IP-raising, tone sandhi and the creation of particles: evidence for PF movement/cyclic Spell-Out

Andrew Simpson (SOAS) & Xiu-Zhi Zoe Wu (USC)

This paper shows how information provided by phonological processes such as tone sandhi can provide key insights into syntactic strctures and movement processes. We focus on the Taiwanese element *kong* (Mandarin *jiang*) 'say' which is grammaticalizing as a complementizer-type particle in an unexpected sentence-final position. Evidence from tone sandhi phenomena indicates that this results from an operation of IP-raising in which the clausal complement of *kong* is raised to its left significantly after the application of tone sandhi rules, hence arguably at PF. The active grammaticalization patterning offers both a clear insight into the creation of clause/sentence-final particles in head-initial languages and also provides strong evidence for the occurrence of large-scale movement at PF, or the alternative idea of 'cyclic Spell-Out'. It is also seen that a derivational rather than a purely representational model of grammar is required to accommodate the patterns found.

0. Introduction

Phonologically reduced grammatical particles are elements which are commonly found in clause- and sentence-final position across a wide range of language types, not just in head-final SOV languages but also in head-initial SVO languages, as illustrated in (1-6) below, Japanese and Burmese being examples of basically SOV-type languages, and Chinese, Thai and English being neutrally SVO:

(1) Taroo-ga kuruma-o kaimashita yo — Japanese
Taroo-NOM car-ACC bought PRT
'Taroo bought a car.'

(2) U-Win-Win-ga beh thwaa –th –leh? — Burmese
U-Win-Win-NOM where go NON-FUT Q-PRT
'Where is U-Win-Win going?'

(3) wo bu qu le — Chinese
I NEG go ASP-PRT
'I'm not going any more.'

(4) ni xihuan zhei-ge ma? — Chinese
you like this-CL Q-PRT
'Do you like this one?'

(5) khun choop lem-nai le? — Thai
you like volume-which Q-PRT
'Which one (book) do you prefer?'

(6) You're going to London eh?

If particles of this type are taken to instantiate functional heads such as Mood (1), Q/C (2,4-6) and Aspect (4) due to their functional roles and interpretations, the clause/S-final positioning of these elements requires explanation in SVO languages but is not unexpected in SOV type languages. In SOV languages traditionally analyzed as being

head-final, it is anticipated that high clausal functional heads such as Mood, Q/C etc should indeed be located in final position in the clause. In head-initial SVO languages however, the occurrence of sentence-final particles is problematic as the high-clausal functional heads which they can be assumed to instantiate are expected to be located in clause-initial positions in line with the general head-initial character of SVO languages. In Chinese the occurrence of sentence-final question particles such as *ma* (example (4)) has led to claims that CP (headed by *ma*) is a head-final projection and that Chinese therefore has to be accepted as being a mixed head-initial and head-final language. Such a view of SVO languages as being exceptionally head-final in certain high clausal projections is rather unsatisfactory and one might wonder if the sentence-final positioning of mood and question particles might be determined by other factors. This paper examines the ongoing creation of a new sentence-final particle in Taiwanese, the element *kong* illustrated in example (7), and shows how the patterning found offers potentially revealing insights into the mechanisms which may underlie the development of sentence-final particles:

(7) A-hui liau-chun A-sin si tai-pak lang **kong**
Ahui thought Asin is Taipei person PRT
'Ahui thought that Asin is from Taipei.'

In contrast with older S-final particles elsewhere where the origin and creation of such elements is often unknown, the source of Taiwanese *kong* is still very clear and the syntax underlying its development is also largely transparent due to interesting patterns of tone sandhi change which occur when *kong* is used. The combination of information available with *kong* is shown to strongly suggest that its sentence-final position in fact results from an operation of clausal raising and that the apparently head-final surface form found with this S-final particle actually hides a fully regular head-initial structure. A study of Taiwanese *kong* therefore indicates that sentence-final particles may not necessarily signal head-final projections and that one should consequently be wary of interpreting the surface position of similar particles elsewhere as being a reliable indication of head-final structures.

The use of tone sandhi as a means to understanding underlying syntactic structure and movement processes also results in the significant conclusion that IP-raising in *kong* structures takes place derivationally *after* the application of the tone sandhi rules, and as these are clearly phonological rules, this seems to indicate that the IP-movement either occurs after Spell-Out at PF or that a cyclic Spell-Out model as proposed in Chomsky (1998) should be adopted. The paradigm furthermore provides a strong argument for a derivational model of grammar and would not seem open to any fully representational account.

The structure of the paper is as follows. In order to develop the analysis of *kong* via tone sandhi phenomena, section 1 will first describe general properties of Taiwanese and the relation of tone sandhi to syntactic structure. Section 2 then begins to examine *kong* and situates it relative to a common process of grammaticalization in which verbs of saying develop into complementizer elements. Pointing to the unusual position of *kong* once its origin is considered, it is then shown how tone sandhi patterns provide strong evidence for an analysis of IP-raising and the conclusion that *kong* in fact selects its complement in the canonical direction of selection, to the right. Continuing with a discussion of why particles frequently occur in sentence-final position when they undergo grammaticalization, section 2 also includes an examination of how the *kong* paradigm constitutes strong evidence for the occurrence

of post-syntactic movement at PF, or alternatively Chomsky's more recent (1998) idea of 'cyclic Spell-Out.' Finally the paper is closed in Appendix I with a related brief extension on how tone sandhi phenomena also offer good support for IP-movement in relative clauses and a Vergnaud/Kaynean analysis of relativization in Taiwanese/Chinese.

1.0 Taiwanese and patterns of tone sandhi

Taiwanese is a variety of SVO Chinese which is very much similar to Mandarin in its basic word order and shows regular head-initial patterns both in the lexical domain and in the functional domain; so, for example, verbs and prepositions consistently take their objects to the right as in (8) and (9), and auxiliaries which can be taken to occur in I^0 similarly have rightward VP complements, as seen in (10):

(8) V-O
be chhe
buy books
'buy books'

(9) P-NP
tui A-sin
to Asin
'to Asin'

(10) Aux/I^0-VP
e lai
will come
'will come'

Concerning the relation of C^0 to IP, the occurrence of sentence-final question particles has been suggested to indicate that C^0 follows IP and that CP is therefore a head-final projection. However, elsewhere non-particle elements equivalent to instantiations of C^0 in other languages (such as *na-si* 'if') occur before the IP, suggesting that CP is also regularly head-initial too:

(11) Comp-IP
[$_{CP}$ na-si [$_{IP}$ A-sin m lai]]…
if Asin neg come
'If Asin is not coming…'

With regard to its phonology and tonal system, Taiwanese is described as having eight tones as listed below:[1]

(12) **Tones in Taiwanese**

tone number	*tonal value/shape*
1	high-level 5-5
2	high-falling 5-1
3	low-falling 2-1
4	low-entering tone (a syllable with a final stop)

5	contour-tone 2-1-4
6	high-falling 5-1 (as tone 2, see fn 1)
7	mid-level 3-3
8	high-entering

In addition to these distinctive tones there are also syllables which do not carry any tone, this sometimes being referred to as "neutral tone" NT. In the phenomenon of *tone sandhi*, the lexically-listed "citation" tone of a syllable undergoes modification according to fully regular rules when preceding some other tone-bearing syllable in the same tone sandhi domain. For example, if a syllable with tone 3 precedes another tone-carrying syllable in the same tone sandhi domain, the tone 3 will change into a tone 2, as illustrated in (13):

(13) **khi3** pak8kiang1 → **khi2** pak8kiang1
go Beijing
'go to Beijing'

Table (14) below shows how the full range of these modifications are made. Note that the changes in tone are not triggered or conditioned by the particular type of tone that the following syllable carries so that a syllable with tone 1 will change its tone to tone 7 no matter whether the following syllable has tone 1, 2 or 3 etc; the essential requirement for tone sandhi to apply is that the following syllable just have some type of lexical tone rather than just 'neutral tone'.[2]

(14) **tone sandhi change in Taiwanese**

(tone.. changes to tone..)		
1	→	7
2	→	1
3	→	2
4	→	8 when the syllable ends in p/t/k;
	→	2 when the syllable ends in a glottal stop
5	→	7 (southern Taiwan);
	→	3 (northern Taiwan)
6	→	1
7	→	3
8	→	4 when the syllable ends in p/t/k;
	→	3 when the syllable ends in a glottal stop

As mentioned just above, tone sandhi may not occur in a syllable if it precedes a syllable which has only neutral tone/no tone. Consequently *zau* in example (15) may not change its tone-2 when occurring before the toneless element *lai*:

(15) **zau2** a-NT → **zau2** a-NT
run already
'already ran'

Similarly, a syllable may not undergo tone sandhi if it occurs sentence-finally. This is due to the fact that tone sandhi is restricted to apply within certain domains and is blocked where a substantial intonational break may occur (as indeed between sentences). In (16) below, the citation tone-2 of sentence-final *ho* may not be

converted into tone-1 even though followed by a syllable (*goan-2*) which does carry tone because this latter syllable occurs in a separate sentence. Note that from this point on for simplicity of representation we will indicate tone sandhi change by means of a simple bolded dot following the relevant syllable. Thus if a syllable is followed by a bolded dot, this indicates that it undergoes tone sandhi change, and if a dot is absent, no tone sandhi change is possible. In (16) sentence-final *ho* is therefore not followed by a dot as no tone sandhi change can occur in sentence-final position:

(16) goa chin• **ho**. goan• lau•pe ma• chin• ho.
I very good my old-father also very fine
'I am fine, and my father is also fine.'

Sentence-internally there would also seem to be other tone sandhi/TS domains relevant for the operation of tonal change, and broadly-speaking every syllable in such a domain will change its tone unless it is the last tone-bearing syllable. Significantly, tone sandhi change in Taiwanese appears to relate to and reveal the underlying syntactic structure in a way which is not found in tone sandhi phenomena in Mandarin, Shanghainese and certain other varieties of Chinese. For present purposes it is important to point out the following three significant generalizations:

(17) **Generalization A: a head and its complement occur in the same TS domain**
The presence of an overt complement consistently triggers tone sandhi change on the selecting head, indicating that a head and its complement are in a single TS domain:

(18) V-NP_{object}
be• [lng•-pun• chhe]
buy two-Cl books
'buy two books'

(19) P-NP
tui• [goan• lau•pe]
to my father
'to my father'

(20) Aux/I-VP
e• lai
will come
'will come'

(21) Comp/C-IP
na•**si•** [A•sin m• lai]…
if Asin neg come
'If Asin is not coming…'

(22) **Generalization B: a head and its Specifier do not occur in the same TS domain**
It is found that a head does not trigger tone sandhi change on the final syllable of its Specifier. Consequently the Specifier of a head constitutes an

independent TS domain. In (23) below, the final syllable of the subject does not change its tone, despite being followed by the tone-bearing head *u* 'have':

(23) [goan• lau•**pe**] u• lng• chhing• kho
my father have two thousand dollar
'My father has two thousand dollars.'

In addition to (23) above with the final syllable of a subject in SpecIP failing to undergo tone sandhi, further examples of Specifiers being isolated TS domains are given in (24) and (25) below, where the DP *tai-oan-oe* 'Taiwanese' occurs as either a moved or base-generated topic relating to the object position. In such a Specifier position, its final syllable *oe* does not undergo any tone sandhi change:

(24) [goan• lau•pe] [tai•oan•**oe**] be• hiao• kong
I old-father Taiwanese not know speak.
'Taiwanese, my father can't speak.'

(25) [tai•oan•**oe**] [goan• lau•pe] be• hiao• kong
Taiwanese I old-father not know speak.
'Taiwanese, my father can't speak.'

(26) **Generalization C: adjuncts are self-contained TS domains**
The final syllable of an adjunct does not undergo tone sandhi even when followed by other tone-bearing syllables. This is illustrated below with the case of a CP adjunct. No tonal change in its final syllable is possible:

(27) [na•si• A•sin m• **khi**], A•hui ma• be• khi
if Asin neg go Ahui also neg go
'If Asin is not going, Ahui will also not go.'

2.0 The syntax of Taiwanese kong

2.1 kong and the grammaticalization of verbs of saying as complementizers

We are now in a position to begin examining the nature of Taiwanese *kong*. The origin of this S-final particle element is still very clear and *kong* has (most arguably) grammaticalized in some way from the fully homophonous general verb of saying *kong* (equivalent to Mandarin *jiang* 'to say, tell'). Elsewhere the element *kong* still occurs as a regular independent verb which can furthermore carry aspectual suffixes such as *–koe* (perfect/experiential aspect, Mandarin *–guo*), as in (28) and (29):

(28) A•-hui kong• A•sin m• lai.
A-hui say A-sin NEG come
'A-hui said A-sin is not coming.'

(29) A•hui u• kong•koe• hit•ku• oe.
A-hui have say-ASP that-CL words
'A-hui had said that sentence before.'

Cross-linguistically it is a well-attested process that such general verbs of communication typically equivalent to English 'to say' may undergo grammaticalization as complementizers when they occur after other more specific verbs of communication or cognitive state such as 'yell', 'whisper', 'think' or 'believe'. Frequently this occurs when a language has serial verb constructions which allow for a sequence of two verbs of communication (one more specific, the second less specific) to become reanalyzed as a sequence of verb + complementizer, schematically as in (30):

(30) Verb1 Verb2 → Verb(1) Complementizer
shout say → shout that

What is of particular interest and relevance here is the position of the verb 'to say' when it becomes grammaticalized as a complementizer. The cross-linguistically common pattern is for the grammaticalized complementizer to occur in the same position that the earlier fully verbal form occurred in. In the many head-initial SVO languages of West Africa and S.E.Asia which show this type of grammaticalization this means that the new complementizer will occur *preceding* its clausal complement. In Thai, for example, the morpheme *waa* is currently both a verb meaning 'to say' as seen in (31) and also grammaticalized as a complementizer preceding its IP complement as in (32). The fact that *waa* may co-occur with verbs of cognition such as *khit* 'think' in (32) no longer with its literal meaning 'to say' is evidence that *waa* has indeed grammaticalized as a complementizer in such positions and is no longer just a verb-in-series. Such an assumption is further supported by the observation that *waa* may now also occur after *nouns* as in (33):

(31) kae waa arai?
you say what
'What did you say?'

(32) khaw book/khit waa Daeng suay
he say/think that Daeng be-pretty
'He says/thinks that Daeng is pretty.'

(33) kham-phaasii waa 'tham bun dai bun'
proverb that do good get good
'the proverb (that) 'If you do good, you will receive goodness.''

In West African Ewe (Heine & Reh 1984:252) it is found that the verb *be* 'say' grammaticalized as a complementizer no longer occurs with the tense-aspect markings or pronoun prefixes which would otherwise be normal for real verbs in serial verb constructions, again indicating rather clearly that a category change from verb to complementizer has taken place. Similarly in Twi (Lord 1993:176) the verb *se* 'say' occurring as a complementizer also now no longer takes verbal affixes such as negation concord which would otherwise occur with verbs-in-series, confirming as with Ewe and Thai that a category change from verb to complementizer has taken place.

In Mandarin Chinese and Cantonese, Hwang (1998) argues that the same type of grammaticalization is taking place, and as Mandarin (34) shows, the verb *shuo* 'to

say' occurs following a verb of cognition. As in Thai (32), this element in (34) no longer has its original verbal meaning of 'saying' but instead appears to be functioning as a general embedding complementizer element:

(34) Zhangsan xiang shuo Lisi bu lai le
Zhangsan think that Lisi NEG come ASP
'Zhangsan thinks Lisi is no longer coming.'

Examples such as (34) are important, as they show that where a complementizer/C^0 is developing in Chinese, it occurs in a pre-IP position and hence conforms with the otherwise head-initial pattern in Chinese preceding its complement. Where suggestions have been made that CP is a head-final projection in Chinese, this has been based on the occurrence of sentence-final question particles and the assumption that Chinese has no other regular instantiations of C^0 equivalent to English 'that'. Here however one finds that a fairly simple equivalent to English 'that' is indeed beginning to occur and significantly it identifies CP as being head-initial and quite regular in its directionality.

A similar pattern also occurs in Taiwanese, and one finds that the verb *kong* occurs following other verbs of communication and verbs of cognition as in (35):

(35) goa• siong• kong• i m• lai
I think KONG he NEG come
'I though that he was not coming.'

Again, as with Thai *waa* and Mandarin *shuo*, the fact that *kong* occurs without its normal verbal meaning of 'saying' with verbs of cognition strongly suggests that it has grammaticalized away from its original verbal source. This is confirmed by the fact that *kong* in such a position cannot occur with any aspectual suffixes, suggesting that *kong* in these instances has indeed undergone a category change from verb to some other non-verbal category and now occurs as a complementizer:

(36) *goa• siong• kong•-**koe** i m• lai
I think KONG ASP he NEG come

This position preceding the embedded IP in (36) is precisely where one would expect to find *kong* occurring as a grammaticalized complementizer, and *kong* as a new C^0 here seems to be fully parallel to Mandarin *shuo*, Thai *waa* and equivalents in other SVO serializing languages. However, in addition to forms such as (36), another arguably more interesting pattern is found with *kong*, as briefly noted in the introduction. For no immediately clear reason, the same element *kong* also seems to occur as a complementizer in clause-*final* position, hence *following* its clausal complement, as in (37) and (7) repeated below:

(7) A•hui liau•chun A•sin si• tai•pak• lang kong•
Ahui thought Asin is Taipei person KONG
'A-hui thought that A-sin is from Taipei.'

(37) goa siong• i m• lai kong•
I think he NEG come KONG
'I think he is not coming.'

As Taiwanese like other varieties of Chinese elsewhere shows evidence of being head-initial (examples (8)-(11)), and *kong* otherwise does occur as a genuine grammaticalized complementizer in clause-initial pre-IP position (as in (35)), this apparent clause-final V-to-C grammaticalization of *kong* is rather strange and seems to go against the general headedness specification of the language. It clearly also does not correspond to any serial verb position from which *kong* could have naturally grammaticalized as a complementizer.

In order to explain the puzzle of clause/sentence-final *kong*, we will shortly suggest that the canonical position of the grammaticalized comeplementizer *kong* is indeed *preceding* its IP complement as in (35) and show that there is certain rather clear evidence from tone sandhi patterns indicating that the unexpected exceptional order in (7) and (37) is one which is actually *derived*, via a process of IP-raising to SpecCP.

2.2 tone sandhi patterns with *kong*

Considering the ordering of C^0 and IP found in (35), one finds quite regular expected patterns of tone sandhi. The C^0 grammaticalized verb *kong* undergoes tone sandhi in its position preceding the IP complement, this caused by a regular head-complement relation, and the final element in the embedded IP *lai* does not undergo tone sandhi. This is fully anticipated as sentence-final elements do not undergo tone sandhi changes (as seen above in (16) and other examples).

Turning to (7) and (37), with the unusual ordering of IP-C^0 in the embedded clause, one now finds two quite unanticipated tone sandhi patterns. The first of these is that the IP-final elmement *lang* in (7) and *lai* in (37) do *not* undergo tone sandhi. If one assumes that the IP is the leftward complement of *kong* in a final C^0 position, this should mean that the IP and the C^0 are in the same tone sandhi domain and it is expected that the head-complement relation should result in tone sandhi occurring between the C^0 and the element left-adjacent to it in this tone sandhi domain, i.e. the final syllable in the IP, yet this doesn't happen.

The second extraordinary tone sandhi patterning in forms such as (7) and (37) is that the *sentence-final* element *kong* does in fact undergo a tone sandhi change. This is very much unexpected as no other elements in sentence-final position are known to undergo tone sandhi, the sentence being a self-contained tone sandhi domain as noted earlier when discussing example (16). Furthermore, the grammaticalization of *kong* might be expected to result in it either maintaining its citation tone2 or simply reverting to a neutral tone/absence of tone as is commonly found in other cases of grammaticalization (e.g. Mandarin, *de, le* and *–zhe*, and various functional elements in Taiwanese). However, instead of this, *kong* undergoes a fully regular tone sandhi change in sentence-final position. Examples such as (7) and (37) need also not be followed by any other sentence for tone sandhi to occur on *kong* and so it would appear that there is nothing following *kong* which could trigger its tonal change.

Both such patterns can now be argued to have a rather simple explanation. Critically, both of the odd patterns observed in (7) and (37) are exactly parallel to those occurring in "regular" examples such as (35) and (38) below where the complementizer *kong* occurs preceding its complement IP:

(38) A•hui liau•chun kong• A•sin si• tai•pak• lang
Ahui thought KONG Asin is Taipei person
'A-hui thought that A-sin is from Taipei.'

In (35) and (38), as just noted, the final syllable in the lower IPs, *lai* and *lang* respectively, do not undergo tone sandhi (as expected), and *kong* preceding its IP complement does undergo tone sandhi (again as expected). Comparing (35)/(38) and (7)/(37) it can therefore be seen that precisely the same tone sandhi patterns occur both when *kong* precedes its complement IP in a regular head-initial C^0 position and when *kong* occurs finally in a rather unusual position:

(39) a. *kong* – IP expected order, expected tone sandhi changes
(i) final syllable in IP does not undergo tone sandhi
(ii) *kong* does undergo tone sandhi

b. IP – *kong* unexpected order, unexpected tone sandhi changes
(i) final syllable in IP does not undergo tone sandhi
(ii) *kong* does undergo tone sandhi

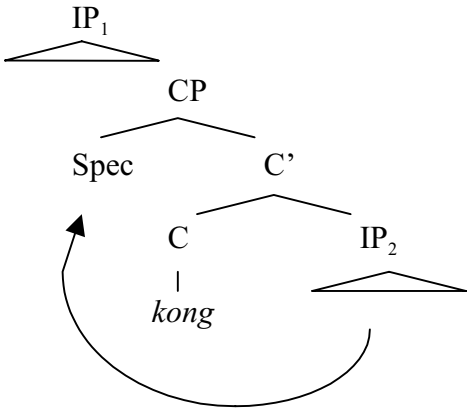
The simple conclusion from such a comparison is that *kong* in its unusual sentence-final position is behaving for tone sandhi purposes exactly as if it occurred in a regular pre-IP position. Syntactically in order to capture this striking parallelism it can now be suggested that IP-*kong* forms such as (7)/(37) are actually the result of an IP-raising operation applying to underlying fully regular *kong*-IP forms *before* they are converted into IP-*kong* sequences. Such a pair of assumptions allows for a very straightforward explanation of the otherwise unanticipated tone sandhi facts, as follows. Prior to IP raising, the final element *lang/lai* in the embedded IP in (7/37) will occur in sentence-final position and *kong* will occur as a regular C^0 preceding an IP-complement. If the tone sandhi rules are applied at this derivational point, the result will be (a) that the final syllable in the IP *lang/lai* does not undergo any tone sandhi change, being in sentence-final position, and (b) that *kong* does undergo tone sandhi, being in a head-position preceding its IP complement.

Observing how the assumption of IP-raising will explain both the unusual tone sandhi patterns in *kong*-final sentences and the odd sentence-final position occupied by *kong*, and further noting that *kong* occurs as a regular CP-initial complementizer in embedded clauses such as (35)/(38), it might naturally be assumed that the hypothesized IP-raising operation applies in the embedded clause in (7)/(37) converting a string such as (35) into (37). Such a derivation is schematically represented in (40):

(35) goa• siong• kong• i m• lai
I think KONG he NEG come
'I though that he was not coming.'

(37) goa siong• $[_{CP}\ [_{IP2}$ i m• lai$]_i$ kong• t_i]

(40)



However, there is actually good reason to believe that this is not exactly how IP_2 and *kong* become re-positioned relative to each other. Although *kong* might seem to bear all the hallmarks of an embedded complementizer grammaticalized from a general verb of communication as in many other languages, further data reveals that *kong* in fact syntactically embeds not just a lower clause but *the entire sentence* in which it occurs sentence-finally.

The evidence that this is so comes in two forms. First of all, in sentences such as (7) and (37) it is possible to have not only a *kong* in sentence-final position, but also a second *kong* in a regular grammaticalized embedded Comp position preceding the embedded IP, as in (41) and (42):

(41) A•hui liau•chun **kong•** A•sin si• tai•pak• lang **kong•**
Ahui thought KONG Asin is Taipei person KONG
'A-hui thought that A-sin is from Taipei.'

(42) goa siong• **kong•** i m• lai **kong•**
I think KONG he NEG come KONG
'I think he is not coming.'

This indicates that the sentence-final *kong* does not originate in an embedded C^0 position, as this position can clearly be filled by a second distinct *kong*. Consequently, the natural assumption to make is that sentence-final *kong* is actually in the matrix C^0 in (7), (37), (41) and (42) and that the entire IP_1 (i.e. the whole sentence consisting of both clauses IP_1 and IP_2) is raised to the Specifier projected by this matrix C^0. Clear confirmation that this is true comes from the fact that it is possible to have a sentence-final *kong* in *single-clause* sentences, as in (43)-(45). This indicates that *kong* here can only possibly be occurring in a matrix Comp as there obviously is no embedded C^0 in such clausal structures:

(43) A•sin m• lai kong•
A-sin NEG come KONG
'A-sin's not coming.'

(44) goa chahng bo• khi• tai•pak• kong•
I yesterday NEG go Taipei KONG
'Yesterday I didn't go to Taipei.'

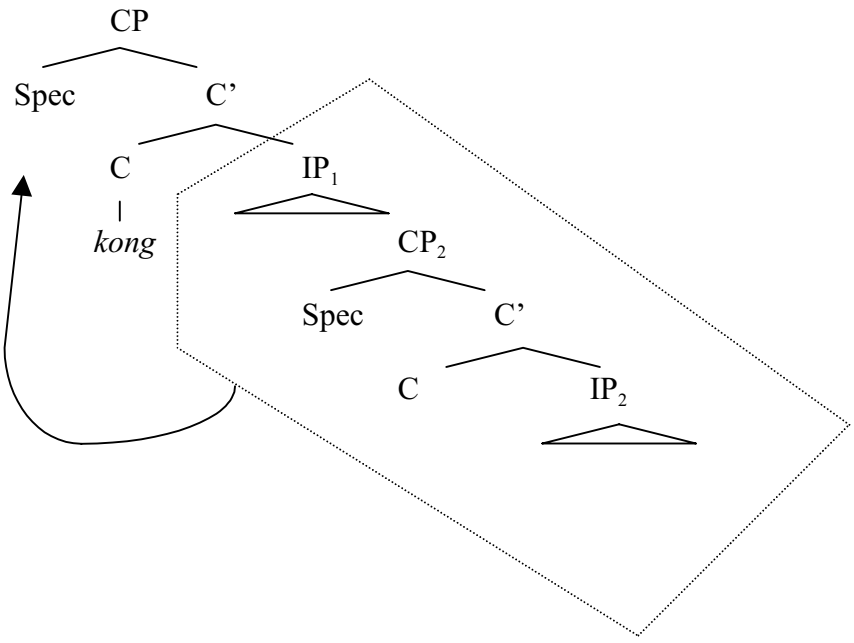
(45) goan• lau•pe si• tai•pak• lang kong•
I father be Taipei person KONG
'My father is from Taipei.'

Furthermore, if one compares (37) which has a single sentence-final *kong* with (35) where *kong* occurs preceding the embedded IP, one finds that the interpretation of the two structures is not fully equivalent. Use of *kong* in (35) essentially adds nothing extra to the meaning of the sentence, much in the way that the optional addition of the English complementizer 'that' adds no extra semantic content when it precedes an embedded clause. Use of sentence-final *kong* however does add some clear extra meaning to the sentences it accompanies in a way which is generally typical of a number of sentence-final particle elements. Though the precise semantic content of sentence-final *kong* is somewhat difficult to pin down and describe, there is most certainly an extra emphatic quality added to a sentence when it is used, and *kong* may be loosely glossed as 'I'm telling you X!' (where X = the content of the sentence). Considering another example, (46) below, the use of S-final *kong* seems to give rise to the interpretation that: "the person referred to has written in his letter saying he is coming, so why do you, the person listening to me (the speaker) think that he will not come?":

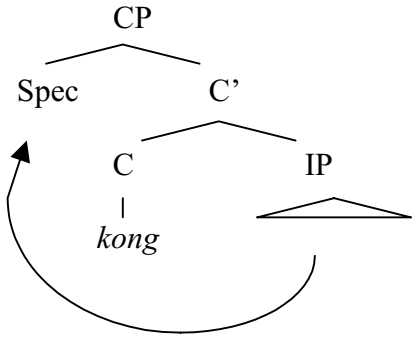
(46) i• e• phoe sia• kong• bin•a•chai beh• lai kong•
he GEN letter write KONG• tomorrow want come KONG
'He wrote that he will come tomorrow.'

As a result of the above observations, it can be suggested that *kong* is indeed a grammaticized C^0 element, but one which critically occurs in matrix clause positions. Quite possibly this restriction results from *kong* being licensed by a speaker-centred propositional attitude (the special emphasis of *kong*) which can only be encoded in matrix clauses where the speaker is the clear source of the information. Assuming *kong* then to be in the matrix C^0, the surface forms found in (37) and (43) can actually be argued to have the underlying derivation and structure indicated in (47) and (48):

(47)



(48)



(47) and (48) will then allow for the basic explanation of the tone sandhi patterns already given. Considering (47) which represents the examples considered in (7) and (37), what needed to be accounted for in (7)/(37) were the two significant facts that (a) the final syllable in IP_2 *lang/lai* does not undergo any tone sandhi change, and (b) that sentence-final *kong* does undergo tone sandhi change. If one assumes that (47) is the underlying structure for (7)/(37) and that the tone sandhi rules apply to (47) before the movement of IP_1 (and IP_2 etc) to $SpecCP_1$, these two patterns are simply explained. The final element *lang/lai* of IP_2 will be in sentence-final position when tone sandhi changes are applied, and so no tonal change will occur in *lang/lai* as there is no tone-bearing syllable following it at this point. As for *kong* in C^0 of the matrix CP_1, it will be followed by its complement IP_1 at the point of tone sandhi application and so this will naturally cause tone sandhi change on *kong*. The conclusion that *kong* is in the matrix C^0 thus essentially alters nothing in the basic account of the unusual tone sandhi patterns in *kong*-final sentences, and the suggestion that there is IP-raising in such forms is seen to account both for the odd tone sandhi changes with *kong* and its non-canonical sentence-final position.

2.3 Motivations for IP-movement

Having observed that there is good evidence for assuming that the IP in *kong* structures undergoes raising to its surface position to the left of *kong*, we would now like to speculate a little on why this raising might occur. Reflecting a little, there might appear to be three potential explanations for the movement and its motivation, all associated with the process of the grammaticalization of *kong*.

Before considering these, it should be noted that because *kong* is assumed to grammaticalize into the *matrix* clause C^0 position, the process of its grammaticalization is actually somewhat different from the grammaticalization of other verbs of saying as *embedded* clause complementizers noted in section 2.1. In the latter cases, it is essentially a serial verb construction consisting of two verbs which is the source of the new sequence of verb and complementizer (as diagrammed in (30)). If S-final *kong* has grammaticalized as a matrix clause C^0, the source of the new IP-*kong* structures can be suggested to have been not a serial verb construction but two-clause structures in which *kong* occurred as the higher clause main predicate 'to say'. Noting that IP-*kong* forms now consistently express the speaker's commitment to the truth of the proposition encoded in the IP, it can furthermore be suggested that *kong* in such source structures had a first person subject:

(49) [goan kong IP]
I say IP
'I tell you that IP'

As subjects in Taiwanese may frequently be phonetically null when their reference is contextually clear, it is likely that there would also have occurred many simple [*pro* kong IP] forms. We would now like to suggest that an important part of the grammaticalization process may have been that this first person subject specification became structurally re-analyzed and absorbed directly into the head element *kong* as an inherent restriction on its use. Such a process of re-analysis has indeed been attested elsewhere in the grammaticalization of quotative complementizers and the creation of evidential morphemes. For example, Harris & Campbell (1995:169) note that the quotative complementizer *metki* in Georgian is only used to quote the words of the speaker and *metki* was formed historically from the sequence *me vtkvi*, where *me* is the pronoun 'I' and *v-tkv-i* is the first person singular subject (*v-*) aorist indicative (*-i*) of the verb 'say', hence meaning 'I said (it)'. Similarly, in many American Indian languages evidential suffixes on verbs have grammaticalized from verbs of seeing and hearing following the collapse of two-clause structures into mono-clausal forms in the way hypothesized for *kong*. For example, in Maricopa (Gordon (1986)) the suffixes *–'yuu* and *–'a* when added to verbs result in the interpretation that the speaker respectively saw or heard the action described:

(50) lima-'yuu
dance-EV
'He danced (I know because I saw it).'

(51) ashvar-'a
sing-EV
'He sang (I know because I heard it).'

The restriction that it is the speaker who has the visual or aural evidence for the truth of the proposition simply results from the fact that these suffixes are derived from the first person singular verbal forms of the verbs *yuu-k* 'to see' and *av-k* 'to hear', the prefix element ['-] being a first person singular marker. As the morphemes *'yuu* and *'a* are synchronically no longer verbs but verbal suffixes or S-/clause-final particles (and do not license any subject of their own), it can be assumed that the first person subject specification has become re-analyzed as an inherent property of these elements as X^0 head categories, restricting their use and resulting in the interpretation that it is the speaker who has the visual/aural evidence for the proposition.

IP-*kong* forms can therefore be suggested to have developed from two-clause structures in a similar way, with *kong* as the higher clause verb undergoing deverbalization and incorporating a first-person speaker-related interpretation from its former syntactic subject. The collapse of a bi-clausal structure into a simplified mono-clausal form in this way intuitively will take place when there is no longer any pressure to see the 'saying' encoded in *kong* etc as instantiating a highlighted discrete event.

Assuming the above we can now outline three different potential explanations for the IP-raising which has accompanied grammaticalization of *kong* as a C^0. A first phonological possibility is that as *kong* has grammaticalized into a particle-like element, like other particles it has become increasingly more dependent and clitic-like

and in need of some kind of (phonological) support. Normally in Chinese such support should critically come from an element to the particle's *left*, as stress in most varieties of Chinese including Taiwanese is phrase-initial and commonly leads to encliticisation rather than the occurrence of proclitics. It can therefore be suggested that the strong tendency for functional clitic-like elements to attach to their left possibly may have triggered movement of the IP complement of *kong* to a position to its left in order to provide *kong* as an enclitic with phonological support.[3]

A second syntactic possibility is that the rightward IP complement of *kong* is raised to the left of *kong* in order to fill the position formerly occupied by its subject when *kong* was still a verb. In other words, one might imagine that IP-raising takes place to satisfy an EPP-like condition on the head instantiated by *kong*; when *kong* previously had a real NP subject, this subject would have satisfied the EPP, but when the subject (by hypothesis) became re-analyzed simply as a first-person specification characterizing *kong* and an NP subject was no longer projected in the syntax, the IP may have been raised to occupy the relevant Spec position. Such a hypothetical sequence of development is represented in (52) below:

(52) **Stage A**: *kong* still a real verb with an NP subject and a clausal complement:

[(goan) kong IP]

Stage/step B: *kong* deverbalizes and loses its NP subject which is incorporated
as a simple restriction on the use and interpretation of *kong* grammaticalized as a C^0

$[_{CP}\ [_{C}$ kong $[_{IP}\ldots]]]$

Stage/step C: in the new structure, the IP is raised to fill the Specifier position to the left of *kong* previously occupied by the NP subject in the bi-claual structure

$[_{CP}\ [_{IP}\ldots]_i\ [_{C}$ kong $t_i\ ]]$

Such a proposal would essentially need to assume that the SpecCP position is in some sense interpreted as being equivalent to the old subject position and in need of filling with some element.[4]

A third possible explanation of the IP re-positioning might relate it to the informational structure present in *kong* sentences. It can be suggested that the IP complement of *kong* essentially represents old information largely presupposed by the participants in the conversation and that the focus of attention and force of *kong* sentences lies in the *assertion* of the proposition by the speaker via the use of *kong*. Earlier in section 2.3 (example 49) it was noted that S-final *kong* seems to add to the proposition expressed in its IP complement the interpretation: 'I'm telling you IP!' or 'Why do/would you doubt IP?' Essentially when S-final *kong* is used, it seems to imply that the hearer may already entertain the proposition expressed in the IP, but be somewhat doubtful of it for no good reason in the speaker's opinion. Use of *kong* then expresses the speaker's strong endorsement of the truth of the proposition, rather similar to the use of 'I'm telling you!' in English (53):

(53) He's gone, I'm telling you!

In S-final *kong* sentences then the proposition encoded in the IP is a possibility which is entertained as true by both speaker and hearer but with different degrees of certainty. In this sense the IP thus represents old information and is consequently topic-like, being assumed in the discourse. Due to this topic-like property and the focus on the asserting act with *kong*, it can therefore be suggested that the IP undergoes obligatory topicalization to sentence-initial position leaving *kong* in clear focus, and that this defocussing, or 'p-movement' in Zubizarreta's (1998) terms, is quite possibly the motivation behind the IP-raising.

To an extent it is difficult to be fully sure which of the above possibilities is triggering the IP-movement, and it is possible that the raising with *kong* might even be caused by a combination and conspiracy of the different factors discussed. However the raising in *kong*-structures is ultimately interpreted, from a wider cross-linguistic perspective, we would like to suggest that the three potential motivations for clausal raising identified with *kong* above represent generally plausible mechanisms which may possibly underlie the creation of S-final particle structures in many other languages. It is well-noted (e.g. Bybee, Perkins & Pagliuca 1994:107) that the grammaticalization of an element frequently does lead to its phonetic reduction and causes phonological dependency and cliticization. Where such an element grammaticalizes from the higher clause predicate in an SVO bi-clausal structure as outlined and comes to have sentential scope as a new propositional attitude, interrogative or evidential-type operator in mono-clausal forms, this might often induce IP-raising to support the new clitic-like particle, especially in those many languages which are stress-initial in their syntactic phrases and favour encliticization of dependent functional elements rather than procliticization. Secondly, elsewhere EPP-type reasons might also result in an IP-clause raising to fill the position previously occupied by the subject of a higher clause verb which has undergone grammaticalization as a new particle.[5] Finally, theme-rheme and topic-focus-type considerations might indeed often play a role in the re-positioning of a particle relative to its clausal complement and give rise to S-final particles in the way described. The availability of a variety of such mechanisms and motivations for IP clausal-raising may then go some way to explaining the otherwise rather surprising observation that S-final particles seem to develop frequently not just in SOV head-final languages but also in a considerable number of SVO head-initial languages.[6]

Concerning Chinese in particular, the *kong*-paradigm investigated here provides new insights into the debate about whether Chinese should be considered to be rather exceptionally head-final in CP constituents due to the occurrence of sentence-final interrogative particles. The patterns found with *kong* suggest that a particle in an apparent sentence-final C^0-position may actually occur in such a surface position due to raising of its IP complement and that underlyingly CP may actually be a fully regular head-initial category (this being further confirmed by the grammaticalization of Mandarin *shuo* as a C^0 in clause-*initial* position as mentioned in section 2.1). Quite generally, what a consideration of *kong* can be argued to show is that one should not automatically assume that the surface linear ordering of a phonologically dependent particle-head and its complement necessarily mirrors the underlying order of these constituents. It is quite possible that the phonetically reduced dependent nature of particles interacting with stress patterns and theme-rheme properties may result in a significant re-alignment of particle and complement. Consequently, in order to arrive at more robust conclusions concerning the base order of high clausal heads and their complements, one should instead arguably focus on

those instantiations of functional heads which are not particle-like and dependent, as only then is it possible to factor out the interfering role that phonology may play in the relative positioning of such elements (and hence in Chinese one should consider the C^0 elements *yaoshi, ruguo* and *shuo* as better indicators of the position of C^0 than particles such as interrogative *ma*).

Finally one might naturally ask about other S-final particles in Taiwanese and whether any tone sandhi patterns similar to those found with *kong* might suggest that other such particles are also derived from underlying head-initial structures. Here it turns out that other obvious particles such as aspectual *a* (as in (54) below) seem to be older and because of this are phonetically more reduced. Critically such particles occur with no tone at all, having neutral tone/NT and consequently nothing can be concluded about their underlying syntax; neither do they provide positive evidence for IP raising, nor do they provide any evidence against it (i.e. only if they had a real lexical tone and this did not undergo tone sandhi change, could this be considered evidence against an IP-raising analysis):

(54) goan• lau•pe be• hiao• kong• tai•oan•oe a (NT)
I old-father not know speak Taiwanese ASP
'My father can't speak Taiwanese any more.'

Over time many S-final particles may grammaticalize and undergo phonetic reduction to such an extent that they lose the ability to carry an earlier lexical tone and consequently will no longer signal any IP-raising. We suggest that this does not mean that there is not any IP-raising with such particles/heads, and only shows that such (hypothetical) raising is no longer made visible by any tone sandhi patterns (just as other varieties of Chinese may also have IP-raising with sentence-final particles but there is no overt tone sandhi evidence to indicate this). Essentially then it is necessary to catch a particle at a particular point in its development in order to be clear about its underlying syntax, and Taiwanese *kong* interestingly seems to allow for this at the moment, having both an obvious source (in the verb 'to say') and still carrying tone and being able to undergo tone sandhi. In the future however, if *kong* reduces further it clearly may lose its tone and so this clue to its underlying syntax will be lost, leaving it like particles in other languages and varieties of Chinese where good evidence for IP-movement with particles is difficult to find; for the moment though *kong* still continues to provide us with interesting insights into the potential ways in which S-final particles may come into existence.[7]

2.4 Evidence for PF movement

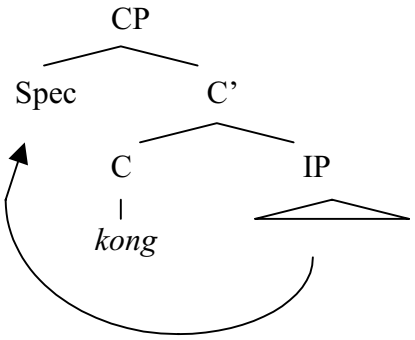
The *kong* paradigm also allows one to draw further significant conclusions concerning the derivational timing of both this particular particle related IP-movement and movement operations in general.

An important part of the IP-raising account of the tone sandhi patterns in *kong*-final sentences has been the suggestion that tone sandhi changes are made at a particular point in the derivation of such sentences when the IP-complement of *kong* in C^0 is still *in situ* and has not yet been raised to SpecCP. Only if the tone sandhi rules are applied at this point can the unusual patterns be given a principled explanation in line with other tone sandhi patterning in Taiwanese. Concerning the essential nature of tone sandhi, given that tone sandhi rules alter the phonetic

interpretation of an element and so apply to specifically to phonetic features, it is most natural to assume that such rules are indeed *phonological rules* and consequently apply in the PF component after Spell-Out. This being so, it will be seen that one seems led to the significant further conclusion that the hypothesized IP-raising operation itself actually has to occur in the PF component too.

The critical sequence in the derivation of *kong*-final sentences is that underlying forms such as (47) repeated below are created in the syntactic component and then *prior to IP-raising* presented for tone sandhi alternation. Assuming that this tone sandhi alternation is a phonological process, under standard assumptions it would seem that it should only take place after a syntactic sequence has entered the PF component. Now, because the IP-raising operation itself has to take place *after* the tone sandhi rules have applied, it then has to be concluded that the IP-raising also occurs in the PF component and hence significantly is an instance of PF rather than syntactic movement.

(47)



Under such assumptions, the derivation of a *kong*-final sentence such as (43) is schematized in (55):

(55) $[_{CP}\ [_{C}$ kong $[_{IP}$ A-sin m lai]

↓

Spell-Out

↓

PF - tone sandhi rules apply (changing the tone on *kong* and maintaining the citation on *lai*)

$[_{CP}\ [_{C}$ kong• $[_{IP}$ A-sin m• lai]]]

↓

IP-raising

$[_{CP}\ [_{IP}$ A-sin m• lai$]_i\ [_{C}$ kong•] t_i]]

The suggestion that movement may perhaps occur in PF has at times been made when certain visible applications of movement appear to have no impact on *interpretation*. In such cases if it is suggested that the relevant movement takes place only in PF after the derivation has left the syntactic component, its effects will not be present in the syntactic string presented for interpretation at LF, and the movement will effectively be semantically vacuous. Arguments for movement at PF are therefore commonly based on the somewhat negative evidence that PF movement is

the only way to explain a lack of otherwise anticipated interpretative effects. In order for such hypothetical PF movement to be generally better supported, one would then hope to find other evidence of a more positive kind indicating that movement may take place after the application of phonology. Here in the tone sandhi patterns of *kong*-final sentences one finds precisely this kind of evidence justifying the suggestion that movement may occur in PF, namely the observation that certain hypothetical movement should only be possible after a clearly *phonological* process has been applied. The tone sandhi patterns investigated here consequently provide interesting strong empirical support for the assumption that certain raising operations which might otherwise be assumed to occur in the syntax actually take place in PF.

2.5 An alternative - cyclic Spell-Out

There is also a second way to interpret the derivational timing of tone sandhi application and IP-raising in *kong*-final sentences, and to avoid the conclusion that there is movement in the PF component it can be suggested that the sequencing of tone sandhi application and IP-raising is due to the occurrence of cyclic Spell-Out (Chomsky 1998) being effected during the course of the derivation.

Chomsky (1998) suggests that there is in fact no single point of Spell-Out where the phonetic features of a sentence are fed off to PF and phonology, but that sub-parts of a derivation may be given phonetic interpretation during the course of a single derivation and before a structure is finally completed. A single syntactic derivation may therefore be spelt-out in a number of successive cycles. The cycles which Chomsky suggests are relevant for cyclic Spell-Out are the "phase" constituents *v*P and CP. Thus the tentative suggestion is that after construction of a *v*P phase, this constituent may be given phonetic interpretation before the derivation continues on syntactically. Similarly once a CP phase is syntactically created, it may also be spelt-out phonetically before the CP is merged into a higher syntactic unit.

The tone sandhi patterns investigated here may be argued to provide evidence in support of such a cyclic Spell-Out approach as an alternative to the possibility that movement occurs at PF.[8] The *kong*-final forms are also able to make somewhat clearer what is arguably to be understood by the idea of cyclic Spell-Out. The critical patterning in *kong*-final sentences in need of some account is the fact that IP-movement seems to have to take place *after* the application of a phonological process, the tone sandhi changes. In section 2.4 it was suggested that this might lead one to the conclusion that the IP-raising therefore takes place at PF. However, if a model incorporating cyclic Spell-Out is adopted and it is assumed that sub-parts of syntactic structure may be given phonetic interpretation mid-way in the course of a derivation, a rather simple second explanation for the sequencing of tone sandhi and IP-raising automatically becomes available. Significantly *kong* is taken to occur grammaticalized in C^0 and hence instantiate the head of a phase-type constituent, CP. It can therefore be suggested that after construction of the phase CP with *kong* in C^0 merged with its complement IP to the right (i.e. $[_{CP}$ kong $[_{IP}$ ….]]), this sequence is given phonetic interpretation and spelt-out in PF. Entering PF the tone sandhi rules will apply to the sequence and cause a tone sandhi alteration in *kong* but not in the final element in the IP, resulting in the surface attested tone sandhi patterns. Following this, the syntactic derivation will then continue, with the IP undergoing raising to a Specifier position to the left of *kong*. On completion of the full syntactic derivation, the sequence will then be spelt-out and will surface with the linear

sequence [[$_{IP}$…] kong]. In such a cyclic Spell-Out approach the IP-raising is therefore a fully regular *syntactic* movement occurring in the syntactic part of the derivation, and movement at PF need not be assumed.

We would like to note that the tone sandhi patterns with *kong* thus interpreted may also allow for a finer understanding of the nature of phase constituents. Chomsky (1998) suggests that there is a distinction between Specifiers that are semantically selected by a head and "extra" Specifiers which it is argued are licensed with the categories C, T and *v* in addition to any selected external argument (EA). Non-selected Specifiers of this second type are taken to host the subject (SpecTP), raised *wh*-phrases (SpecCP), and shifted objects (Spec*v*P). Projections of the "core functional categories" C, T and *v* are accordingly schematized as in (56), with H being the head, YP its complement, EA a semantically selected Specifier, and XP the extra non-selected Specifier:

(56) [XP [(EA) H YP]]

The outer Specifier XP is furthermore suggested to be a position which is critically visible to syntactic heads which occur higher than a CP or *v*P phase, allowing for an element in XP to enter into Agreement relations with such a higher position (and also raise to satisfy EPP requirements of a higher head). Elements inside the inner pair of square brackets in (56) are taken to be invisible to higher positions due to the opacity of phases ("phase impenetrability"). The outer Specifier is therefore a position which is in a sense importantly not inside the phase proper and not part of the phase's core. Turning back to Taiwanese and *kong*-sentences now, it has been suggested that the IP complement of *kong* raises to a Specifier associated with *kong* after the sequence *kong*-IP has been spelt-out. Such a Specifier (SpecCP) is not semantically selected and is therefore of the extra "outer" type just described (XP in (56)). It can therefore now be argued that the input to cyclic Spell-Out may quite possibly be the inner core of phases consisting in the head of a phase, its complement YP and any external argument Specifier (EA), but not necessarily a phase's outer phase-peripheral Specifier XP. Such a Spec position is perhaps created only after the inner core of a phase has been sent to Spell-Out. In *kong* sentences then, the inner core of the phase headed by C^0 (*kong*) is created resulting in the linear sequence [kong IP/TP] and then this is spelt-out phonetically, critically also undergoing tone sandhi alteration at this point. Following Spell-Out of the inner core of the phase, an outer Specifier position is created and the IP (TP) complement of *kong* is moved to this position. Finally the full and final syntactic structure is presented to the phonological component again where the linear order [IP/TP kong] is pronounced.[9] This sequencing is now diagrammed in (57) below (using example (43) again):[10]

(57) **Syntactic creation of the inner core of phase headed by C^0 *kong*:**

↓

[kong [$_{IP/TP}$ A-sin m lai]]

↓

Spell-Out of the inner core + application of tone sandhi rules:

↓

[kong• [$_{IP/TP}$ A-sin m• lai]]

↓

Syntactic raising of the output of mid-derivational Spell-Out → IP/TP raising to outer phase-peripheral Spec of the phase CP:

↓

$[_{CP}\ [_{IP/TP}$ A-sin m• lai $]_i$ kong• $t_i\ ]$

Final syntactic form is pronounced (as above)

The Taiwanese *kong* paradigm thus might seem to add interesting empirical support for the idea of cyclic Spell-Out as an alternative to movement at PF and also allows one to understand more precisely what may be involved in cyclic Spell-Out. In cyclic Spell-Out the linguistic items present in the (syntactically created) inner core of a phase will be given phonetic interpretation and allow phonological processes to apply to the cycle. Following application of the phonology, the sequence will then (normally) be subject to further syntactic operations combining the inner core of the phase with an outer Specifier and merging this further into a higher phase constituent. Phonological processes will then essentially continue to apply on completion of each successive phase until a final syntactic form is created and presented for pronunciation. A cyclic Spell-Out approach thus allows for the interleaving of syntax and phonology and the incremental creation of syntactic and phonetic forms largely in tandem. Here to the extent that the tone sandhi patterns found with *kong* may seem to require such an analysis, the idea of cyclic Spell-Out might indeed seem to be interestingly justified.[11]

3.0 Concluding remarks

This paper began as a simple investigation of the syntax underlying the element *kong* and set out to answer the question of why it is that not only descriptively head-final languages but also head-initial languages develop sentence-final particles. It was suggested that in SVO languages such elements may be created when bi-clausal structures collapse into mono-clausal forms with the verb in the higher clause grammaticalizing as a functional head in the new simplified structure. Increased dependency following grammaticalization combined with the frequent stress-initial property of syntactic phrases and possibly also considerations of theme-rheme structure were then argued to potentially lead to an inversion of the original linear order and cause raising of the functional head's clausal complement to its Specifier position. Critically in the case of *kong* such raising is revealed in the patterns of tone sandhi still found and supported by knowledge of the origin of *kong* as a verb selecting a clausal complement. Combined with other patterns of grammaticalization seen with Mandarin *shuo*, this led to the conclusion that Chinese is not exceptionally head-final in CP constituents, and that the position of particles relative to their complements may in general not be a good indication of underlying head-complement directionality. Having thus provided a possible explanation for the occurrence of sentence-final particles in head-initial languages, the paper then showed how the *kong*-paradigm may be interpreted as providing strong evidence either for the possibility of movement at PF, or for Chomsky's recent proposals for cyclic Spell-Out. Probing the latter approach, the paper finally showed how the patterns with *kong* may reveal more about the operation of cyclic Spell-Out and its application to phases. Throughout sections 2.4 and 2.5 it was also seen how the necessary ordering of tone sandhi application and IP-movement can only be captured in a derivational rather than a representational model of language.

Appendix: relative clauses in Taiwanese and tone sandhi phenomena

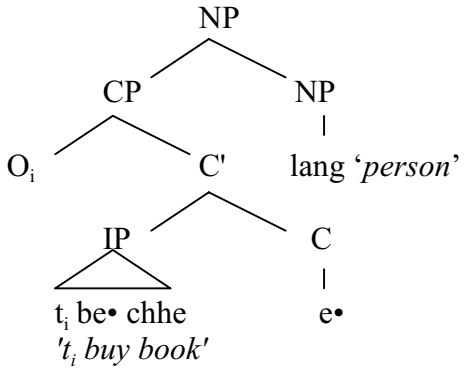
In this final appendix here we would like to briefly discuss how tone sandhi phenomena provide interesting clues to the underlying syntax of relative clauses in Taiwanese and Chinese, in a way quite similar to the investigation of tone sandhi in *kong*-final sentences. Essentially the patterns found will indicate that a Vergnaud/Kaynean analysis of relativization in Chinese is strongly supported and that a further case of a supposed head-final category in Chinese turns out to be head-initial on closer inspection also as a result of IP-raising.

Relative clauses in Chinese and Taiwanese show a highly-marked word-order pattern just as *kong*-final sentences do, and Chinese is possibly the only language anywhere which has been noted to have both a basic surface V-O order and the order of relative clauses occurring before the head-noun (Dryer 1992).[12] Example (58) shows a simple relative clause structure in Taiwanese with an IP-like clause followed by a functional element *e* and then the head-noun:

(58) $[_{IP}$ be• chhe] e• $[_{NP}$ lang]
buy book E person
'the person who bought the book'

In earlier analyses such as Ning (1993) and Chiu (1995) relative clauses in Chinese are assigned the structure in (59) (here using Taiwanese words rather than the Mandarin which appears in these works). In (59) the functional element *e* (Mandarin *de*) is analyzed as a relativizing complementizer in C^0 similar to English 'that', and the CP relative clause is taken to be left-adjoined to the final NP. Note that such CPs are assumed to be head-final categories:

(59)



Patterns of tone sandhi in Taiwanese suggest however that this cannot be the right kind of structure for relative clauses. There are basically two important tone sandhi properties which are in need of some explanation in such structures, as noted in (60):

(60) **Property (a)**
The final element in the IP preceding *e* does *not* undergo tone sandhi (so *chhe* 'book' in (59) does not change its tone).

Property (b)
The functional element *e does* undergo tone sandhi.

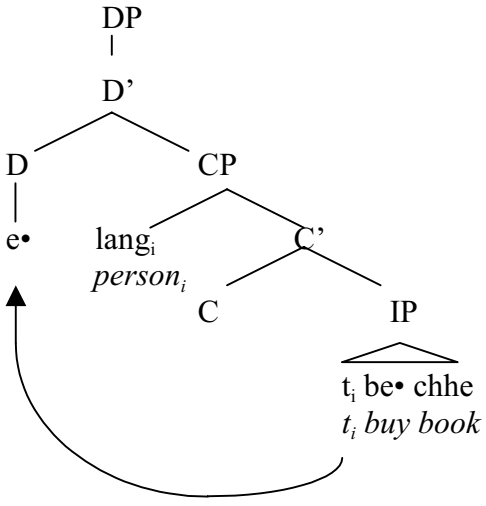
Property (a) is unexpected. If the IP-clause is a leftward complement to the functional element *e* in C^0, then this IP and *e* in C^0 should constitute a single tone

sandhi domain (and tone sandhi patterns in example (21) in section 1.0 show that a C^0 and its IP complement do indeed occur in the same tone sandhi domain). This being so, it is expected that the C^0-element *e* in (59) should be able to trigger tone sandhi on the adjacent element to its left in the IP-clause - the NP *chhe* 'book', but this is not possible.

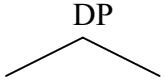
Property (b) is also unexpected and it is not at all clear how the NP can trigger tone sandhi on the adjacent element to its left (i.e. the C^0 *e*). In (59) *e* is analyzed as a C^0 heading a CP which is adjoined to the NP *lang* 'person'. Earlier in (26) Generalization C pointed out that tone sandhi elsewhere does not take place between an adjunct and the element it is adjoined to and that adjuncts are self-contained tone sandhi domains. It is therefore quite unexpected that the head-noun/NP *lang* 'person' is able to cause a tone sandhi change on the adjacent element to its left because this element *e* is contained within a (CP) adjunct. There are consequently good reasons to doubt the correctness of the standard analysis of relative clauses in Chinese in (59).

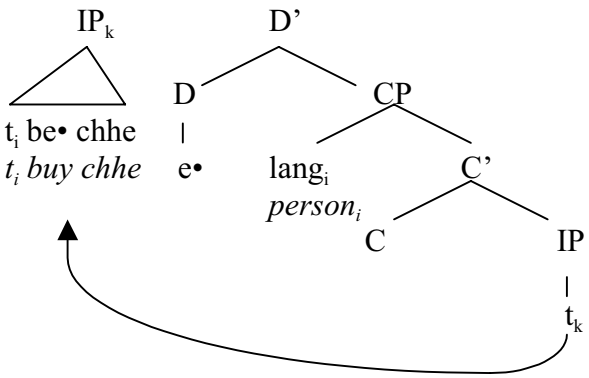
In Simpson (forthcoming) and Wu (in preparation) a rather different Vergnaud-Kaynean-analysis of relative clauses in Chinese is argued for principally to explain the very marked word order property of Chinese relative clauses. In these accounts it is suggested that the functional element *de* (Taiwanese *e*) is actually an instantiation of D^0, a grammaticalized determiner derived from an early demonstrative *zhi* now bleached of any former definiteness value.[13] It is also suggested that this element is an *enclitic* determiner of the type found in Romanian and various other languages which Grosu (1988) analyzes as attracting an element to its Spec-position for phonological support. The surface order of simple relative clause such as (58) then results from an underlying structure which is fully head-initial and regular with *e* in D^0 selecting a rightward CP complement as in (61), basically following ideas in Kayne (1994). The relativized NP is then taken to raise to SpecCP as Vergnaud (1985) and Kayne (1994) suggest, and this is followed by raising of the IP to SpecDP to satisfy the clear enclitic properties of *e*, as in (62). In this approach, the markedness of the word-order in Chinese relative clauses is simply attributed to the particular enclitic properties of the lexical item *e* in D^0:

(61)



(62)





Returning to the Taiwanese tone sandhi, a Kaynean-analysis of relative clauses of this kind is able to explain all of the tone sandhi patterns observed without any of the difficulties found in the more traditional analysis in (59).

Property (a) is easily explained. The IP raises to a Specifier position (SpecDP) and because of Generalization A in (17) that tone sandhi does not occur between a Specifier and its head, it is fully anticipated that the head *e* will not cause any tone sandhi change on the element left-adjacent to it and final in the IP *chhe.*

Property (b) is also straightforwardly accounted for. In (62) the head-noun/NP *lang* is suggested to be part of the CP *complement* to *e* in D^0, which is a significantly different syntactic relation to the adjunction relation assumed in (59). As seen elsewhere, complements instantiated by overt lexical material do trigger tone sandhi changes on the preceding selecting head. Because *e* in D^0 therefore occurs with a lexically-instantiated complement to its right, this quite naturally results in the otherwise unexplained tone sandhi change on *e*.

Consequently, tone sandhi evidence adds interesting support for a Vergnaud/Kaynean analysis of relative clauses in Taiwanese/Chinese and once again provides arguments against analyses of a CP being head-final in Chinese, as proposed in (59) in Ning (1993) and Chiu (1995). As in the case of *kong*-sentences, the element taken to be a head-final C^0 is a heavily grammaticalized element which is considerably reduced and now phonetically enclitic. Hypothesizing that historically it may be related to an early demonstrative and therefore most likely to be in D^0, an investigation of the tone sandhi patterns then suggests that the surface ordering of this functional head and the IP clause preceding it is quite possibly due to the occurrence of IP-raising triggered to provide phonological support for the grammaticalized head as an enclitic. As with the analysis of *kong* sentences, tone sandhi then turns out to be a revealing and potentially powerful tool in the examination of underlying syntactic structure, and once more leads to the conclusion that the surface position of grammaticalized and reduced particle-like heads relative to their heavier complements may perhaps frequently be distorted by non-syntactic phonological reasons.

[1] Two of the 'eight' traditionally recognized tones, tone2 and tone6 are actually identical in phonological terms - both are high-falling 5-1.

[2] Basically it can be seen that a high-level tone will become a mid-level tone and a mid-level tone will turn into a low-falling tone - in both cases a "down-grading" of the tonal value. If a tone occurs at the bottom end of the tonal system, tone sandhi will convert it back up to the top of the tonal group - e.g. a low-falling tone will become a *high*-falling tone (and a low-entering tone becomes a high-entering tone, and high-entering tone becomes a low-entering tone).

[3] See here Grosu (1988) and Giusti (1997) for clear evidence that dependent enclitic definite determiners in Romanian attract elements to D^0/SpecDP in order to support them phonologically and hence that this kind of attraction for phonological support is indeed attested elsewhere.

[4] A question here is if there is some EPP condition on C^0 why it would not be possible to fill the SpecCP position with an empty expletive *pro*. Such an element might be argued to be available and occur in structures such as (i) and (ii) below:

(i) *pro* haoxiang [$_{IP}$ Zhangsan bu lai] le
seems Zhangsan not come ASP
'It seems that Zhangsan is not coming now.'

(ii) *pro* xia yu le
fall rain ASP
'Its started to rain.'

[5] Note that if raising is triggered either to satisfy EPP-type conditions or to provide phonological support for a new enclitic particle, one might expect that it would be possible to raise just an element from within the IP to satisfy such EPP/phonological requirements rather than the entire IP complement clause. This does in fact seem to happen in certain cases. For example, in Cantonese a bi-clausal structure with the V-V unit *teng-gong* 'hear-say' occurring as the higher clause predicate has arguably undergone re-analysis with *teng-gong* grammaticalizing as a simple evidential head selecting a rightward IP clausal complement. In such structures the subject of the rightward IP can for many speakers naturally be raised to the left of *teng-gong* as in (i). Here the raised NP is not interpreted as the subject of *teng-gong* as *teng-gong* has undergone deverbalization in its grammaticalization (no longer occurring with aspectual suffixes) and therefore does not project any subject:

(i) kui teng-gong m wooi lai
he apparently NEG will come
'Apparently he will not come.'

Cantonese *teng-gong* is not yet phonologically dependent in the way that other grammaticalized particles are, and it does not absolutely require the raising of an element to its left. However if *teng-gong* does become a typically enclitic functional head over time, one can see that the need for phonological support to its left could be satisfied by raising of the subject of its complement IP (as in (i)) rather than the whole IP. Why some languages would raise just an NP in this way and others a full IP-clause is not fully clear, but it might relate to the distinction between old and new information. In *kong* structures the complement of *kong* is essentially old information and so naturally topicalized, while with *teng-gong* the content of its IP-complement is new information and so might resist topicalization.

[6] As well as it being common for languages of all types to have S-final particles, many languages also have second-position/P2 particles (enclitics). Old Japanese, for example, had a series of enclitic-particles which occurred in second position in the sentence, the *kakari-musubi* elements *zo, namu* (emphatic) and *ya, ka* (interrogative), as in the following examples from Shibatani (1996):

(i) tsuki zo/namu kiyoki
moom EMPH clear
'The moon, oh it is clear!'

(ii) mizu ya/ka nagaruru
water Q flow
'Does the water flow?'

These particles which have sentential scope/force later re-positioned as S-final particles (as in the modern Japanese Q-particle *ka*). Although the origin of the particles is not known, quite generally the occurrence of particles in second position might occur when a higher clausal performative verb in a bi-clausal structure (such as e.g. '(I) ask/confirm/doubt' etc) grammaticalizes as a new de-verbalized head in a mono-clausal structure, and an element from the rightward complement of the new functional head raises to its left to provide it with phonological support as an enclitic, essentially as discussed with Cantonese *teng-gong* in endnote 5. In the case of the Old Japanese P2 particles this would require assuming an underlying SVO structure as a source, which may be contentious (though see Whitman, forthcoming, for interesting discussion). However, such a developmental sequence would certainly seem plausible for other SVO languages with P2 particles with sentential scope (it can also be noted that the second-position occurrence of the interrogative clausal heads *ya/ka* in Old Japanese is difficult to explain in a strictly head-final analysis of Japanese and so a head-initial approach to Old Japanese is actually not so unmotivated).

[7] The best evidence for an IP-raising analysis we are aware of elsewhere is an interesting argument suggested by Dominique Sportiche (1996 class lectures) relating to French and English intonation

questions, which accounts for patterns in the licensing of NPIs. Sportiche notes that in English and French it is possible to form yes/no questions both by means of subject-auxiliary inversion (SAI), and with a particular type of sentence-final intonation (a rising pattern):

(ii) Did you see John? *subject-auxiliary inversion (SAI)*

(iii) You saw John? *rising intonation only*

However, whereas SAI questions license NPIs, intonation-questions do not:

(iv) Did you see anyone?

(v) *You saw anyone?

Sportiche makes the standard assumption that an NPI is licensed if the Q-morpheme in C is able to c-command the NPI, and this occurs straightforwardly in SAI questions such as (iv), with the auxiliary raised to an initial C^0. It is then suggested that NPI licensing fails in intonation questions such as (v) because the entire IP actually undergoes raising to SpecCP, so destroying the c-command relation between the Q-morpheme in C^0 and the NPI in the IP in SpecCP.

[8] Our thanks to both Joseph Aoun and David Pesetsky for pointing out to us how the tone sandhi patterns might be considered evidence for cyclic Spell-Out.

[9] It can be assumed that such an end-of-derivation re-presentation of the completed syntactic form to the phonology will not result in any second application of tone sandhi rules and that tone sandhi alterations occur only once to any phase.

[10] Supposing one wished to avoid the conclusion that the input to cyclic Spell-Out is just the inner core of the phase, one might attempt to suggest that the IP/TP in *kong* sentences is raised to SpecCP before the phase is spelled-out and it is in fact a *copy* of the IP/TP (later deleted) which triggers tone sandhi on *kong*. Such an idea however leads to the expectation that copies of movement in other constructions should be able to trigger tone sandhi changes, but this is not at all possible. For example, where an object DP is topicalized to SpecCP as in example (25) one would expect that a copy of the object should be able to trigger tone sandhi on the verb, but this cannot happen. Consequently it seems that one must accept that it is the inner core of a phase which is the input to cyclic Spell-Out and not the inner core + outer Specifier. Possibly though there might be variation here relating to the type of movement involved. Earlier it was suggested that the IP-raising in *kong* sentences is perhaps triggered either to provide leftward phonological support for *kong* or alternatively for theme-rheme defocussing reasons ('p-movement'). Noting that Chomsky (1998) suggests that certain types of stylistic movement may actually not be feature-driven, if the IP/TP-raising with *kong* were to be of this type and not triggered by any features, it might turn out that there is a divide between feature-driven and non-feature-driven movement with regard to the outer Specifier of a phase and Spell-Out. Possibly where the creation and filling of such a Specifier is feature-driven, this might have to take place before the relevant phase is spelled-out, whereas with non-feature movement it might be that Spell-Out of the phase may occur before the Specifier is created (as here with *kong*).

[11] The idea of cyclic Spell-Out proposed in Chomsky (1998) is anticipated in interesting work in Bresnan (1971) where data and arguments relating to A'-dependencies and sentential stress lead to conclusions similar to those made here with *kong*. Bresnan convincingly shows that *wh*-phrases and other elements which appear raised in surface forms in fact behave as if they were *in situ* for purposes of sentential stress assignment. For example, whereas sentential stress will normally be placed on the final element in a sentence, in *wh*-questions and relative clauses it is placed on a raised *wh*-phrase, as in (i) with 'what books' receiving the sentential stress:

(i) John asked what BOOKS Helen had written.

It is argued that in order to explain the stress on the *wh*-phrase and the lack of stress on the sentence-final verb, sentential stress must be assigned when the *wh*-element is *in situ* in sentence-final object position prior to raising to SpecCP. As sentential stress is a phonological rule and this must apply before syntactic raising of the *wh*-phrase to SpecCP, it is concluded that phonological rules apply to each transformational cycle in syntax before further syntactic operations in higher cycles, and that phonology will therefore be interwoven with syntax in a single derivation (i.e. there is cyclic phonological Spell-Out). Here we can point out two significant points in the *wh* data Bresnan presents. First of all, if sentential stress as a phonological rule is naturally applied to a CP constituent, then importantly it applies to the CP *before* the SpecCP position is created by raising of the *wh*-phrase (i.e. sentential stress applies to the object *wh*-phrase in its *in situ* position before any raising). This would seem to endorse the suggestion here that the CP phase input to Spell-Out and phonology is the inner core of the CP phase without its external outer Specifier. Furthermore, if *wh*-movement is naturally feature-driven movement one cannot invoke the hypothetical distinction between feature-driven and non-feature-driven movement mentioned in footnote 10. Significantly two of the strongest sets of evidence for cyclic Spell-Out then both suggest that the CP input to mid-derivational Spell-Out is the

inner core of the CP and excludes its external Specifier position. Secondly, Bresnan's patterns might seem to suggest that *wh*-movement does not in fact proceed via Spec*v*P, for the following reason. If sentential stress applies to the CP phase when the CP (without its Specifier) is phonetically spelt-out, the *wh*-phrase clearly has to still be *in situ* in CP-final object position at this point for the stress to be correctly assigned to it. Consequently the *wh*-phrase will not be in any intermediate Spec*v*P position. Following Spell-Out of the CP, the *wh*-phrase will immediately raise to SpecCP. In Chomsky (1998) it is suggested that phases such as *v*P are opaque, and that elements which need to raise higher out of *v*P will first have to raise to Spec*v*P in order to be visible to heads higher than *v*P. *Wh*-movement is then expected to need to go via Spec*v*P and should raise to Spec*v*P as soon as the *v*P phase is completed. The sentential stress phenomena noted above however might now seem to suggest that this is actually not the case. In order to maintain the basic phase-based assumptions in Chomsky (1998), one therefore seems led to either of two possible explanations. One possibility might be to suggest that sentential stress is applied not to a CP phase as input, but instead to the inner core of a *v*P phase before a *wh*-phrase is raised to Spec*v*P. This will certainly capture much of the data, but it is still questionable whether sentential stress can indeed really be re-interpreted as non-sentential *v*P stress, and other aspects of sentential stress might make such an explanation difficult to maintain. A second possibility might be to relativize the notion of opacity and phase impenetrability and suggest that a *v*P phase is only opaque and impenetrable for A-type dependencies but not for A'-dependencies. Consequently *wh*-phrases *in situ* inside *v*Ps would still be visible to a higher C^0 probe and would not have to raise to an intermediate Spec*v*P position before raising higher. When the (inner-core) CP phase is then spelt-out and sentential stress is applied, a *wh*-object may legitimately still remain in its full *in situ* position.

[12] The universal tendency noted across languages is that languages which have their relative clauses preceding the relativized head-noun are descriptively head-final O-V languages, as for example Japanese, Korean and Turkish:

(i) Universal Tendency: if RC-N → O-V
Chinese: RC-N but V-O (single exception)

[13] Examples such as (i) from Pulleyblank (1995) show the early occurrence of *zhi* as a demonstrative:

(i) zhi er chong you he zhi
these two worm again what know
'And what do these two worms know?' (Zhuangzi 1.10)

References

Bresnan, Joan. 1971. Sentence Stress and Syntactic Transformations. *Language* XLVII.2, 257-281. Reprinted in: *Syntactic Argumentation*, (1979) 233-257, edited by D. J. Napoli & E. Rando, Georgetown University Press.

Bybee, Joan, Perkins, Revere & William Pagliuca. (1994) *The Evolution of Grammar*. University of Chicago Press, Chicago & London.

Chiu, Bonnie. 1995. An Object Clitic Projection in Mandarin Chinese. *Journal of East Asian Linguistics 4.2*: 77-117.

Chomsky, Noam. 1998. *Minimalist Inquiries: the framework*. manuscript, MIT.

Dryer, Matthew. 1992. The Greenbergian word order correlations. *Language* 68/1, 81-138

Giusti, Giuliana. 1997. The categorial status of determiners. in *The New Comparative Syntax* ed: L. Haegeman, Longman, London.

Gordon, Lynn. 1986. The development of evidentials in Maricopa. *Evidentiality: the linguistic encoding of epistemology*. 75-88, edited by W. Chafe. Ablex, Norwood, New Jersey.

Grosu, Alex. 1988. On the distribution of genitive phrases in Roumanian. *Linguistics* 26:931-49

Harris, Alice & Lyle Campbell. 1995. *Historical syntax in a cross-linguistic perspective*. Cambridge University Press.

Heine, B. & M. Reh. 1984. *Grammaticalization and reanalysis in African languages*. Helmut Buske Verlag, Hamburg.

Hwang, Jya-Lin. A comparative study on the grammaticalization of saying verbs in Chinese. *Proceedings of the 10th North American Conference on Chinese Linguistics*, 585-597, edited by Chao Fen Sun. GSIL, USC.
Kayne, Richard. 1994. *The Antisymmetry of Syntax.* MIT Press
Lord, Carol. 1993. *Historical change in serial verb constructions*. John Benjamins, Amsterdam.
Ning, Chunyan. 1993. *The overt syntax of relativization and topicalization*. PhD dissertation, University of California, Irvine.
Pulleyblank, Edward. 1995. *Outline of Classical Chinese Grammar.* UBC Press, Vancouver
Shibatani, M. (1996) *The languages of Japan*. Cambridge University Press.
Simpson, Andrew. (forthcoming) On the Status of Modifying DE and the Syntax of the Chinese DP. *On the Formal to Chinese Languages.* eds C.S.Liu & S.W.Tang, CSLI, Stanford.
Vergnaud, Jean-Roger. 1985. *D□pendances et niveaux de repr□sentation en syntaxe.* John Benjamins, Amsterdam.
Whitman, John. (to appear) Kayne 1994: p. 143, fn. 3. *The Minimalist Parameter*, edited by G. Alexandrova. John Benjamins, Amsterdam.
Wu, Xiu-Zhi Zoe. (in preparation). *The development of functional categories in Mandarin Chinese.* PhD dissertation, University of Southern California.
Zhuangzi, cited according to Zhuangzi yinde (A concordance to Chuang Tzu), Harvard-Yenching Institute Sinilogical Index Series, Supplement No. 20 Peiping: Yenching University Press, 1947.
Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order.* MIT Press, Cambridge, Mass.

Andrew Simpson,
SOAS, University of London,
Russell Sq,
London WC1H OXG
England
as4@soas.ac.uk

Xiu-Zhi Zoe Wu
Dept. of East Asian Languages
and Cultures, USC,
University Park Campus, THH226A
Los Angeles, CA 90089-0357, USA
xiuzhiwu@usc.edu