the other *-aa4* SFPS have a “strength” reading? And why do not the non-*aa4* particles display the q-effect?

In answering these questions, it is helpful to make the further observation that all the tone 4 particles where tone 4 seems to have a strengthening effect are “suggestive” and “noteworthiness” particles; it may be the case that only in such contexts is “strength” a relevant notion. Supposing, for now, that this is indeed the case, we conclude that we actually can isolate tone 4, the low boundary tone, as an MMU, as argued above: the speaker orientation cue or the strengthener in Law's terms.

This leaves us with the question where the q-semantics comes from in the (C)*aa4*-particles that have it. What is clear is that it cannot be ascribed to the speaker oriented low boundary tone itself. At best, the low boundary tone is compatible with the [+q] semantics we are dealing with here, which is always confirmation seeking. The speaker does have some sense of grasp of the situation, or it may be similar to what we see in *wo4*: sudden awareness may prompt a reaction. In any case, the low boundary tone cannot be held responsible for the q-semantics of the (C)*aa4*-particles. So where *does* it come from?

As far as we can see, there are three possible answers. The first is that there is a second tone **4**, functioning as a q-element; the fact that this element is not compatible with “suggestive”-particles must then be ascribed to semantic incompatibility. A second possible explanation is that there is a second *aa*-particle, a [+q] particle, in which case the tone **4** in the [+q]-(C)*aa4*-particles is the low boundary tone. Finally, it may be the case that there is a separate q-particle, *aa4*_[+q], not further dissectable.

Because as far as we can make out at this point, there is no evidence to decide between these options, we adopt the third one, as this option not only reflects our ignorance most directly, but also avoids the postulation of a second *aa* and a second low tone, for which there is no independent evidence (but see Li, 2006a for a different view).

Granting all this, the discussion so far leads us to isolate two MMUS, the low boundary tone, i.e., the L-strengtheners, and *aa4*_[+q].

2.5.3. *Tone 2 particles and discussion of some of our basic assumptions*

As reported above, Law (1990) derives tone **2** by combining tone **3** with H. For us, this is not a straightforward option, since for us, going for the strongest possible hypothesis, tone **3** is always a default tone, which presumably will only be added at the very end of the derivation, some time in PF. However, we need to realize that the idea that tone **3** is a dummy in all cases is only a hypothesis. Let us discuss some aspects of tone **2** and a number of correlations.

First of all, when we look at the distribution of tone **1** and tone **2**, we observe that no family has both tone **1** and tone **2** particles. In other words, family-wise, they are in complementary distribution. The second observation to make with respect to tone **2** particles is that aside from the *h*-particles and *waa2*, which are all tag-like adhortative-interrogative SFPS, the only other two tone **2** particles are members of the *g*-family and they are not question SFPS.

ge2: assertion with reservation, uncertainty, surprise, etc.

gaa2: assertion with disbelief, surprise, etc.

haa2, **he2**, **ho2**: adhortative-negotiative-interrogative

waa2: used in (echo-)questions to emphasize the interrogativity, with adhortative touch

The *h-waa*-particles can be seen as incorporating the hearer orientation of the high boundary tone. However, it must be noted, and emphasized, that none of the tone **1** particles which we saw as having incorporated the forward looking hearer orientation of the high boundary tone have the tag-like adhortative sense to them, which suggests that the *h-waa*-particles may be principally different from tone **1** particles. Although *ge2* and *gaa2* are not question SFPS, they are also different from the tone **1** particles: they are forward looking in that they signal the presence of a *but*-clause, either explicit or implicit, but they do not expect input from the hearer.

The core question of this section is: what is tone **2**? Can it be associated with the high boundary tone?

Looking at these matters from the distributional angle again, we only need to take three families into consideration: the *aa*-particles, the *l*-particles and the *g*-particles, because the other families have no tone **1** or tone **2** particles, or they have nothing but. The fact that tone **2** is in complementary distribution with tone **1** in that, of these three families, no family has both tone **2** and tone **1** particles can be interpreted as confirmation for the idea that there is a difference

between the families that have tone **1**, that is the *aa*-group and the *l*-group, and the one that has tone **2**, the *g*-group. The difference could be that the base-particle in the *g*-group is *ge3* or *g3*, including the tone (where *e* may still be a dummy), while it carries no tone in the other families. In other words: Law is right in claiming that both tone **2** and tone **1** particles involve a non-segmental high tone *SFP*, the difference being that in the former the *SFP* is attached to a base with tone **3** while in the latter the base particle has no tone at all.

The problem with this approach is that, if we take the other tone **2** particles into consideration as well, we would have to assume that they have been derived on the basis of a tone **3** particle that never materializes: the *h*-particles and *waa2* have no close counterpart with tone **3**.

To get out of this dilemma, taking the difference between the tone **2** particles in the *g*-family and all the other ones (noticed above) into consideration, there are two possibilities. The first is that we postulate a separate non-segmental *SFP* tone **2**, which is part of the *h*-particles as well as *waa2*.³⁷ For *ge2* and *gaa2*, we would follow the derivation mentioned earlier: we take *g(e)3* to be the base particle and association with the high boundary tone (forward looking) leads to the tone **2** contour.

The second possibility would be to take the tag-like nature of the *h*-particles and *waa2* literally and treat them as belonging to a different cycle in the derivation of the sentence: they are only added to the sentence when it is completely done and the rising tone may be an inherent feature of adhortative-interrogative tags or it may be question intonation. The derivation of the *ge2* and *gaa2* would be the same as above: *g(e)3* + H. For reasons that will be made clear in section 3.1, we prefer this second possibility.

G(e)3 would be the second MMU which includes a tone; the other one is *aa4*_[+q]. The postulation of *g(e)3* as the base particle for the *g*-family does not run into problems with the derivation of the other members of the family, with the possible exception of *gaa4*, which, on the basis of our conclusion so far must be a combination of *g(e)3* and *aa4*. We will have to assume that that **4** overrides **3**, but why that would be the case, we don't know, other than pointing at the fact that **3** is anyhow more unmarked than other tones.

In sum, this section has led to the postulation of one new MMU: *g(e)3*, and to the proposal that the *h*-particles as well as *waa2* are tags, which are added to the sentence after the derivation of it has been completed.

2.5.4. Tone 5 particles

There are just three tone **5** particles:

aa5: seeking confirmation, implying speaker's high confidence

wo5: noteworthiness; reporting hearsay news

le5: seeking re-assertion, almost rhetorically

Following Li (2006a), we propose that tone **5** is another non-segmental *SFP*, this time one conveying evidentiality. As Rooryck (2001) records, markers of evidentiality represent two notions: the source of the information and the reliability of the information. Recalling the descriptions of the three tone **5** particles provided in previous sections, summarized above, we see that *wo5* represents the first category (source), as it reports hearsay information, while the

³⁷ Note that Law (1990:86ff) also proposes a separate non-segmental particle for what she calls "echo questions" (see footnote 33). Her echo particle is another H, not a rising particle, as in the text.

other two, *aa5* and *le5*, convey that the speaker is convinced that his information is reliable. Here is a summary in dissection:

aa5: *aa* (‘smooth-alert’) + 5 (‘evidentiality’)

wo5: *o* (‘noteworthiness’) + 5 (‘evidentiality’)

le5: *l/n* (‘evaluative’) + [default vowel] + 5 (‘evidentiality’)

2.6. Conclusion: the MMUS and how they combine into the SFPS

In our discussions above, we have isolated 13 MMUS, which are listed below; some are more speculative than others.³⁸

Initials: **g(e)3**: asserting relevance
 l: indicating realization of state
 m: yes/no question marker
 n/l: evaluative marker
 z: marking restriction

Rhymes: **aa**: smooth-alerting
 e: suggestive (for those who have it)
 o: marking noteworthiness
 aa4: [+q]
 [e: default]

Codas: **k**: emotion intensifier

Tones: **1**: indicating ‘forward-looking’, marking ‘hearer-orientation’
 4: marking ‘speaker-orientation’
 5: evidential marker
 [3: default]

On the basis of our MMUS, the most common Cantonese SFPS are assembled in the following manner, formally and semantically:

aa1: *aa* (‘smooth-alert’) + *1* (‘forward looking’)

aa3: *aa* (‘smooth-alert’) + [default tone]

aak3: *aa* (‘smooth-alert’) + *k* (‘emotion intensifier’) + [default tone]

aa4: (i): *aa* (‘smooth-alert’) + 4 (‘speaker orientation’)
 (ii): *aa4* (‘[+q]’)

aa5: *aa* (‘smooth-alert’) + 5 (‘evidentiality’)

wo3: *o* (‘noteworthiness’) + [default tone]

wo4: *o* (‘noteworthiness’) + 4 (‘speaker orientation’)

wo5: *o* (‘noteworthiness’) + 5 (‘evidentiality’)

³⁸ Interestingly, Law (190, 141) also ends up with a list of 13 of what she calls “underlying particles”: three tonal ones, five toneless segmental ones (*la*, *a*, *lo*, *e*, *za*) and five segmental ones with an inherent tone (*ge3*, *a5*, *wo4*, *le/ne1*, *le5*). These “underlying particles” are very different from our MMUS. Note, however, that Law’s underlying particles only underly 26 particles, not including, for instance, y/n-question particle *me1*.

ge2: *g3* ('asserting relevance') + *l* ('forward looking')

ge3: *g3* ('asserting relevance') + [default vowel]

gaa2: *g3* ('asserting relevance') + *aa* ('smooth-alert') + *l* ('forward looking')

gaa3: *g3* ('asserting relevance') + *aa* ('smooth-alert') + [default tone]

gaak3: *g3* ('asserting relevance') + *aa* ('smooth-alert') + *k* ('emotion intensifier') + [default tone]

gaa4: *g3* ('asserting relevance') + *aa4* ('[+q]')

laa1: *l* ('realization of state') + *aa* ('smooth-alert') + *l* ('forward looking')

laa3: *l* ('realization of state') + *aa* ('smooth-alert') + [default tone]

laak3: *l* ('realization of state') + *aa* ('smooth-alert') + *k* ('emotion intensifier') + [default tone]

laa4: *l* ('realization of state') + *aa4* ('[+q]')

lo1: *l* ('realization of state') + *o* ('noteworthiness') + *l* ('forward looking')

lo3: *l* ('realization of state') + *o* ('noteworthiness') + [default tone]

lok3: *l* ('realization of state') + *o* ('noteworthiness') + *k* ('emotion intensifier') + [default tone]

lo4: *l* ('realization of state') + *o* ('noteworthiness') + 4 ('speaker orientation')

n/le1: *n/l* ('evaluative') + [default vowel] + *l* ('forward looking')

le4: *n/l* ('evaluative') + [default vowel] + 4 ('speaker orientation')

le5: *n/l* ('evaluative') + [default vowel] 5 ('evidentiality')

ze1: *z* ('restriction') + [default vowel] + *l* ('forward looking')

zek1: *z* ('restriction') + [default] + *k* ('emotion intensifier') + *l* ('forward looking')

zaa3: *z* ('restriction') + *aa* ('smooth-alert') + [default tone]

zaa4: *z* ('restriction') + *aa4* ('[+q]')

me1: *m* + [default vowel] + *l* ('forward looking')

SFPS in Chart 1 not included here are the *e*-particles, the *h*-particles, *waa2*, *gwaa3* and *bo3*, for different reasons.³⁹ This does not mean that we are absolutely sure about all other details in this list. For instance, *zek1* and even *ze1* are problematic.⁴⁰ Also, it must be acknowledged that it cannot be proven that *laa3* and *zaa3* and their cognates involve the *aa*-element ('smooth-alert'). In previous sections, we noted that there is no argument against the claim that *aa* is part of these elements, but we have to admit that there is no argument in favor of it either, unless we regard the claim that these elements are never perceived as blunt as one. However, *aa* in these elements may also simply be a default vowel; for our basic approach, it does not matter, but it would complicate matters in raising questions, such as what determines

³⁹ The *e*-particles: because our informants do not have them; the *h*-particles and *waa2*: because they are tags, and because it is unclear whether they are, aside from the tonal aspect, further dissectable; *gwaa3*: because the internal structure is not clear: if it is a modificational element, as suggested above (section 2.4), it is different from the other particles; the same applies to *maa3* from footnote 23; *bo3* simply because we do not know what to do with it.

⁴⁰ An interesting feature of *ze1* and *zek1* is that their tone 1 is actually realized as 53. This is in contrast with, e.g. *laa1* and *lo1*, two clear high boundary tone cases, which are always realized as 55. *Ze1* and *zek1* resemble *tim1* and *sin1* from footnote 1, however, which have a 53 contour. Possibly significantly, *sin1* and *tim1* are indissectable into smaller MMUS. They could have an inherent tone, which could be a falling tone, or it could be 55, which gets pulled down by the low boundary tone. We leave the question open whether *ze1* and *zek1* are also indissectable, reminding ourselves of the fact that the *z*- in *ze1* is very likely the same as that in *zaa3*.

the choice for which default vowel. These and similar details aside, we think that most of the above is correct.

In the following section, we try to give the MMUS a place in the structure of the sentence.

3. The structural mapping of Cantonese SFPS

3.1. Clustering and ordering

In this section, we will relate our findings so far to the structure of the CP domain. What we have done in the previous sections is isolate MMUS and describe their semantics. In as far as these units combine into SFPS, we have also looked at the order in which the MMUS occur. In this subsection, we briefly look at the ordering within, but also beyond, individual SFPS. Subsequently, in the next subsection, we label them and relate them to more general theories on the structure of CP.

Most works on SFPS in Cantonese mention the possibility of clustering them, and Leung (1992), Law (1990), Yau (1980) and Cheung (1972) discuss the phenomenon extensively. We have isolated some 13 different MMUS. The combinatory possibilities are remarkably restricted. As Matthews and Yip (1994:245) have also observed (and the lists in Leung, Law, Yau and Cheung show the same thing), clusters of more than four or five of these elements are rare, even if we take the boundary tones into account as well. And when we realize that, if we cluster four or five, *tim1* and *ge3* are practically always among them, we see how few elements can actually co-occur. Here are two sentences from the published literature.

- (51) a. keoi⁵ lo²-zo² dai⁶-jat¹ ming⁴ tim¹ ge³ laa³ wo³ (M&Y, 345)
 3s take-PRF first place
 ‘and she got first place too, you know’
 b. keoi⁵ zing⁶-hai⁶ dak¹ ni¹-go³ lei⁵ ge³ ze³ me¹? (Cheung, 1972:194)
 3s only get this-CL daughter
 ‘does she only have this daughter?’

We need to dwell briefly on the subject of the quality of the vowel in the SFPS here. As we already noted in footnote 11, SFPS that occur in non-final position in particle clusters are generally not realized in their citation form. For instance, we find schwas instead of *e* (that is, [ɛ]) and *aa* ([a:]). Cheung (1972:171) observes that the schwa is actually not part of the vowel inventory of Cantonese and that the SFPS constitute the only domain in which we find it in the language. The *ge3* in (51a) is definitely pronounced with a schwa by our native speaker consultants.⁴¹ It is definitely not pronounced with [a:]. We asked our consultants to pronounce the *ge3* in (51a) and similar examples with [a:], but they are unable to do so. Law (1990:180) observes that *aa* is always in final position, but this is not always true: preceding *me1*, for instance, *gaa3* is certainly possible.⁴²

⁴¹ Law (1990:192 footnote 2): “The particle [ge3] is often pronounced with a schwa, ‘kə’, especially when it occurs in particle clusters.” Also her p. 179.

⁴² In the literature, we occasionally find examples with, e.g. *gaa3* in the middle of a cluster. Fung (2000:97, 98) has *gaa3 laa3*, Leung (1992:2) has *gaa3 zaa3*, but our informants find these hard to pronounce as indicated. Law (2004:71) speculates that sequences like *zaa3 laa3* and *laa3 zaa3* are ill-formed because of a ban on identical vowels in the vowel tier.

The sentence in (51b) may be seen as providing another example of a non-citation pronunciation. *Ze3* does not figure in [Chart 1](#) and was not discussed in [section 2](#), presumably because it does not exist—more precisely, because it does not exist *in isolation*. The same is true for *le3*, which, though our native consultants do not have it as an individual SFP, occasionally shows up in non-final position in their sentences (*le3 me1*). But: how do they sound? More systematic, machine-assisted research on spontaneous speech is absolutely necessary, but we notice the following. Although, to our ears, the tone is always either absent or **3** (in any case, it is never high or low or rising or falling), the vowel of *ze3* and *le3* in intermediate position in clusters can be any of the following three: [ɛ], [a] (very short) or [ə]. Notice that [a:] in these particular cases (i.e., preceding *me1*) is also an option.⁴³

It is possible to view these non-citation forms as “reduced” versions of the particles-in-isolation. However, we prefer to view them as confirmation of our approach: the forms are really just the onset with a minimal default vowel to enhance pronounceability. In fact, [Cheung \(1972:171\)](#) observes that in some cases no vowel needs to be pronounced: [kə lə] is commonly realized as [klə].

If all this is true, then the sentence in (51a) stacks four of the MMUS we isolated in the previous section (that is, not counting *tim1*):

g3 [+default vowel]—l—aa [+default tone]—o [+default tone]⁴⁴

In (51b) we see four MMUS:

g3 [+default vowel]—z [+default vowel] + m [+default vowel] + high boundary tone.

Moving from clustering to ordering, let us now determine the order of the 13 MMUS of [section 2.6](#) in two steps.⁴⁵ First (this section), we look at straightforward empirical co-occurrence facts (sentences, such as the ones in (51), the conclusions drawn in [section 2](#), empirical work done by Law and others and additional work with informants; see also [Li, 2006a](#)), after which we take more general considerations into account as well (next section). As was mentioned earlier on, the combinatory possibilities are actually quite limited. It is certainly not possible to put all elements in one sentence, so we have to look at the ordering among subsets. For ease of explication, the MMUS are presented in different subgroups in (52). (The dashes indicate that we take the elements separated by them to occupy different positions in the structure, for semantic reasons, with no relative order implied, the commas indicate that they may occupy the same position.)

- (52) i. g3—l—z
 ii. aa—o
 iii. n/l—m, aa4
 iv. -k—5—1, 4

⁴³ [Leung \(1992:128\)](#) suggests that in cases in which schwa and the vowel are in free variation, the speaker's choice for the former expresses his casualness and lack of seriousness and solemnity.

⁴⁴ Alternatively, *laa3-wo3* is a bisyllabic particle; many sources list such a particle. See [Law \(1990:19–20\)](#), why it is a bisyllabic particle rather than a cluster.

⁴⁵ Except *e*: since our informants do not have it, we cannot experiment with it.

What is clear is that $g(e)3$, if present, is always first. If z and l are present, they follow $g3$ immediately. Z and l do not co-occur, as a result of which it is impossible to determine their relative order. We take l to precede z in our linear ordering, for reasons that will become clear in the next subsection.⁴⁶

The elements in (52ii) definitely follow the ones in (52i). The relative order of aa and o is harder to assess, because they never co-occur, unless laa^3-wo^3 in (51a) is not the bisyllabic SFP, that it seems to be (see footnote 44).

With respect to the elements in (52iii), purely on the basis of co-occurrence facts, it is impossible to decide where they are exactly. The element m , a [+q] element, has to follow aa in (52ii) (as we can have $laa^3 me^1$, $zaa^3 me^1$), but whether it follows or precedes evaluative n/l cannot be made out as they never co-occur; the same applies to $aa4$. Indeed, as to evaluative n/l , it is even unclear where it is relative to aa and o of (52ii), as it does not frequently appear together with other SFPs at all.

The elements in (52iv) are possibly all suprasegmental: for the tonal MMUS this is uncontroversial, but Stephen Matthews (p.c.) suggests that the sound we represent with $-k$ is actually a glottal stop, and as such, not segmental. Aside from $zekl$, which is unusual anyway, all checked SFPs have the default tone **3**, which is consistent with this idea.⁴⁷ After all, the non-segmentals express different things. Note that we established that tone **5** and the boundary tones are not the same type of elements. Tone **5** is an MMU in its own right, while the boundary tones are reanalyzed chunks of intonation. If we take the terms “boundary tones” seriously, they should be put at the very end, past coda emotion intensifier $-k$, and this does not apply to tone **5**. The fact that only the last element in a series of elements can bear a tone other than **3** suggests that the tones come last in the ordering.

In sum, purely empirically, all we know is that $g3$ is first, that z and l follow before aa and that the tones and $-k$ come last. We also know that m follows aa . There does not seem to be an empirical base for ordering the remaining elements. We will take more general considerations into account to give all elements a place in the structure in the next section.

Before we do that (in the next subsection), we return, briefly, to the h -particles and $waa2$ and the question whether or not they, as a whole, are tags, which are added to the sentence after the derivation of it has been completed. If they are, we expect that they can be preceded by SFPs with other than tone **3** and by checked SFPs. This expectation is borne out, at least for the tones: Fāng (2003:147) has an example in which $haa2$ is preceded by $ge2$:

- (53) gam^3 $ngok^3$ gau^2 ge^2 $haa^{2?}$
 so hard do GE2 HAA2
 ‘why is it so hard, huh?’

⁴⁶ It is not clear to us why z and l should not be able to appear together (unless one puts them under the same node; cf. Tang, 1998), but they do not, in either order ($z-l$, $l-z$), as is clear from the co-occurrence lists provided in the literature and confirmed by our consultants. One of our informants very marginally accepts the following example, with z preceding l , but our other informants do not like it.

(i) ^{2/3%} $keoi^5$ giu^3 - zo^2 $loeng^5$ bui^1 zau^2 ze^3 laa^3
 3S call-PRF two cup alc.bev. ZE3 LAA3
 ‘she only ordered two glasses of wine’

⁴⁷ Particles with $-k$ are always at the fringe, according to Fung (2000:10), but Law (1990:182) does not agree: in her view, the coda simply gets suppressed when it is followed by other segmental particles.

3.2. Labeling and the structure of CP

We now take our final step: assigning the different MMUS a position in the structure, basing ourselves on the content descriptions we have given in section 2 and the linear order in as far as we have been able to establish it on empirical grounds in section 3.1. We will do so against the background of the discussion on the expanded CP domain as initiated by Hoekstra and Zwart (1994), Zwart (1993) and Rizzi (1997).⁴⁸ Naturally, we take the conclusions on adverbial and functional heads in Cinque (1999) into consideration as well. The top projection in Rizzi's (1997:297) CP is ForceP (to type the clause; Cheng, 1991) and the lowest one, connecting the CP to the IP is FinP. In between, we have a FocP and several TopPs. Our structure will most likely go beyond the ForceP, as we have elements expressing speech act information as well as some that play a primarily discoursal role.

Most of our MMUS (actually, all but *g3* and *l*) can be given a label quite straightforwardly in view of the characterizations given above, while for *g3* and *l* others the labeling is less straightforward (and for space limitations we cannot discuss them extensively). Let us look at them one by one, while at the same time putting them in the right order, mainly on the principle that like elements group together. We go inside out, or from bottom to top.

g3—This is the lowest element in the structure. In Sybesma (2004), it is proposed to occupy a head in the C-domain the specifier position of which contains a tense related operator. Let's say it is in **FinP**.

l—For *le*, *laa3*'s counterpart in Mandarin, Sybesma (1997) proposes that it performs a function similar to that of T in languages, such as Dutch and English: by explicitly establishing a link with the speech moment, it anchors the sentence to the time axis of the real world. The projection headed by *le* is called **DeikP** in Sybesma (1997), for “deictic”, to represent this reference-linking property.

z—‘only’: **FocP**.

aa—*Aa* performs a discourse function (smoothing the sentence into the discourse, thus alerting the hearer to the relevance of its content), but syntactically it is very closely related to the core of the sentence in that deletion of the element leads to the feeling that the sentence is incomplete. Let's put it in **DiscourseP**.

o—**Mood_{InformativeP}**.

n/l—**Mood_{EvaluativeP}**.

5—**Mood_{EvidentialityP}**.

m, **aa4**—for *yes/no*-questions: **ForceP**.

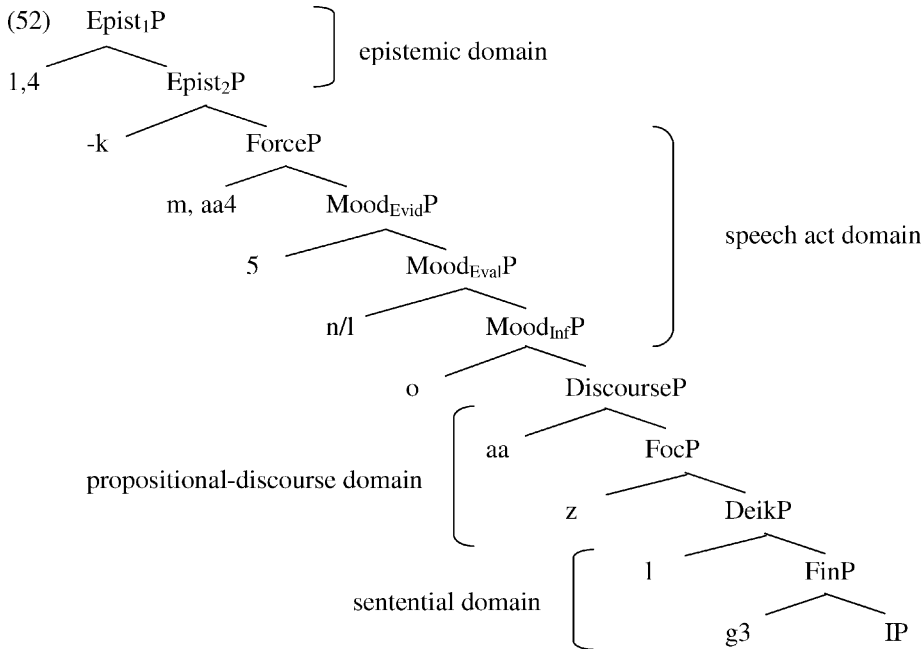
k—emotional involvement: **EpistemicP**

1, **4**—speaker/hearer orientation: **EpistemicP**

If we transform this into a tree structure, we get the following, abstracting away from headedness⁴⁹:

⁴⁸ For the evaluation of some of these ideas in the face of Chinese, see Paul (2004).

⁴⁹ The hierarchy in (52) supposedly reflects the eventual linear order. In this paper, we do not go into the question of headedness. We do assume that (most of) the particles head their own projection (see footnote 50), but whether these projections are head final or head initial is of no concern to us. In any case, hierarchically, *g3* is the lowest, the tonal elements are the highest. For discussion on matters of headedness in the context of Chinese SFPS, see Gasde and Paul (1996), Sybesma (1999), Simpson and Wu (2002), Hsieh (2006) and Hsieh and Sybesma (2006).



In the literature, specifically dealing with the syntactic aspect of the SFPs in Cantonese, the SFPs are generally divided into two groups. Law (1990) puts *ge3*, *laa3*, *laak3* and *lo3* in COMP and all others in SpecCP (except the tonal SFPs, which are attached at the end of the sentence). Law (2004:62) has two nodes, SFP2 (with *zaa3*, *tim1* and *laa3*) and SFP1 (for the rest; for her, *ge3* is not in the CP). SFP1 is ForceP and SFP2 is between Rizzi's higher TopP and FocusP. Finally, Tang (1998:42; 2002) distinguishes “inner particles” and “outer particles”, the former being “realizations of T” (viz., *laa3*, *lei4* and *zaa3*), the latter acting as “typing particles”.

The structure in (52) gives reason to distinguish more than two layers, echoing Fung's (2000) comment that SFPs operate in different domains, viz., the sentential, propositional, discourse, speech act and epistemic domains (where we take “epistemic” as relating to the speaker's attitude). Interestingly, these domains are quite neatly separated in the structure: FinP and DeikP representing the sentential domain, FocP and DiscourseP representing the propositional-discourse domain, MoodP and ForceP for the speech act domain, and, finally, the EpistemicPs for the epistemic domain.⁵⁰

This structure, though it is the culmination of the work reported on in this paper, should be taken as a point departure for comparative work on SFPs, not just within Sinitic, but also with the many other languages with rich arrays of such elements.⁵¹

⁵⁰ It goes without saying that it will always be possible to put suggestive *e* in the structure as well. The same applies to *b* or *bo*. The *h*-particles and *waa2* do not fit in, as they are mini-sentences in themselves, attached to other, longer sentences. For *gwaa3* (and non-interrogative *maa3*), see footnote 39: they head a modificational projection scoping over ForceP (in the spirit of Li, 2006a).

⁵¹ Li (2006a) shows that Mandarin may already give reason to change the order of some of the elements and projections involved, though her work on the minority language Zhuang confirms the ordering of Cantonese (Li, 2006b).

Acknowledgements

Results of the research reported here were presented at the Dix-huitièmes Journées de linguistique d'Asie Orientale, June 2004, Paris, and at the 11th SoY Colloquium, June 2005, Leiden. Our main informants have been Lisa Cheng, Leo Wong and Joanna Sio: we are very grateful for their willingness to help us over and over again with all the subtleties involved in SFPS. Thanks also to all other native speakers of Cantonese who occasionally helped us with the data. We are also obliged to Paul Law, Thomas Lee and Stephen Matthews and especially Lisa Cheng for discussing all aspects of the paper with us. The two reviewers for *Lingua* also gave valuable comments, especially the one who wrote four pages of detailed notes, comments and questions, which we have benefited from tremendously, as is apparent from the many references to them in footnotes; thanks a lot! The research reported here was conducted in the context of R.P.E. Sybesma's "Vernieuwingsimpuls"-project on syntactic variation in southern China, co-funded by the Dutch Organization for Scientific Research NWO, Universiteit Leiden (main sponsors) and the International Institute for Asian Studies, IIAS.

Appendix A

Notes to Chart 1. *The goal of these notes is to give information about which particle is mentioned in which source as an indication (be it only a weak one) of how general a particle may be assumed to be (note that none of the sources claims to be exhaustive; Fung only deals with particles with initials z-, l-, g-). The notes mention some particles that can be found in the literature but are not treated in this paper (and are, thus, not found in the chart). These notes furthermore provide additional information, which is relevant to our treatment of the particles in this article. The numbers correspond to the numbers in the chart. References should be self-evident; "M&Y" stands for "Matthews and Yip 1994".*

1. **e**: Only Law (1, 3, 4) and Yau (1, 4) include *e* as a separate particle in their inventories. Our native speaker consultants don't have it in their speech.
2. **aa**: Law and M&Y fully agree on the possible forms of *aa*: 1, 3, 4, 5. Cheung seems to only have 1, 3, 4, but this is not entirely clear. Y. Law only has 3 and 4. Yau has 1, 2, 3, 4, 5. Fāng 1, 3, 4, 5, 6, the only time we have come across a particle with a tone 6. In this paper, we only treat 1, 3, 4, 5. Yau, Law, Fāng and M&Y have *aak3*.
3. **(w)o** and **bo**: We are assuming that *wo* is *o* with $\emptyset$ -onset. Yau, M&Y, Law, Fāng and Y. Law have 3, 4, 5; Cheung has 4 and 5, and a third one, of which he says explicitly that it is higher than 33, lower than 55. Law has *bo3* and treats it as a variant of *wo3*; Fāng (67, 75) says that *wo3* is a weak form of *bo3*. Yau, Cheung, M&Y, Y. Law also have *bo3*, not as a variant of *wo3*, but as a separate particle.
4. **be**, **bo**: See note 3 above. M&Y have *be* as variant, no-one else mentions *be*.
5. **ge**: Everybody agrees (except that Y. Law does not list any *ge*). Cheung notes that 2 is under the influence of the intonation.
6. **gaa**: Yau, M&Y, Fāng and Fung have 2, 3, 4. Law only has 3, and so does Cheung, but the latter probably would have *gaa4*, just not as separate particle (see note 9). Fāng, Yau, M&Y and Law have *gaak3*.
7. **he**, **haa**, **ho**: Fāng has *he2*, *haa2*, *ho2*. M&Y have *haa3*, *ho3* (p. 343).

8. **le**: Fung has 3, 4, 5, she does not mention 1 in any form. Yau and Law have 1, 4, 5, where for Law 1 has a variant form with *n* instead of the initial *l*; Y. Law only has *le/ne1*. M&Y p. 340 list no *l/ne*, but they mention *le/ne1* and *le5* on p. 341. Cheung has *ne1*, but under *lo-la* he also mentions *le4*, *le5* (p. 174, 173, respectively). Fāng has 1, 3, 4, 5, though 1 does not occur very often, in contrast to *ne1*. None of the native speakers we consulted has *le3*.
9. **laa**: Yau, Fung, Fāng and M&Y have 1, 3, 4. Y. Law and Cheung only have 1 and 3, but Cheung treats all *Caa4*-particles as combinations with *aa4*, and does not list them separately (on p. 176 he has an example with *laa4*). Law is the only one with also 5. All except Y. Law have *laak3* (Cheung: ex. p. 189).
10. **lo**: Yau, Y. Law, M&Y, Fāng and Fung all have 1, 3, 4. Cheung only has 3, 4, Law has an additional 5. All have *lok3*.
11. **me**, **maa**: All (except Fung; see note above) have *maa3* and *me1* (M&Y on p. 341).
12. **ze**: Fung also has *ze4* (no one else does); she says (p. 69) that *ze4* is infrequently used; this is confirmed by our informants who don't have it in their speech. M&Y also have *ze3*, but no one else does and neither do our speakers know how to use it; M&Y always slash it with *ze1* ('je3/je1'). All sources have *zek1* in their inventory.
13. **zaa**: Fung and Law also have *zaa5*, which Fung calls "unproductive". Fung is the only one with *zaak1*, which she also calls "unproductive". Yau and Fāng have 3, 4; Law has 1, 3, 5; M&Y have 2, 3, 4, Y. Law only has 3, Cheung only has 4.
14. *gwaa3* is mentioned in Yau, Y. Law, Law, M&Y, Cheung, Fāng and *waa2* in Y. Law, Fāng.

References

- Alleton, V., 1981. Final particles and expressions of modality in modern Chinese. *Journal of Chinese Linguistics* 9/1, 91–114.
- Boyle, E.L., 1970. *Cantonese. Basic Course*, vol. I. Foreign Service Institute, Washington, DC.
- Caysac, G., 1952. *Introduction à l'étude de dialecte cantonnais*, second ed. Imprimerie de Nazareth, Hong Kong.
- Chan, M., 2002. Chinese: gender-related use of sentence-final particles in Cantonese. In: Hellinger, M., Bussman, H. (Eds.), *Gender Across Languages: The Linguistic Representation of Women and Men*, vol. 2. John Benjamins, Amsterdam, pp. 57–72.
- Chan, Y.K., 1949. *Everybody's Cantonese*. Man Sang printers, Hong Kong.
- Chao, Y.R., 1947. *A Cantonese Primer*. Harvard University Press, Cambridge, MA.
- Chao, Y.R., 1968. *A Grammar of Spoken Chinese*. University of California Press, Berkeley.
- Cheng, L.L.S., 1991. On the typology of wh-questions. Doctoral Dissertation. MIT.
- Cheung, K., 1986. The phonology of present-day Cantonese. Doctoral Dissertation. University of London.
- Cheung, S.H. 張洪年, 1972. *Xiānggǎng Yuèyǔ yǔfǎ de yánjiū* [Cantonese as spoken in Hong Kong]. The Chinese University of Hong Kong, Hong Kong.
- Chu, C.C., 2002. Relevance theory, discourse markers and the Mandarin utterance-final particle *a/ya*. *Journal of the Chinese Language Teachers Association* 37/1, 1–42.
- Cinque, G., 1999. *Adverbs and Functional Heads. A Cross-Linguistic Approach*. OUP USA, New York.
- Fabb, N., 1992. Reduplication and object movement in Ewe and Fon. *Journal of African languages and Linguistics* 13/1, 1–39.
- Fāng, Xiǎoyàn 方小燕, 2003. *Guǎngzhōu fāngyán jù mò yǔ qì zhù cí* [Sentence final modal particles in the Guangzhou dialect]. Jinán University Press, Guangzhōu.
- Fung, R.S.-Y., 2000. Final particles in standard Cantonese: semantic extension and pragmatic inference. Doctoral Dissertation. Ohio State University.
- Gasde, H.-D., Paul, W., 1996. Functional categories, topic-prominence and complex sentences in Mandarin Chinese. *Linguistics* 24, 263–294.
- Herman, R., 2000. Phonetic markers of global discourse structure in English. *Journal of Phonetics* 28, 466–493.

- Hoekstra, E., Zwart, C.J.W., 1994. De structuur van de CP. Functionele projecties voor topics en vraagwoorden in het Nederlands. *Spektator* 23, 191–212.
- Hsieh, F.F., 2006. Atoms and Particles. Ms. MIT.
- Hsieh, F.F., Sybesma, R., 2006. About Chinese Sentence Final Particles and why CP Moves. Ms. Leiden University and MIT.
- Kirsner, R., Van Heuven, V., 1996. Boundary tones and the semantics of the Dutch final particles *hè*, *hoor*, *zeg* and *joh*. In: Cremers, C., Den Dikken, M. (Eds.), *Linguistics in the Netherlands 1996*. John Benjamins, Amsterdam, pp. 133–146.
- Kirsner, R., Van Heuven, V., Van Bezooijen, R., 1994. Interaction of particle and prosody in the interpretation of factual Dutch sentences. In: Bok-Bennema, R., Cremers, C. (Eds.), *Linguistics in the Netherlands 1994*. John Benjamins, Amsterdam, pp. 107–118.
- Kwok, H., 1984. Sentence Particles in Cantonese. Center of Asian Studies, Hong Kong.
- Law, S., 1990. The syntax and phonology of Cantonese sentence-final particles. Doctoral Dissertation. Boston University.
- Law, Y.K.A., 2004. Sentence-final particles in Cantonese. Doctoral Dissertation. University College London.
- Lee, Th.H.T., Yiu, C., 1998. Final “de” and “ge3”—a nominalization analysis for cleft sentences in Mandarin and Cantonese. Ms. based on a paper presented at the 1998 Linguistics Society of Hong Kong Annual Research Forum.
- Leung, Chung-sum, 梁仲森, 1992. *Xiānggǎng Yuèyǔ yǔzhùcí de yánjiū* [A study of the utterance particles in Cantonese as spoken in Hong Kong]. M. Phil. Thesis. Hong Kong Polytechnic University.
- Leung, W., 2005. Two historical sources of the final particle *wo* in Cantonese. Paper presented at the 13th annual meeting of the International Association of Chinese Linguistics. Leiden, The Netherlands.
- Li, B., 2006a. Chinese final particles and the syntax of the periphery. Doctoral Dissertation. Leiden.
- Li, B., 2006b. Final particles in Zhuang. Paper presented at the 14th SoY Linguistics Colloquium. Leiden.
- Lucas, A., Xie, H., 1994. There is practically one universal Chinese grammar”: A propos de *xian1* 先 en mandarin et *sin1* en cantonais. *Cahiers de linguistique Asie Orientale* 23, 189–206.
- Luke, K.K., 1990. Utterance Particles in Cantonese Conversation. John Benjamins, Amsterdam.
- Mài, Yún 麦耘, 1993. *Guǎngzhōuhuà “sin1” zài fēnxī* [Analysing Cantonese *sin1* once more.]. In: Zhèng, D., Zhōu, X. (Eds.), *Guǎngzhōuhuà yánjiū yǔ jiàoxué* [Research and teaching of Cantonese]. Zhōngsān Dàxué chūbǎnshè, Guǎngzhōu, pp. 63–73.
- Matthews, S., Yip, V., 1994. *Cantonese. A Comprehensive Grammar*. Routledge, London.
- Paul, W., 2004. Low IP area and left periphery in Mandarin Chinese. *Recherches linguistiques de Vincennes*, vol. 33.
- Pierrehumbert, J., 1980. The phonology and phonetics of English intonation. Doctoral Dissertation. MIT.
- Pierrehumbert, J., Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P., Morgan, J., Pollack, M. (Eds.), *Intentions in Communication*. MIT Press, Cambridge, MA, pp. 271–312.
- Rizzi, L., 1997. The fine structure of the left periphery. In: Haegeman, L. (Ed.), *Elements of Grammar*. Kluwer, Dordrecht, pp. 281–337.
- Rooryck, J., 2001. Evidentiality, part I. *Glott International* 5/4, 125–133 (Part II, *Glott International* 5/5, 161–168).
- Schubiger, M., 1965. English intonation and German modal particles: a comparative study. *Phonetica* 12, 65–84.
- Simpson, A., Wu, Z., 2002. IP-raising, tone-sandhi and the creation of S-final particles: evidence for cyclic spell-out. *Journal of East-Asian Linguistics* 11/1, 67–99.
- Steedman, M., 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31/4, 649–689.
- Sybesma, R., 1997. The deictic function of TP and sentence-le in Mandarin Chinese. Paper presented at NACCL9. University of Victoria.
- Sybesma, R., 1999. Overt wh-movement in Chinese and the structure of CP. In: Wang, H.S., Tsao, F.F., Lien, C.F. (Eds.), *Selected papers from the fifth International Conference on Chinese Linguistics*. The Crane Publishing Co., Taipei, (paper presented in 1996), pp. 279–299.
- Sybesma, R., 2004. Exploring Cantonese tense. In: Cornips, L., Doetjes, J. (Eds.), *Linguistics in the Netherlands 2004*. John Benjamins, Amsterdam, pp. 169–180.
- Tang, S.-W., 1998. Parametrization of features in syntax. Doctoral Dissertation. University of California, Irvine.
- Tang, S.-W. 鄧思穎, 2002. *Yuèyǔ jù mò zhùcí de bùduìchèn fēnbù*. [The a-symmetric distribution of sentence final particles in Cantonese]. *Zhōngguó Yǔwén Yánjiū* 2, 75–84.
- Tang, S.-W. 鄧思穎, 2006. *Yuèyǔ yíwèn jù mò “sin1” zì de jǔfǎ tèdiǎn*. [The syntactic properties of sentence final “sin1” in interrogative sentences in Cantonese]. *Zhōngguó Yǔwén* 2006/3, 225–232.
- Wong, L., Sybesma, R., Tsai, W.T.D., 2005. The Prosodic, Syntactic and Semantic Properties of Cantonese *tim1*. Ms. Leiden and Tsinghua.
- Wu, Guo, 2005. The discourse function of the Chinese particle *ne* in statements. *Journal of the Chinese Language Teachers Association* 40/1, 47–82.

- Yau, S.C., 1980. Sentential connotations in Cantonese. *Fāngyán* 1980/1, 35–51.
- Yip, M., 2002. *Tone*. Cambridge University Press, Cambridge.
- Yip, V., Matthews, S., 2000. *Basic Cantonese. A Grammar and Work Book*. Routledge, London.
- Yiu, C., 2001. Cantonese final particles “lei”, “zyu” and “laa”: an aspectual study. M. Phil. Thesis. Hong Kong University of Science and Technology.
- Zwart, C.J.W., 1993. *Dutch syntax. A minimalist approach*. Doctoral Dissertation. Groningen.