

*Proceedings of the Twenty-Sixth Meeting of the
Austronesian Formal Linguistics Association (AFLA)*

Ileana Paul
(ed.)



AFLA XXVI
The Twenty-Sixth Meeting of the
Austronesian Formal Linguistics Association

University of Western Ontario
May 24-26, 2019

**OPTIONAL WH-MOVEMENT AND TOPICALIZATION IN
EASTERN CHAM**

Kenneth Baclawski Jr.
University of California, Berkeley

Table of Contents

Paul, Ileana	<i>Preface</i>	i
Baclawski, Kenneth, Jr.	<i>Optional wh-movement and topicalization in Eastern Cham</i>	1-17
Chang, Henry Y.	<i>Tsou Exclamatives in comparative syntax</i>	18-35
Chen, Tingchun	<i>Raising-to-object in Amis</i>	36-53
Collins, James N. and Peter Schuelke	<i>Roviana fronting and the relationship between syntactic and morphological ergativity</i>	54-70
Erlewine, Michael Yoshitaka, Theodore Levin and Coppe van Urk	<i>The typology of nominal licensing in Austronesian voice system languages</i>	71-87
Finer, Daniel and Hasan Basri	<i>Clause truncation in South Sulawesi: Restructuring and nominalization</i>	88-105
Hopperdietzel, Jens	<i>Pseudo noun incorporation and differential object marking: object licensing in Daakaka</i>	106-123
Hsieh, Henrison	<i>On the structure of Tagalog non-DP extraction</i>	124-141
Kroeger, Paul	<i>Marking accessible information in Kimaragang</i>	142-158
Kroeger, Paul and Kristen Frazier	<i>Crossed-control in Malay/Indonesian aslong-distance passivization</i>	159-174
Macaulay, Benjamin	<i>The prosodic structure of Paze</i>	175-191
Ono, Hajime, Koichi Otaki, Manami Sato, 'Ana Heti Veikune, Peseti Vea, Yuko Otsuka and Masatoshi Koizumi	<i>Relative clause processing in Tongan: an effect of syntactic ergativity on the object preference</i>	192-208
Paillé, Mathieu	<i>V=nya in colloquial Malay</i>	209-226
Tollan, Rebecca	<i>Subjecthood and unmarkedness in Niuean</i>	227-247
Travis, Lisa DeMena and Diane Massam	<i>What moves, why, and how: the contribution of Austronesian</i>	248-264

PREFACE

The 26th Annual Meeting of the Austronesian Formal Linguistics Association (AFLA 26) was held on May 24-26, 2019 at the University of Western Ontario (Canada). The programme consisted of 24 presentations in addition to four plenary talks by Juliette Blevins, Vera Hohaus, Marian Klamer and Becky Tollan. This volume includes 13 papers from the conference.

As conference organizer, I received generous support from a variety of sources. Financial support came from the Social Sciences and Humanities Research Council of Canada (SSHRC), Research Western, the Joint Fund (Research Western, SOGS, SGPS), the Theoretical and Applied Linguistics Lab, the Canadian Linguistic Association, the Faculty of Arts and Humanities, the Graduate Program in Linguistics and three departments (French Studies, Modern Languages and Literatures, and Anthropology). The conference would not have been possible without the student volunteers (Sonia Masi, William Tran, Caylen Walker and Kang Xu), plus several others who helped out at the registration desk. Finally, I am grateful to the Department of French Studies for administrative support.

Many thanks to the abstract reviewers, to all those who attended, and to Mitcho Erlewine, who helped develop the current stylesheet.

Ileana Paul
University of Western Ontario

OPTIONAL *WH*-MOVEMENT AND TOPICALIZATION IN EASTERN CHAM*

Kenneth Baclawski Jr.
University of California, Berkeley
kbaclawski@berkeley.edu

What at first glance appear to be topicalization and optional *wh*-movement in Eastern Cham (Austronesian: Vietnam) are in fact instances of *discourse connected*- or DC-movement. DC is argued to be an \bar{A} -feature that marks a property of discourse structure (i.e. rhetorical relations). Namely, DC-marked phrases must be previously mentioned in a prior sentence that the current sentence explains or elaborates upon. Distributional and pragmatic evidence support this analysis of DC over information structural phenomena such as topicalization, clefting, and D-linking. This provides evidence that syntax must have an interface with discourse structure.

1. Introduction

This paper argues that what appear to be topicalization and optional *wh*-movement in Eastern Cham are in fact instances of *discourse connected*- or DC-movement. DC is claimed to be an \bar{A} -feature, in terms of syntax, that enforces a specific relation between two sentences in a discourse. Descriptively, this paper aims to promote rhetorical relations as a possible diagnostic for syntactic movement operations.

Consider the sentences in (1a–b). Objects are postverbal in an unmarked context (1a), but they can also appear at the left edge of the clause (1b). There is no obvious meaning difference between the two, as reflected in the English gloss. However, (1b) is infelicitous in an out-of-the-blue context. Many movement phenomena with this property are attributed to topicalization, broadly, syntactic movement that marks previously mentioned referents about which a sentence is organized (e.g. Reinhart 1981, among many others).¹

*My sincere thanks to the Cham people of Ninh Thuận province, Vietnam, especially to Sakaya and Sikhara (Hamu Ligaih), without whom this work would not be possible. All mistakes are my own. Thanks to Peter Jenks and Line Mikkelsen, for their extensive help in the development of this work. Thanks also Seth Yalcin, reviewers and audiences at AFLA 26, the 2019 LSA Annual Meeting, UC Berkeley, the University of Oslo, University of Geneva, and elsewhere for their helpful feedback and comments.

¹Eastern Cham examples largely follow IPA conventions. In line with the Chamic linguistic tradition, open circles underneath consonants indicates falling, breathy tone or register on following vowels. Abbreviations used: ANIM=animate; CLF=(numeral) classifier, COMP=complementizer, EMPH=emphatic, EX.COP=existential copula, POL=polite, PROG=progressive, ROOT=root modal, VN=Vietnamese loanword/code switch.

- (1) a. kǎw tɔʔ bǎŋ ʔɔʔ ni
 1SG PROG eat mango this
 ‘I am eating this mango.’
- b. ʔɔʔ ni kǎw tɔʔ bǎŋ
 mango this 1SG PROG eat
 ‘I am eating this mango.’

Next, consider the sentences in (2a–b). *Wh*-phrase objects are typically pronounced in situ (2a), but they can also appear at the left edge of the clause as well (2b). Again, there is no obvious difference in interpretation, while (2b) is disfavored in out-of-the-blue or discourse-initial contexts.

- (2) a. hi tɔʔ bǎŋ ket
 2SG PROG eat what
 ‘What are you eating?’
- b. ket hi tɔʔ bǎŋ
 what 2SG PROG eat
 ‘What are you eating?’

Despite the formal similarity between (1) and (2), phenomena like the latter have been attributed to syntactic operations driven by a *wh*-feature. Examples include optional *wh*-movement and concealed pseudoclefts. Under an optional *wh*-movement analysis, (2b) is true *wh*-movement, but the relevant features or lexical items are not always generated in the structure, sometimes resulting in (2a) (cf. Denham 2000). In the Austronesian context, this has been proposed for languages like Malay (Denham 2000, 248). Under a concealed pseudocleft analysis, (2b) covertly contains a pseudocleft, as proposed for languages like Malagasy (Paul 2001, Potsdam 2006, and others; cf. also Cheng 1997 on a clefting analysis of Malay).

This paper argues instead that the movement operations in (1–2) are in fact identical; they are both driven by the same \bar{A} -feature. A closer examination of the syntax of the constructions suggests that a cleft/pseudocleft analysis is untenable for the Eastern Cham case, and a closer examination of their discourse properties suggests that topicality is insufficient to account for the pragmatics. Instead, phrases can undergo this \bar{A} -movement only if they are *discourse connected*. As laid out in (3) informally, DC-marking has two components: previous mention and a rhetorical relation such that the anaphor’s sentence explains or elaborates upon the antecedent’s.

- (3) DC (informal): A phrase can be marked as DC iff it is previously mentioned in a sentence that the current sentence explains or elaborates

Section 2 gives some background on the Eastern Cham language and a brief look at general information structural phenomena. Section 3 examines the syntactic properties of DC-movement and concludes it is an \bar{A} -movement operation. Evidence is also found that it is distinct from hanging topic and cleft constructions in the language. Section 4 then presents the pragmatics of apparent topicalization (1b) and analyzes it as DC-movement. Section 5 examines the apparent optional *wh*-movement case (2b) and analyzes it as DC-movement of *wh*-phrases. Other information structural phenomena such as topicalization and D-linking are found to make incorrect predictions in Eastern Cham. Finally, an account is sketched by which a phrase can be interpreted both as DC and *wh*. Section 6 concludes.

2. Background on Eastern Cham

Eastern Cham is Malayo-Polynesian language spoken in south-central Vietnam largely in the provinces of Ninh Thuận and Bình Thuận. The community has a population of about 160,000, according to the 2009 Vietnamese census. The number of fluent speakers of Eastern Cham is likely closer to 120,000 individuals, among whom there is quasi-universal bilingualism with Vietnamese (Baclawski Jr. 2018, 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. 2018). Phonetic variation is abstracted in the data here through the use of broad phonological transcription.

Data for this paper were collected by the author in Vietnam from 2015–2019 with six consultants, all native speakers of Eastern Cham born and raised in the Cham villages of Phan Rang, Vietnam. Among these six, each data point has been checked by at least two. 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). 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 sound changes. The example below in (4) gives a basic example of an Eastern Cham sentence. The first line reflects Cham script, through a romanization known as Rumi.

- (4) Adei palaik tijuh abaoh tamókai nan hu.
 ṭɛj mlɛʔ cǔh pǒh tamkaj nən hu
 younger.sibling drop 7 CLF watermelon that ROOT
 ‘You[younger sibling] can drop those 7 watermelons.’

In (4), the historical causative *pa-* is crystallized as a short *m-* (cf. *leʔ* ‘to fall’).² Like Vietnamese and other Mainland Southeast Asian languages, kinship terms may function as pronominals (here, *ṭɛj* ‘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).

²The reduction of unstressed syllables such as **pa-* is complex and requires thorough phonetic research.

3. The Syntax of DC-movement

This section examines the distributional properties of DC-movement and concludes that it is an \bar{A} -movement operation alongside clefting and relativization in the language. DC-movement of *wh*-phrases shares all these characteristics as well. At the same time, DC-movement is found to have distinct properties from other constructions such as presentational clefts and hanging topics. Five pieces of evidence will be discussed: presence of the complementizer, preposition-drop, island sensitivity, weak crossover, and locality effects. First, DC-movement, clefting and relativization are all optionally marked by the complementizer *po* (5).³ This follows if these constructions involve \bar{A} -movement to a Spec-CP position, and *po* spells out a C-head.

- (5) a. *pu* (*po*) *hi* *ʔa* *ɸu* *hu*
 Phú COMP 2SG invite ROOT
 ‘Phú, you can invite.’ *DC-movement*
- b. *t^hěj* (*po*) *kăw* *ʔa* *t^hěj* *hu*
 who COMP 1SG invite ROOT
 ‘Who can I invite?’ *DC-movement of wh-phrase*
- c. *pu* (*po*) *hi* *ʔa* *ɸu* *tɔʔ* *păʔ* *teh*
 Phú COMP 2SG invite EX.COP at there
 ‘Phú, who you invited, is over there.’ *Relative clause*

Second, these constructions all exhibit preposition- or p-drop. When a phrase is extracted from an argument prepositional phrase, the preposition is neither pied-piped nor pronounced in its base position ((6); cf. Wang 2007; Sato 2011 on Indonesian). There is likely a prosodic explanation for p-drop in this case. For example, the preposition *ka* could be too weak to be pronounced without a complement in situ. Note that p-drop results in these constructions generally being limited to nominals in Eastern Cham.

- (6) a. *nǐʔ* *năn* *kăw* *plěj* *han* *ni* *ka* *nǐʔ* *năn*
 child that 1SG give cake this to
 ‘That child, I [will] give this cake to.’ *DC-movement*
- b. *t^hěj* *hi* *plěj* *han* *ni* *ka* *t^hěj*
 who 2SG give cake this to
 ‘Who [will] you give this cake to?’ *DC-movement of wh-phrase*
- c. *hu* *t^ha* *jaŋ* *nǐʔ* *kăw* *plěj* *han* *ni* *ka* *t^ha* *jaŋ* *nǐʔ*
 EX.COP one CLF.PERSON child 1SG give cake this to
 ‘There is a child I [will] give this cake to.’ *Presentational cleft*

³This use of the form *po* is specific to the author’s fieldwork and is unattested or even explicitly argued against existing in prior literature (Thurgood 2005, 508).

These two diagnostics distinguish DC-movement from a hanging topic construction. Hanging topics, as in (7), are marked by a pause. They also have no category restriction, cannot be marked by the complementizer, and do not exhibit p-drop. Note that the hanging topic binds a null pronoun in (7a), resulting in a surface similarity to DC-movement.

- (7) a. *pu_i \ \ (*po) hi ʔa Ø_i hu*
 Phú PAUSE COMP 2SG invite 3.ANIM ROOT
 ‘Phú. You can invite him.’ *Hanging topic*
- b. *mij oim⁴⁵ \ \ (*po) zut jum ʔja ce ni*
 with straw(VN) PAUSE COMP friend drink water tea this
 ‘With [a] straw, you[friend] drink this tea.’ *Hanging topic*

Third, in line with \bar{A} -movement cross-linguistically, DC-movement, clefts, and relativization are sensitive to island constraints, such as complex DPs (8a–c). Note that (8a) and (8c) become grammatical if the moved phrase is pronounced in its base position. The *wh*-phrase in (8b) cannot even be pronounced in its base position, owing to the syntax of *wh*-in situ in Eastern Cham.

- (8) a. **meʔ kǎw kǎw plěj do²¹ bǎŋ po meʔ-kǎw ɲǎʔ*
 mother 1SG 1SG buy stuff(VN) eat COMP make
 INTENDED: ‘My mother, I buy the food she makes.’ *DC-movement*
- b. **jaŋ hlěj hi plěj do²¹ bǎŋ po jaŋ-hlěj ɲǎʔ*
 person which 2SG buy stuff(VN) eat COMP make
 INTENDED: ‘Which person do you buy the food they make?’
DC-movement of wh-phrase
- c. **hu t^ha jaŋ kǎw plěj do²¹ bǎŋ po t^ha-jaŋ ɲǎʔ*
 EX.COP one CLF.PERSON 1SG buy stuff(VN) eat COMP make
 INTENDED: ‘There is a person who I buy the food they make.’
Presentational cleft

Fourth, \bar{A} -movement gives rise to weak crossover effects, again in line with \bar{A} -movement cross-linguistically (e.g. Postal 1971). 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 (9). The direct object can bind a pronoun within the indirect object, but not vice versa (b), as indicated by the ungrammaticality of the *i* index. The same goes for *wh*-phrases, which can in theory bind pronouns, but not in the configuration in (9c).

- (9) a. kǎw m̄jan lǎj? nǐ? mjaw nǎn_i ka po ju_i
 1SG return CLF.ANIMAL cat that to owner 3.ANIM
 ‘I returned that cat to its owner.’
- b. kǎw m̄jan lǎj? ɓɔp⁴⁵ ju_{*i/j} ka pu_i
 1SG return wallet(VN) 3.ANIM to Phú
 ‘I returned his wallet to Phú.’
- c. hi m̄jan lǎj? ɓɔp⁴⁵ ju_{*i/j} ka t^hɛj_i
 2SG return wallet(VN) 3.ANIM to who
 ‘Who did you return their wallet to?’

If an indirect object is \bar{A} -moved over a direct object, a crossover context arises. For example, the DP *t^ha sǐt pu mǐn* ‘only Phú’ crosses over the direct object in (10a), which contains a coreferential 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. This is despite the fact that the binder now presumably c-commands the pronoun. Likewise, the DC-moved *wh*-phrase *t^hɛj* cannot bind the pronoun within the direct object either (10b).

- (10)a. *t^ha sǐt pu mǐn_i* kǎw m̄jan lǎj? ɓɔp⁴⁵ ju_{*i/j} ka *t^ha sǐt pu mǐn*
 only Phú EMPH 1SG return wallet(VN) 3.ANIM
 ‘I only returned Phú his wallet.’ *DC-movement*
- b. *t^hɛj* hi m̄jan lǎj? ɓɔp⁴⁵ ju_{*i/j} ka *t^hɛj_i*
 who 2SG return wallet(VN) 3.ANIM
 ‘Who did you return their wallet to?’ *DC-movement of wh-phrase*

Fifth, DC-movement is subject to locality effects, as with \bar{A} -movement in general. In Eastern Cham, locality effects arise in the form of path containment effects (Pesetsky 1982, 309). Path containment effects arise for phenomena like *wh*-movement in English when there are multiple movement paths, or in an Agree framework, multiple probes and multiple Agree relations (Chomsky 2000). The descriptive generalization is that movement paths must be nested; paths cannot cross. To illustrate, in (11) two phrases are DC-moved, *han ni* ‘this cake’ and *nǐ? mɛj sǐt nǎn* ‘that little girl’. 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 as in (11a). If the paths are crossed, however, as in (11b), the resulting sentence is strongly and consistently ungrammatical.

- (11)a. [han ni] [niʔ məj sít nǎn] tʰu:ŋm̄³¹² ʔa niʔ məj sít 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. *[niʔ məj sít nǎn] [han ni] tʰu:ŋm̄³¹² ʔa niʔ məj sít 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.’

Baclawski Jr. and Jenks (2016) argue for the related language Moken that these effects are due to multiple C-heads each with a syntactic probe that operates via locality. The first probe in (11) searches and finds the closest phrase, *niʔ məj sít nǎn* ‘that little girl’. Only then can the next probe search and find *han ni* ‘this cake’. Crossed paths are illicit, because they require one probe to look past the most local phrase. It is worth noting that identical effects arise in embedded clausal peripheries (12), implying that the pattern is not specific to the matrix clausal periphery.

- (12)a. kǎw hniŋ [han ni] [niʔ məj sít nǎn] tʰu:ŋm̄³¹² ʔa
 1SG think cake this child female small that Thuận(VN) invite
 niʔ məj sít nǎn maj bǎŋ han-ni
 come eat
 ‘This cake, I think Thuận invited that little girl to come eat.’
- b. *kǎw hniŋ [niʔ məj sít nǎn] [han ni] tʰu:ŋm̄³¹² ʔa
 1SG think child female small that cake this Thuận(VN) invite
 niʔ məj sít nǎn maj bǎŋ han-ni
 come eat
 INTENDED: ‘This cake, I think Thuận invited that little girl to come eat.’

As in Moken, multiple *wh*-phrases can be DC-moved in Eastern Cham, and the same path containment effects arise. This is unexpected under an optional *wh*-movement account, as *wh*-movement is subject to Superiority effects, which would predict the opposite grammaticality pattern (i.e. that (13a) is ungrammatical and (13b) grammatical). Instead, these facts can be explained if there are multiple separate probes in the clausal periphery.

- (13)a. [kət] [tʰɛj] tʰu:ŋm³¹² ʔa tʰɛj maj bǎŋ kət
 what who Thuận(VN) invite come eat
 ‘Who did Thuận invited to come eat what?’
- b. * [tʰɛj] [kət] tʰu:ŋm³¹² ʔa tʰɛj maj bǎŋ kət
 who what Thuận(VN) invite come eat
 INTENDED: ‘Who did Thuận invited to come eat what?’

It should also be noted that the same path containment effects also arise if one *wh*- and one non-*wh*-phrase are moved to the clausal periphery (i.e. *han ni* ‘this cake’ and *tʰɛj* ‘who’). This can be explained if both phrases are possible goals for the same probe, resulting in locality effects. This would not be expected if movement of the *wh*-phrase is syntactically independent from movement of the non-*wh*-phrase.

Finally, multiple DC-movement presents evidence against a clefting analysis of DC-movement. Multiple clefts cannot coexist in the same clausal periphery in Eastern Cham, as in (14a–b). If the examples above were all due to some kind of cleft, it would have to be explained why multiple clefts are only sometimes licit.

- (14)a. * hu tʰa kleh han hu tʰa jaŋ nujh kǎw ʔa
 EX.COP one piece cake EX.COP one CLF.PERSON person 1SG invite
 tʰa-jaŋ-nujh maj bǎŋ tʰa-kleh-han
 come eat
 INT.: ‘There is a piece of cake there is a person I invited to come eat.’ *Cleft*
- b. * hu kət hu tʰɛj kǎw ʔa tʰɛj maj bǎŋ kət
 EXIST what EXIST who 2SG invite come eat
 INTENDED: ‘Who is that what is it that you invited to come eat?’ *Cleft*

Taken together, these pieces of evidence lead to the conclusion that DC-movement is an \bar{A} -movement operation. It is also a unified phenomenon. Both *wh*- and non-*wh*-phrases can be DC-moved. Furthermore, neither is due to clefting or hanging topics. The next sections examine the pragmatics of DC. DC is found to mark not information structure, but discourse structure.

4. The Pragmatics of DC-movement

This section finds that what appears to mark topic in Eastern Cham in fact marks discourse structural pragmatics. An informal statement of DC is given in (15), repeated from the introduction. For a phrase to be DC-marked, it must have an antecedent in the discourse, and there must also be a rhetorical relation such that the anaphor’s sentence explains or elaborates upon the antecedent’s. This is comparable to the analysis of clitic dislocation in Catalan introduced by López (2009).⁴

⁴DC-marking in Eastern Cham differs from clitic dislocation in Catalan in at least two ways: in Eastern Cham, pronouns cannot be DC-moved, while *wh*-phrases can. The converse occurs in Catalan.

- (15) DC (informal): A phrase can be marked as DC iff it is previously mentioned in a sentence that the current sentence explains or elaborates

A positive example of DC-marking is seen in (16). From the perspective of (16b), the phrase *ʔiŋ ʔɔŋ nǎn* ‘that frog’ is previously mentioned. Additionally, (16b) is naturally interpreted as an elaboration on the event of cooking in (16a). As a result, *ʔiŋ ʔɔŋ nǎn* ‘that frog’ is free to be moved to the left edge of the clause. In this example and all others in this paper, DC-movement is optional; the phrase may just as well be pronounced in situ.

- (16)a. t^hu:ŋm³¹² ʔɔʔ ŋǎʔ ʔiŋ ʔɔŋ nǎn
 Thuận(VN) PROG make frog that
 ‘Thuận is cooking that frog.’
- b. ʔiŋ ʔɔŋ nǎn ju ʔɔʔ ŋǎʔ ʔiŋ ʔɔŋ nǎn ŋi lo
 frog that 3.ANIM PROG make delicious very
 ‘That frog, he is cooking very well [Lit: deliciously].’

Two examples of the failure of DC-marking are given in (17c–d). In (17c), there is no previous mention, as *kiwʔ* refers to a different kind of frog. In (17d), the sentence is interpreted as a sequence of events, not an explanation or elaboration. As a result, movement of the respective phrases to the left edge of the clause is infelicitous. These sentences become felicitous if those phrases are pronounced in their base positions.

- (17)c. #kiwʔ nǎn ju ʔɔŋ ʔɔʔ ŋǎʔ kiwʔ nǎn
 frog that 3.ANIM also PROG make
 INTENDED: ‘That [other kind of] frog, he is also cooking.’
- d. #ʔiŋ ʔɔŋ nǎn ju ʔɔʔ ɓǎŋ ʔiŋ ʔɔŋ nǎn
 frog that 3.ANIM PROG eat
 INTENDED: ‘[Now], he is eating that frog.’

It is worth noting that consultants, when presented with (17c), report that it is only felicitous in the presence of additional discourse, such as *What did Thuận do with that frog?* In this case, (17c) is now interpretable as an elaboration on that prior discourse. Similarly, (17d) becomes felicitous if there is prior discourse about a set of kinds of frogs.

Based on the one positive example of DC-marking so far (16), 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. *ḥǎŋ nǎn kḥəh lo*
 door that beautiful very
 ‘That door is very beautiful.’
- b. *ḥǎŋ kǎw cǐh ḥǎŋ pan ɕaw*
 door 1SG paint color blue
 ‘The door, I painted blue.’

DC-marking also requires true previous mention; only full nominals can undergo DC-movement. Pronouns categorically cannot. For example, the third-person animate pronoun *nu* cannot in any context be DC-moved, even if it is stressed.

- (19) **nu kǎw hu kɔʔ ɲɰ mǐŋ pjoj*
 3.ANIM 1SG EX.COP meet yesterday
 INTENDED: ‘I did meet him/HIM yesterday.’

Discourse connectedness, thus, requires previous mention and explanation or elaboration. By the same token, many of the above examples could be analyzed as some kind of topicalization. Many of the English glosses, for example, can be paraphrased by an *As for X...* construction (cf. Reinhart 1981). However, certain phrases claimed not to be topicalizable are DC-movable in Eastern Cham. *Wh*-phrases behave like foci in many languages, and are non-topical if one assumes a strict divide between topic and focus. Yet, they can be DC-moved in Eastern Cham. Similarly, phrases with focus associators such as *only* can be DC-moved, as in (10a) above. Downward-entailing quantifiers such as *less than* are also predicted not to be topics by Ebert (2009) and others. Yet, *kiʔ hən mi jaŋ* ‘less than five people’ can be DC-moved in (20b).

- (20)a. *hi ʔa lo nujh lɛj*
 2SG invite many person Y/N.Q
 ‘Did you invite many people?’
- b. *kiʔ hən mi jaŋ kǎw ʔa kiʔ hən mi jaŋ maj pǎk ni*
 few exceed five person 1SG invite come here
 ‘I invited less than five people to come here.’

There are several complications in this example. First, there is a set-subset relation between the antecedent and anaphor, which is expanded upon at the end of this section. Second, the two sentences form a question-answer pair. Still, (20b) elaborates upon the question in (20a), as it is an elaborating answer. In theories of rhetorical relations (e.g. Asher and Lascarides 2003, 332), answers can elaborate upon questions if they provide more information than requested. Another way to conceptualize the relation is that (20b) elaborates upon the answer *No, I did not invite many people*.

DC-movement is also distinct from contrastive topic (CT). Following Büring (2003) and Constant (2014), CTs arise from complex answers to Questions Under Discussion. The example below, (21a), essentially raises a multiple *wh*-question, as *züt* refers to multiple addressees in the context (i.e. *Which person did which of you*

friends invite?). In this context, if the CT is inside the predicate, the existential copula *hu* appears preverbally. If the subject is a CT, *hu* appears sentence-initially. This is the preferred strategy of CT-marking in Eastern Cham, not DC-movement.

- (21)a. züt ʔa jaŋ hlěj maj pǎʔ ni
 friend invite which person come in this
 ‘Which person did you[*friends*] invite to come here?’
- b. cěj hu ʔa t^hu:ŋm³¹²_{CT} maj pǎʔ ni...
 self EX.COP invite Thuận(VN) come here
 ‘I[*myself*] invited THUẬN to come here...’ Object CT
- c. hu cěj_{CT} ʔa t^hu:ŋm³¹² maj pǎʔ ni...
 EX.COP self invite Thuận(VN) come here
 ‘I[*MYSELF*] invited Thuận to come here...’ Subject CT

With these facts in mind, (22) begins to formalize a semantics of DC-marking, though it remains a heuristic. In Discourse Representation Theory (DRT; Kamp and Reyle 1993), each sentence is represented as a Universe, a tuple of referents and predicates therein. Segmented DRT (Asher and Lascarides 2003) adds rhetorical relations to this formalization. In SDRT, explanation and elaboration are subtypes of *discourse subordination*, a broad class of rhetorical relations that expand upon a prior sentence and keep it open for further discussion (cf. Asher and Vieu 2005 for diagnostics). In a discourse subordination relation, the superordinate sentence is the one being explained or elaborated upon. (22a) defines U^\uparrow as any superordinate Universe to the current sentence.

- (22)a. DC (formal, heuristic): Let U^\uparrow be the Universe of any superordinate sentence
 b. $\llbracket \text{DC} \rrbracket = \lambda x : x \in U^\uparrow .x$

Given that only nominals can be DC-moved in Eastern Cham, one possible analysis of DC-marking is that there is a particle that merges with DPs and marks them as DC, much like focus particles and Q-particles for *wh*-phrases (Cable 2010). (22b) gives a possible denotation for the DC-particle. It combines with an individual and returns that individual as an identity function provided it satisfies the presupposition that that individual is mentioned in U^\uparrow . If that presupposition fails, the resulting sentence has no interpretation, ensuring that the DC-particle must combine with a phrase in the appropriate discourse anaphoric configuration.

It is worth noting that DC-marking must also allow for set-subset relations, set-superset relations, and bridging. For example, the antecedent *t^haŋ nǎn* ‘that house’ can license the DC-marked phrase *hǎŋ* ‘door’ (23). This must also be the case for a variety of other examples throughout this paper, such as (20) above on the downward-entailing quantifier ‘less than’.

- (23)a. *tʰaŋ nǎn k̚həh lo* b. *h̚ǎŋ kǎw cih h̚ǎŋ pan ɕaw*
house that beautiful very door 1SG paint color blue
‘That house is very beautiful.’ ‘The door, I painted blue.’

DC-movement, thus, marks a discourse structural configuration in Eastern Cham. Information structural notions such as topic and contrastive topic do not account for the data or are marked separately in the language.

5. Optional *wh*-movement as DC-movement of *wh*-phrases

Finally, this section examines the pragmatics of DC-marking a *wh*-phrase, how it compares with D-linking, and how the semantics of DC can coexist with that of *wh*. First, (24) presents a positive example of a *wh*-phrase undergoing DC-movement in context. The prior sentence (24a) mentions multiple pots being cooked. Sentence (24b) elaborates upon that prior sentence by asking which pot is being cooked at that very moment.

- (24)a. *məŋ mi kǎw tǔʔ ʔiŋ ʔəŋ tʰa k̚əʔ h̚əŋm̄ kiwʔ tʰa k̚əʔ*
look father 1SG boil frog 1 pot with kiek 1 pot
‘Look at my father boil one pot of frog and one of kiek.’
b. *jǎʔ ni k̚əʔ k̚et oŋ nǎn t̚əʔ ŋǎʔ k̚əʔk̚et nǎn*
now pot what old.man that PROG make that
‘Now, what pot is that old man making [working on]?’

Recall the informal characterization of DC pragmatics repeated below. In (24b), *pots* is in some way previously mentioned, and the sentence is interpreted as elaborating on (24a).

- (25) DC (informal): A phrase can be marked as DC iff it is previously mentioned in a sentence that the current sentence explains or elaborates

As in the previous section, previous mention and elaboration or explanation are crucial. If a different set of things is questioned, such as the set of things the old woman is cooking (26c), the *wh*-phrase cannot be DC-moved. Second, if the sentence is interpreted as being in a sequence of events, DC-movement is also infelicitous. For example, in (26d), the old man has finished cooking the pots of frogs and is now in the process of eating. Here, DC-movement is again infelicitous. Note that these sentences are felicitous if the *wh*-phrase is pronounced in situ.

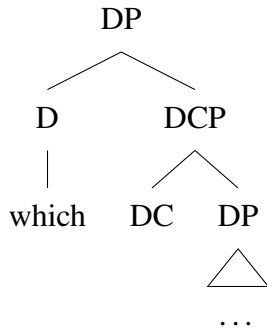
- (26)c. jǎʔ ni ket muʔ nǎn tɔʔ ɲǎʔ ket nǎn
 now what old.woman that PROG make that
 ‘Now, what is that old woman making [working on]?’
- d. #jǎʔ ni kɔʔ ket oŋ nǎn tɔʔ bǎŋ ket ket nǎn
 now pot what old.man that PROG eat that
 ‘Now, what pot is that old man eating?’

When consultants are presented with contexts like (26), they accept them only with additional prior discourse or different rhetorical relations. For example, (26d) is accepted if the old man is tasting one pot to see if it is done. This enforces an elaboration relation with (24a).

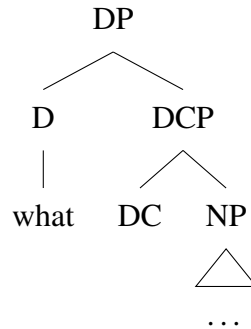
Descriptively, *wh*-phrases can be DC-marked, but how can DC-marking co-exist with *wh*-semantics? Under Alternative Semantics, a *wh*-phrase is a set of alternative possible answers and has no regular semantic interpretation, so it is not immediately clear how it can be previously mentioned (Rooth 1992). I posit that DC-marking occurs inside the DP, before the *wh*-D-head is merged. In general, *wh*-phrases have either a DP or NP complement. These complements have regular DP and NP denotations; it is the D-head that initiates the computation of alternatives. It is the complement that combines with the DC-particle, as represented in Figure 1.

Figure 1: DC-marking of *wh*-phrases

(a) D-linked *wh*-phrase



(b) Bare *wh*-phrase



Under this analysis, DC-marked *wh*-phrases can be of the form *which of the X*, such that *X* is DC-marked, or *which one*, such that *one* is DC-marked. This analysis introduces a complication. In Figure 1, the DC-particle combines with an NP, not a DP. Therefore, there must be at least two DC-particles, one that combines with an individual and one that combines with a property (27). The same DC pragmatics apply, as Universes consist of not only individuals but also properties and predicates.

- (27)a. DC (formal, heuristic): Let U^\uparrow be the Universe of any superordinate sentence
 b. $\llbracket DC_A \rrbracket = \lambda x : x \in U^\uparrow . x$
 c. $\llbracket DC_B \rrbracket = \lambda P : P \in U^\uparrow . P$

This analysis predicts that D-linked *wh*-phrases of the form *which X* and bare *wh*-phrases of the form *what* can be DC-moved in Eastern Cham. This prediction is borne out (28).

- (28) { *nĩʔ hlěj / nĩʔ kət / kət* } hi jɔŋ *nĩʔ hlěj/nĩʔ kət/kət*
 animal which / animal what / what 2SG raise
 ‘Which animal/What animal/What do you raise?’

It is worth further examining the interaction between DC-marking and D-linking. D-linked *wh*-phrases are those whose set of alternatives is saliently shared by the speaker and addressee (Pesetsky 1987; Comorovski 1996). It has been argued that *wh*-phrases can be topicalized if they are D-linked (e.g. Pan 2014). Given the broad similarity between DC and topicality, it is possible that D-linking is related as well. However, as with topicality, D-linking is insufficient for diagnosing DC-movement in Eastern Cham. This is because D-linking independently exists in the language.

Recall from Section 3 above on weak crossover that *wh*-phrases can bind a pronoun. This is true for D-linked *wh*-phrases of the form *which X*, but not true for other *wh*-phrases (29a–c).

- (29)a. *nĩʔ hlěj_i* hi jɔŋ *ɲu_i*
 animal which 2SG raise
 ‘Which animal do you raise?’
 b. ??*nĩʔ kət_i* hi jɔŋ *ɲu_i*
 animal what 2SG raise
 ‘What animal do you raise?’
 c. ??*kət_i* hi jɔŋ *ɲu_i*
 what 2SG raise
 ‘What do you raise?’

According to Pesetsky (1987) and many others, *wh*-phrases of the form *which X* are obligatorily D-linked. Every other kind of *wh*-phrase, such as *what X* and *what* are only D-linked in a context that makes the alternative set explicit. This is borne out in Eastern Cham. A nearly identical sentence to (29) above becomes felicitous across the board if the set of animals is explicitly specified (30).

- (30)a. *pih nĩʔ thun* pǎʔ ni *nĩʔ hlěj_i* hi cəh jɔŋ *ɲu_i*
 all animal here animal which 2SG like raise
 ‘[Of] all the animals here, which animal do you like to raise?’

- b. pih nĩʔ thun pǎʔ ni nĩʔ kət_i hi jəŋ ɲu_i
 all animal here animal what 2SG like raise
 ‘[Of] all the animals here, what animal do you like to raise?’
- c. pih nĩʔ thun pǎʔ ni kət_i hi jəŋ ɲu_i
 all animal here what 2SG like raise
 ‘[Of] all the animals here, what do you like to raise?’

D-linking, thus, is active in Eastern Cham grammar. However, it is orthogonal to DC-movement. As seen above in (28), DC-movement of all types of *wh*-phrase is possible, regardless of whether they are D-linked.

6. Conclusion

To conclude, DC-movement in Eastern Cham arises under a specific discourse structural configuration. The moved phrase must be mentioned in a prior sentence that the current sentence explains or elaborates upon. Furthermore, both *wh*- and non-*wh*-phrases can be DC-moved. Information structural phenomena, such as topicalization, clefting, and D-linking have all been shown to be insufficient in accounting for DC-movement. This provides evidence that there must be an interface between syntax and discourse structure.

Discourse structure is more typically marked by prosody and *cue phrases* (cf. Grosz and Sidner 1986). For example, *After that* cues a sequence of event relation, and *For example* cues an elaboration. DC-movement is similar to these cue phrases in that it cues explanation or elaboration. Its form, however, is a syntactic operation, not an independent adverbial.

In terms of syntax, this paper has proposed that there is a DC-particle alongside other particles responsible for \bar{A} -movement operations, such as focus particles and Q-particles for *wh*-movement (Cable 2010). At least in Eastern Cham, the DC-particle can coexist with the Q-particle, both marking the same phrase. More is needed to be said about the place of DC within the general realm of \bar{A} -features (cf. Aravind 2018 on the \bar{A} -feature hierarchy).

This paper has also sketched a possible semantic account of DC, but it remains to be fully formalized. One possible avenue is the Question Under Discussion framework (Roberts 1998). However, to formalize DC this way, the notion of questions and sub-questions will have to be broadened to include explanation and elaboration. This has been incorporated into some implementations of QUD in corpus linguistics (Riester et al. 2018), but it has yet to be formalized. Instead, a question can only be subordinate to another if it contributes in some way to the answer of the superordinate question (cf. Constant 2014). This would undergenerate the Eastern Cham examples of DC-movement. More promising is an event semantic approach, as explanation and elaboration are taken to be relations between eventualities in theories like SDRT (Asher and Lascarides 2003).

References

- Aravind, Athulya. 2018. Licensing long-distance wh-in-situ in Malayalam. *Natural Language & Linguistic Theory* 36:1–43.
- Asher, Nicholas, and Alex Lascarides. 2003. *Logics of conversation*. Cambridge University Press.
- Asher, Nicholas, and Laure Vieu. 2005. Subordinating and coordinating discourse relations. *Lingua* 115:591–610.
- Baclawski Jr., Kenneth. 2017. Clause-final particles and focus in Eastern Cham. In *Proceedings of the Linguistic Society of America*, volume 2.18, 1–12.
- Baclawski Jr., Kenneth. 2018. Diglossia and change from below in Eastern Cham. *Asia-Pacific Language Variation* 4:73–102.
- Baclawski Jr., Kenneth, and Peter Jenks. 2016. Clefts and Anti-Superiority in Moken. *Journal of the Southeast Asian Linguistics Society* 9:81–96.
- Brunelle, Marc. 2009. Diglossia and monosyllabization in Eastern Cham: A sociolinguistic study. In *Variation in Indigenous Minority Languages*, ed. James Stanford and Dennis Preston, 47–75. John Benjamins.
- 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. 1997. *On the typology of wh-questions*. New York: Garland.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Springer.
- Constant, Noah. 2014. Contrastive topic: Meanings and realizations. Doctoral Dissertation, University of Massachusetts, Amherst.
- 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.
- Grosz, Barbara, and Candace Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12:174–204.
- Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer.
- López, Luis. 2009. *A derivational syntax for information structure*. Oxford University Press.
- 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* 110:707–727.
- Pesetsky, David. 1982. Paths and categories. Doctoral Dissertation, MIT.
- Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In *The Representation of (in)definiteness*, ed. Eric J. Reuland and Alice G.B. ter Meulen, 98–129. Cambridge, MA: MIT Press.

- Postal, Paul M. 1971. *Crossover phenomena*. New York: Holt, Rinehart and Winston.
- Potsdam, Eric. 2006. More concealed pseudoclefts in Malagasy and the Clausal Typing Hypothesis. *Lingua* 116:2154–2182.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27:53–94.
- Riester, Arndt, Lisa Brunetti, and Kordule De Kuthy. 2018. Annotation guidelines for questions under discussion and information structure. In *Information structure in lesser-described languages: Studies in prosody and syntax*, ed. Evangelia Adamou, Katharina Haude, and Martine Vanhove, 403–443. John Benjamins.
- Roberts, Craige. 1998. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 6:1–69.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Sato, Yosuke. 2011. P-stranding under sluicing and repair by ellipsis: Why is Indonesian (not) special? *Journal of East Asian Linguistics* 20:339–382.
- 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 *The Oxford handbook of comparative syntax*, ed. Guglielmo Cinque and Richard Kayne, 806–838. Oxford University Press.
- Thurgood, Graham. 1996. Language contact and the directionality of internal drift. *Language* 72:1–31.
- Thurgood, Graham. 1999. *From ancient Cham to modern dialects: Two thousand tears of language contact and change*. University of Hawai'i Press.
- Thurgood, Graham. 2005. A preliminary sketch of Phan Rang Cham. In *The Austronesian languages of Asia and Madagascar*, ed. K. Alexander Adelaar and Nikolaus Himmelmann, 489–512. Routledge.
- Wang, Chyan-an Arthur. 2007. Sluicing and Resumption. In *Proceedings of the 37th North East Linguistic Society*, 239–252. Amherst, MA: GLSA.