

# Eastern Cham optional *wh*-movement is discourse connected-movement

## Abstract

Eastern Cham (Austronesian: Vietnam) exhibits apparent optional *wh*-movement, which presents a potential counterexample to Cheng's (1991, 1997) Clausal Typing Hypothesis. This paper shows that the movement of *wh*-phrases is due to a discourse structural property we call *discourse connectedness* (DC). A phrase can be marked as DC if it is previously mentioned in a sentence that the current sentence explains or elaborates upon. In Eastern Cham, *wh*-phrases are moved to left periphery only if they satisfy this condition. This paper proposes a model of DC-marking that applies both to *wh*- and non-*wh*-phrases. Distributional evidence supports this model over other *wh*-phenomena such as pseudoclefts and D-linking. Syntactically, DC-movement is argued to be an  $\bar{A}$ -movement operation driven by a probe on C. Evidence from locality effects shows that this probe searches only for a feature that indexes DC, not *wh*. We argue that syntax must be sensitive to certain aspects of discourse structure and that the notion of context must be enriched enough to track rhetorical relations between sentences and the referents within them.

## Keywords

Syntax, *wh*-movement, Agree, information structure, Austronesian languages

# 1 Introduction

It is widely known that there is cross-linguistic variation in the surface position of *wh*-phrases: they can be moved to the left periphery (*wh*-movement), or they can remain in their base position (*wh*-in-situ). Cheng (1991, 1997) proposes the Clausal Typing Hypothesis (CTH), which holds that languages can only make use of one of these strategies: *wh*-in-situ with an accompanying *wh*-particle, or *wh*-movement (1); there are no mixed *wh*-movement/in-situ languages (Cheng 1997: 28). In more recent formalizations, every language is predicted to have exactly one Agree mechanism between C and *wh* (cf. Roussou & Vlachos 2011 and references therein).

- (1) Clausal Typing Hypothesis: Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle in C<sup>0</sup> is used or else fronting of a *wh*-word to the Spec of C<sup>0</sup> is used, thereby typing a clause through C<sup>0</sup> by Spec-head agreement. (Cheng 1997: 22)

Eastern Cham (Austronesian: Vietnam) presents a potential counterexample to the CTH. There is a yes-no question particle *lěj* (2a), hence Eastern Cham has a *wh*-particle in the sense of Cheng, and *wh*-phrases are in-situ in the basic case ((3a); *wh*- and corresponding phrases bolded throughout). Thus, the CTH predicts Eastern Cham to be a *wh*-in-situ language. However, *wh*-phrases may appear in the left periphery (3b).

- (2) hi tɔʔ bǎŋ vǎʔ lěj  
2SG PROG eat ITER Y/N.Q  
'Are you eating more/again?'

- (3) a. hi tɔʔ bǎŋ **keʔ**  
2SG PROG eat what  
'What are you eating?'
- b. **keʔ** hi tɔʔ bǎŋ keʔ  
what 2SG PROG eat  
'What are you eating?'

On the one hand, if Eastern Cham represents a true counterexample to the CTH, it would constitute an instance of optional *wh*-movement. Denham (1998, 2000) and others have proposed that optional *wh*-movement is in fact feasible based on Minimalist framework assumptions, pointing to the related language Malay as a possible case (Denham 2000: 248; cf. conversely Cheng & Rooryck 2000 on optional in-situness). On the other hand, Eastern Cham would not constitute a true counterexample if either (3a) or (3b) do not represent basic *wh*-questions (see Denham 1998: 19). Within Austronesian, apparent *wh*-movement has been argued to be due to concealed clefts or pseudoclefts (Cheng 1991 on Bahasa Indonesian; Paul 2001, Potsdam 2006 on Malagasy), or focus-movement (cf. Jensen 2013: 112 on the Chamic language Jarai). In these analyses, it is either the *wh*-feature itself that motivates the derived position of the *wh*-phrase, or a feature like focus, which indexes an inherent property of *wh*-phrases if it is assumed that all *wh*-phrases are foci.

This paper argues that Eastern Cham is not a counterexample to the CTH. There is a pragmatic distinction between (3a) and (3b), which is encoded in the grammar. What makes Eastern Cham different from Malagasy, Bahasa Indonesian, and Jarai is that the pragmatic distinction is not related to a phrase's *wh*-ness (i.e. *wh*- or focus-features). It

is due to *discourse connectedness* (DC), which imposes restrictions on the rhetorical structure of a discourse and where a given phrase is previously mentioned in that structure. Specifically, a phrase can be marked as DC if it is previously mentioned in a sentence that the current sentence is interpreted as explaining or elaborating upon (4). This definition is expanded in Section 3. *Wh*-phrases in derived positions, as in (3b) above, must be DC-marked.

- (4) DISCOURSE CONNECTEDNESS (DC): For individual  $x$  in sentence  $\phi$ ,  $x$  can be DC-marked  $m$  iff it is previously mentioned in sentence  $\psi$  and  $\phi$  is interpreted as explaining or elaborating on  $\psi$

DC-marking is not limited to *wh*-phrases. Moved phrases like *ʔʔ ni* ‘this mango’ in (5b) must also be DC-marked. The moved phrases in (3b) and (5b) share DC pragmatics, and the movement operations share syntactic properties. For these reasons, we conclude that the syntactic movement operation in (3b) and (5b) should be treated as a unitary phenomenon: DC-movement.

- (5) a. *kăw ʔɔʔ ɓăŋ ʔʔ ni*      b. *ʔʔ ni<sub>DC</sub> kăw ʔɔʔ ɓăŋ ʔʔ-ni*  
 1SG PROG eat mango this      mango this 1SG PROG eat  
 ‘I am eating this mango.’      ‘This mango, I am eating.’

We will show that DC-movement is an instance of  $\bar{A}$ -movement, driven by a probe on C. Assuming an Agree-based model of  $\bar{A}$ -movement, the C-probe searches for a specific feature (here, [DC]); the closest phrase is moved to Spec-CP due to an EPP-feature on that head (cf. Chomsky 2000). The DC-feature itself is introduced via a phonologically null DC-particle that merges with phrases, on analogy with focus- and Q-particles (e.g. Cable 2010). *Wh*-phrases are shown to be compatible with DC-marking only if the DC-particle merges below the D-head that introduces the computation of alternatives (in the sense of Rooth 1992).

In the end, the CTH is supported by Eastern Cham; apparent *wh*-movement is due to a syntactic feature not based on *wh* at all. Furthermore, DC provides a model by which *wh*-phrases can be marked for their discourse-pragmatic properties and could account for topicalization-like phenomena, especially those in which *wh*-phrases can apparently be topicalized. In a broader sense, it adds to the typology of  $\bar{A}$ -features and presents a means for syntax to be sensitive to discourse structure.

The paper proceeds as follows. Section 2 gives some relevant background on the Eastern Cham language and establishes that DC-movement is an instance of  $\bar{A}$ -movement by comparing it with  $\bar{A}$ -movement operations in general. Section 3 presents our proposal for the pragmatics of *discourse connectedness* and extends it to *wh*-phrases. Additionally, Section 3.3 examines the relation of DC, topicality, and Discourse/D-linking. The notion of topicality is shown to overgenerate the set of phrases that can be DC-marked, and it misses the discourse structural generalization. D-linking is shown to be distinct from DC-marking in Eastern Cham, as only a subset of DC-marked phrases can be D-linked, namely those with resumptive pronouns. Section 4 establishes the syntactic commonalities between the DC-movement of *wh*- and non-*wh*-phrases, based on the DC-movement of multiple phrases in the same clause and resulting locality effects. Finally, Section 5 shows that

DC-movement is orthogonal to the syntax of in-situ *wh*-phrases. In-situ *wh*-phrases must enter an Agree relation with  $C_Q$ , but they do not intervene on DC-movement.

## 2 Background

Section 2.1 presents brief background on the Eastern Cham language and methodology. Section 2.2 outlines general characteristics of  $\bar{A}$ -extraction and concludes that DC-movement is an instance of  $\bar{A}$ -movement.

### 2.1 The Eastern Cham language

Eastern Cham is an Austronesian language in the Malayo-Polynesian branch spoken across Vietnam, mostly in Cham villages in the Ninh Thuận and Bình Thuận provinces along the South-Central coast. According to the 2019 Vietnam Census, there are about 74,000 Cham people in Ninh Thuận and Bình Thuận, 19,000 in Ho Chi Minh City and neighboring Đồng Nai, where there are urban Eastern Cham communities, and smaller populations elsewhere (General Statistics Office of Vietnam 2020).<sup>1</sup> In these communities, there is near-universal bilingualism with Vietnamese (Baclawski Jr 2018b: 76). Owing to language contact and a prominent quasi-diglossia in the community, there is widespread inter- and intra-speaker variation (cf. Brunelle 2009; Baclawski Jr 2018b). Phonetic variation is abstracted in the data here through the use of broad phonological transcription.

Data for this paper were collected by the authors in Vietnam from 2015–2018 with six major consultants, among whom each data point has been checked by at least two. Of the six consultants, one was an older male Cham scholar, while the other five were young adults who had attended university. All were native speakers of Eastern Cham, born and raised in the Cham villages of Phan Rang, Vietnam. Despite their time away for university and bilingualism with Vietnamese, each consultant reported daily use of Eastern Cham and were readily able to produce the sentences and discourses elicited. Certain data points have been checked with some of 20 other consultants with a range of ages and schooling. No significant differences have been found regarding the core syntactic and pragmatic claims of this paper, with the exception of a small number of Vietnamese-dominant speakers who calqued Vietnamese syntax.

Fieldwork was conducted in the Cham villages near Phan Rang, Vietnam, in informal settings such as cafes, in order to encourage natural, colloquial Eastern Cham speech. In recording sessions, one of the younger consultants acted as translator, using a combination of English and Vietnamese. This translator was instructed to encourage natural, colloquial Eastern Cham speech. The data were elicited through grammaticality and felicity judgment tasks. Individual sentences were constructed, pieced together into dis-

---

<sup>1</sup>The Vietnam Census lists a total of about 175,000, but this figure includes the closely related Western Cham and Haroi. For example, 15,000 almost certainly represent Western Cham people in the Tây Ninh and An Giang provinces along the border with Cambodia, and 30,000 are reported in the Bình Định and Phú Yên provinces where Haroi is spoken.

courses, then assessed for their cultural acceptability and naturalness before assessed for pragmatic felicity. The data for this research are archived through the California Language Archive at the University of California, Berkeley, with data from 2018–2019 in a prearchival status.

In modern colloquial speech, Eastern Cham shares many typological characteristics with languages of Mainland Southeast Asia: it is a largely morphologically isolating SVO language with a tone or register system and generally monosyllabic roots (cf. Thurgood 1996, 1999; Brunelle 2009; Brunelle and Phú 2018). Concurrently, there is an ancient script tradition dating back at least to the 9th century CE that preserves a stage of the language before many subsequent developments. In the interlinearized examples throughout this paper, IPA transcription follows the Chamic linguistic tradition (e.g. Moussay 2006; Brunelle and Phú 2018). Open circles underneath consonants reflect a falling, breathy register on the following vowel, reflecting a historical devoicing sound change that led to tonogenesis/registerogenesis. There is a vowel length distinction in both Eastern Cham and Vietnamese indicated by a short vowel mark on short vowels (however, no short vowels in Vietnamese appear in this paper).

The example below in (6) gives a basic example of an Eastern Cham sentence.<sup>2,3</sup> Like Vietnamese and other Mainland Southeast Asian languages, kinship terms may function as pronominals (here, *těj* ‘younger sibling’ functioning as an addressee). Throughout this paper, these uses are indicated by brackets in sentence glosses. Finally, Eastern Cham is generally right-branching, with certain exceptions seen elsewhere in Southeast Asia, such as DP-final demonstratives and predicate-final modals and aspect markers. These exceptions have been argued to maintain right-branching through NP and predicate-fronting (cf. Simpson 2005 on the former; Baclawski Jr 2017, Simpson 2001 on the latter).

- (6) *těj*                      *młɛ?* *çũh* *pɔh*                      *tamkaj*                      *nǎn hu*  
 younger.sibling drop 7 CLF.ROUND watermelon that ROOT  
 ‘You[younger sibling] can drop those 7 watermelons.’

Eastern Cham exhibits basic characteristics of *wh*-in-situ languages. In an out-of-the-blue context, *wh*-phrases must remain in their base position (7a). Based just on this observation, it is conceivable that the *wh*-phrase does move, but to a low position, such as the right edge of the *vP*. Such an analysis has been proposed for certain dialects of Spanish (Uribe-Etxebarria 2002) and Hindi-Urdu (Manetta 2011: 87). Eastern Cham *wh*-phrases, however, are truly in-situ, as illustrated in (7b). The *wh*-phrase is bounded on the right by the indirect object and the root modal *hu* (which denotes either ability or

<sup>2</sup>The following abbreviations are used: ANIM = animate, CLF = (numeral) classifier, COMP = complementizer, COP = copula, EMPH = emphasis particle, EXIST = existential closure, EXP = experiential aspect, ITER = iterative aspect, NEG = negation, OBJ = object, PERF = perfective aspect PL = plural, POL = polite, PROG = progressive aspect, PRT = particle, ROOT = root modal (i.e. circumstantial/abilitative & deontic/permissive), Q.WH = *wh*-question, SG = singular, Y/N.Q = polar question particle. (VN) indicates a vocabulary item from Vietnamese (i.e. code-switching or names), pronounced in line with the Southern dialect of Vietnamese (cf. Hoàng 1989; Brunelle 2014), indicated by the corresponding transcription.

<sup>3</sup>In (6), the historical causative *pa-* is crystallized as a short *m-* (cf. *łɛ?* ‘to fall’). A small number of historically trisyllabic roots have been reduced to disyllabic roots such as *tamkaj* ‘watermelon’.

permission). Note that the modal *hu* triggers predicate raising (cf. Baclawski Jr 2017), so the *wh*-phrase is not moving to the edge of or outside of the predicate.

- (7) CONTEXT: Out-of-the-blue.
- a. *hi tɔʔ bǎŋ ɰeʔ / \*ɰeʔ hi tɔʔ bǎŋ ɰeʔ*  
 2SG PROG eat what what 2SG PROG eat  
 ‘What are you eating?’
- b. *kǎw [vP plɛj ɰeʔ ka nǐʔ sǐt nǎn] hu vP*  
 1SG give what to child little that ROOT  
 ‘What can I give to that little child?’

Cheng (1991, 1997) makes the typological observation that *wh*-in-situ languages have overt polar question particles, but *wh*-movement languages do not. Indeed, there is a polar question particle in Eastern Cham, *lɛj*, that is not found in *wh*-questions.

- (8) *hi tɔʔ bǎŋ vɔʔ lɛj*  
 2SG PROG eat ITER Y/N.Q  
 ‘Are you eating more/again?’

Given that Eastern Cham appears to be a regular *wh*-in-situ language, Cheng’s Clausal Typing Hypothesis predicts that there should not be true *wh*-movement.

## 2.2 $\bar{A}$ -extraction

As described in the introduction, *wh*-phrases can be moved to the left periphery in Eastern Cham. This operation (DC-movement) is an instance of  $\bar{A}$ -movement that shares a variety of characteristics with other  $\bar{A}$ -movement operations like relativization and presentational cleft constructions.<sup>4</sup> Note that throughout, the DC-movement of non-*wh*-phrases shares the same characteristics. First, DC-movement is unbounded, as expected for  $\bar{A}$ -movement in general (Chomsky 1973). A DC-marked phrase in an embedded clause can be moved to its immediate left periphery (9a) or undergo successive cyclic movement to a higher left periphery (9b).

- (9) a. *kǎw hnǐj han ni<sub>DC</sub> t<sup>h</sup>uŋm<sup>312</sup> ʔa nǐʔ mɛj sǐt nǎn*  
 1SG think cake this Thuận(VN) invite child female small that  
*maj bǎŋ han-ni*  
 come eat  
 ‘Thuận invited that little girl to come eat this cake.’ (DC-movement)
- b. *han ni<sub>DC</sub> kǎw hnǐj t<sup>h</sup>uŋm<sup>312</sup> ʔa nǐʔ mɛj sǐt nǎn*  
 cake this 1SG think Thuận(VN) invite child female small that  
*maj bǎŋ han-ni*  
 come eat  
 ‘Thuận invited that little girl to come eat this cake.’ (DC-movement)

<sup>4</sup>Baclawski Jr (2019: 47) presents an additional argument for  $\bar{A}$ -movement based on the complementizer *po*. However, the complementizer does not align with other work on Eastern Cham, so it is excluded here.

Second, prepositions are dropped when argument preposition phrases are  $\bar{A}$ -moved (cf. ‘p-drop’: Hoonchamlong 1991: 134 on Thai; Wang 2007 on Mandarin Chinese; Sato 2011 on Indonesian). For example, in (10a), the preposition *ka* is obligatory with in-situ indirect objects in ditransitive constructions. When these phrases are  $\bar{A}$ -moved, however, *ka* cannot appear in any position (10b–c). We follow previous analyses of p-drop in positing that there is a pronunciation rule that deletes prepositions whose complements have been  $\bar{A}$ -moved. The p-drop also has the effect of restricting  $\bar{A}$ -movement to nominal arguments; adjuncts cannot be  $\bar{A}$ -moved, as described below for hanging topics.

- (10) a. hi plěj han năn \*(ka) t<sup>h</sup>ěj  
 2SG give cake that to who  
 ‘Who did you give that cake to?’
- b. t<sup>h</sup>ěj<sub>DC</sub> hi plěj han năn (\*ka) t<sup>h</sup>ěj  
 who 2SG give cake that to  
 ‘Who did you give that cake to?’ (DC-movement)
- c. hu t<sup>h</sup>a nǐ? měj kăw plěj han ni (\*ka) nǐ? měj  
 EX.COP one child female 1SG give cake this to  
 ‘There is a girl who I [will] give this cake to.’ (Cleft)

Third,  $\bar{A}$ -movement is consistently sensitive to strong islands, such as complex DPs (11a–b), in line with  $\bar{A}$ -movement cross-linguistically. While there is known to be inter-speaker variation with regard to grammaticality judgments of island constraints (e.g. Szabolcsi 2006), the facts above appear to be robust for Eastern Cham. These examples reflect the consistent judgments of six consultants. One consultant accepted in-situ and moved *wh*-phrases in strong and weak islands, and another in weak, but not strong islands. Impressionistically, these last two consultants were often permissive with judgments in general, and we will set them aside and focus on the majority pattern reported here.

- (11) a. \*jaŋ hlěj hi plěj dow<sup>21</sup> bǎŋ jaŋ hlěj ŋǎ?  
 person which 2SG buy stuff(VN) eat make  
 INTENDED: ‘Which person do you buy the food they make?’ (DC-movement)
- b. \*hu t<sup>h</sup>a jaŋ hi plěj dow<sup>21</sup> bǎŋ t<sup>h</sup>a jaŋ ŋǎ?  
 EX.COP one person 2SG buy stuff(VN) eat make  
 INTENDED: ‘Which person do you buy the food they make?’ (Cleft)

Fourth,  $\bar{A}$ -movement gives rise to weak crossover effects, again in line with  $\bar{A}$ -movement cross-linguistically, or minimally the absence of A-movement (Ruys 2004). Weak crossover occurs when a DP cannot move over a coreferential pronoun, even though that pronoun does not c-command the base position of the DP. The base order of arguments in Eastern Cham ditransitive predicates is direct object–indirect object, as in (12).<sup>5</sup> The direct object can bind a pronoun within the indirect object, but not vice versa (b).

<sup>5</sup>Objects can be shifted to result in other relative orders, which can be diagnosed by clause-final modals and aspect markers (Baclawski Jr 2017).

- (12) a. *kǎw mjan [nǎʔ mjaw nǎn]<sub>i</sub> ka po ju<sub>i</sub>*  
 1SG return CLF.ANIMAL cat that to owner 3.ANIM  
 ‘I returned that kitten to its owner.’
- b. \**kǎw mjan ɓɔp<sup>45</sup> ju<sub>i</sub> ka pu<sub>i</sub>*  
 1SG return wallet(VN) 3.ANIM to Phú  
 INTENDED: ‘I returned his wallet to Phú.’

If an indirect object is  $\bar{A}$ -moved over a direct object, a crossover context arises. For example, the DP *t<sup>h</sup>a sīt pu mǎn* ‘only Phú’ crosses over the direct object in (13a), which contains a pronoun. The ungrammaticality of the *i* index on the pronoun indicates that coreference is impossible; the pronoun can only refer to someone else in the context. By contrast, there is no crossover in (13b), as *t<sup>h</sup>a sīt mjaw mǎn* ‘only the cat’ always c-commands the pronoun. Thus, no weak crossover effect obtains; the pronoun may corefer with the  $\bar{A}$ -moved DP.

- (13) a. *t<sup>h</sup>a sīt pu<sub>i</sub> mǎn<sub>DC</sub> kǎw mjan ɓɔp<sup>45</sup> ju<sub>\*i/j</sub> ka t<sup>h</sup>a sīt pu mǎn*  
 only Phú EMPH 1SG return wallet(VN) 3.ANIM to  
 ‘I only returned Phú his wallet.’ (DC-movement)
- b. *t<sup>h</sup>a sīt mjaw<sub>i</sub> mǎn<sub>DC</sub> kǎw mjan t<sup>h</sup>a sīt mjaw mǎn ka po ju<sub>i/j</sub>*  
 only cat EMPH 1SG return to owner 3.ANIM  
 ‘I only returned the cat to its owner.’ (DC-movement)

$\bar{A}$ -movement in Eastern Cham, then, can be unified based on five characteristics: unboundedness, p-drop, restriction to DPs, island sensitivity, and weak crossover effects. These contrast with a hanging topic construction. Hanging topics are marked prosodically by a pause. Distributionally, they lack p-drop, restriction to DPs, and island sensitivity (14a–b). A full analysis of this hanging topic construction is left for future research, though it is likely that hanging topics are base generated in the left periphery.

- (14) a. *pu<sub>i</sub> // hi ʔa hu*  
 Phú 2SG invite ROOT  
 ‘Phú, you can invite him.’ (Hanging topic)
- b. *mǎŋ owŋm<sup>45</sup> // jūt nǔm ʔja ce ni*  
 with straw(VN) friend drink water tea this  
 ‘With [a] straw, you[friend] drink this tea.’ (Hanging topic)

This section has presented evidence that DC-movement in Eastern Cham is an instance of  $\bar{A}$ -movement. The following section examines the pragmatics of *discourse connectedness* and concludes that it unifies *wh*- and non-*wh*-phrases.

### 3 Discourse connectedness

This section proposes that DC-movement imposes a pragmatic restriction on the moved phrase in the form of *discourse connectedness* (DC), whether it is a *wh*- or non-*wh*-phrase.



For an individual to be DC-marked, it must be previously mentioned in the discourse, and the discourse must be structured in such a way that the current sentence is interpreted as explaining or elaborating upon the sentence where the individual was previously mentioned (15). The previous mention component of this characterization is found not to be a surface identity constraint, but one of referential accessibility along the lines of definiteness. The elaboration or explanation component is argued to reflect the broad category of rhetorical relation known as *discourse subordination*.

- (15) DISCOURSE CONNECTEDNESS (DC): For individual  $x$  in sentence  $\phi$ ,  $x$  can be DC-marked iff it is previously mentioned in sentence  $\psi$  and  $\phi$  is interpreted as explaining or elaborating on  $\psi$

This characterization takes after López’s (2009: 47) account of [+Anaphor] as the feature underlying clitic right-dislocation, especially of pronouns, and one of the features underlying clitic left-dislocation in Catalan. However, López’s (2009) [+Anaphor] is incompatible with *wh*-phrases, whereas Eastern Cham DC is incompatible with pronouns. See Baclawski Jr (2019: 167) for additional comparison between Catalan clitic right-dislocation and Eastern Cham DC-movement. For the purposes of this paper, the characterization in (15) will be used as a heuristic. See Baclawski Jr (2019: 39) for a formal account of DC, defined in terms of events in a discourse and their respective participants, and implemented via a presupposition that checks whether the phrase the DC-particle combines with satisfies the relevant conditions.

Section 3.1 examines the basic case of DC-marking with non-*wh*-phrases. The previous mention and discourse subordination components of DC are explored, as are the kinds of phrases that can be DC-marked. In addition to DPs that denote individuals, NPs that denote properties are found to be capable of DC-marking as kinds vis-à-vis Chierchia’s (1997) ‘down’ operator. This correctly predicts that generics and quantifier phrases may be DC-marked. Section 3.2 extends DC-marking to *wh*-phrases. In this case, a subpart of the *wh*-phrase, namely its restrictor, must be DC-marked in order for DC-marking not to clash with the computation of *wh*-alternative sets. Finally, Section 3.3 compares DC with the information structural notions of topicality and Discourse/D-linking. Topicality undergenerates DC-marking in Eastern Cham, as phrases argued to be anti-topical, such as downward-entailing quantifier phrases, can be DC-marked. D-linking overgenerates, as only a subset of DC-marked phrases exhibit the properties of D-linking: those with resumptive pronouns. Additionally, both information structural notions fail to capture the discourse structural generalization of DC. Taken together, DC represents a pragmatic phenomenon that must make reference to the rhetorical structure of discourse.

### 3.1 Non-*wh*-phrases

Discourse connectedness has two requisite components: previous mention and an explanation or elaboration relation between two sentences in a discourse.<sup>6</sup> For an illustration

---

<sup>6</sup>For the purposes of this section, the relevant unit of discourse will be the sentence. Therefore, for ease of exposition, the term ‘sentence’ will be used in the description of discourse relations and discourse structure, in line with much of the literature (cf. Webber 1988: 2). A more appropriate descriptor of the

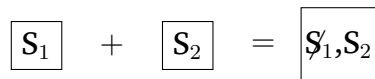
of these components, consider (16). The phrase *?iŋ ?əŋ nəŋ* ‘that frog’ in (16b) is previously mentioned in (16a).<sup>7</sup> Note that DC-movement is consistently optional. Throughout this section, in-situ phrases are felicitous in every context.

- (16) a. t<sup>h</sup>uŋm<sup>312</sup>    ʈəʔ    ŋǎʔ    ?iŋ ?əŋ    nəŋ  
 Thuận(VN)    PROG    make    frog    that  
 ‘Thuận is cooking that frog.’
- b. ?iŋ ?əŋ    nəŋ<sub>DC</sub>    ɲu    ʈəʔ    ŋǎʔ    ?iŋ ?əŋ    nəŋ    ŋi    lo  
 frog    that    3.ANIM    PROG    make    be.delicious    very  
 ‘That frog, he is cooking very well [Lit: deliciously].’

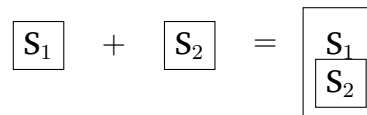
Together, these two sentences are also in an elaborating discourse relation. According to theories of rhetorical relations and discourse structure, sentences can have relations between one another that reflect the focus of attention and the flow of information. Two major categories of discourse moves are *coordinating* and *subordinating discourse relations* (cf. Grosz & Sidner 1986 on ‘satisfaction-precedence’ and ‘dominance’; Fabricius-Hansen & Ramm 2008 and references therein on these terms).<sup>8,9</sup> In a coordinating discourse relation, a sentence is added to the discourse and supplants the prior as the focus of attention (Figure 1a). By contrast, in a subordinating discourse relation, a sentence is interpreted as expanding upon a prior sentence, such that both remain active points of attention. Subordinating discourse relations result in hierarchical relations, reflected in Figure 1b, such that the prior sentence is superordinate and the current sentence subordinate.

Figure 1: Models of sentence relations in a discourse

(a) Coordinating discourse relation



(b) Subordinating discourse relation



A full background on discourse relations and theories of discourse structure is beyond the scope of this paper. Of relevance for this section, these theories have identified a set of subordinating discourse relations that provide diagnostics for discourse connectedness (cf. Asher & Lascarides 2003: 44 and Asher & Vieu 2005 on discourse subordination; Mann & Thompson 1988 on nucleus-satellite relations). Prototypical subtypes of dis-

unit of discourse in theories of discourse structure may be ‘logical form’ (e.g. Asher & Lascarides 2003: 110). However, Baclawski Jr (2019) argues that discourse connectedness is computed between events.

<sup>7</sup>The pronoun *ɲu* can also be thought of as previously mentioned in (16a). However, Eastern Cham has no means of DC-marking pronouns, matrix subjects (see later this section).

<sup>8</sup>Coordination and subordination in the discourse sense are theoretically and descriptively distinct from coordination and subordination in the syntactic sense (cf. Fabricius-Hansen & Ramm 2008).

<sup>9</sup>To be sure, there are exceptional discourse moves that do not fit into either of these categories, such as background information and corrections. In the remainder of this section, a broad distinction will be used between subordinating and non-subordinating relations, the latter of which includes coordinating and exceptional relations.

course subordination include any kind of explanation or elaboration.<sup>10</sup> In other words, if a sentence is interpreted as an elaboration or explanation of another, those two sentences are in a subordinating discourse relation (17). In this section, a down arrow  $\Downarrow$  will be used to indicate a subordinating discourse relation between a superordinate and subordinate sentence (and  $\nabla$  to the absence of discourse subordination).

(17) DISCOURSE SUBORDINATION ( $\Downarrow$ ): Sentence  $\phi \Downarrow$  sentence  $\psi$  iff  $\psi$  is interpreted as an elaboration or explanation of  $\phi$  and  $\phi$  remains open for further discussion after  $\psi$

Returning to the Eastern Cham example repeated below, (18b) is interpreted as an elaboration on (18a). In the elicitation of this context, it was made clear that the speaker was observing an act of cooking and then commenting on the cooker’s effectiveness within that event. Subsequent discourse may continue to discuss the effectiveness of the cooking, or return to the more general act of cooking. In other words, both (a) and (b) remain open for further discussion. Therefore, (a  $\Downarrow$  b).

(18) a. t<sup>h</sup>uŋm<sup>312</sup> tɔʔ ɲǎʔ ʔiŋ ʔɔŋ nəŋ  
 Thuận(VN) PROG make frog that  
 ‘Thuận is cooking that frog.’

b. ʔiŋ ʔɔŋ nəŋ<sub>DC</sub> ju tɔʔ ɲǎʔ ʔiŋ-ʔɔŋ-nəŋ ɲi lo  
 frog that 3.ANIM PROG make be.delicious very  
 ‘That frog, he is cooking very well [Lit: deliciously].’ (a  $\Downarrow$  b)

Why, then, is *ʔiŋ ʔɔŋ nəŋ* ‘that frog’ DC in (18b)? It has satisfied both conditions of DC: the phrase must be previously mentioned, and there must be a subordinating discourse relation. More precisely, the phrase must be previously mentioned in a superordinate sentence. Throughout this paper, the subscript *DC* will be used when a phrase satisfies these conditions.

DC-movement is infelicitous when the DC conditions above are not satisfied.<sup>11</sup> In (16b’), which continues from (16a) above, *ʔiŋ ʔɔŋ nəŋ* ‘that frog’ is previously mentioned, but (b’) is not interpreted as an elaboration or explanation. Instead, it describes the event of eating. (16b’’) presents an example where there is no suitable antecedent.<sup>12</sup> Accordingly, DC-movement is infelicitous.

(16) b’.# ʔiŋ ʔɔŋ nəŋ ju tɔʔ ɲǎŋ ʔiŋ-ʔɔŋ-nəŋ  
 frog that 3.ANIM PROG eat  
 INTENDED: ‘[Now,] he is eating that frog.’ (a  $\nabla$  b’)

<sup>10</sup>Explanation and elaboration correspond with ‘causal’ and ‘constitutive explanation’, respectively. A causal explanation explains why something is the case, while a constitutive explanation expands upon the parts or organization of things such that they can lead to a causal explanation (Salmon 1984; Ylikoski 2013). For the remainder of this paper, we use ‘explanation’ and ‘elaboration’ as a shorthand for these concepts.

<sup>11</sup>Note that here and throughout this section, the sentences are all felicitous if the bolded phrase is left in-situ. Thus, the sentence as a whole is not degraded in each context.

<sup>12</sup>The word *ʔiŋ ʔɔŋ* refers to a species of true frog common in rice paddies, possibly *Fejervarya limnocharis*. The word *kɪwʔ* is a more general term for frog. For ease of exposition, the former is translated as ‘frog’ and the latter as ‘kiép’.

b''.#kǐw? nǎn ju      çin tɔ?    ɲǎ?    kǐw?nǎn  
 kiép that 3.ANIM also PROG make

INTENDED: 'That kiép[other kind of frog], he is also cooking.' (a ↓ b'')

If consultants are presented with (16b'–b''), they may accept them, but only if additional discourse is present, for example, if a speaker had asked prior about different kinds of frogs. In these cases, superordinate sentences are added such that the moved phrase can satisfy the DC conditions.

Based on just the positive example of DC-marking above (16b), one might hypothesize that only an entailment relation is necessary between the antecedent and anaphor sentences. This is not the case, however. Any two sentences can license DC-marking, as long as one is interpreted as an explanation or elaboration, as in (18).

(18) a. ɬǎŋ nǎn kʰəh      lo  
 door that be.beautiful very  
 'That door is very beautiful.'

b. ɬǎŋ<sub>DC</sub> kǎw cǐh    ɬǎŋ pan    çaw  
 door 1SG paint      color blue  
 'The door, I painted blue.'

(a ↓ b)

The examples above involve single-speaker narration, but DC-movement also arises in conversations. For example, (19) illustrates two different answers to the same polar question. If an answer elaborates upon a question by being overinformative or underinformative, the answer is discourse subordinate to the question (cf. Asher & Lascarides 2003: 320 on 'q-elaboration' and 'partial answers'). The answer in (19b) can also be interpreted as an explanation of a covert answer to the polar question (i.e. 'Do you want to eat mango? [No. Why?] I already ate mango.'). Accordingly, DC-movement is felicitous in (19b), as ʔǎ? 'mango' is mentioned in the superordinate question. By contrast, direct answers have no such subordinating relation with their respective questions. In (b'), ʔǎ? 'mango' cannot be DC-moved, as there is no superordinate sentence in the discourse.

(19) a. hi hu      ɲŋ    ɬǎŋ ʔǎ?    lǛj  
 2SG EX.COP want eat mango Y/N.Q  
 'Do you want to eat mango?'

b. ʔǎ?<sub>DC</sub> kǎw ɬǎŋ ʔǎ? jə  
 mango 1SG eat      already  
 'I already ate mango.'

(a ↓ b)

b'.#ʔǎ?      kǎw ɲŋ    ɬǎŋ ʔǎ?  
 mango 1SG want eat  
 'I want to eat mango.'

(a ↓ b')

Thus far, all the examples have involved a surface identity relation between the DC-marked phrase and its antecedent. However, the previous mention component of DC is not one of surface identity, but semantic equivalence or accessibility. For example,

generics can be DC-marked if the previous mention is either a generic as well, or an individual referent (20).

- (20) a.  $t^h u\eta m^{312}$   $t\grave{o}?$   $\eta\check{a}?$   $\eta i\eta$   $\eta\grave{o}\eta$   $n\grave{a}n$   
 Thuận(VN) PROG make frog that  
 ‘Thuận is cooking that frog.’
- b.  $\eta i\eta$   $\eta\grave{o}\eta_{DC}$   $ju$   $\eta\check{a}?$   $\eta i\eta$   $\eta\grave{o}\eta$   $\eta i$   $lo$   
 frog 3.ANIM make be.delicious very  
 ‘Frog, he cooks very well [Lit: deliciously].’ (a ↓ b)

Bridged definites also license DC-marking. In (21),  $t^h a\eta$   $n\grave{a}n$  ‘that house’ is a sufficient previous mention for  $\eta\check{a}\eta$  ‘[the] door’.

- (21) a.  $t^h a\eta$   $n\grave{a}n$   $k^h\grave{a}h$   $lo$   
 house that be.beautiful very  
 ‘That house is very beautiful.’
- b.  $\eta\check{a}\eta_{DC}$   $k\check{a}w$   $c\grave{i}h$   $\eta\check{a}\eta$   $p\grave{a}n$   $\zeta\grave{a}w$   
 door 1SG paint color blue  
 ‘The door, I painted blue.’ (a ↓ b)

It is worth noting that pronouns cannot be DC-moved in Eastern Cham, even though they could in theory satisfy the DC conditions. For example, the pronoun  $ju$  cannot be DC-moved in (22b), regardless of whether it is stressed, even though it refers to an individual in a superordinate sentence. Perhaps pronouns in Eastern Cham are subject to distinct discourse requirements from DC. For instance, perhaps they require *discourse coordination*, in contrast with discourse subordination (cf. Fabricius-Hansen & Ramm 2008, and references therein). Further work is needed to test the role of pronouns.

- (22) a.  $hi$   $t^h\check{a}w$   $n\grave{u}j\eta$   $n\grave{a}n$   $l\check{e}j$   
 2SG know person that Y/N.Q  
 ‘Do you know that person?’
- b.  $\ast\{ju/NU\}$   $k\check{a}w$   $k\grave{o}?$   $\eta\eta$   $m\check{i}\eta$   $p\grave{o}j$   
 3.ANIM 1SG meet in yesterday  
 INTENDED: ‘I met him/HIM yesterday.’ (a ↓ b)

Finally, as mentioned in Section 2.2, DC-marking is limited to nominals in Eastern Cham. We have seen DC-marking of nominals that refer to individuals (e.g. ‘that frog’) and those that refer to properties (e.g. ‘mango’). We posit that a generic like  $\eta\check{a}?$  ‘mango’ is interpreted as a kind, through Chierchia’s (1997) ‘down’ operator that transforms a property of type  $\langle e, t \rangle$  to a corresponding kind of type  $e$  (23a). By doing so, a unified semantics of the DC-particle can be used: when it combines with an individual, it checks if that individual satisfies the DC conditions; when it combines with a property, it checks if the corresponding kind satisfies the DC conditions.

- (23) For property P,  $\uparrow P$  denotes the corresponding kind (after Chierchia 1997: 77)

In addition to generics, nominals that denote properties can be DC-marked in the form of quantifier phrases, where the DC-marked phrase is a subpart of a DP. A quantifier phrase with the universal quantifier *p̥ih* can be DC-moved, as in (24b), even though it is the restrictor *nĩ? t<sup>h</sup>un* ‘animal’ that is DC-marked. Baclawski Jr (2019: 107) explores inventory forms and partitive constructions, where a subpart of a DP is DC-marked and moved to the left edge of the DP itself.

- (24) a. *l̥əm nĩ? t<sup>h</sup>un p̥ă? ni...*  
of animal at this  
‘Of the animals here...’
- b. *p̥ih nĩ? n̥n<sub>DC</sub> k̥ăw j̥ɔŋ p̥ih-nĩ?-n̥n*  
all CLF.ANIMAL that 1SG raise  
‘I raise all of them.’ (a ↓ b)

In sum, DC-marking of non-*wh*-phrases requires a discourse subordination rhetorical structure where the DC-marked phrase is accessible to a phrase in a superordinate sentence. Nominals that refer to both individuals and properties can be DC-marked, which allows for generics and quantifier phrases to be DC-moved.

### 3.2 *Wh*-phrases

This section demonstrates that *wh*-phrases can be DC-marked under the same conditions as in the previous section, repeated below as (25). It is shown that DC-marking is compatible with *wh*-phrases, as long as it is the restrictor of the *wh*-phrase that is DC-marked, not the whole *wh*-phrase itself.

- (25) DISCOURSE CONNECTEDNESS (DC): For individual *x* in sentence  $\phi$ , *x* can be DC-marked iff it is previously mentioned in sentence  $\psi$  and  $\phi$  is interpreted as explaining or elaborating on  $\psi$

First, (26) illustrates the DC conditions applied to a *wh*-phrase. In (26a), two individuals are mentioned, a pot of frog and a pot of *kiép*. In (26b), the *wh*-phrase *k̥ɔ? k̥e?* ‘what pot’ is interpreted as a set containing those two pots (i.e. *which of those two pots*). In (26b’), the bare *wh*-phrase *k̥e?* ‘what’ can also be DC-moved, which refers either to the contextually salient set of pots, or maximally to the superset of non-human entities.<sup>13</sup>

- (26) a. *mi k̥ăw t̥ɔ? t̥ũ? ʔiŋ ʔɔŋ t<sup>h</sup>a k̥ɔ? h̥ɔŋm̥ k̥iŋ? t<sup>h</sup>a k̥ɔ?*  
father 1SG PROG boil frog 1 pot with kiép 1 pot  
‘My father is boiling one pot of frog and one of kiép.’
- b. *j̥ă? ni k̥ɔ? k̥e?<sub>DC</sub> oŋ n̥n t̥ɔ? ŋă? k̥ɔ?k̥e? n̥n*  
now pot what old.man that PROG make that  
‘Now, what pot is that old man making [working on]?’ (a ↓ b)

<sup>13</sup>Sentence-final *n̥n* in these examples marks clause-level deixis and does not form a constituent with the *wh*-phrase.

b'. jǎʔ ni kɛʔ<sub>DC</sub> oŋ nǎn tɔʔ ŋǎʔ kɔʔ kɛʔ nǎn  
 now what old.man that PROG make that  
 ‘Now, what is that old man making [working on]?’ (a ↓ b’)

In terms of discourse structure, (26b) is naturally interpreted as an elaborating question on (26a). In this context, the father is in the process of cooking two pots on a stove, but in that moment is stirring one of them. Here, the speaker asks for an elaboration of the cooking event: within the broader event of cooking, which pot is he working on right now? When this context is made explicit, DC-movement of *kɔʔ kɛʔ* ‘what pot’ is accepted. This is because there is some previous mention of the phrase in a superordinate sentence (26a).

(26b’) illustrates the absence of discourse subordination. Here, the question is asked after the cooking has been completed and the father has transitioned to eating. The speaker is unclear which kind of frog the father is eating in that moment. In this context, DC-movement of *kɔʔ kɛʔ* ‘what pot’ is rejected. (26b’’) illustrates the absence of previous mention. If the *wh*-phrase refers to a different set of pots, here the set of pots the grandmother is cooking, the result is infelicitous. Therefore, there must be discourse subordination and previous mention.

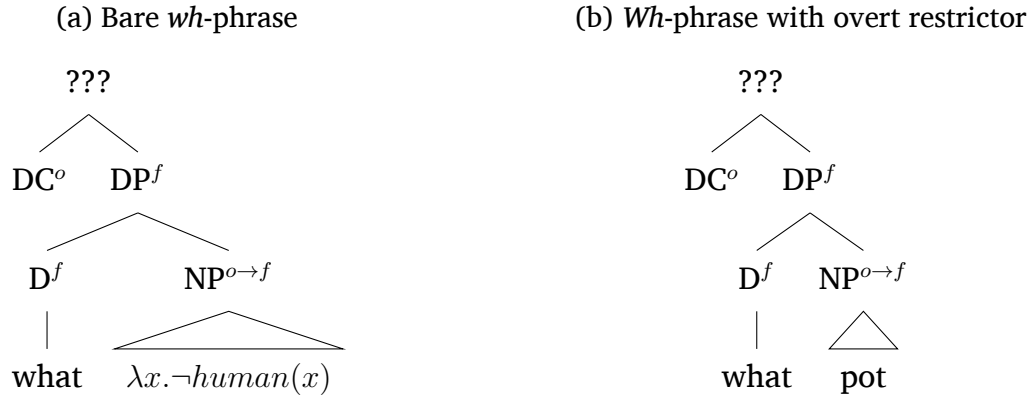
(25) b'.#jǎʔ ni kɔʔ kɛʔ oŋ nǎn tɔʔ bǎŋ kɔʔ kɛʔ nǎn  
 now pot what old.man that PROG eat that  
 INTENDED: ‘Now, what pot is that old man eating?’ (a ↓ b’)

b''.#jǎʔ ni kɛʔ muʔ nǎn tɔʔ ŋǎʔ kɛʔ nǎn  
 now what old.woman that PROG make that  
 INTENDED: ‘Now, what is that old woman making [working on]?’ (a ↓ b’)

When presented with (26b’–b’’), speakers accept them only if the father is eating from the pots as part of the event of cooking (i.e. tasting to check if the food is done), or if there were prior discourse about multiple people and their cooking. These metalinguistic judgments are instructive, as the speakers are adding additional discourse material, which add superordinate sentences that can then license DC-marking.

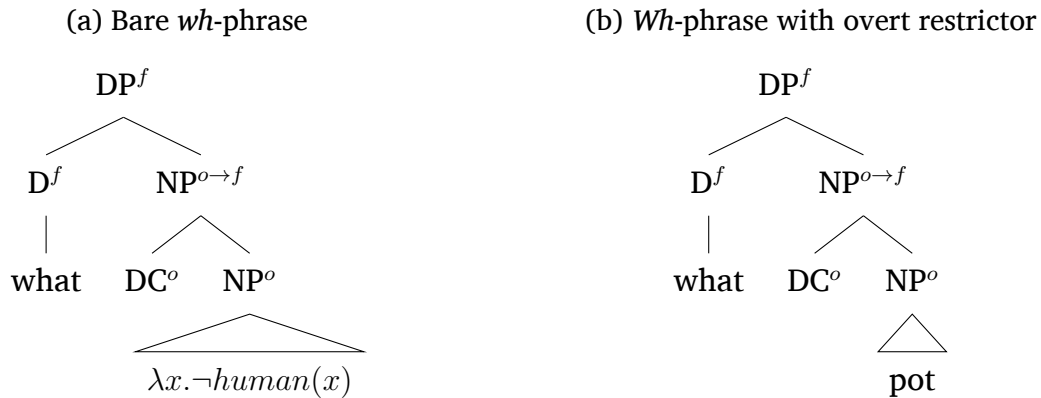
If one assumes an Alternative Semantic account of *wh*-phrases (Rooth 1992), an incompatibility arises between *wh*-phrases and DC-marking. For most of a derivation, *wh*-phrases do not have ordinary semantic values. They only function as sets of alternatives along a focus dimension of meaning, until they combine with a focus operator that allows them to be interpreted. The DC conditions described above assume that a phrase has an ordinary semantic interpretation, rendering it impossible to determine if *wh*-phrases can be DC-marked. In Figure 2a–b, ordinary semantic values are represented with <sup>o</sup> and focus semantic values are represented by <sup>f</sup>. We assume that bare *wh*-phrases like *what* are considered to be specified to take a null property as their NP complement (e.g.  $\lambda x. \neg human(x)$ ). The structure of *who* can then be split into a *wh*-determiner and an NP-restriction (Figure 2a). A *wh*-phrase with an overt restrictor is shown in Figure 2b. In both cases, there is no way for a DC-particle to combine with the top DP.

Figure 2: Incompatibility between DC-particle and *wh*-phrases



DC-marking is compatible with the restrictor of *wh*-phrases. In Figure 2a–b, the NP restrictor must eventually be interpreted along the focus dimension of meaning to combine with the D-head. Before then, however, it does have an ordinary semantic interpretation. We propose that DC-particles can combine with *wh*-restrictors, before the D-head initiates the computation of alternatives. Figure 3 presents this analysis. Figure 3a is analogous to *which one* in English such that *one* is DC-marked. Both Figure 3a–b are analogous to DC-marked quantifier phrases, as shown in Section 3.1 above.<sup>14</sup>

Figure 3: Compatibility between DC-particle and *wh*-restrictors



The DC-marking mechanism above predicts that focussed phrases can also be DC-marked, as they have comparable structure to *wh*-phrases. This prediction is borne out. In (26b), the generic *paw* ‘water buffalo’ is DC-marked, while the whole focus phrase including the focus particle *t<sup>h</sup>a sīt* and emphatic *mīn* is DC-moved. The NP, like *wh*-phrases is DC-marked

<sup>14</sup>DC-marking of individuals is also possible within *wh*-phrases. Discourse/D-linked *wh*-phrases have been argued to contain an embedded DP that refers to a contextual antecedent (e.g. Boeckx & Grohmann 2004), which must always refer to a plurality, as the *wh*-phrase needs a set of alternative answers greater than one (cf. Evans 1980; Heim 1982 on E-type pronouns). Section 3.3 shows DC-marking of D-linked *wh*-phrases that likely should be analyzed as DC-marking of an individual referent along these lines.



before the focus particle introduces the computation of alternatives.<sup>15</sup>

- (26) a. ɭəm nĩʔ t<sup>h</sup>un pǎʔ ni...  
of animal at this  
‘Of the animals here...’
- b. t<sup>h</sup>a sīt p̚aw mĩn<sub>DC</sub> kǎw jɔŋ t<sup>h</sup>a sīt p̚aw mĩn  
only water.buffalo EMPH 1SG raise  
‘I only raise water buffalo.’ (a ↓ b)

Thus, *wh*-phrases can be DC-marked, as long as their restrictors can combine with DC-particles before they are transformed to sets of alternatives along the focus dimension of meaning. The next section turns to topicality and D-linking to examine if DC can be reduced to information structural notions. DC is shown not to be reducible to topicality or D-linking.

### 3.3 Topicality & D-linking

Because of the conceptual similarity between DC, topicality, and Discourse/D-linking, it is worth investigating to what extent they overlap. A DC-marked *wh*-phrase, as laid out in the previous section, requires previous mention in a superordinate sentence in the discourse. Topics are typically described in terms of old information or previous mention and aboutness (i.e. the phrase about which a sentence is organized; e.g. Reinhart 1981, among many others). D-linked *wh*-phrases are usually characterized as alternative sets saliently shared by the speaker and addressee (Pesetsky 1987; Comorovski 1996: 2; but cf. Wiltschko 1997 for problems). A growing literature acknowledges that D-linked *wh*-phrases behave syntactically like topics in a variety of languages (e.g. Polinsky 2001; Grewendorf 2012). D-linking has also been explicitly argued to condition *wh*-ex situ in languages like Mandarin (Pan 2014). Topicality and D-linking are information structural notions, in that they make reference to shared information at a point in time and the organization of information a given sentence. This is distinct from DC, which makes reference to the organization of multiple sentences in a discourse.

Our analysis of DC predicts that it is orthogonal to topicality and D-linking, as DC does not require contextual salience, and it uniquely imposes a discourse structural requirement. Based on language-internal evidence, DC-marking in Eastern Cham is orthogonal to both. First, indefinites, such as bare *wh*-phrases and downward entailing quantifiers are generally argued to be anti-topical, in that they cannot be topicalized in many languages (e.g. Ebert 2009). By these diagnostics, DC-phrases do not align with topics, and DC-movement cannot be described purely as topicalization. Bare *wh*-phrases can be DC-moved, as described in the previous section, as can downward entailing quantifiers like *kiʔ həŋ*<sup>33</sup> ‘less than’ (27).

<sup>15</sup>Note that the focus operator *t<sup>h</sup>a sīt* literally translates to ‘one small’. For many speakers, the form is no longer decomposable, as it has coalesced to *cīt*. The emphatic particle *mĩn* occurs with the focus operator *t<sup>h</sup>a sīt* and seems to indicate the right edge of its scope.

- (27) a. hi ʔa lo nujh lěj  
 2SG invite many person Y/N.Q  
 ‘Did you invite many people?’
- b. kiʔ hən<sup>33</sup> mi jaŋ<sub>DC</sub> kăw ʔa kiʔ hən<sup>33</sup> mi jaŋ maj pă ni  
 few exceed(VN) five person 1SG invite come here  
 ‘I invited less than five people to come here.’ (a ↓ b)

Turning to D-linking, Pesetsky (1987: 107) argues that the form of a *wh*-phrase determines its D-linking specification. *Wh*-phrases of the form *which X* are taken to be obligatorily D-linked (i.e. lexically specified as such; 28a). Bare *wh*-phrases are optionally D-linked, in that a D-linked reading can be coerced, given an appropriate context (28b). *What X* is typically described as non-D-linked, with a D-linked reading only salvageable in very specific contexts, such as with an overt partitive ((28c); Pesetsky 1987: fn.36; Wiltschko 1997: 113). And finally, *wh-the-hell* are described as ‘aggressively non-D-linked’, never D-linked ((28d); cf. den Dikken & Giannakidou 2002).

- (28) CONTEXT: Some people<sub>i</sub> entered the room...
- a. Which (ones)<sub>i</sub> did Antonia talk to? [D-linked]
- b. ʔWho<sub>i</sub> did Antonia talk to? [Optionally D-linked]
- c. #What ones<sub>i</sub> did Antonia talk to? [Non-D-linked]
- d. \*Who the hell<sub>i</sub> did Antonia talk to? [Aggressively non-D-linked]

DC-movement of *wh*-phrases in Eastern Cham does align with some basic predictions of D-linking. It is infelicitous out-of-the-blue. Aggressively non-D-linked *wh*-phrases cannot be DC-moved (29a). DC-moved *wh*-phrases are also often translated as D-linked *wh*-phrases in both English and Vietnamese (e.g. 29b). However, these data points can also be explained by discourse connectedness. DC phrases require antecedents in the discourse, something that out-of-the-blue contexts and aggressively non-D-linked *wh*-phrases lack. As for (29b), perhaps the closest translation equivalent of Eastern Cham DC-movement in English and Vietnamese is D-linking.

- (29) a. {\*} hi tɔʔ ŋăʔ {mbroj kɛʔ}  
 2SG PROG do crazy what  
 ‘What the hell are you doing?’
- b. kɛʔ hi tɔʔ bǎŋ  
 what 2SG PROG eat  
 ‘Which one [LIT.: what] are you eating?’

However, DC-movement is not sensitive to the form of *wh*-phrase in the way described above. Phrases of the form *which X*, *what X*, and bare *wh*-phrases can all be DC-moved. When prompted with (30b–b’), consultants regularly provide superordinate discourse contexts in which individuals or kinds of animals are mentioned. If DC-marking involved only D-linking, contexts should more easily license the form in (30b) than those in (30b’–b’’).

- (30) a. kǎw ɓoh lo      nǐ? t<sup>h</sup>un pǎ? tɛh  
 1SG see many animal at there  
 ‘I see many animals over there.’
- b. nǐ?              hlɛ́j<sub>DC</sub> hi jɔŋ nǐ? hlɛ́j  
 CLF.ANIMAL which 2SG raise  
 ‘Which animal do you raise?’ (a ↓ b)
- b'. nǐ?              kɛ?<sub>DC</sub> hi jɔŋ nǐ? kɛ?  
 CLF.ANIMAL what 2SG raise  
 ‘What animal do you raise?’ (a ↓ b')
- b''. kɛ?<sub>DC</sub> hi jɔŋ kɛ?  
 what 2SG raise  
 ‘What do you raise?’ (a ↓ b'')

There is a separate phenomenon that does track the form of *wh*-phrase: resumptive pronouns. Resumptive pronouns may occupy the base position of DC-moved *wh*-phrases. When prompted with (31), consultants consistently accept resumptive pronouns with *which* X (31b), but not with *what* X or bare *wh*-phrases (31b'–b'').

- (31) a. kǎw ɓoh lo      nǐ? t<sup>h</sup>un pǎ? tɛh  
 1SG see many animal at there  
 ‘I see many animals over there.’
- b. nǐ?              hlɛ́j<sub>DC,i</sub> hi jɔŋ ju<sub>i</sub>  
 CLF.ANIMAL which 2SG raise 3.ANIM  
 ‘Which animal do you raise?’ (a ↓ b)
- b'.<sup>??</sup> nǐ?              kɛ?<sub>DC,i</sub> hi jɔŋ ju<sub>i</sub>  
 CLF.ANIMAL what 2SG raise 3.ANIM  
 INTENDED: ‘What animal do you raise?’ (a ↓ b')
- b''.<sup>??</sup> kɛ?<sub>DC,i</sub> hi jɔŋ ju<sub>i</sub>  
 what 2SG raise 3.ANIM  
 INTENDED: ‘What do you raise?’ (a ↓ b'')

In a subset of contexts where a D-linked reading is strongly coerced, such as the partitive-like context in (32), though, all three forms of *wh*-phrases are accepted with resumptives. This aligns precisely with Pesetsky’s (1987) characterization of D-linking in English.

- (32) a. pih nǐ? t<sup>h</sup>un pǎ? ni // nǐ?              hlɛ́j<sub>DC,i</sub> hi cəh jɔŋ ju<sub>i</sub>  
 all animal at this CLF.ANIMAL which 2SG like raise 3.ANIM  
 ‘[Of] all the animals here, which animal do you like to raise?’
- b. pih nǐ? t<sup>h</sup>un pǎ? ni // nǐ?              kɛ?<sub>DC,i</sub> hi cəh jɔŋ ju<sub>i</sub>  
 all animal at this CLF.ANIMAL what 2SG like raise 3.ANIM  
 ‘[Of] all the animals here, what animal do you like to raise?’

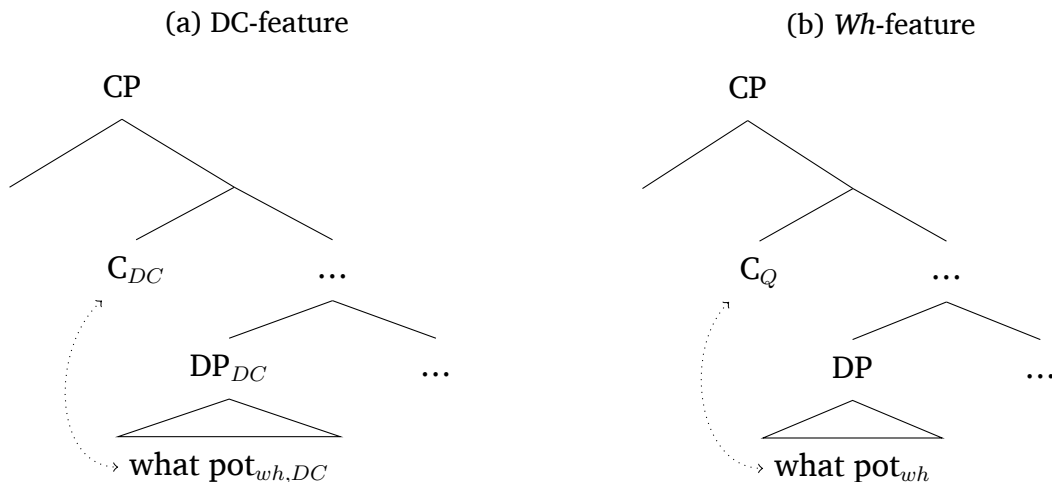
- c. pih nĩ? t<sup>h</sup>un pǎ? ni // k̄e?<sub>DC,i</sub> hi cəh jəŋ ju<sub>i</sub>  
 all animal at this what 2SG like raise 3.ANIM  
 ‘[Of] all the animals here, what do you like to raise?’

Why should resumptive pronouns be sensitive to D-linking? As described in the previous section, D-linked *wh*-phrases contain an index that refers to the contextual antecedent. This renders the *wh*-phrase truly anaphoric, not unlike anaphoric definites (Schwarz 2009). It follows that a pronoun can bear the same index. By contrast, non-D-linked *wh*-phrases lack such an index and do not require the kind of anaphoric relation that licenses pronouns. In other words, the contexts that license *what X* and bare *wh*-phrases do not necessary license coreferential pronouns.

## 4 Syntax of multiple DC-movement

We have seen that DC-movement of *wh*- and non-*wh*-phrases share basic  $\bar{A}$ -movement properties (Section 2.2) and that DC-marked *wh*- and non-*wh*-phrases share pragmatic properties (Section 3). We propose that these facts are best captured by an Agree mechanism between a probe on C specified for a DC-feature and a phrase bearing that DC-feature thanks to the DC-particle (Figure 4a). Note that this analysis assumes that the DC-feature introduced by the DC-particle can percolate up to DP, represented by the <sub>DC</sub> subscript on DP, such that the entire *wh*-phrase can be moved. By contrast, in principle, it is possible that the movement of *wh*-phrases is only superficially similar to the movement of non-*wh*-phrases. For example, the former could be due to optional *wh*-movement or a pseudocleft construction, in which case it is the *wh*-feature that is responsible for the derived position of the phrase, perhaps through an Agree relation as in Figure 4b.

Figure 4: DC- and *wh*-feature Agree mechanisms



This section provides additional evidence in favor of the DC-feature approach over an approach that appeals to the *wh*-feature. Data from the DC-movement of multiple phrases in the same clause provide two forms of evidence: DC-movement is distributionally dissimilar

from clefts, and locality effects arise that imply that DC-marked *wh*- and non-*wh*-phrases bear the same syntactic feature. First, multiple phrases can undergo DC-movement to the same left periphery (33c). Multiple DC-movement is possible in a context where each moved phrase is previously mentioned in a superordinate sentence and is most felicitous when there are multiple superordinate sentences with one antecedent in each. For example, (33c) is interpreted as an explanation of (33b) and (33b) is of (33a). Multiple clefts are ungrammatical, regardless of context (33b').

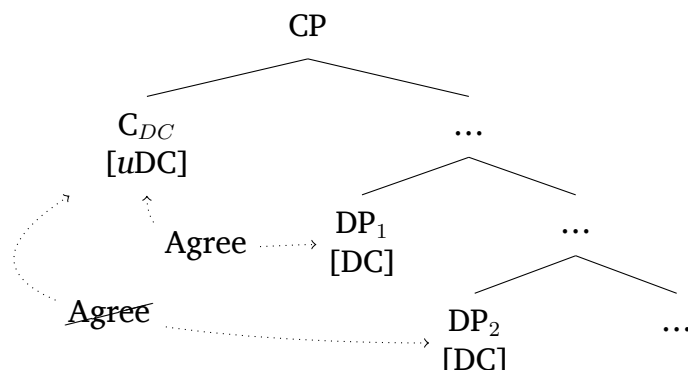
- (33) a. nǎj měj sīt nǎn puǰ pɛ lo  
 child female small that happy very  
 'That little girl is very happy.'
- b. nǎj měj sīt nǎn ɕǝʔ ʈi bǎŋ **han ni**  
 child female small that PROSP eat cake this  
 'That little girl is about to eat this cake.' (a ↓ b)
- c. **han ni** nǎj měj sīt nǎn kǎw ʔa nǎj-měj-sīt-nǎn maj  
 cake this child female small that 1SG invite come  
 bǎŋ **han ni**  
 eat  
 'This cake, I invited that little girl to come eat.' (b ↓ c)
- c'.\*hu t<sup>h</sup>a **kleh han** hu t<sup>h</sup>a jaŋ nuǰh kǎw ʔa  
 EXIST one piece cake EXIST one CLF.PERSON person 1SG invite  
 t<sup>h</sup>a-jaŋ-nuǰh maj bǎŋ t<sup>h</sup>a **kleh han**  
 come eat  
 INTENDED: 'There is a piece of cake there is a person I invited to come eat.'

Multiple *wh*-phrases can also be DC-moved in similar contexts (34c). Again, multiple clefts are ungrammatical regardless of context (34). Therefore, the movement of *wh*-phrases in (34c) is structurally more similar to that in (33c) above than to cleft constructions.

- (34) a. mǎj jaŋ nǎn puǰ pɛ lo  
 5 person that happy very  
 'Those five people are very happy.'
- b. mǎj jaŋ nǎn ɕǝʔ ʈi bǎŋ **dow**<sup>21</sup> **bǎŋ ni**  
 5 person that PROSP eat stuff(VN) eat this  
 'Those five people are about to eat this food.' (a ↓ b)
- c. **keʔ** t<sup>h</sup>ǝj hi ʔa maj bǎŋ  
 what who 2SG invite come eat  
 'Who did you invite to come eat what?' (b ↓ c)
- c'.\*hu **keʔ** hu t<sup>h</sup>ǝj hi ʔa maj bǎŋ  
 EXIST what EXIST who 2SG invite come eat  
 INTENDED: 'Who is that what is it that you invited to come eat?'

Second, multiple DC-movement provides an environment to examine locality effects. Locality effects arise when structural closeness determines which phrases can interact with syntactic probes (e.g. Chomsky 2000). For example, in Figure 5, DP<sub>1</sub> is structurally closer to the C-probe than DP<sub>2</sub>. Therefore, when the C-probe searches, it can Agree with the closest phrase bearing the appropriate feature, here DP<sub>1</sub>. It cannot Agree with DP<sub>2</sub>.

Figure 5: Predicted locality effect



In Eastern Cham, locality effects arise in multiple DC-movement in the form of path containment effects (Pesetsky 1982: 309). To illustrate, in (35) two phrases are DC-moved, *han ni* ‘this cake’, bolded throughout this section, and *nǐ? měj sīt nǎn* ‘that little girl’, underlined throughout. Each phrase has a movement path, or chain from its base position to its position derived by movement. The resulting sentence is grammatical if one movement path is completely contained within the other. In (35a), the movement path of the underlined phrase is contained within that of the bolded phrase. If the paths are crossed, however, as in (35b), the resulting sentence is strongly and consistently ungrammatical.<sup>16</sup>

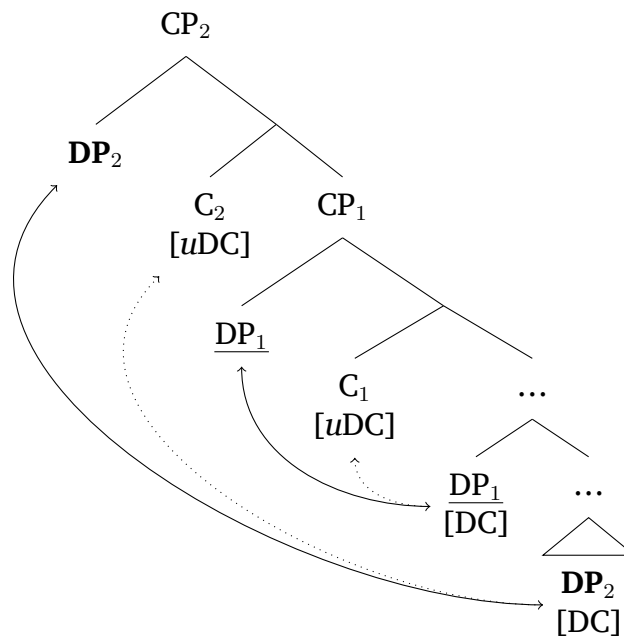
- (35) a. **han ni** nǐ? měj sīt nǎn t<sup>h</sup>uŋm<sup>312</sup> ʔa nǐ? měj sīt nǎn  
 cake this child female small that Thuận(VN) invite  
 maj bǎŋ **han ni**  
 come eat  
 ‘This cake, Thuận invited that little girl to come eat.’
- b. \*nǐ? měj sīt nǎn **han ni** t<sup>h</sup>uŋm<sup>312</sup> ʔa nǐ? měj sīt nǎn  
 child female small that cake this Thuận(VN) invite  
 maj bǎŋ **han ni**  
 come eat  
 INTENDED: ‘This cake, Thuận invited that little girl to come eat.’

Path containment effects like these occur when there are multiple syntactic probes, which are each constrained by locality, or structural closeness (e.g. Pesetsky 1982: 309 on

<sup>16</sup>Path containment effects also arise in multiple DC-movement to embedded peripheries. For example, if (35a–b) were embedded under a matrix clause, the same grammaticality judgments obtain. This implies that this effect is not only an effect of the matrix left periphery (e.g. interference with hanging topics).

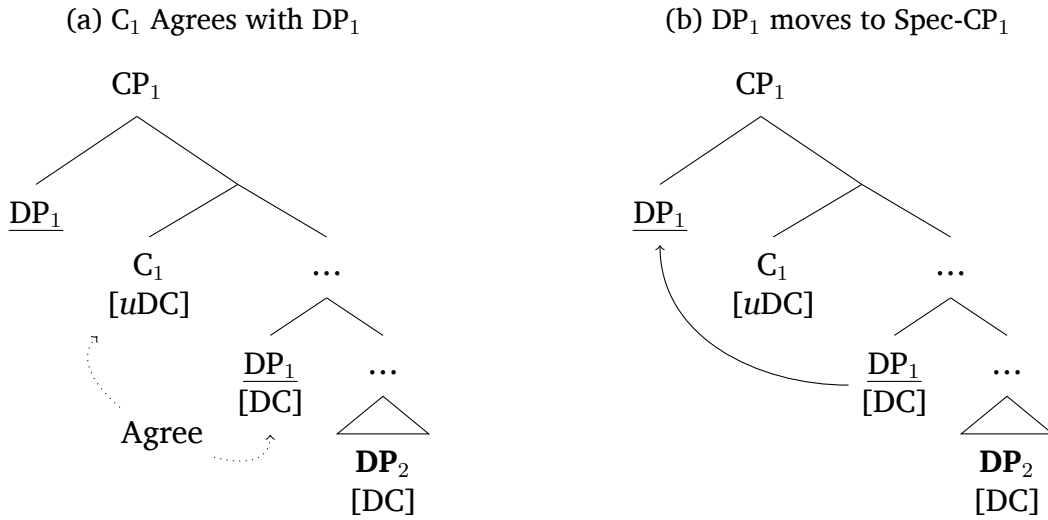
English *wh*-movement), and the same phrase head cannot host both probes. Baclawski Jr & Jenks (2016) analyze a similar phenomenon in Moken (Austronesian: Thailand) with two CPs each with a C-probe (cf. Rizzi 1997: 297 on multiple iterated TopicPs). An analysis of (35a) is depicted in Figure 6. This analysis is elaborated upon in the following figures. For ease of exposition, the bolded DPs correspond with the bolded phrases in the interlinearized Eastern Cham examples, and the underlined DPs with the underlined phrases. Also note that the trees have been abbreviated only to include CP projections and the relative structural hierarchy of the base positions of the two DPs.

Figure 6: Path containment derivation (cf. 35a)



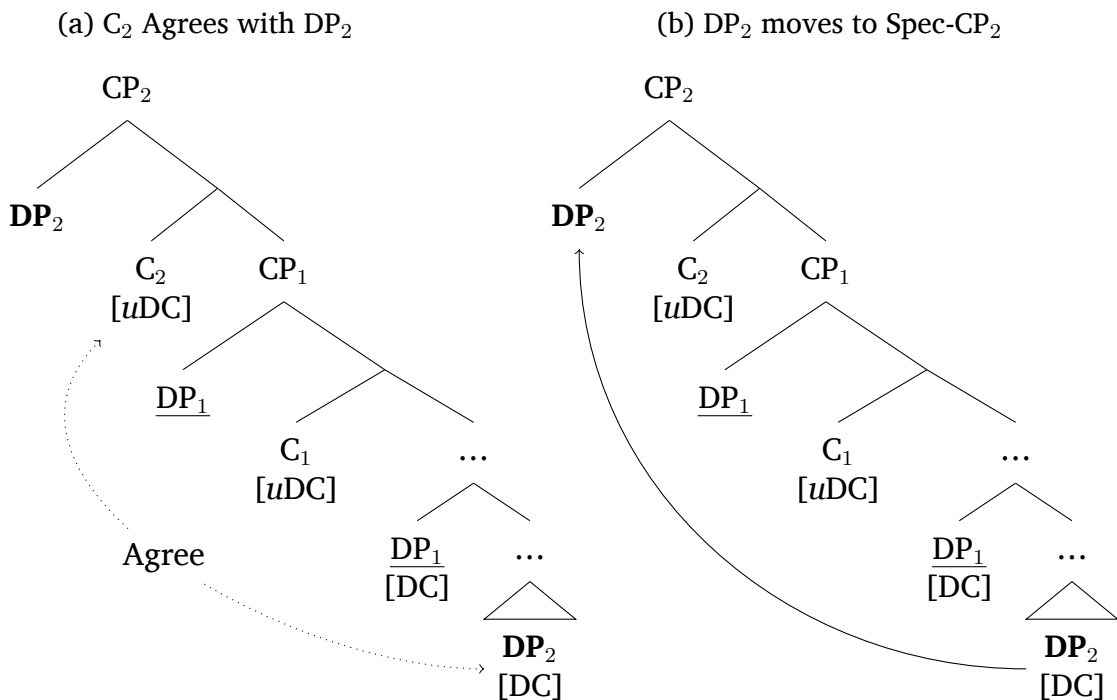
The first step of this derivation is the merge of  $C_1$ .  $C_1$  has a probe that searches for a phrase bearing the DC-feature. As shown in Figure 7a,  $C_1$  Agrees with the structurally closest such phrase,  $DP_1$ . It is that DP that moves to Spec- $CP_1$  (Figure 7b), assuming that DC-probes bear an EPP feature, attracting phrases to their specifiers.

Figure 7: Projection of CP<sub>1</sub>



With the DC-probe satisfied and DP<sub>1</sub> moved, C<sub>2</sub> is merged, along with another DC-probe on C<sub>2</sub>. At this point, C<sub>2</sub> Agrees with the next closest DP, DP<sub>2</sub> (Figure 8a). It is that phrase that is moved to Spec-CP<sub>2</sub> (Figure 8b). This scenario guarantees a path containment effect, because the innermost probe must Agree with the structurally highest DP, and the outermost probe with the lowest DP.

Figure 8: Projection of CP<sub>2</sub>





An explanandum in this analysis is why the probe in  $C_2$  cannot Agree with  $DP_1$ . There are at least two possible explanations for this restriction. First is criterial freezing. Rizzi (2010) and others have proposed that certain  $\bar{A}$ -movement operations disallow phrases from participating in subsequent syntactic movement operations. Criterial freezing in particular has been proposed for topicalization and other left peripheral movement operations. If Eastern Cham DC-movement were to result in criterial freezing, then we would not predict a DC-moved DP to be movable beyond Spec-CP. Second is (specifier-to-specifier) anti-locality, which has been proposed to be a general constraint on syntactic movement (e.g. Erlewine 2016). According to anti-locality, a phrase in the specifier of an XP must cross at least one other phrasal projection if moved. Movement from Spec-XP to the specifier of the immediately dominating YP is impossible. Both criterial freezing and anti-locality predict that the phrase in Spec-CP<sub>1</sub> cannot move to Spec-CP<sub>2</sub> in the path containment derivations above.

When multiple *wh*-phrases are DC-moved, the same path containment effect arises. The movement path of one phrase must be completely contained within that of the other (36a). When paths are crossed, the resulting sentence is consistently ungrammatical (36b). This is unexpected for *wh*-movement, as it represents an Anti-Superiority effect (cf. Baclawski Jr & Jenks 2016 on Moken).

- (36) a. **keʔ** t<sup>h</sup>ɛj t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj maj bǎŋ **keʔ**  
 what who Thuận(VN) invite come eat  
 ‘Who did Thuận invite to come eat what?’
- b.\* t<sup>h</sup>ɛj **keʔ** t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj maj bǎŋ **keʔ**  
 who what Thuận(VN) invite come eat  
 INTENDED: ‘Who did Thuận invite to come eat what?’

Typically, *wh*-movement is thought to be driven by a single C-probe. In multiple *wh*-questions, the probe proceeds by locality. If multiple phrases are *wh*-moved to the left periphery, as in Romanian or Bulgarian, the opposite, crossed path order obtains (cf. Richards 1997 on tucking in). Superiority effects are known to be violable in matrix clauses (cf. Bošković 2002: 353 on Serbo-Croatian). However, the path containment effect persists in embedded clauses in Eastern Cham (37).

- (37) a. hi hnɲj **keʔ** t<sup>h</sup>ɛj t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj maj bǎŋ **keʔ**  
 2SG think what who Thuận(VN) invite come eat  
 ‘Who do you think Thuận invited to come eat what?’
- b.\* hi hnɲj t<sup>h</sup>ɛj **keʔ** t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj maj bǎŋ **keʔ**  
 2SG think who what Thuận(VN) invite come eat  
 INTENDED: ‘Who do you think Thuận invited to come eat what?’

Not only do *wh*-phrases also exhibit path containment effects, but they can also intervene on DC-movement of a non-*wh*-phrase. In (38), one *wh*- and one non-*wh*-phrase are DC-moved to the left periphery. Again, the resulting sentence is grammatical if the paths are nested (38a) and ungrammatical if they are crossed (38b).

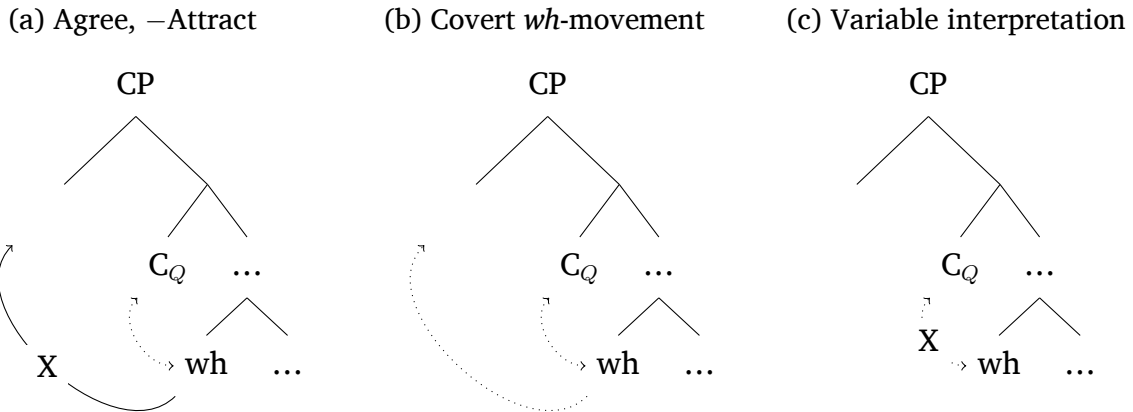
- (38) a. **han ni** nǎʔ měj sǐt hlěj t<sup>h</sup>uŋm<sup>312</sup> ʔa nǎʔ měj sǐt hlěj  
 cake this child female little which Thuận(VN) invite  
 maj bǎŋ **han ni**  
 come eat  
 ‘Which little girl did Thuận invite to come eat this cake?’
- b.\*nǎʔ měj sǐt nǎn **han hlěj** t<sup>h</sup>uŋm<sup>312</sup> ʔa nǎʔ měj sǐt nǎn  
 child female little which cake this Thuận(VN) invite  
 maj bǎŋ **han hlěj**  
 come eat  
 INTENDED: ‘Which cake did Thuận invite that little girl to come eat?’

This path containment effect implies that there must be overlap in the features that drive the movement of *wh*- and non-*wh*-phrases. We take this overlap to be the DC-feature itself. However, the overlap could also be due to a higher-order feature overlap. For example, if non-*wh*-phrases bear a DC-feature and *wh*-phrases only bear a *wh*-feature, a probe specified for a generalized  $\bar{A}$ -feature could enter an Agree relation with either (in the sense of Aravind 2017, 2018). Such an analysis (where *wh*-phrases only ever bear *wh*-features) would miss the generalization that moved *wh*-phrases exhibit DC pragmatics. Additionally, it would predict that in-situ *wh*-phrases would give rise to the same locality effects seen above. The following section examines in-situ *wh*-phrases and finds that this prediction is not borne out.

## 5 In-situ *wh*-phrases

This section examines how in-situ *wh*-phrases are interpreted in Eastern Cham. They are shown to be truly in-situ and to require an Agree relation with  $C_Q$ , but in a way that does not interact with DC-movement. Therefore, DC-movement is driven by a DC-feature in a way that is separate from *wh*-in-situ. There are multiple structures that surface as *wh*-in-situ cross-linguistically. Figure 9 presents three of these. First, in-situ phrases can Agree with C, but C does not attract them to Spec-CP (cf. Cable 2010: 85 on Q-adjunction; Hagstrom 1998 on Japanese). Second, C can both Agree with in-situ phrases and Attract them to Spec-CP. This last step is covert, however, resulting in a surface in-situ word order (cf. Cable 2010: 86 on Q-projection; Kishimoto 2005 on Sinhala). Third, there can be no Agree relation between C and in-situ *wh*-phrases at all. Instead, *wh*-phrases are interpreted as variables underneath  $C_Q$  (cf. Tsai 2009 on Vietnamese).

Figure 9: Types of *wh*-in-situ



These structures can be differentiated by applying movement diagnostics to in-situ *wh*-phrases. Evidence from island constraints and intervention effects in Eastern Cham point to the Agree, but not Attract structure in Figure 9a. First, *wh*-phrases are ungrammatical within syntactic islands. For instance, in a complex NP, the existence of an in-situ *wh*-phrase leads to ungrammaticality (39a), even though no overt movement has taken place. This indicates that in-situ *wh*-phrases do enter into an Agree relation with C, under the assumption that Agree is bounded by islands. It should be noted that this derivation is licit in the absence of a *wh*-phrase (39b).

- (39) a. \*hi plěj dɔw<sup>21</sup> bǎŋ tʰɛj ɲǎ?  
 2SG buy stuff(VN) eat who make  
 INTENDED: ‘You buy the food that who makes?’
- b. kǎw plěj dɔw<sup>21</sup> bǎŋ mɛ? kǎw ɲǎ?  
 1SG buy stuff(VN) eat mother 1SG make  
 ‘I buy the food that my mother makes.’

Second, intervention effects indicate that *wh*-phrases cannot covertly move to Spec-CP in Eastern Cham. According to Beck (1996, 2006), intervention effects arise when a focus operator intervenes between C and a *wh*-phrase such as *which soup* (40a). In this schema, the  $C_Q$ -head introduces a  $\sim$ -operator that interprets the *wh*-alternative set (cf. Rooth 1992), while the focus operator *only* functions as the intervener. Intervention effects arise because the focus operator cannot interpret the *wh*-alternative set.

- (40) a. ✓Intervention: [ $Q$  [ $\sim_C$  [ONLY<sub>C</sub> T ...which soup]]]  
 b. ✗Intervention: [which soup [ $Q$  [ $\sim_C$  [ONLY<sub>C</sub> T ...~~which~~ soup]]]]

Movement, including covert movement, is known to obviate intervention effects (cf. Kotek 2014, 2017). If a *wh*-phrase can move out of the intervention configuration, no effects arise (40b), as the *wh*-phrase no longer must be interpreted under the scope of the intervening focus operator.

In Eastern Cham, intervention effects categorically do arise with in-situ *wh*-phrases. In (41a), there is a *wh*-phrase, *?ja paj hlěj* ‘which soup’ under the scope of ‘only’. If covert *wh*-movement were possible, this sentence would be expected to be grammatical. Since the sentence is ungrammatical, we conclude that in-situ *wh*-phrases have no means of moving out of the scope of intervening focus operators. Similarly, the focus operator *çij* ‘also’ gives rise to an intervention effect in (41b).

(41) a. \*<sup>t</sup>h<sup>a</sup> sīt ?aj t<sup>h</sup>uŋm<sup>312</sup> kri ?ja paj hlěj mĩn  
 only older.sibling Thuận(VN) like soup which EMPH  
 INTENDED: ‘Which soup does only Thuận like to eat?’

b. CONTEXT: We can go to all the restaurants, but there are some Kenny cannot.  
 \*ken ni çij naw bǎŋ pǎ? ja<sup>21</sup> haŋ<sup>21</sup> hlěj hu  
 Kenny also go eat at restaurant(VN) which ROOT  
 INTENDED: ‘Which restaurant can Kenny also go eat at?’

Overt movement, by contrast, does alleviate intervention effects. In (42), DC-movement of the *wh*-phrases allows them to escape the scope of the focus operators. As predicted, the resulting sentences are grammatical, as the intervention configuration has been avoided. Note that the specific type of movement does not matter; any movement of a *wh*-phrase allows it to obviate intervention effects.

(42) a. ?ja paj hlěj<sub>DC</sub> t<sup>h</sup>a sīt ?aj t<sup>h</sup>uŋm<sup>312</sup> kri ?ja paj hlěj mĩn  
 soup which only older.sibling Thuận(VN) like EMPH  
 ‘Which soup does only Thuận like to eat?’

b. ja<sup>21</sup> haŋ<sup>21</sup> hlěj<sub>DC</sub> ken ni çij naw bǎŋ ja<sup>21</sup> haŋ<sup>21</sup> hlěj hu  
 restaurant(VN) which Kenny also go eat ROOT  
 ‘Which restaurant can Kenny also go eat at?’

When *wh*-phrases are c-commanded by certain other operators, non-interrogative indefinite readings obtain. Eastern Cham *wh*-phrase forms are ‘indeterminates’ (Kuroda 1965; Kratzer & Shimoyama 2002), as is commonly attested in East and Southeast Asia (e.g. Cheng 1991 on Mandarin Chinese; Tsai 2009 on Vietnamese). These contexts include the scope of negation (43a) and the antecedent of conditionals (43b). This further demonstrates that *wh*-phrases cannot covertly move out of the scope of operators, given that interrogative readings are impossible in these contexts.

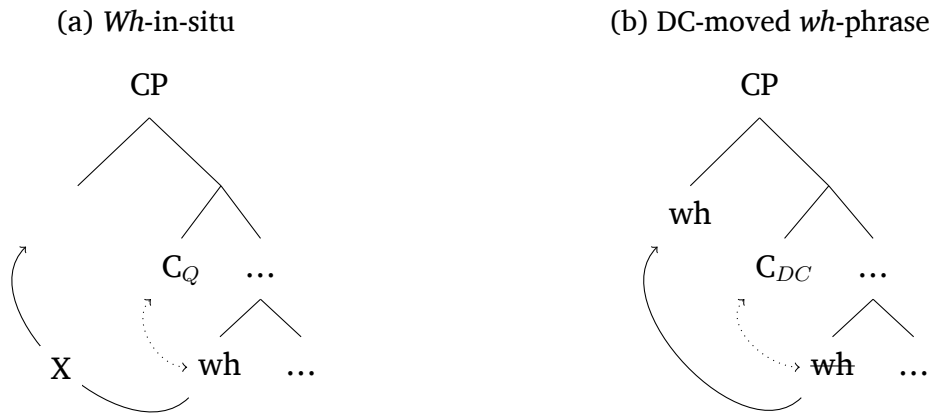
(43) a. hi hu bǎŋ ke? o  
 2SG EXIST eat what NEG  
 ‘You didn’t eat anything.’ / \*‘What didn’t you eat?’ (Negation)

b. t<sup>h</sup>ěj nũm ka<sup>21</sup> fej<sup>33</sup> hǎ? hǎ? nɔŋ nujh nǎn  
 who drink coffee(VN) 1SG.POL 1SG.POL be.angry person that  
 ‘If someone drinks my coffee, I will be angry at them.’ (Conditional)

Together the island and intervention effects point to an Agree, but not Attract model of Eastern Cham *wh*-in-situ (Figure 10a). Under DC-movement, by contrast, a *wh*-phrase

Agrees with C and moves to Spec-CP (Figure 10b).

Figure 10: Positions of Eastern Cham *wh*-phrases



Turning back to DC-movement, if there were any featural overlap between the DC-movement of *wh*-phrases and *wh*-in-situ, we would expect an interaction such that in-situ *wh*-phrases can intervene on DC-movement. In English, any *wh*-phrase is a candidate for *wh*-movement. Accordingly, a *wh*-phrase cannot be moved across a structurally higher in-situ *wh*-phrase (44a). For example, the object *what* cannot move across the subject *who*. Here, the <sup>\*PL</sup> notation indicates that the question loses its paired list and single answer readings.

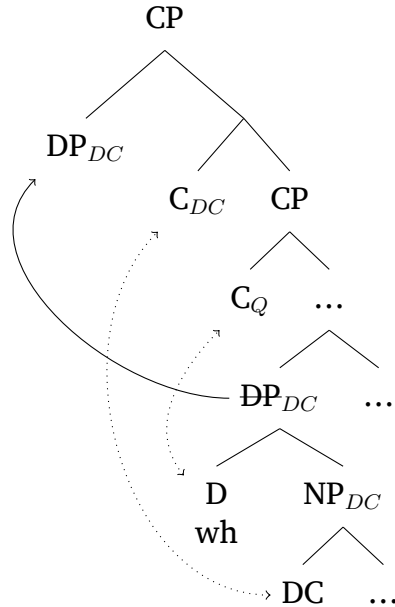
(44) <sup>\*PL</sup>What<sub>A:wh</sub> C<sub>uA</sub> did who<sub>A:wh</sub> buy? (Pesetsky 2000: 15–16)

In-situ *wh*-phrases in Eastern Cham do not exhibit such an interaction. In-situ phrases never show any signs of competing for DC-movement. An object DC-phrase (45a) or *wh*-phrase (45b) can be DC-moved over an in-situ *wh*-subject. This is expected under our analysis of DC-movement, as an in-situ *wh*-feature is never visible to C<sub>DC</sub>.

- (45) a. **han ni**<sub>DC</sub> C<sub>uDC</sub> t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj<sub>wh</sub> maj ɓǎŋ **han ni**  
 cake this Thuận(VN) invite who come eat  
 ‘Who did Thuận invite to come eat this cake?’
- b. **keʔ**<sub>DC,wh</sub> C<sub>uDC</sub> t<sup>h</sup>uŋm<sup>312</sup> ʔa t<sup>h</sup>ɛj<sub>wh</sub> maj ɓǎŋ **keʔ**  
 what Thuận(VN) invite who come eat  
 ‘Who did Thuận invite to come eat what?’

We conclude that DC-movement of a *wh*-phrase proceeds as in Figure 11. First, the in-situ *wh*-phrase enters into an Agree relation with C<sub>Q</sub>, as all in-situ *wh*-phrases do (or, C<sub>Q</sub> Agrees with the Q-particle, as per Cable 2010). This ensures that the *wh*-phrase is interpretable. Next, C<sub>DC</sub> probes for a DC-feature and moves the *wh*-phrase if it bears that feature. If the DC-particle has merged with the *wh*-restrictor, then the DC-feature must have percolated up to the top DP. In the absence of a DC-feature, the phrase cannot move.

Figure 11: Eastern Cham DC-movement of a *wh*-phrase



This analysis maintains Cheng’s (1997) Clausal Typing Hypothesis. There is only one Agree mechanism between  $C_Q$  and *wh*. No exceptions or additional mechanisms related to *wh* are needed to account for DC-movement.

## 6 Conclusion

To conclude, Eastern Cham upholds the Clausal Typing Hypothesis, despite exhibiting apparent optional *wh*-movement. Instead, *wh*-phrases can undergo DC-movement, an  $\bar{A}$ -movement operation driven by a feature that indexes *discourse connectedness*. This DC-feature is introduced by a lexical item in the same vein as focus particles and Q-particles (cf. Cable 2010). In Eastern Cham, this lexical item is phonologically null, but our analysis predicts it may be overt in other languages. This result expands the typology of possible  $\bar{A}$ -features and possible explanations for apparent optional *wh*-movement.

DC also leads to the conclusion that syntax can be sensitive to a range of relations between sentences in a discourse. The phenomenon described here cannot be formalized in a model of discourse like Questions Under Discussion (QUD; Roberts 1998). In a QUD model, questions can be subordinate to broader questions if those broader questions remain unanswered (cf. Büring 2003; Constant 2014 on contrastive topic). However, DC-marking in Eastern Cham does not require an open question; in fact, DC-movement is dispreferred in those environments (cf. Baclawski Jr 2018a; Baclawski Jr 2019: 180). Instead, DC requires *discourse subordination*, a looser relation between sentences. In order to account for DC-movement, then, a model of discourse must allow for such loose

relations (e.g. Onea’s (2013, 2016) Potential Questions).

Given that DC requires such a loose connection between sentences, it is worth asking why DC should be marked at all. As noted above, DC-movement is generally optional in Eastern Cham; DC-phrases can remain in-situ in nearly any context.<sup>17</sup> In order to account for this optionality, we turn to ‘cue phrases’. Cue phrases in languages like English are often adverbials that mark discourse relations (cf. Grosz & Sidner 1986: 196). For example, *After that* cues a sequence of event relation known as *Narration*. *After that* is not obligatory in a *Narration* context. When present, it increases the likelihood that a sentence is interpreted as in a *Narration* relation. When absent, the same interpretation can still be made, albeit with an increase in ambiguity (46). In (46), a *Narration* reading is unavoidable with the cue phrase. Without it, the second sentence may be interpreted as an elaboration on the event of the dog’s trotting, for instance, or perhaps an explanation for it.

(46) The dog trotted down the street. (**After that,**) it barked at a cat on the corner.

We posit that Eastern Cham DC-movement functions in a similar way to cue phrases (cf. Eckardt & Fränkel 2012 on the additive *too*). Merging a  $C_{DC}$ -probe is comparable to merging a cue phrase. If the C-probe is merged, it must merge a DC-phrase in its specifier position, and the resulting sentence must be interpreted as in a subordinating discourse relation. Furthermore, the DC-moved phrase identifies the superordinate sentence in the discourse (i.e. the one that mentions the phrase itself). Without the C-probe, that same interpretation can be computed. However, there is increased ambiguity. Perhaps the sentence will be interpreted with a non-subordinating discourse relation or as subordinate to a different sentence in the discourse than intended. DC-movement, then, functions like a cue phrase, but is marked solely by a syntactic probe, resulting in word order permutation.

## References

- Alcalá, Carlos Rubio. 2014. *Syntactic constraints on topicalization phenomena*. Universitat Autònoma de Barcelona dissertation.
- Aravind, Athulya. 2017.  $\bar{A}$ -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://doi.org/10.1016/j.lingua.2003.09.017>.

<sup>17</sup>Optionality is a general problem for formal accounts of topicalization-like phenomena (cf. Erteschik-Shir 2007: 56; Horvath 2010: 1364; Alcalá 2014: 131).

- Baclawski Jr, Kenneth. 2017. Clause-final particles and focus in Eastern Cham. In *Proceedings of the Linguistic Society of America*, vol. 2.18, 1–12. <https://doi.org/10.3765/plsa.v2i0.4066>.
- Baclawski Jr, Kenneth. 2018a. Contrastive topic in Eastern Cham. *Berkeley Formal Papers in Linguistics* 1(1). <https://escholarship.org/uc/item/3hp0s863>.
- Baclawski Jr, Kenneth. 2018b. Diglossia and change from below in Eastern Cham. *Asia-Pacific Language Variation* 4(1). 73–102. <https://doi.org/10.1075/aplv.17003.bac>.
- Baclawski Jr, Kenneth. 2019. *Discourse connectedness: The syntax–discourse structure interface*. University of California, Berkeley dissertation.
- 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>.
- Beck, Sigrid. 1996. *Wh-constructions and transparent logical form*. Universität Tübingen dissertation.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14. 1–56. <https://doi.org/10.1007/s11050-005-4532-y>.
- Boeckx, Cedric & Kleanthes K. Grohmann. 2004. SubMove: Towards a Unified Account of Scrambling and D-Linking. In David Adger, Cécile de Cat & George Tsoulas (eds.), *Peripheries*, 241–257. Springer. [https://doi.org/10.1007/1-4020-1910-6\\_10](https://doi.org/10.1007/1-4020-1910-6_10).
- Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33(3). 351–383.
- Brunelle, Marc. 2009. Diglossia and monosyllabization in Eastern Cham: A sociolinguistic study. In James Stanford & Dennis Preston (eds.), *Variation in indigenous minority languages*, 47–75. John Benjamins. <https://doi.org/10.1075/impact.25.04bru>.
- Brunelle, Marc. 2014. Vietnamese (Tiếng Việt). In Paul Sidwell & Mathias Jenny (eds.), *The Handbook of Austroasiatic Languages*, 907–953. Brill.
- Brunelle, Marc & Phú, Văn Hân. 2018. Colloquial Eastern Cham. In Alice Vittrant & Justin Watkins (eds.), *The Mainland Southeast Asia linguistic area*, 522–557. Mouton de Gruyter.
- Büring, Daniel. 2003. On d-trees, beans, and b-accent. *Linguistics and Philosophy* 26. 511–546.
- Cable, Seth. 2010. *The grammar of Q*. Oxford University 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.
- Cheng, Lisa L.S. & Johan Rooryck. 2000. Licensing wh-in situ. *Syntax* 3(1). 1–20.
- Chierchia, Gennaro. 1997. Partitives, reference to kinds and semantic variation. In Aaron Lawson (ed.), *Proceedings of SALT VII*, 73–98.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. MIT Press.
- Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Springer.
- Constant, Noah. 2014. *Contrastive topic: Meanings and realizations*. University of Massachusetts, Amherst dissertation.
- den Dikken, Marcel & Anastasia Giannakidou. 2002. ‘Aggressively non-D-linked’ wh-phrases as polarity items. *Linguistic Inquiry* 33(1). 31–61.



- Denham, Kristin E. 1998. *A Minimalist account of optional wh-movement*. University of Washington dissertation.
- Denham, Kristin E. 2000. Optional wh-movement in Babine-Witsuwit'en. *Natural Language & Linguistic Theory* 18. 199–251.
- Ebert, Cornelia. 2009. *Topicalized quantifiers: A treatment of exceptional wide scope phenomena*. Springer.
- Eckardt, Regine & Manuela Fränkel. 2012. Maximize Presupposition and discourse management. *Lingua* 122. 1801–1818.
- Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34(2). 429–479. <https://doi.org/10.1007/s11049-015-9310-z>.
- Erteschik-Shir, Nomi. 2007. *Information structure*. Oxford University Press.
- Evans, Gareth. 1980. Pronouns. *Linguistic Theory* 11(2). 337–362.
- Fabricius-Hansen, Cathrine & Wiebke Ramm. 2008. *Editor's introduction: Subordination and coordination from different perspectives*. Cathrine Fabricius-Hansen & Wiebke Ramm (eds.). J. Benjamins. 1–30. <https://doi.org/10.1075/slcs.98.01fab>.
- Grewendorf, Günther. 2012. Wh-movement as topic movement. In Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *Functional heads: The cartography of syntactic structures, Volume 7*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199746736.003.0004>.
- Grosz, Barbara & Candace Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3). 174–204.
- Hagstrom, Paul Alan. 1998. *Decomposing questions*. MIT dissertation.
- Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. University of Massachusetts, Amherst dissertation.
- Hoàng, Thị Châu. 1989. *Tiếng Việt trên các Miền Đất Nước [The Vietnamese Language in the Regions of the Country]*. Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.
- Hoonchamlong, Yuphaphan. 1991. *Some issues in Thai anaphora: A government and binding approach*. University of Wisconsin-Madison dissertation.
- Horvath, Julia. 2010. Discourse features, syntactic displacement, and the status of contrast. *Lingua* 120. 1346–1369. <https://doi.org/10.1016/j.lingua.2008.07.011>.
- Jensen, Joshua. 2013. *The Structure of Jarai clauses and noun phrases*. The University of Texas at Arlington dissertation.
- Kishimoto, Hideki. 2005. *Wh-in-situ and movement in Sinhala questions*. *Natural Language and Linguistic Theory* 23. 1–51. <https://doi.org/10.1007/s11049-004-6574-0>.
- 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>.
- Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In Yukio Otsu (ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, 1–25. Tokyo: Hituzi Syobo.
- Kuroda, S.Y. 1965. *Generative Grammatical studies in the Japanese language*. Massachusetts Institute of Technology dissertation.
- López, Luis. 2009. *A derivational syntax for information structure*. Oxford University Press.

- Manetta, Emily. 2011. *Peripheries in Kashmiri and Hindi-Urdu: The syntax of discourse-driven movement*. John Benjamins.
- 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>.
- Moussay, Fr. Gérard. 2006. *Grammaire de la langue Cam [Grammar of the Cham Language]* (Mission Étrangères de Paris [Foreign Missions of Paris]). Paris: Les Indes Savant.
- 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.
- Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 110(10). 707–727. [https://doi.org/10.1016/S0024-3841\(00\)00044-9](https://doi.org/10.1016/S0024-3841(00)00044-9).
- 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.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. MIT Press.
- Polinsky, Maria. 2001. *Information structure and syntax: Topic, discourse-linking, and agreement*. Third Workshop on Spoken and Written Texts. Austin: Texas Linguistic Forum.
- Potsdam, Eric. 2006. More concealed pseudoclefts in Malagasy and the Clausal Typing Hypothesis. *Lingua* 116. 2154–2182. <https://doi.org/10.1016/j.lingua.2005.07.003>.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27. 53–94.
- Richards, Norvin. 1997. *What moves where in which language?* MIT dissertation.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar: A handbook of generative syntax*, 281–327. Kluwer. [https://doi.org/10.1007/978-94-011-5420-8\\_7](https://doi.org/10.1007/978-94-011-5420-8_7).
- Rizzi, Luigi. 2010. On some properties of criterial freezing. In E. Phoevos Panagiotidis (ed.), *The Complementizer Phrase*, 17–32. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199584352.003.0002>.
- Roberts, Craige. 1998. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 6. 1–69. <https://doi.org/10.3765/sp.5.6>.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116. <https://doi.org/10.1007/BF02342617>.
- Roussou, Anna & Christos Vlachos. 2011. Introduction to optionality in wh-movement. *Linguistic Analysis* 31(1–2). 3–10.
- Ruys, E.G. 2004. A note on weakest crossover. *Linguistic Inquiry* 35(1). 124–140. <https://doi.org/10.1162/002438904322793365>.
- Salmon, Wesley C. 1984. *Scientific explanation and the causal structure of the world*. Princeton University Press.

- Sato, Yosuke. 2011. P-stranding under sluicing and repair by ellipsis: Why is Indonesian (not) special? *Journal of East Asian Linguistics* 20(4). 339–382. <https://doi.org/10.1007/s10831-011-9082-3>.
- Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts, Amherst dissertation.
- Simpson, Andrew. 2001. Focus, presupposition and light predicate raising in East and Southeast Asia. *Journal of East Asian Linguistics* 10. 89–128.
- 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>.
- Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Everaert, Henk Van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds.), *The Blackwell companion to syntax*, 479–531. Blackwell. <https://doi.org/10.1002/9780470996591>.
- Thurgood, Graham. 1996. Language contact and the directionality of internal drift. *Language* 72(1). 1–31. <https://doi.org/10.2307/416792>.
- Thurgood, Graham. 1999. *From ancient Cham to modern dialects: Two thousand years of language contact and change*. University of Hawai'i Press.
- Tổng Cục Thống Kê [General Statistics Office of Vietnam]. 2020. *Kết Quả Toàn Bộ: Tổng Điều Tra Dân Số Và Nhà Ở Năm 2019 [Completed Results of the 2019 Viet Nam Population and Housing Census]*. Retrieved on December 19, 2020. <https://www.gso.gov.vn/en/data-and-statistics/2020/11/completed-results-of-the-2019-viet-nam-population-and-housing-census/>.
- Tsai, Cheng-Yu Edwin. 2009. *Wh-dependency in Vietnamese and the syntax of wh-in-situ*. National Tsing Hua University dissertation.
- Uribe-Etxebarria. 2002. In-situ questions and masked movement. *Linguistic Variation Yearbook* 2. 217–257. <https://doi.org/10.1075/livy.2.09uri>.
- 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>.
- Wiltschko, Martina. 1997. D-linking, scrambling, and Superiority in German. *Groninger Arbeiten zur Germanistischen Linguistik* 41. 107–142.
- Ylikoski, Petri. 2013. Causal and constitutive explanation compared. *Erkenn* 78. 277–297. <https://doi.org/10.1007/s10670-013-9513-9>.