Optional *wh*-movement is discourse-connected movement in Eastern Cham

Kenneth Baclawski Jr.*

Abstract. Eastern Cham (Austronesian: Vietnam) exhibits apparent optional *wh*-movement, which shares properties with apparent topicalization. This paper demonstrates that it is not true *wh*-movement, but *discourse connected-*, or DC-movement. DC requires a phrase to have an antecedent in a prior sentence and for the antecedent's sentence and the anaphor's sentence to be in a particular discourse structural configuration. Data from complex DP's, specifically partitives, inventory forms, and close appositives demonstrate that DC is a property of referential indices that bind DP's. The incompatibility of *wh*-phrases and topicality is then explained as the inability of *wh*-phrases to supply referential indices on their own.

Keywords. syntax; wh-movement; information structure; pragmatics; Austronesian; Southeast Asian languages

1. Introduction. Ā-movement operations to the left periphery seem to exhibit variation as to whether they are optional or obligatory. *Wh*-movement in languages like English is taken to be obligatory (1a). In situ *wh*-phrases (setting aside *do*-support) are typically argued to mark echo questions, with special pragmatic and prosodic properties, labelled 'EQ' below. Topicalization, however, appears to be optional (1b). Moved and in situ topics can have comparable pragmatic interpretations and prosody.

(1) a. {**Which soup**} do you like {**which soup**$_{EQ}$}?

b. {**THE GAZPACHO**}, I like {**THE GAZPACHO**}…

In terms of syntax, these movement operations have been distinguished in terms of Agree (cf. Chomsky 2005, 2008). *Wh*-movement involves an Agree relation between C and the *wh*-phrase. Topicalization, or any movement operation with information structural or discourse-related effect, has been argued to be triggered just by an Edge Feature, in the absence of an Agree relation (contra the Cartographic enterprise: Rizzi 1997, among others, who do posit topic features).

Author: Kenneth Baclawski Jr., University of California, Berkeley (kbaclawski@berkeley.edu).

In some languages, such as Eastern Cham (Austronesian: Vietnam), however, *wh*-phrases appear to be optionally Ā-moved.[1] *Wh*-phrases can be moved to the left periphery (2a), much like topicalization (2b), with only a discourse-related effect. This raises a problem: how can *wh*-movement be optional if it involves an Agree relation with C? Denham (2000) argues for an analysis with optional projection of CP. Cheng (1991, 1997) and others argue through the Clausal Typing Hypothesis that optional *wh*-movement does not exist: the *wh*-phrases must have been moved for some other reason, such as clefting.

(2)   a.   *Hagait hâ daok mbeng? / Hâ daok mbeng hagait?*
          {k̥ɛɪt} hɨ   tɔʔ   ɓăŋ   {k̥ɛɪt}
          what   2SG   PROG   eat   what
          'What are you eating?'
     b.   *Ing-aong ni kau daok mbeng. / Kau daok mbeng ing-aong ni.*
          {ʔiŋ ʔɔŋ ni} kăw   tɔʔ   ɓăŋ   {ʔiŋ ʔɔŋ ni}
          frog   this 1SG   PROG   eat   frog   this
          'This frog, I am eating.'

In this paper, optional *wh*-movement is argued not to be true *wh*-movement, but the same movement operation as topicalization (cf. also Mathieu 2004; Pan 2014 for similar analyses). However, this raises a second problem: how can *wh*-phrases be topicalized? Owing to their interrogative or focal nature (assuming a dichotomy between topic and focus), some have claimed that *wh*-phrases cannot be topics (e.g. Cable 2008; Cruschina 2012; 158). This paper proposes a solution to this problem with two parts. First, the movement operations in (2a–b) are due not to topicality, but *discourse connectedness* (DC), a property defined in terms of discourse structure. DC is not incompatible with focus, as it is not an information structural notion. Second, it is shown that DC marks not a whole DP, but its *referential index*. *Wh*-phrases can be pied-piped if they have a DC-marked referential index. However, they themselves cannot directly be marked as DC, affirming the intuition that *wh*-phrases should not be referential.

The remainder of this paper is structured as follows. First, Section 2 demonstrates that there is no phrasal *wh*-movement in Eastern Cham, only covert movement of the Q-particle. Second, locality effects indicate that topicalization and the movement of *wh*-phrases are triggered by the same syntactic feature (Section 3). Section 4 presents a new analysis of both movement operations as discourse connected-, or DC-movement. Section 5 presents data from partitives, and other complex DP's that elucidates the role of referential indices in DC-marking and their interaction with *wh*-phrases. Section 6 concludes.

---

[1]Data for this paper were collected by the author in Vietnam from 2015–2018 with 33 native speaker consultants born and raised in the Cham villages near Phan Rang, Vietnam. Eastern Cham is an SVO language with little bound morphology. For Eastern Cham examples, the first line represents a Romanization of Cham script known as Rumi. The second, IPA, in line with the Chamic linguistic tradition. Open circles underneath consonants indicates falling, breathy tone or register on the following vowel. The following abbreviations are used CLF = numeral classifier, COMP = complementizer, EMPH = emphatic particle, FEM = feminine proper noun, NEG = negation particle, PROG = progressive aspect marker, ROOT = root modal, SG = singular, (VN) = loanword or code-switching in Vietnamese.

2. Covert Q-movement. Before addressing whether there is optional *wh*-movement in Eastern Cham, it is necessary to understand the underlying mechanism by which *wh*-phrases are interpreted. In the basic case, *wh*-phrases are in situ in Eastern Cham, much like many languages of East and Southeast Asia (cf. Cheng 1991, 1997). Cable (2010) makes a finer-grained distinction between types of *wh*-in situ languages: those that involve covert phrasal movement to C and those that involve movement of just the Q-particle.[2] At issue is whether *wh*-phrases show effects of moving at LF to the left periphery.

Eastern Cham *wh*-phrases do involve some relation with C, as they are sensitive to island constraints. As shown in (3a), an in situ *wh*-phrase is ungrammatical within a relative clause island, even though no overt movement has occurred.[3] The same context is grammatical without a *wh*-phrase (3b). This implies that there exists some Agree relation between C and *wh*. Note that *wh*-phrases are generally acceptable in situ in embedded clauses.

(3) a. **Hâ blei đồ mbeng baow thei ngap?*
         *hɨ plěj đò̰   ɓǎŋ po   **thěj** ŋǎʔ?
         2sg buy stuff(VN) eat COMP who make
         INTENDED: 'You buy the food that who makes?'
     b. *Kau blei đồ mbeng baow amaik kau ngap.*
         kǎw plěj đò̰   ɓǎŋ po   mɛʔ   kǎw ŋǎʔ?
         1sg buy stuff(VN) eat COMP mother 1sg make
         'I buy the food that my mother makes.'

However, intervention effects demonstrate that in situ *wh*-phrases cannot covertly move. In (4a), a *wh*-phrase is under the scope of *tha çěj* 'only', an intervention effect trigger in Eastern Cham. Kotek (2014, 2017) argues that intervention effects arise when a *wh*-phrase cannot escape the scope of an intervention trigger through overt or covert movement. The ungrammaticality of (4a) indicates that *wh*-phrases have no option to covertly move. As predicted, overt movement is sufficient to obviate intervention (4b).

(4) a. **Sa drei sa-ai Thuận takrâ aia bai halei min?*
         *tha çěj ʔaj   thṳ̀ən kɨ **ʔjæ p̥aj hlɛj** mǐn
         only older.sibling Thuận like soup which EMPH
         INTENDED: 'Which soup does only Thuận like to eat?'
     b. *Aia bai halei sa drei sa-ai Thuận takrâ min?*
         ʔjæ p̥aj hlɛjᵢ   tha çěj ʔaj   thṳ̀ən kɨ ~~ʔjæ p̥aj hlɛj~~ mǐn
         soup which only older.sibling Thuận like   EMPH
         'Which soup does only Thuận like to eat?'

---

[2] Cf. also true *wh*-in situ languages like Vietnamese (Tsai 2009), where *wh*-phrases have been argued to be interpreted in situ as variables.

[3] The complementizer *po* is necessary to distinguish between a relative clause reading and a serial verb construction.

Thus, Eastern Cham *wh*-phrases are not interpreted by covert phrasal movement. They must be interpreted by movement of the Q-particle. It should be noted that this Q-particle is null in Eastern Cham, but perhaps overt in closely related languages like Moken (Baclawski Jr. & Jenks 2016). Given that in situ *wh*-phrases are interpreted in this way, it would be surprising for there also to be optional phrasal movement.

3. Locality effects. In this section, apparent *wh*-movement is shown to be featurally identical to topicalization by locality effects. Both are sensitive to island constraints (5a–b). It should be noted that there is a hanging topic construction, not shown here, that is distinguished by a pause and resumptive pronoun, not seen in the kind of topicalization examined in this paper.

(5) a. **Thei, hâ blei đồ mbeng baow ngap?*
         ***thĕj** hɨ plĕj độ ɓăŋ po ~~thĕj~~ ŋă?
         who 2SG buy stuff(VN) eat COMP make
         INTENDED: 'You buy the food that who makes?'

    b. **Amaik kau, kau blei đồ mbeng baow ngap.*
         ***mɛʔ kăw** kăw plĕj độ ɓăŋ po ~~mɛʔ kăw~~ ŋă?
         mother 1SG 1SG buy stuff(VN) eat COMP make
         INTENDED: 'I buy the food that my mother makes.'

Locality effects arise when there are multiple phrases in the left periphery. In general, when there are multiple phrases in the left periphery, the movement chains must be nested in the sense of Pesetsky (1982). (6a) gives an example of nested paths: the movement path of *han ni* 'this cake' does not cross that of *nɪ̆ʔ mːɛj thɪ̆t năn* 'that little girl'. However, in (6b), the paths cross, and the result is ungrammatical. Baclawski Jr. & Jenks (2016) analyze the same effect in Moken (Austronesian: Thailand) with multiple probes in the left periphery.

(6) a. *Ahar ni anâk kumei sit nan Thuận da-a mai mbeng.*
         han ni nɪ̆ʔ mːɛj thɪ̆t năn*ᵢ* thùən ʔa ~~nɪ̆ʔ mːɛj thɪ̆t năn~~ maj
         cake this child woman small that Thuận invite come
         ɓăŋ ~~han ni~~
         eat
         'This cake, Thuận invited that little girl to come eat.'

    b. **Anâk kumei sit nan ahar ni Thuận da-a mai mbeng.*
         *nɪ̆ʔ mːɛj thɪ̆t nan **han ni** thùən ʔa ~~nɪ̆ʔ mːɛj thɪ̆t năn~~ maj
         child woman small that cake this Thuận invite come
         ɓăŋ ~~han ni~~
         eat
         INTENDED: 'This cake, Thuận invited that little girl to come eat.'

When there is one topic and one *wh*-phrase in the left periphery, the same effect holds. If the paths are nested (7a), the result is grammatical. If they are crossed (7b), it is strongly and consistently ungrammatical. This follows if there are two probes in the left periphery both searching for the same feature.

(7) a. *Ahar ni anâk kumei sit halei Thuận da-a maj mbeng?*
   han **ni** n̆ĭʔ mːɛj thĭt hlɛj thừən ʔa ~~n̆ĭʔ mːɛj thĭt halei~~
   cake this child woman small which Thuận invite
   maj ɓăŋ ~~han ni~~
   come eat
   'Which little girl did Thuận invite to come eat this cake?'

   b. **Anâk kumei sit nan ahar halei Thuận da-a maj mbeng?*
   *n̆ĭʔ mːɛj thĭt năn **han** **hlɛj** thừən ʔa ~~n̆ĭʔ mːɛj thĭt năn~~
   child woman small that cake which Thuận invite
   maj ɓăŋ ~~han halei~~
   come eat
   INTENDED: 'Which cake did Thuận invite that little girl to come eat?'

One hypothesis is that these probes are searching for a general Ā-feature in the sense of Aravind (2017, 2018), and others. Such an Ā-feature could include topic and *wh*. Aravind (2017) uses these general features to explain similar path containment effects in English. A general Ā-feature would not account for Eastern Cham, though, because in situ *wh*-phrases do not intervene on movement of topics or other *wh*-phrases (8a–b).[4] *Wh*-phrases, in situ or not, are taken to be assigned a general Ā-feature, and in other languages such as English and Malayalam, they do in fact intervene on movement (Aravind 2017). Something else is driving the movement in Eastern Cham.

(8) a. *Ahar ni Thuận da-a thei mai mbeng.*
   han **ni** thừən ʔa thĕj maj ɓăŋ ~~han ni~~
   cake this Thuận invite who come eat
   'Who did Thuận invite to come eat this cake?'

   b. *Hagait, Thuận da-a thei mai mbeng?*
   ꞵ̣ɛt thừən ʔa thĕj maj ɓăŋ ꞵ̣ɛt
   what Thuận invite who come eat
   'Who did Thuận invite to come eat what?'

Topicalization and the movement of *wh*-phrases in Eastern Cham are instances of the same movement operation from a featural standpoint. Both are driven by the same syntactic feature. However, this raises a problem. If that feature is [topic], it would require *wh*-phrases to be able to be marked as topical. This is problematic under certain views of information structure. Cruschina (2012), and others argue that there is an incompatibility between topic and focus or interrogativity.

4. Discourse connectedness. This section begins to answer how *wh*-phrases can appear to be topical by redefining both categories. What seems to be topicalization is in fact *discourse connected-* or DC-movement, and what appears to be optional *wh*-movement is DC-movement of a *wh*-phrase. *Discourse connectedness*, first observed by López (2009) (as [+anaphora]) for clitic right-dislocation in Catalan, has two components (9).

---

[4]Note that these same facts obtain for movement out of embedded clauses. Superiority effects do not obtain in local or non-local contexts (cf. Baclawski Jr. & Jenks 2016 on Moken).

First, the phrase in question must be a discourse anaphor, an anaphor whose antecedent is in a prior sentence in the discourse. Second, the antecedent's sentence must *discourse subordinate* the anaphor's sentence (10).

(9) DISCOURSE CONNECTEDNESS (DC): For antecedent $x \in$ sentence ϕ and anaphor $y \in$ sentence ψ, y is DC iff ϕ *discourse subordinates* ψ

(10) DISCOURSE SUBORDINATION: Sentence ϕ discourse subordinates sentence ψ if ψ is interpreted as a subpart of the focus space denoted by ϕ, but not vice versa

Discourse subordination was first defined in terms of focus spaces and attention states by Grosz & Sidner (1986) (as 'dominance'). Subsequently, it has been defined in terms of embedded eventualities (cf. Asher & Lascarides 2003; Asher & Vieu 2005 in Segmented Discourse Representation Theory), or in terms of discourse intentions (cf. Mann & Thompson 1988 in Rhetorical Structure Theory, as 'nucleus-satellite relations'). Informally, it involves two sentences, or discourse units, one of which elaborates on the other and leaves it open for future discussion. Why should discourse subordination be relevant to an anaphor? According to the Right-Frontier Constraint (Polanyi 1985; Webber 1988), discourse subordination is one of multiple conditions that render an antecedent accessible for anaphora. Discourse connectedness restricts discourse anaphora to just this one condition, resulting in a stronger, more specific anaphoric relation.

To illustrate, consider the discourse in (11a–b). (11a) contains an antecedent 'frog'. (11b) has an anaphor, identical to its antecedent. That anaphor can be DC-moved, because the two sentences are in a subordinating discourse relation. One elaborates upon the other. The same anaphoric relation is present in (b′). However, DC-movement is illicit, because the sentence is not in a subordinating relation. Instead, it is interpreted as a separate event.

(11) a. CONTEXT: 'Look at Thuận cook frog$_i$.'
   b. *Ing-aong nyu ngap bingi ralo.*
      {ʔiŋ ʔɔŋ$_i$} ɲu   ŋăʔ {} ŋ̊:i   lo
      frog   3.ANIM make   be.delicious very
      'He makes[cooks] frog very well [LIT: deliciously].'   ✓Subordination(a,b)
   b′. *Sơn ngap ing-aong hu o.*
      {#} sɤn thằw   ŋăʔ {ʔiŋ ʔɔŋ$_i$} o
           Sơn know.NEG   make frog   NEG
      'Sơn does not know how to make[cook] frog.'   ✗Subordination(a,b′)

The same holds for *wh*-phrases. The antecedent in brackets in (12a) defines a set of two pots. Both (12b,b′) have anaphors that refer to that set in the form of *wh*-phrases. However, only (12b) is in a subordinating discourse relation with (12a), as it is interpreted within the same event. By contrast, (12b′) is interpreted as a separate event, and DC-movement is infelicitous.

(12) a. CONTEXT: 'Look at my father boil [one pot of fish and one pot of frog]$_i$.'
b. *Urak ni, gaok hagait ong nan daok ngap nan?*
ja ni {ḳɔʔ ḳɛɪt$_i$} oŋ nǎn t̥ɔʔ ŋǎʔ {} nǎn
now pot what old.man that PROG make that
'Now, what pot is that old man making [working on]?' ✓Subordination(a,b)

b′. *Urak ni, ong nan daok mbeng gaok hagait nan?*
ja ni {#} oŋ nǎn t̥ɔʔ ɓǎŋ {ḳɔʔ ḳɛɪt$_i$} nǎn
now old.man that PROG eat pot what that
'Now, what pot is that old man eating?' ✗Subordination(a,b′)

Topicality does not directly account for the pattern above. If topic is defined in terms of old information (e.g. Chafe 1976), it is insufficient to distinguish between the (b) and (b′) sentences. Both have identical prior information states (i.e. (a)). Topic is also defined in terms of aboutness, as in the phrase around which a sentence is organized (e.g. Reinhart 1981). Some analyses preclude *wh*-phrases from being topics altogether, which would not account for the Eastern Cham data.[5] Otherwise, both the (b) and (b′) sentences can be construed as being about the set of pots (cf. Casielles-Suarez 2004; 23; López 2009; 27 for other arguments against old information and aboutness topic). Nevertheless, aboutness has some explanatory power: only the (b) examples can be paraphrased with 'As for X…', at least in the English glosses (cf. Reinhart 1981). Discourse connectedness provides a concrete way to encode the pattern above, which can be embedded within broader theories of discourse (i.e. Segmented Discourse Representation Theory or Rhetorical Structure Theory).

Encoding anaphoric relations between discourse moves has some precedent in the literature. According to Büring (2003), Constant (2014), and others, contrastive topics require a discourse anaphoric relation within a complex Question Under Discussion (QUD). There must be an antecedent and an anaphor within a subpart, or strategy, to answer a broader QUD. This is comparable to discourse connectedness, as both involve leaving a prior sentence open for further discussion. However, these accounts of CT are more restricted: they require open Questions Under Discussion and an element of contrast. In Eastern Cham DC-movement, contrast is never present, and open QUD's are not required.[6] Other models of discourse have posited less restricted relations between sentences. Onea (2013, 2016) posits that certain phenomena are sensitive to Potential Questions, any question that is compatible with the presuppositions of the prior sentence. Discourse connectedness is more restricted, requiring discourse subordination, though an analysis could be construced within the Potential Question framework, especially following Onea (2016), which integrates Potential Questions with Segmented Discourse Representation Theory.

As detailed above, discourse connectedness is defined in terms of discourse structure, not pure topicality. DC is not necessarily incompatible with *wh*-phrases. It is widely accepted that *wh*-phrases, especially discourse-, or D-linked *wh*-phrases, can have

---

[5]López (2009) also restricts the ability of *wh*-phrases to be assigned [+anaphora] due to the fact that they cannot be clitic right-dislocated in Catalan.
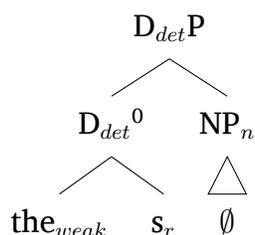
[6]Cf. Baclawski Jr. (2018) on the realization of contrastive topic in Eastern Cham.

discourse antecedents in some way (cf. Pesetsky 1987; Comorovski 1996). DC simply requires that a phrase have an antecedent and their respective sentences have a certain type of relation in the discourse. The following section examines these antecedents in more detail.
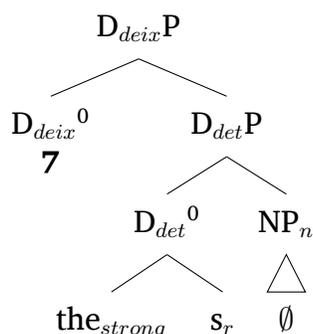
5. Referential indices. This section examines complex DP's, such as partitives, and finds that only *referential indices* are DC-marked. It has been argued that definite DP's of various kinds contain syntactic instantiations of their domain restrictions. These domain restrictions are generally filled by indices supplied by the context. For example, Patel-Grosz & Grosz (2017) propose the structures in Figure 1 for personal and demonstrative pronouns (cf. also Schwarz 2009). Demonstrative pronouns are distinguished from personal pronouns by an additional DP projection, which merges a referential index from the context set, here, **7**, for example.

Figure 1: Pronoun structure, as per Patel-Grosz & Grosz (2017; (7–8))

(a) Personal pronoun

$D_{det}P$
D_{det}^0 NP_n
the$_{weak}$ s_r $\emptyset$

(b) Demonstrative pronoun

$D_{deix}P$
D_{deix}^0 $D_{det}P$
7
D_{det}^0 NP_n
the$_{strong}$ s_r $\emptyset$

With this in mind, it has been argued that overt pronouns or other DP's can specify the reference of the referential indices. As evidenced by Huang et al. (2009; 298) and Jenks (2018), close appositives such as (13a) in Mandarin Chinese are structurally singular DP's in which the element on the left overly marks the referential index that binds the definite DP on the right. Eastern Cham has a similar close appositive (13b). Note that the root modal *hu* is clause-final and has scope over the predicate.

(13) a. wo xihuan [**Zhangsan**, **Lisi** na   ji-ge   guai haizi]$_{DP}$.
        I   like   Zhangsan Lisi those several-CLF good children
        'I like those good kids like Zhangsan and Lisi.'   MANDARIN
        (Huang et al. 2009; 299)

    b. *Hâ da-a Phú saong Hoa, dua urang nan hu.*
        hɨ ʔa   [**pu hɔŋ** ŋ=**hwa** k̥wa raŋ   năn]$_{DP}$ hu
        2SG invite Phú with FEM=Hoa 2   person that   ROOT
        'You can invite Phú and Hoa, those two people.'

Two additional constructions have similar properties. Partitives are formed in Eastern Cham by moving a noun that originates lower in the DP to a higher position (14a). The noun that moves corresponds with the one that marks the antecedent set in the

discourse, much like the referential index of definites (cf. Enç 1991). Inventory forms are minimally different from partitives in that the moved noun is indefinite (14b). The moved noun again corresponds with the one that marks the antecedent set in the discourse and has been equated with a DP-internal topic (cf. Simpson 2005).

(14) a. *Hâ blei pa-aok ni, tajuh abaoh hu.*
    hɨ   plɛj [ʔɔ̆ʔ **ni** çuh pɔh ~~ʔɔ̆ʔ ni~~]ₒₚ hu
    2SG buy mango this 7   CLF   ROOT
    'You can buy seven of these mangoes.'
  b. *Hâ blei pa-aok, tajuh abaoh hu.*
    hɨ   plɛj [ʔɔ̆ʔ   çuh pɔh ~~ʔɔ̆ʔ~~]ₒₚ hu
    2SG buy mango 7   CLF   ROOT
    'You can buy mangoes, seven.'

Each of these constructions represents a singular DP. They can be moved to the subject position of unaccusatives, as evidenced in (15a–b). The moved noun is also a DP, as it can be subextracted from a complex DP to subject position in unaccusatives, stranding the remainder of the DP (15c; cf. Wang 2007 and others on Japanese).

(15) a. *Phú saong Hoa, dua urang nan laik trun.*
    [**pu hɔŋ ŋ = hwa** k̥wa raŋ   năn]ₒₚ lɛ̆ʔ tʃŭn
    Phú with FEM = Hoa 2   person that   fall down
    'Phú and Hoa, those two people fell.'
  b. *Pa-aok ni, tajuh abaoh laik trun.*
    [ʔɔ̆ʔ **ni** çuh pɔh]ₒₚ lɛ̆ʔ tʃŭn
    mango this 7   CLF   fall down
    'Seven of these mangoes fell down.'
  c. *Pa-aok ni laik trun tajuh abaoh.*
    [ʔɔ̆ʔ   **ni**]ₒₚ lɛ̆ʔ tʃŭn [~~ʔɔ̆ʔ ni~~ çuh pɔh]ₒₚ
    mango this   fall down   7   CLF
    'Seven of these mangoes fell down.'

These complex DP constructions are relevant to discourse connectedness and *wh*. First, only the referential index DP can be DC-moved. In (16a), the referential index is subextracted from a partitive. Neither the rest of the partitive cannot be subextracted (16b), nor the whole partitive itself (16c) can be DC-moved. The same pattern obtains with close appositives and inventory forms. This indicates that it is really the referential index, not the whole DP that is marked as DC.

(16) a. *Pa-aok ni, kau blei tajuh abaoh.*
    ʔɔ̆ʔ   **ni** kăw plɛj [~~ʔɔ̆ʔ ni~~ çuh pɔh]
    mango this 1SG buy   7   CLF
    'Of these mangoes, I boughts seven.'
  b. *Tajuh abaoh, kau blei pa-aok ni.*
    *çuh pɔh kăw plɛj [ʔɔ̆ʔ   ni   ~~çuh pɔh~~]
    7   CLF 1SG buy mango this
    INTENDED: 'Seven, I bought these mangoes.'

9

c. *Pa-aok ni, ajuh abaoh, kau blei.
   *ʔɔ̌ʔ   ni   cuh   pɔh   kăw   plɛj   [ʔɔ̌ʔ ni cuh pɔh]
   mango this 7   CLF   1SG   buy
   INTENDED: 'Seven of these mangoes, I bought.'

Second, the referential index DP conversely cannot consist of a *wh*-phrase. In general, a partitive can contain a *wh*-phrase, as in (17a). However, the referential index DP cannot (17b). Note that the question is askable when prompted in elicitation, and the gloss in English is grammatical.

(17) a. *Hâ blei pa-aok ni hadom ki?*
        hɨ   plɛj   [ʔɔ̌ʔ   ni   tom   ki?]
        2SG buy   mango this how.many kg
        'How many kilograms of these mangoes did you buy?'
     b. **Hâ blei abaoh kayau halei sa ki?*
        hɨ   plɛj   [pɔh zaw   hlɛj   tha ki?]
        2SG buy   fruit   which one kg
        INTENDED: 'Which fruit did you buy one kilogram of?'

These facts can be explained if it is the referential index that is marked as discourse connected, not the whole DP. That is why only the referential index DP can be DC-moved. This also accounts for the incompatibility of *wh*-phrases and referentiality: *wh*-phrases can be DC-moved, if they are pied-piped by a DC-marked referential index; however, they cannot supply that referential index on their own.

6. Conclusion. What seems to be optional *wh*-movement in Eastern Cham is in fact discourse connected-, or DC-movement. This result upholds the Clausal Typing Hypothesis and provides a way to account for movement operations that target both *wh*-phrases and topics. Both can be DC-moved so long as they are discourse anaphors in a particular discourse structure configuration. Yet, there remains an incompatibility: DC targets referential indices, and *wh*-phrases cannot supply referential indices themselves. Instead, they can be pied-piped by a DC-marked index.

Many questions remain for future research. First, even though discourse connectedness restricts the contexts DC-movement can occur in, the movement operation remains optional in some sense. Is this true optionality, perhaps susceptible to Optimality Theoretic constraints, or only apparently optional, the result of different derivations that coincidentally receive similar interpretations (cf. Müller 2003 on pseudo-optionality)? Second, this paper has only presented a syntactic analysis of DC; how is DC interpreted, and how can it mark sets of alternatives like *wh*-phrases? Third, how is DC comparable to other information structural notions like D-linking? Finally, how is DC-movement of *wh*-phrases possible in languages like Eastern Cham, but not in others like Catalan?

References

Aravind, Athulya. 2017. Ā-interactions and feature geometries. In Claire Halpert, Hadas Kotek & Coppe van Urk (eds.), *A Pesky Set: Papers for David Pesetsky*, 333–342. MIT Press.

Aravind, Athulya. 2018. Licensing long-distance wh-in-situ in Malayalam. *Natural Language & Linguistic Theory* 36. 1–43. https://doi.org/10.1007/s11049-017-9371-2.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

Asher, Nicholas & Laure Vieu. 2005. Subordinating and coordinating discourse relations. *Lingua* 115. 591–610. https://dx.doi.org/10.1016/j.lingua.2003.09.017.

Baclawski Jr., Kenneth. 2018. Contrastive topic in Eastern Cham. *Berkeley Formal Papers in Linguistics* 1(1). https://escholarship.org/uc/item/3hp0s863.

Baclawski Jr., Kenneth & Peter Jenks. 2016. Clefts and Anti-Superiority in Moken. *Journal of the Southeast Asian Linguistics Society* 9. 81–96. https://openresearch-repository.anu.edu.au/handle/1885/105175.

Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26. 511–546.

Cable, Seth. 2008. Wh-fronting in Hungarian is not focus fronting. Manuscript. ling.auf.net/lingbuzz/000674/v1.pdf.

Cable, Seth. 2010. *The Grammar of Q*. Oxford University Press.

Casielles-Suarez, Eugenia. 2004. *The syntax-information structure interface: Evidence from Spanish and English*. Routledge.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 25–55. New York: Academic Press.

Cheng, Lisa L.S. 1991. *On the typology of wh-questions*: MIT dissertation.

Cheng, Lisa L.S. 1997. *On the typology of wh-questions*. New York: Garland. Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22. https://doi.org/10.1162/0024389052993655.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguis-tic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. MIT Press. https://doi.org/10.7551/mitpress/9780262062787.001.0001.

Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Springer.

Constant, Noah. 2014. *Contrastive topic: Meanings and realizations*: University of Massachusetts, Amherst dissertation.

Cruschina, Silvio. 2012. *Discourse-related features and functional projections*. Oxford University Press.

Denham, Kristin E. 2000. Optional wh-movement in Babine-Witsuwit'en. *Natural Language & Linguistic Theory* 18. 199–251.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1). 1–25.

Grosz, Barbara & Candace Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3). 174–204.

Huang, James C.-T., Y.-H. Audrey Li & Yafei Li. 2009. *The syntax of Chinese*. Cambridge Unviersity Press.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49(3). 501–536. https://doi.org/10.1162/ling_a_00280.

Kotek, Hadas. 2014. *Composing questions*: MIT dissertation.

Kotek, Hadas. 2017. Dissociating intervention effects from superiority in English wh-questions. *The Linguistic Review* 34(2). 397–417. https://doi.org/10.1515/tlr-2017-0005.

López, Luis. 2009. *A derivational syntax for information structure*. Oxford University Press.

P. Mann, William & Sandra Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(3). 243–281. https://doi.org/10.1515/text.1.1988.8.3.243.

Mathieu, Eric. 2004. The mapping of form and interpretation: The case of optional wh-movement in French. *Lingua* 114(9–10). 1090–1132. https://doi.org/10.1016/j.lingua.2003.07.002.

Müller, Gereon. 2003. Optionality in optimality-theoretic syntax. In Lisa Cheng & Rint Sybesma (eds.), *The Second Glot International State-of-the-Article Book*, 289–322. Walter de Gruyter. https://doi.org/10.1515/9783110890952.289.

Onea, Edgar. 2013. *Potential questions in discourse and grammar*: University of Göttingen Habilitation thesis.

Onea, Edgar. 2016. *Potential questions at the semantics-pragmatics interface*. Brill.

Pan, Victor Junnan. 2014. Wh-ex-situ in Mandarin Chinese: Mapping between information structure and split CP. *Linguistic Analysis* 39. 371–413.

Patel-Grosz, Pritty & Patrick Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48(2). 259–297. https://doi.org/10.1162/ling_a_00243.

Pesetsky, David. 1982. *Paths and categories*: MIT dissertation.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland & Alice G.B. ter Meulen (eds.), *The representation of (in)definiteness*. 98–129. Cambridge, MA: MIT Press.

Polanyi, Livia. 1985. A theory of discourse structure and discourse coherence. *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistics Society (CLS)*.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27. 53–94.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: A handbook of generative syntax*. 281–327. Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7.

Schwarz, Florian. 2009. *Two Types of Definites in Natural Language*. University of Massachusetts, Amherst dissertation.

Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In

Guglielmo Cinque & Richard Kayne (eds.), *The Oxford handbook of comparative syntax*, 806–838. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195136517.001.0001.

Tsai, Cheng-Yu Edwin. 2009. *Wh-dependency in Vietnamese and the syntax of wh-in-situ*: National Tsing Hua University dissertation.

Wang, Chyan-an Arthur. 2007. Sluicing and resumption. In *Proceedings of the 37th North East Linguistic Society*, 239–252. Amherst, MA: GLSA.

Webber, Bonnie L. 1988. Discourse deixis: Reference to discourse segments. In *Proceedings of the 26th Annual Meeting on Association for Computational Linguistics*, 113–122. https://doi.org/10.3115/982023.982037.