UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Discourse connectedness: The syntax--discourse structure interface

Permalink
https://escholarship.org/uc/item/0349w4fh

Author
Baclawski Jr, Kenneth Paul

Publication Date
2020

Peer reviewed|Thesis/dissertation

Discourse connectedness: The syntax–discourse structure interface

by

Kenneth Paul Baclawski Jr.

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Peter Jenks, Chair
Associate Professor Line Mikkelsen
Professor Eve Sweetser
Associate Professor Seth Yalcin

Fall 2019

Discourse connectedness: The syntax–discourse structure interface

Abstract

Discourse connectedness: The syntax–discourse structure interface

by

Kenneth Paul Baclawski Jr.

Doctor of Philosophy in Linguistics

University of California, Berkeley

Associate Professor Peter Jenks, Chair

This dissertation argues for the existence of a new Ā-feature, *discourse connected* (DC), which grammatically encodes a constraint on the relation between the constituent it attaches to and discourse relations to previous sentences. Connectives like *That's because* and *For example* encode the rhetorical relations of explanation and elaboration. DC encodes these relations as well, but by Ā-movement of a phrase to the left edge of a clause or noun phrase, specifically the movement of a phrase that is previously mentioned in the sentence that the current one is explaining or elaborating upon. I argue that there must be a DC-feature in the lexicon on par with other Ā-features, such as *wh*. Given that DC encodes a discourse structural constraint, there must be a syntax–discourse structure interface.

One consequence of this dissertation is descriptive: a range of phenomena in the Eastern Cham language (Austronesian: Vietnam) are found to be instances of DC-marking. These phenomena include what appear on the surface to be topicalization, optional *wh*-movement, partitives, and inventory forms (e.g. *bread, three loaves*). I argue that these phenomena in Eastern Cham must be analyzed in terms of DC, not information structure or other previously proposed analyses for comparable constructions in other languages. As a result, multi-sentence discourses that control for discourse structure should be used as diagnostics for constructions that might be licensed by DC or something similar.

This dissertation adds a new Ā-feature, DC, to the typology of Ā-features. DC-movement shares a variety of characteristics with Ā-movement more broadly. It exhibits sensitivity to syntactic islands, weak crossover, and locality effects. As with other Ā-features in some cases, there is a parallelism between CP and DP: a phrase can be DC-moved to the left edge of either a clause or a noun phrase. Similarly, *wh*-phrases can undergo secondary movement to the left edge of the noun phrase in some languages. Additionally, this dissertation examines the position of DC among other Ā-features, following recent work on the hierarchy of Ā-features (Aravind 2017, 2018). Despite its apparent optionality,

DC-movement is argued to be obligatory, like *wh*-movement. DC is also found to be independent from other features such as *wh*; the same phrase can be marked as both DC and *wh*. Unlike *wh*, DC can only be checked once on a respective phrase in a derivation, and the movement of multiple DC-phrases to the edge of the same clause exhibits Path Containment Effects (Pesetsky 1982), not the tucking-in pattern observed in multiple *wh*-movement (Richards 1997).

DC is argued to provide evidence for the need for a dynamic event semantics that allows the events introduced throughout the discourse to be tracked. This dissertation proposes that the DC-feature is introduced by a DC-particle on analogy with focus particles and the Q-particle on *wh*-phrases (Cable 2010). The DC-particle is shown to introduce a presupposition that checks the participants of two events in a discourse: the current event and an event in a prior sentence inferred by a subordinating discourse relation (i.e. explanation or elaboration).

The dissertation proceeds as follows. Chapter 1 introduces DC, along with the ongoing debate on the existence of pragmatic features in syntax. Additionally, the concept of *hierarchical discourse constraints* (HDCs) is introduced as a heuristic to understand that position of DC in comparison with *linear information states* (LISs), exemplified by information structural notions like old information topic. Chapter 2 examines the basic case of DC-movement, topicalization, and demonstrates that DC must be an Ā-feature. In Chapter 3, *wh*-phrases are shown to be able to be DC-marked in Eastern Cham. An analysis is proposed in which DC-particles and Q-particles can be present in the same DP. Chapter 4 turns to DC-marking inside a noun phrase and finds a CP/DP parallelism: phrases can be DC-moved to the left edge of a noun phrase. Despite these cases involving movement inside a noun phrase, the event semantic interpretation of DC is affirmed: DC is only computed between the larger events of which the noun phrase is a part. Chapter 5 concludes. Additionally, it analyzes clitic right-dislocation in Catalan as DC-movement with minor syntactic and semantic differences from DC-movement in Eastern Cham (cf. López 2009). Contrastive topic is also examined and proposed to be a second hierarchical discourse constraint, different from DC.

To Mom & Dad

Elise & Indy

Nana & Mohammad

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation is the result of kindness, support, and hard work from so many people. First and foremost, I owe a huge debt of gratitude to all the Cham people of Vietnam. Sakaya (Dr. Trương Văn Món) was instrumental in helping me start fieldwork in Vietnam. Hamu Ligaih (Sikhara, Đàng Thanh Quốc Thuận) has been there by my side for almost five years as a research assistant, fellow Cham scholar, and friend. And above all, Haji Mohammad Saleh Thiên (Thiên Sanh Phân) was my first Cham teacher. While I have so many fond memories of the kindness of the Cham people, not to mention the wonderful food, I think most about a small moment with Mohammad. In 2014, I had just started studying the Eastern Cham language, and I was trying to construct some complicated sentences with difficult pragmatics. Mohammad kept shaking his head, saying, "No…No…Why would you want to say that?" But then, I tried one particular sentence in one particular context, and his eyes lit up. He said, "Now, you're doing it! Now you're speaking Cham!" It turns out that particular sentence led me down the path towards this dissertation, and after five years I am only starting to appreciate all the nuances of that one sentence.

Next, I am so grateful for my committee, Peter Jenks, Line Mikkelsen, Seth Yalcin, and Eve Sweetser. I have been working on this Eastern Cham data with Peter and Line for years now, and our discussions have been fascinating and thought-provoking throughout the whole process. They have always given me detailed and careful feedback, even when I am starting to feel that I have written every possible permutation of words. Peter has been a wonderful advisor because of his enthusiasm, energy, and passion for linguistics. I would leave our meetings with many new ideas and a renewed spark every time. I am also thankful that Peter was willing to explore such new topics as information structure and discourse structure with me in order to arrive at this project. Line first inspired me to explore the connection between syntax and discourse structure. Since then, she has been a great grounding force in my academic life, always giving astute comments, driving me to be careful and rigorous in my analysis, and helping me find my voice in giving presentations. Seth has really transformed this dissertation, even though we relatively recently started working together. He dove with me into the puzzle of discourse connectedness and brought a significant amount of clarity to the project. Finally, Eve has been a great friend and mentor, who has given me thoughts from numerous different angles on my research. I am happy this dissertation has somewhere in its pages an initial foray into Eastern Cham gesture.

I am proud to have been able to be a part of the Berkeley Linguistics department for almost seven years now. It has been a wonderful community full of dear friends. I will list just a few here. If you are reading this and wondering why you don't see your name, rest assured, I will remember you seconds after I click the submit button, and I will feel mortified. To my cohort, Anna Jurgensen, Spencer Lamoureux, Hannah Sande, and Alison Zerbe. In no particular order, to Sarah Bakst, Nico Baier, Emily Clem, Emily Cibelli, Meg Cychosz, Ginny Dawson, Jevon Heath, Edwin Ko, Schuyler Laparle, Myriam

other Clivias around it had red fruit as well, but this one lasted days, weeks, months, long after the other Clivias' fruits had dropped or been eaten. Every day, I walked by, wondering how long it would hold onto that fruit. The rainy season started, and the little Clivia would get battered by rain, still holding onto its fruit. The other Clivias soon started flowering again, but not this one. I wondered what was going through the Clivia's mind. Why have the other plants moved on already? When is my fruit finally going to fall? Will I ever flower again?

I lost track of the Clivia, but as the years have gone on, the more I thought about it. On the day I left Berkeley, I went back to La Loma to check on it. It was a bright, sunny day, with a gorgeous view of the Bay and San Francisco. Sure enough, the Clivia was still there at the left foot of the bench. The fruit had long since gone, and it was just a mess of green leaves. It seemed to be doing fine.

Chapter 1

Introduction

This dissertation argues for the existence of a new Ā-feature, *discourse connected* (DC), which grammatically encodes a constraint on the relation between the constituent it attaches to and discourse relations to previous sentences. There are two components of this claim: first, that DC is an Ā-feature, and second, that it encodes a discourse relational constraint. This introduction will unpack those two claims in turn.

To start, what exactly is *discourse connectedness*? To illustrate, consider the following mini-discourse in Eastern Cham, an Austronesian language spoken in Vietnam and the primary empirical focus of this dissertation. Observe that *ʔiŋ ʔɔŋ năn* 'that frog' is moved to the left edge of the clause in (1.1b). Next, observe that *ʔiŋ ʔɔŋ năn* 'that frog' is previously mentioned in (1.1a), which also represents the basic word order in Eastern Cham. Finally, observe that (1.1b) elaborates on (1.1a), saying more about Thuận's cooking of that frog.

(1.1)   a. tʰuːɲ͡m³¹²   ʈɔʔ   ŋăʔ   **ʔiŋ ʔɔŋ   năn**
         Thuận(VN)   PROG   make   frog   that
         'Thuận is cooking that frog.'

   b. **ʔiŋ ʔɔŋ   năn** ɲu   ʈɔʔ   ŋăʔ   ~~ʔiŋ ʔɔŋ năn~~ ɲi   lo
      frog   that   3.ANIM   PROG   make   be.delicious   very
      'That frog, he is cooking very well [Lit: deliciously].'
      (a) is explained or elaborated upon by (b)

This dissertation argues that the movement of *ʔiŋ ʔɔŋ năn* 'that frog' in (1.1b) is an Ā-movement operation driven by a syntactic probe searching for some feature. That feature is called *discourse connected* (DC), and it is only assigned to phrases that are mentioned in a previous sentence in the discourse, such that the current sentence explains or elaborates upon that previous one. Previous mention alone is not sufficient. If the discourse is continued instead by (1.1b′), movement of *ʔiŋ ʔɔŋ năn* 'that frog' is infelicitous. This will be argued to be the case because the sentence moves on a new event, eating the frog, and is not interpreted as explaining or elaborating upon (1.1a).

(1.1)   b′. #jăʔ ni **ʔiŋ ʔɔŋ năn** ɲu   t̪ɔʔ   ɓăŋ ʔ̶i̶ŋ̶ ̶ʔ̶ɔ̶ŋ̶ ̶n̶ă̶n̶
        now   frog   that 3.ANIM PROG eat
        INTENDED: 'Now, that frog, he is eating.'
        (a) is <u>not</u> explained or elaborated upon by (b)

To take a step back, what exactly is an Ā-feature? This dissertation is concerned with syntactic movement, where a phrase is pronounced and/or interpreted in a higher structural position than its base position, as in the (b) sentences above. A-movement typically refers to movement operations to argument positions. Ā-movement typically refers to movement operations beyond argument positions, often marking discourse or information structural features, most prototypically *wh* (i.e. *who, what,...*; Chomsky 1977). In the Minimalist program (Chomsky 1995), Ā-features are the building blocks of those Ā-movement operations. A phrase can bear one or more such features under the appropriate circumstances.

Ā-features are known to share a variety of characteristics, from unboundedness to sensitivity to island constraints, weak crossover, and locality effects. This dissertation argues that DC shares all of these characteristics and should therefore be thought of as an Ā-feature, alongside *wh* and others like topic and focus. A further shared characteristic is argued to be a parallelism between CP and DP (i.e. clauses and noun phrases). Under the Agree framework (Chomsky 2000), Ā-movement results from a syntactic probe entering an Agree relation with a phrase bearing a certain feature. If the probe also bears an EPP feature, the phrase is moved to the specifier position of the probe. As has been proposed for *wh* (cf. Cable 2010 on the Q-particle) and focus (cf. Rooth 1992 on alternative semantics), the DC-feature is argued to be introduced by a DC-particle that can be merged with NPs or DPs. The DC-feature corresponds with syntactic probes on C^O and D^O that search for the DC-feature. In cases like Figure 1.1, that results in movement of the phrase to the specifier of C.

Figure 1.1: DC-movement

(a) Agree with C

CP
├── C
│   [uDC]
│   [EPP]
└── ...
    └── ... DCP
        ├── DC
        │   [DC]
        └── DP
            ʔiŋ ʔɔŋ năn
            'that frog'

(b) Move to Spec-CP

CP
├── DCP
│   ├── DC
│   │   [DC]
│   └── DP
│   ʔiŋ ʔɔŋ năn
│   'that frog'
└── C
    [uDC]
    [EPP]
    └── ...
        └── ... DCP
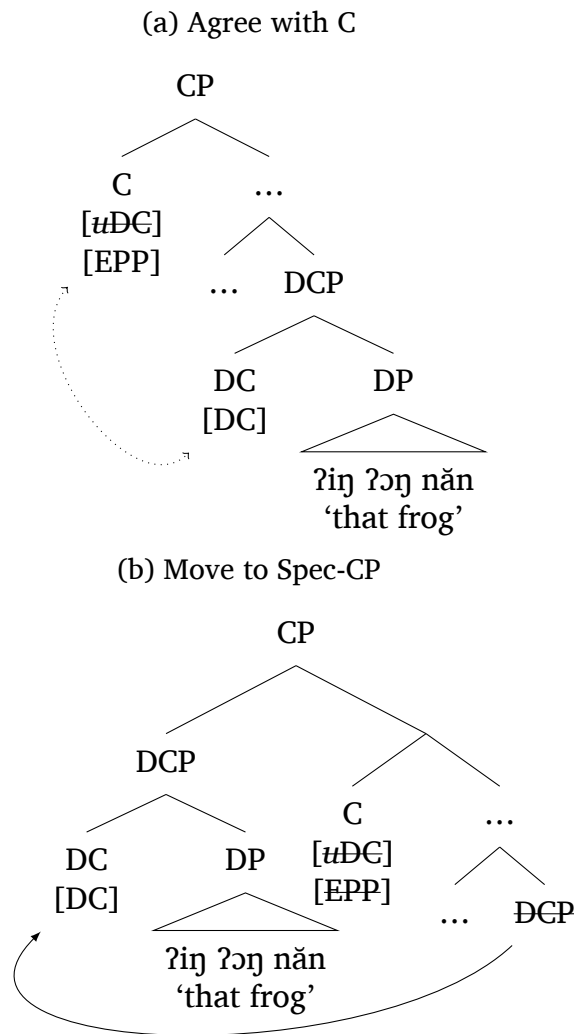
What licenses the DC-particle to begin with? This dissertation proposes that the DC-particle combines with phrases that satisfy the following DC conditions. The DC conditions place requirements on the phrase itself, the current sentence, a prior sentence in the discourse, and the semantic events those sentences introduce. First, the current sentence must be in a subordinating discourse relation with a prior sentence. According to theories of discourse relations such as Segmented Discourse Representation Theory (Asher & Lascarides 2003) and Rhetorical Structure Theory (Mann & Thompson 1988), explanation and elaboration exemplify a type of rhetorical relation between sentences called *discourse*

subordination.[1] Second, the current and prior sentence must introduce semantic events that correspond with the event relation inferred by that subordinating discourse relation. In theories of discourse relations, explanation infers a causal relation between two events and elaboration infers a subtype relation between two events (Asher & Lascarides 2003). Third, the phrase that the DC-particle combines with must be a participant in both of the aforementioned semantic events. Chapter 2, Sections 2.1.1–2.1.3 describe the DC conditions in more detail and give relevant background on discourse relations and event semantics.

Chapter 2, Section 2.1.4 proposes the three sets in (1.2a–d) and the meaning of the DC-particle in (1.2e). The set $\mathcal{E}_c$ contains the events so far added to the discourse context by its constituent sentences. The set $\mathcal{E}_e$ contains the events inferred by the explanation and elaboration relations that the current sentence is involved in. The set $\mathcal{P}_e$ contains the participants of a given event e. The DC-particle must combine with a phrase interpreted as an individual, indicated by the x variable.

(1.2) a. Let $\mathcal{E}_c$ be the set of events introduced in a discourse

  b. Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

  c. Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

  d. Let $\mathcal{P}_e$ be the set of participants in event e

  e. $[\![\mathrm{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

Then, the DC-particle introduces a presupposition that checks the DC conditions with regard to that individual. In prose, the presupposition requires that there be an event introduced by a prior sentence in the discourse, the current sentence must explain or elaborate upon that prior sentence, and the individual denoted by the phrase the DC-particle combines with must be a participant both in the event in the current sentence and that prior event. Chapter 2, Section 2.1.4 argues the need for a dynamic event semantics, without which $\mathcal{E}_c$ cannot be defined or checked.

Why should one study *discourse connectedness* in the first place? There are a variety of implications for syntax, the syntax–pragmatics interface, and pragmatics. Descriptively, DC presents a new possible analysis for a range of linguistic phenomena. The following phenomena in Eastern Cham will be analyzed in terms of DC-marking: topicalization (1.3a), optional *wh*-movement (1.3b), inventory forms (1.3c), and partitives (1.3d). Finally, these phenomena can feed other phenomena, such as an inventory form feeding quantifier float, resulting in a DC-marked subject (1.3e).

---

[1]See Fabricius-Hansen & Ramm (2008) for comparable notions in other theories and how it compares with subordination at the clause level.

(1.3)   *Manifestations of DC-marking in Eastern Cham:*

    a. **ʔiŋ ʔɔŋ năn**$_{DC}$ ɲu   ṭɔʔ   ŋăʔ   ~~ʔiŋ ʔɔŋ năn~~ ɲi   lo
       frog   that   3.ANIM PROG make   be.delicious very
       'That frog, he is cooking very well [Lit: deliciously].'   (Topicalization)

    b. **ḳeʔ**$_{DC}$ hɨ   ṭɔʔ   ɓăŋ   ~~ḳeʔ~~
       what 2SG PROG eat
       'What are you eating?'   (Optional *wh*-movement)

    c. kăw plĕj [$_{DP}$ **ʔɔ̆ʔ**$_{DC}$ çŭh pɔ̥h   ~~ʔɔ̆ʔ~~]
       1SG ɓuy   mango 7   CLF.ROUND
       'I bought mangoes, seven.'   (Inventory form)

    d. kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̆ʔ**   **ni**$_{DC}$] çŭh pɔh   ~~ʔɔ̆ʔ ni~~]
       1SG ɓuy   mango this   7   CLF.ROUND
       'I bought seven of these mangoes.'   (Partitive)

    e. **hɔŋm**$_{DC}$ lĕʔ tʃŭn [$_{DP}$ ~~hɔŋm~~ çŭh pɔh   ]
       papaya fall down   7   CLF.ROUND
       'Papaya, seven fell down.'   (Inventory form feeding quantifier float)

Next, the study of the specific syntactic and pragmatic properties of DC adds to a growing typology of Ā-features and informs the syntax–pragmatics interface. Aravind (2017, 2018) and others have proposed a feature geometry for Ā-features in order to account for shared properties between *wh*-movement and other Ā-movement operations like topicalization. Understanding the set of features as a whole is imperative for a full account of Ā-feature geometry. Hence, the place of DC as such a feature should be better understood.

DC also necessitates an interface between syntax and discourse structure. As mentioned above, the DC conditions place a requirement on two sentences such that they be in a subordinating discourse relation. This dissertation argues that DC can only be characterized in terms of these discourse relations. At the same time, DC is argued to be a syntactic Ā-feature. Therefore, there must be syntactic features that index discourse structure; there must be a syntax–discourse structure interface. Section 1.1 presents a heuristic for understanding the discourse requirement of DC: hierarchical discourse constraints. Section 1.2 discusses the ongoing debate around the existence of pragmatic features in syntax. In the end, it lays out under what syntactic and pragmatic circumstances DC would require a syntactic featural representation.

1.1   Hierarchical discourse constraints

This section introduces a contrast between *linear information constraints* (LICs) and *hierarchical discourse constraints* (HDCs), which will be used as a heuristic for understanding

DC, which is an HDC. To begin, imagine a discourse containing three sentences, S_1, S_2, and S_3. How do speech participants track such a discourse? Perhaps they track discourse as a linear string of updates, as in Figure 1.2a, where *t* stands for time. There is only one possible string for these three sentences, ordered by when each sentence was uttered.

Figure 1.2: Mental representations of discourse

(a) Linear string

(b) Hierarchy of discourse relations
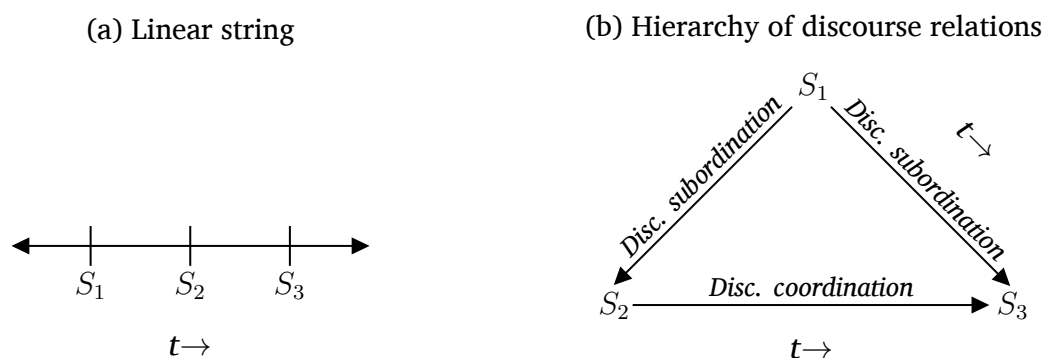
Or, perhaps the sentences can have an articulated hierarchical structure between them. One of multiple such hierarchical structures between three sentences is represented in Figure 1.2b. In this figure, discourse is represented as a directed graph, with subordinating discourse relations (i.e. explanation and elaboration) as downward arrows and coordinating discourse relations (i.e. narration) left-to-right. What separates hierarchical structure from linear strings is that subordinating discourse relations create hierarchical levels; the sentence being explained or elaborated upon is on a higher level in the discourse than the sentence doing the explaining or elaborating. In the hierarchy represented here, both S_2 and S_3 elaborate upon or explain the first sentence, and there is some kind of coordinating relation between them, such as a narration, or sequence of events.

I hypothesize that speech participants simultaneously track discourse both as a linear string and as a hierarchy of discourse relations.[2] For example, the same sequence of three sentences can be represented simultaneously as Figure 1.2a and Figure 1.2b. I posit this hypothesis because there is evidence for linear and hierarchical constraints on English pronoun resolution.

A linear information constraint (LIC) is a constraint on linear representations of discourse. For example, there could be an LIC on anaphora such that the antecedent has to be the first eligible referent to the left. For a pronoun in S_3, the LIC "look to the left" (or, linear closeness) picks out the referent y as an antecedent, and does not pick out x, as shown in Figure 1.3a. English pronouns, in some cases, find their reference via an LIC like this. According to Zeldes (2018) and others, linear closeness is a strong factor in predicting the reference of pronouns, and Baclawski Jr & Yang (2019) and others find a

---

[2]If the discourse only contains coordinating discourse relations, the linear and hierarchical representations are identical, at least in terms of the directed graph representation.

similar effect in experimental settings. In these studies, the LIC can outweigh whatever constraints are placed on the hierarchical discourse structure.

Figure 1.3: Discourse conditions

(a) Linear information state (LIC): closest phrase before S_3

(b) Hierarchical discourse constraint (HDC): closest phrase above S_3
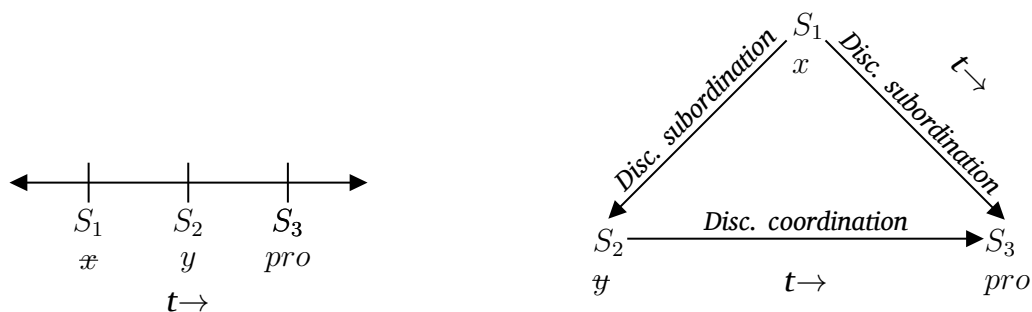
A hierarchical discourse constraint (HDC) is a constraint on the hierarchy of discourse relations. For example, there could be an HDC on anaphora such that the antecedent has to be the first eligible referent *above* the current sentence. For a pronoun in Figure 1.3b, the HDC "look up" picks out x as an antecedent, and does not pick out y. Put another way, S_2 is passed over in the search for an antecedent, because it is not above S_3. A similar HDC to "look up", termed the Right-Frontier Constraint (Polanyi 1985, 1988) or Veins Theory (Cristea, Ide & Romary 1998), has been argued to constrain the reference of English pronouns in corpora and experimental settings.[3,4] The HDC "look up" also reflects the discourse structural requirement of *discourse connectedness*.

The distinction between LICs and HDCs is meant to be a heuristic for understanding *discourse connectedness*. Chapters 2–4 argue that DC is a hierarchical discourse constraint represented syntactically by an Ā-feature. The existence of DC raises the question as to whether there are other HDCs. Chapter 5, Section 5.2 argues that contrastive topic reflects a different HDC from *discourse connectedness*. This dissertation will not make any claims about LICs. Perhaps, certain kinds of topic, such as old information topic, can be thought of in terms of an LIC.

---

[3]More precisely, the Right-Frontier Constraint states "look one left or up".

[4]How it is that English pronouns have both LICs and HDCs is an open question. Holler & Irmen (2007) and Baclawski Jr & Yang (2019) find evidence that the LIC is the basic case and that ambiguity or competition between possible antecedents gives rise to the HDC.

1.2   Pragmatic features in syntax

This dissertation argues for a syntactic feature which marks as its primary exponent DC, a hierarchical discourse constraint. It is not straightforward to claim that syntactic features can index such a constraint, however. This section details arguments for and against the inclusion of pragmatic features in syntax and arguments in favor of DC as a true Ā-feature.

In Chomsky's (2000) framework, there are two ways to derive movement, schematized in Figure 1.4a–b.[5] A syntactic head, X, can have an EPP-feature, which attracts the closest phrase to Spec-XP (Figure 1.4a). It is presumed that movement in this case proceeds by locality, so β would not undergo such movement. Otherwise, the head can have a probe that initiates an Agree relation with the closest phrase bearing the relevant feature, F (Figure 1.4b; e.g. Miyagawa 2009, shown by the dotted arrow).[6] In this case, it is β that undergoes movement to Spec-XP, as α lacks the feature. A non-local phrase such as β should not be able to be moved merely by an EPP-feature as in Figure 1.4c.

Figure 1.4: Syntactic movement mechanisms

(a) EPP (b) Agree + EPP (c) EPP + D-effect
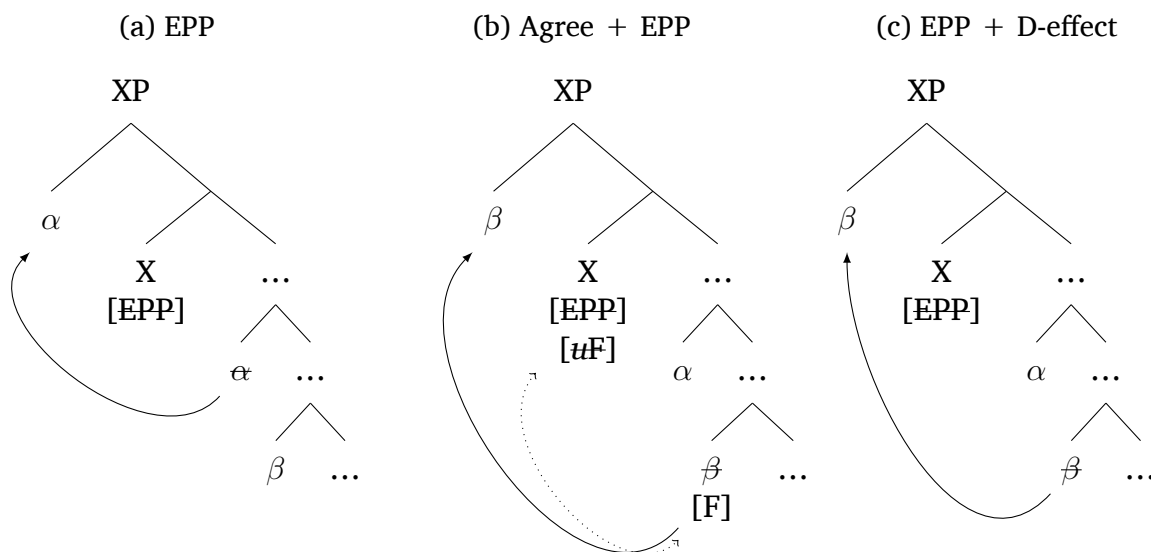
There are two major approaches to syntactic movement operations that appear to mark pragmatics or information structure, such as topicalization and focus-movement

---

[5]This section assumes an altruistic view of syntactic movement, though comparable issues arise under greed- and labeling-based views (cf. Zyman 2018 and references therein).

[6]In the notation of features, *u* is used throughout to indicate probes in a general sense, not necessarily unvalued or uninterpretable features (cf. Chomsky 2000, 2001). Edge features and probes could be substituted, for example, with the [•F•] and [∗F∗] notation of Müller (2010).

(e.g. Rizzi 1997). First, there are feature-based approaches that allow for syntactic features that mark topic, focus, and the like. Under these approaches, topicalization and focus-movement involve a probe and Agree relation, as in Figure 1.4b (or comparable formalisms). Second, there are non-syntactic approaches that ban syntactic features from marking pragmatics and information structure at all (e.g. Horvath 2010). These approaches require a different mechanism for deriving movement. One mechanism is to allow EPP-features to result in movement of a non-local phrase such as β in Figure 1.4c. Chomsky (2001, 2008) argues that a non-local phrase can undergo movement as long as it has some effect on the eventual interpretation, termed a Discourse- or D-effect.[7]

Feature-based approaches make the following predictions. If there are features in the lexicon that mark pragmatics, there could be overt morphemes that mark the same pragmatics, along with morphological phenomena that pertain to those features. Sections 1.2.1–1.2.2 present evidence that these predictions are borne out specifically for the information structural category topic. Another prediction made by a feature-based approach is that the syntactic movement operation in question should have a predictable relation between form and meaning. In other words, the moved phrase should only have the pragmatic interpretation indexed by the syntactic feature that triggers the movement.

Non-syntactic approaches make different predictions. If there are no pragmatic features in the lexicon, then no morphemes or morphological phenomena should consistently mark specific pragmatics. Additionally, there should be no consistent relation between form and meaning. The derivation in Figure 1.4c is allowed, on this account, if there is any effect on the outcome; any pragmatic interpretation should suffice (e.g. Mikkelsen 2015; Bossi & Diercks 2019). Finally, advanced in these approaches is that pragmatics is computed at the level of a sentence, after the syntactic derivation is complete. Therefore, syntactic features marking those pragmatics would violate Chomsky's (1995) Inclusiveness Condition. Section 1.3 outlines these arguments and how what we have already seen about DC provides counterarguments in favor of a feature-based approach for DC-marking.

It is worth noting that there is an intermediate approach that some phenomena are best explained as EPP-driven movement with a D-effect, as in Figure 1.4c above, while others are best analyzed as feature-driven movement as in Figure 1.4b. This intermediate approach would proceed phenomenon by phenomenon. This dissertation concludes that DC-movement in Eastern Cham should be analyzed via a DC-feature in the syntax, but it is compatible with this intermediate approach as well.

1.2.1 Topic morphology

First, the existence of morphemes that mark topic provides evidence that there should be syntactic features that mark pragmatics. This section focuses on topic, as no languages

---

[7]Slioussar (2007) argues that this D-effect can take the form of information structural templates like topic-comment and focus-presupposition.

with overt morphemes marking DC have been described or will be found in this dissertation. Topicality refers broadly to old information or previously mentioned information around which a sentence is organized (e.g. Reinhart 1981). If a morpheme exclusively marks topics and has systematic syntactic effects, then it follows that there exists a feature in the lexicon that this morpheme is the exponent of (e.g. Cruschina 2012: 32). Topic morphemes have been argued to exist in a variety of languages, such as the Gbe languages (Aboh 2004a; Aboh & Essegbey 2010), Somali (Frascarelli & Puglielli 2007), and Japanese (Miyagawa 2017). For example, in Gungbe, extracted topics are marked by *yà*. Topic markers are typically associated with syntactically moved topics. It is important that topic markers are distinct from focus markers (in Gungbe, *wὲ*), as it shows that they do not mark a more general category such as Ā-movement.

(1.4) Ùn sὲ ɖɔ̀ **Súrù yà** lέsì lɔ́ wὲ é ɖà gànjí
    1SG hear COMP Suru TOP rice DET[deixis] FOC 3SG cook well
    'I heard that, as for Suru, he cooked THE RICE really well.'   GUNGBE
    (Aboh & Essegbey 2010: (53))

Not only do topic morphemes imply that topic is a category relevant to the lexicon, they also put topic on similar footing as *wh*-movement. Cable (2010) and others have argued for the existence of Q-particles, which can be overtly marked in some languages. This is the function claimed for the form *sá* in Tlingit. As with the topic marker in Gungbe, the Q-particle in Tlingit is associated with *wh*-phrases that are moved.[8]

(1.5) **Aadóo yaagú sá** ysiteen ~~aadóo yaagú sá~~?
    who boat Q you.saw.it
    'Whose boat did you see?' (Cable 2010: 32)   TLINGIT

An alternative possible analysis is that topic markers are epiphenomenal; they mark some orthogonal category. Horvath (2010) argues that focus-marking in Hungarian is actually performed by an exhaustivity operator that makes no indepedent reference to focus. A similar analysis could be argued if topic markers only marked *contrastive* topics. Perhaps there is an independent contrast operator that makes no reference to topicality (e.g. Molnár 2006). This would predict that topicalization always cooccurs with something else, like contrast. However, this is not the case in Gungbe, where contrast is associated neither with topic nor focus marking (Ameka 2007).

One argument generally levied against a feature-based approach to topicalization is optionality. Cross-linguistically, topicalization appears to be optional in contexts where it would be predicted to occur (i.e. the phrase could be left in situ; Erteschik-Shir 2007; Horvath 2010). Topic morphemes provide insight into this optionality. In situ phrases are not generally marked by topic morphemes in languages like Gungbe. Under a feature-based approach, this implies that if a topic-feature is merged in a derivation (via an overt

---

[8]In situ *wh*-phrases with Q-particles are interpreted as *wh*-indefinites (Cable 2010: 93).

or covert morpheme), then topicalization occurs. If the feature is not merged, then no topicalization occurs.

Eastern Cham does not have an overt morpheme that marks DC, but it will be argued that there is a covert DC-marking morpheme, the DC-particle. Much like the Gungbe topic marker, the Eastern Cham DC-particle not only marks a phrase as DC, but also has systematic syntactic effects, namely movement or base generation of the phrase in the CP or DP periphery. Chapters 2–3 find that there is a consistent relation between DC-marking and DC-pragmatics. Furthermore, DC-marked phrases have different syntactic properties than do phrases left in situ.

1.2.2 Morphological effects of topic

A second argument that topic should be a marked by a syntactic feature comes from morphological effects, namely anti-agreement. Anti-agreement refers to the absence or syncretism of morphological agreement forms in certain circumstances. According to Baier (2018), these circumstances all involve Ā-features. To illustrate, in the northern Italian dialect Fiorentino, *wh*-movement triggers anti-agreement (Baier 2018: 82). In general, verbs agree in person, number, and gender with their subjects. When a subject like *quante ragazze* 'how many girls' is *wh*-moved, however, the verb displays default third person singular masculine agreement (1.6a). The expected agreement morphology is ungrammatical (a′). Topicalization also triggers anti-agreement, as shown with clitic left-dislocation of *la Maria* in (1.6b). Again, the expected third person singular feminine agreement is ungrammatical (b′).

(1.6) a. Quante   ragazze gli   ha   parlato con   te
          how.many girls   3SG.M have.3SG spoken   with you
          'How many girls have spoken to you?'   FIORENTINO

      a′. *Quante   ragazze le   hanno   parlato con   te
           how.many girls   3PL   have.3PL spoken   with you
          INTENDED: 'How many girls have spoken to you?' (Brandi & Cordin 1989: 124)

      b. La Maria, gli   è   venuto,   non la   Carla
          the Maria 3SG.M be.3SG   come.PTCP   not the Carla
          'Maria has come, not Carla'

      b′. *La Maria, l'   è   venuta,   non la   Carla
           the Maria 3SG.F be.3SG   come.PTCP.F   not the Carla
          INTENDED: 'Maria has come, not Carla' (Brandi & Cordin 1989: 139)

Baier (2018: 3) analyzes these effects via generalized Ā-features. Generalized Ā-features have been proposed in work on the hierarchical feature geometry of Ā-features.

Aravind (2017, 2018) posits that Ā-features are arranged in a hierarchical feature geometry underneath a top node, the generalized Ā-feature (Figure 1.5; cf. also Starke 2001, Rizzi 2004, Abels 2012, van Urk 2015 for similar ideas). If a syntactic probe is specified for a generalized Ā-feature, for example, then it can Agree with any phrase bearing a topic, *wh*, or other Ā-feature. Other levels are possible, such as an Operator feature, which includes focus, *wh*, and others.[9]

Figure 1.5: An Ā-feature hierarchy (simplified from Aravind 2017: (44))

[Ā]

[Op]   [Topic]   …

[*wh*]   [Focus]
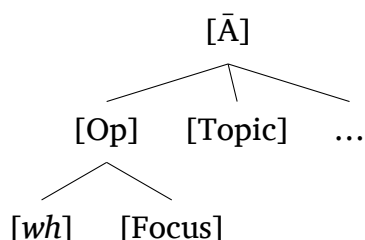
Anti-agreement effects are derived by a morphological rule that impoverishes φ-features resulting from agreement in the context of any generalized Ā-feature. This is possible in a Distributed Morphology (DM) framework where the same features that drive the syntactic derivation are then manipulated through morphological operations (Halle & Marantz 1993). This kind of morphological parallelism between topic and *wh* is expected if there is a syntactic feature marking topic. Under a non-syntactic approach, such a featural analysis of anti-agreement cannot be used, as topicalization is only derived by an EPP-feature. Instead, it would have to be argued that anti-agreement is an effect of *movement*.

However, this is not the case. Anti-agreement effects arise even in the absence of overt movement in the presence of Ā-features. Baker (2008: 172) shows that *wh*-phrases and foci trigger anti-agreement in Tundra Nenets (Uralic: Siberia), even though they remain in situ. Typically, subject and object agreement is marked on Tundra Nenets verbs. This agreement marker is lost if the subject or object is a *wh*-phrase or focus. (1.7) gives an example of anti-agreement with an object *wh*-phrase. The regular agreement pattern in (1.7b) is ungrammatical. Note that the object is in situ in this example, due to SOV ordering and a null subject.

(1.7)   a. ŋəmke-m   taxabta°?
             what-ACC   break

             'What did he break?'   TUNDRA NENETS

---

[9]Cf. Baier (2018: 194) for a discussion of Dinka anti-agreement that evokes a generalized Ā-feature or an Op-feature in embedded and matrix clauses, respectively, providing more articulated evidence that the Ā-hierarchy is needed.

    b. *ŋəmke-m   taxabtaˀ-da?
       what-ACC break-3SG>SG.OBJ
      INTENDED: 'What did he break?' (Nikolaeva 2014: 204)

Further examination shows that in situ foci and *wh*-phrases are truly in situ in Tundra Nenets. In situ *wh*-phrases and foci are grammatical in syntactic islands, such as the relative clause island in (1.8). This implies that the in situ phrases do not covertly move to Spec-CP. Instead, Baier (2018) argues that the presence of a generalized Ā-feature in any phrase is sufficient to trigger anti-agreement.

(1.8) [$_{DP}$ [$_{CP}$ xībʹa-h   xada-wiˀ   ] ti-m   ] məne-ca-nˀ?
          who-GEN kill-PERF.PART   reindeer-ACC   see-INTER-2SG
    LIT.: 'You saw the reindeer killed by whom?' (Nikolaeva 2014: 311)

Topic should be able to be marked by a syntactic feature, as topicalization gives rise to the same anti-agreement effects as moved and in situ *wh*-phrases and foci. In situ topics never appear to give rise to anti-agreement effects. Why in situ topics do not Agree in the same way as in situ *wh*-phrases and foci is an important question. However, just by virtue of being in the same class of Ā-operations, anti-agreement constitutes that certain instances of topicalization should be regarded as feature-driven. This provides further evidence for the feature-based approach to pragmatics that will be espoused in this dissertation for DC. However, there are no specific anti-agreement effects that will be reported for Eastern Cham.

1.3 Non-syntactic approaches

Finally, this section presents several arguments that have been put forth in favor of eliminating pragmatic features from syntax, i.e. the non-syntactic approach. Facts about discourse connectedness that we have already seen provide counterarguments for each and lead to a feature-based approach for the syntax of DC. First, as mentioned in Section 1.2.1, information structural movement operations like topicalization are typically optional. In other words, topics can remain in situ in a given context in a way that *wh*-phrases cannot in languages with *wh*-movement. This fact leads some to argue that topicalization cannot reflect a true Ā-movement operation, on the logic that an Agree relation should not optionally hold between a probe and goal (e.g. Erteschik-Shir 2007; Horvath 2010, using different formalisms).

    Such optionality has been argued not to exist for *wh*-phrases via Cheng's (1991, 1997) Clausal Typing Hypothesis (CTH). In a modern Agree framework, the CTH amounts to a claim that a language should only have one kind of Agree relation between C and *wh*, be it one that gives rise to *wh*-movement or *wh*-in situ (cf. Roussou & Vlachos 2011, and references therein). *Wh*-phrases should not be able to variously move or stay in situ in

the basic case in a given language. The CTH has been maintained by various authors (e.g. Potsdam 2006 on Malagasy; cf. Cable 2010 on adjunction and projection of the Q-particle) and will be maintained for Eastern Cham in Chapter 3 (cf. also Baclawski Jr forthcoming(a), Baclawski Jr forthcoming(b)). Nevertheless, others have argued that optional *wh*-movement not only empirically exists cross-linguistically, but is theoretically implementable in Minimalist syntax (Denham 1998, 2000). One mechanism argued to derive optionality is that the entire probe itself (i.e. the relevant CP) may or may not be merged. Therefore, optionality may even be present for *wh*.

There are several means by which the optionality of DC-movement can be derived, explored in more detail in Chapter 2, Sections 2.1.4 and 2.4. The wholesale optionality of the DC-particle and probe, as proposed for optional *wh*-movement by Denham (2000), is perhaps tenable. However, the lack of the presupposition introduced by the DC-particle would presumably violate Maximize Presupposition, a principle which holds that among alternative possible derivations, the one with the strongest presuppositions should be chosen (e.g. Schlenker 2012). Alternately, the apparent lack of DC-movement could in fact reflect covert DC-movement (i.e. movement of the phrase at PF). For example, the movement of *ʔiŋ ʔɔŋ năn* 'that frog' represented in Figure 1.6a could be covert. Or, the same surface order could be derived by the merging of a DC-particle with a phonologically null index in Spec-CP, as shown in Figure 1.6b. Here, the index is only accidentally coreferential with *ʔiŋ ʔɔŋ năn* 'that frog', and it is the index that is DC-marked. Chapter 2, Section 2.1.3 argues that referential indices can be DC-marked within DPs. These derivations uphold Maximize Presupposition, as there is still a DC-particle merging with a phrase and checking it for the DC conditions. This dissertation will not find evidence to distinguish between the analyses of optionality mentioned above, but it is important to note that the apparent optionality of DC-movement can be maintained in a DC-feature analysis.

Figure 1.6: DC-movement

(a) DC-movement to Spec-CP

```
              CP
          /        \
       DCP          \
      /    \       /    \
    DC      DP    C      ...
   [DC]    /  \  [uDC]   /  \
         ʔiŋ ʔɔŋ năn [EPP] ... DCP
         'that frog'
```

(b) In situ phrase

```
              CP
          /        \
       DCP          \
      /    \       /    \
    DC      DP    C      ...
   [DC]    /\    [uDC]   /  \
          /__\         ...   DP
           ∅                /  \
                      ʔiŋ ʔɔŋ năn
                      'that frog'
```
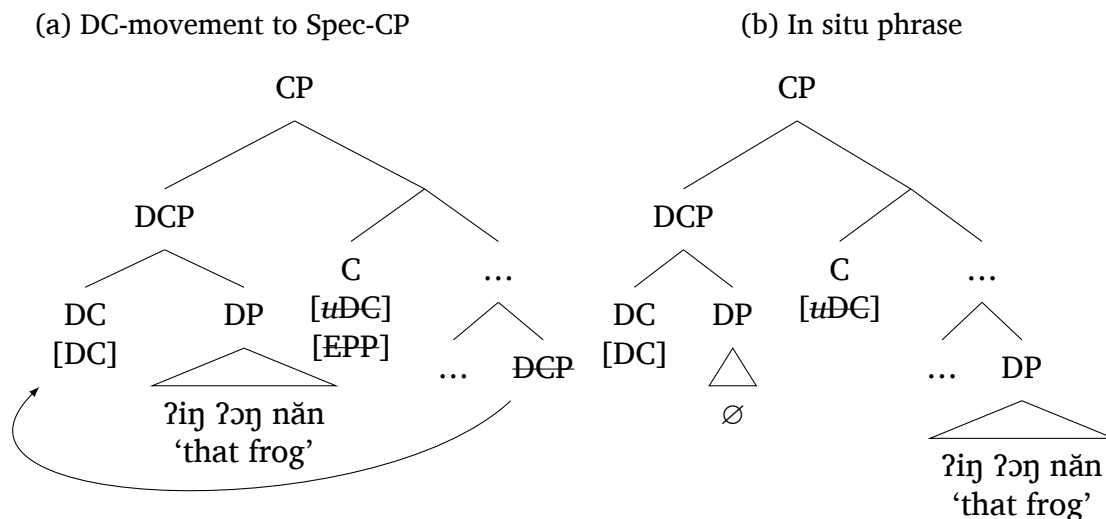
Second, some phenomena have been given non-featural analyses, as their pragmatic interpretations are not specific. As described in the introduction to Section 1.2, Chomsky (2001, 2008) argues that the EPP-feature triggers movement of the most local phrase in the typical case. However, an EPP-feature can trigger movement of a non-local phrase, provided there is some resulting 'Discourse-effect'. A prediction of this account is that the Discourse-effect must be non-specific; a range of possible interpretations such as topic and focus could be expected. Some phenomena such as scrambling appear to have this kind of effect (e.g. Yang 2009). It is worth noting that the generalized Ā-feature (cf. Section 1.2.2) could provide a feature-based analysis of these effects as well.

DC-movement will be found not to correspond with a nonspecific discourse effect; DC-movement in Eastern Cham is only associated with the pragmatics of DC. As shown in Chapters 2 and 3, the main requirement of DC-movement is that a phrase satisfies the DC conditions. Topicality is shown not to be sufficient in accounting for the data. Contrastive topic is marked by a different construction in the language. Focus and *wh* appear to be orthogonal to DC, as foci and *wh*-phrases may be DC-marked if they themselves satisfy the DC conditions. This is not expected under a Discourse-effect account and is more indicative of Agree-based movement.

Finally, the Inclusiveness Condition has been argued to preclude syntactic features that mark pragmatics. Chomsky (1995: 225) introduces the Inclusiveness Condition as a means of constraining what kinds of syntactic features can enter a syntactic derivation. According to this condition, no features can be added to a derivation after the numeration, or the prior stage at which lexical items are called from the lexicon. This would preclude

features like topic on the assumption that topic is only calculated at the level of the sentence. This assumption follows from various theories of information structure that posit topic and focus as categories that split sentences into topic–comment and focus–presupposition units, respectively. More recent formulations of minimalist syntax have done away with numerations and lexical arrays that require Inclusiveness (broadened to the No Tampering Condition in Chomsky 2008: 138, a.o.). Nevertheless, the Inclusiveness Condition continues to be a prominent argument that influences the syntax-information structure literature (López 2009: 100; Horvath 2010; a.o.; cf. also counterarguments from Aboh 2010, a.o.).

Discourse connectedness at first appears to involve an inherent violation of Inclusiveness, as it pertains to a relation between two sentences in a discourse. However, DC is syntactically introduced by a lexical item, the DC-particle; DC is not a feature inherent on the DP itself. Horvath (2010) argues for an analysis of Hungarian contrastive focus-marking, in which contrastiveness is marked by a phonologically null exhaustive identification operator. In both Hungarian and Eastern Cham, an Inclusiveness violation does not occur, as no feature needs to be added over the course of the derivation. The relevant lexical item itself is either present or absent.

This section has presented some potential arguments in favor of a non-syntactic approach to syntactic phenomena that mark pragmatics or information structure. In all cases, the arguments do not necessarily pertain to DC-movement. Hence, a feature-based approach can be upheld. The rest of this dissertation expands upon a range of evidence that DC must be an Ā-feature in Eastern Cham alongside *wh*.

1.4 Language background & methodology

This dissertation is focused on Eastern Cham, an Austronesian language in the Malayo-Polynesian branch spoken in south-central Vietnam. 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 2018c: 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 2018c).

Data for this dissertation were collected by the author in Vietnam from 2014–2019 with six major consultants, among whom each data point has been checked by at least two. Of the six consultants, one is an older male Cham scholar, while the other five are young adults who attended university. All are 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, all consultants 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 dissertation, 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 cafés, 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 discourses, then assessed for their cultural acceptability and naturalness before being assessed for pragmatic felicity. The data for this research are archived through the California Language Archive at the University of California, Berkeley.

Eastern Cham has 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 Eastern Cham communities, there is a quasi-diglossia situation between colloquial, everyday speech and formal speech, which is closely connected with Cham script (Brunelle 2005, 2009; Baclawski Jr 2018c). Much of the existing linguistic research on Eastern Cham focuses on formal speech (e.g. Bùi 1996; Thurgood 2005; Moussay 2006). Brunelle and Phú (2018) introduce the study of colloquial Eastern Cham (cf. Doris Blood 1961; David Blood 1967; Alieva 1991, 1994 for earlier work on more colloquial speech in Eastern Cham).

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). In what follows, some relevant notes on Eastern Cham phonology, morphology, syntax, and sociolinguistic variation are given, which make reference to the example sentence (1.9) below. In the interlinearized examples throughout this dissertation, IPA transcription follows the Chamic linguistic tradition (e.g. Moussay 2006; Brunelle and Phú 2018).[10] There is a vowel length distinction in both Eastern Cham and Vietnamese indicated by a short vowel mark on short vowels. Open circles underneath consonants reflect a falling, breathy register on the following vowel, reflecting a historical devoicing sound change that led to tonogenesis/registerogenesis. Vowel-initial words such as *ʔaj* 'older sibling' are transcribed with initial glottal stops. The glottal stops are not intended to carry theoretical significance. Instead, they provide a way to mark falling, breathy register on initial vowels. For instance, the vowel in *ʔḁ*

---

[10]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, including tone numbers (cf. Hoàng 1989), indicated by the corresponding transcription.

'invite' has a falling, breathy register, which originates from a devoicing of the initial consonant in *daʔa*, resulting in *t̥aʔa*, then deletion of the initial syllable.

(1.9) ʔaj   mlɛ̆ʔ c̥ŭh pɔh   tamkaj   năn hu
       older.sibling drop 7   CLF.ROUND watermelon that ROOT
       'You[older sibling] can drop those 7 watermelons.'

A large number of roots in Eastern Cham underwent monosyllabization, wherein the first syllable in a disyllabic root is reduced to a consonant or completely deleted (David Blood 1967; Brunelle 2009; Baclawski Jr 2018b). Monosyllabization also resulted in the reduction or deletion of grammatical markers. In (1.9), the historical causative *pa-* is crystallized as a short *m-* in *mlɛ̆ʔ* 'drop' (cf. *lɛ̆ʔ* 'to fall'), though most causative roots do not appear to retain any reflex of *pa-*. A small number of historically trisyllabic roots have been reduced to disyllabic roots such as *tamkaj* 'watermelon', which derives from *tamɨkaj*.

Both Eastern Cham and Vietnamese exhibit two-word collocations that are not decomposable into two separate morphemes. In Eastern Cham, these include basic forms such as *ʔiŋ ʔɔŋ* 'frog' and elaborate expressions such as *puj pɛ* 'happy', which involve semi-regular reduplication and complex meaning differences from the one-word correlate (e.g. Peterson 2010). In Vietnamese, two-word collocations are indicative of Sino-Vietnamese roots, such as *tak³⁵ ja²¹⁴* 'author'. These forms are glossed as one morpheme throughout this dissertation.

As for syntax, Eastern Cham exhibits right-branching predicate and noun phrase ordering effects widely seen in other Mainland Southeast Asian languages. The Eastern Cham noun phrase is largely right-branching: numerals precede classifiers, which precede nouns. The one exception is that demonstratives are DP-final, whereas right-branching would predict them to precede numerals. Chapter 4 gives a more detailed background and argues that the Eastern Cham DP is underlyingly right-branching. Movement of a numeral phrase constituent strands demonstratives at the right edge of the DP, as in the DP highlighted in (1.10a). Simpson (2005) argues that movement of the numeral phrase also accounts for noun phrase ordering effects in a range of other Mainland Southeast Asian languages.

(1.10)   a. [c̥ŭh pɔh   tamkaj   năn ~~c̥ŭh pɔh tamkaj~~]
                 7   CLF.ROUND watermelon that
             'those 7 watermelons'

         b. [mlɛ̆ʔ c̥ŭh pɔh   tamkaj   năn hu
                drop 7   CLF.ROUND watermelon that ROOT
             ~~mlɛ̆ʔ c̥ŭh pɔh tamkaj năn~~]

             'can drop those 7 watermelons'

Similarly, Eastern Cham predicates are largely right-branching, with many auxiliaries preceding verbs, then objects. Exceptions include some aspect markers and the modal *hu*, which appear at the right edge of the predicate, in contrast with the majority of aspect markers that precede the verb. Baclawski Jr (2017) argues that Eastern Cham predicates are underlyingly right-branching. Predicate-final aspect markers and *hu* appear at the right edge due to movement of a constituent underneath them to a higher position, as in (1.10b). Simpson (2001) argues for a similar predicate raising operation in other Mainland Southeast Asian languages.

Like Vietnamese and other Mainland Southeast Asian languages, kinship terms may function as pronominals (here, *ʔaj* 'older sibling' functioning as an addressee). Throughout this dissertation, these uses are indicated by brackets in sentence glosses, such as 'you[older sibling]'. Kinship terms coexist with null pronouns, bare classifiers such as *tʰa jaŋ* 'one person [LIT.: one CLF.PERSON]', and true pronouns such as *kăw* '1SG', *hḻ̆aʔ* '1SG.POL', *hɨ* '2SG', and *ɲu* '3.ANIM'. As for the latter, there is only one third person pronoun, and it only refers to animate referents.

Eastern Cham does not have a large set of inherited modals. In the author's fieldwork, the only two inherited modals found are *pʰaj* 'must' (which appears to be a nativized borrowing of Vietnamese *faj²¹⁴* 'must') and *hu* 'can'. The first modal, *pʰaj* is a necessity root modal, which encompasses circumstantial and deontic modalities (cf. Kratzer 1991 on modal space). The second, *hu*, is a possibility root modal. The latter, *hu*, appears often in this dissertation and is trascribed as ROOT. Other modals, especially epistemics, are adverbials borrowed from Vietnamese.

Serial verb constructions are prevalent in Eastern Cham, though they will not be mentioned further in this dissertation. Serial verb constructions are sequences of verbs within one predicate that are distinct from control and raising in that they do not select each other (Cleary-Kemp 2015). In Eastern Cham, serial verb constructions include allatives, directionals, resultatives, and purposives in Cleary-Kemp's (2015) typology. One example of an allative serial verb construction is given in (1.11), which is comparable to English *go eat*.

(1.11) hɨ   naw ɓăŋ ḵeʔ
      2SG go   eat what
   'What did you go eat?'

Eastern Cham also exhibits *wh*-in situ. *Wh*-phrases such as *ḵeʔ* 'what' are pronounced in their base position, as with many languages in East and Southeast Asia. Chapter 3 gives more detailed background on *wh*-phrases. Seemingly optional *wh*-movement presents a core case of DC-marking in Eastern Cham.

As for sociolinguistic variation, Eastern Cham IPA throughout this dissertation will be represented conservatively, with a broad transcription. Sociolinguistic variation will be abstracted when possible to the most common forms in colloquial speech as per Baclawski Jr's (2018c) results. Many disyllabic roots alternate with monosyllabic roots in formal

speech (e.g. *lĭmin ∼min* 'elephant'). However, the monosyllabic forms are predominant in colloquial speech and will be used here. The same goes with an alternation between *s ∼tʰ*, the former of which is vanishingly rare in colloquial speech. Similarly, regarding *l/r ∼n* in coda position, the former are almost never found. In these cases, the attested colloquial speech forms will be used in cited examples, which in nearly all cases correspond with the elicited content.

More complicated is the variation involving onset *r/j* and coda *ŋ*. Onset *r* is pronounced as [r] in about 40% of the tokens reported in Baclawski Jr (2018c), though there is evidence that it indexes formal speech, as it is used in hypercorrect forms. More common are pronunciations that range from [j] to [ʐ] and [z]. While each of these variants have been attested for nearly all words elicited in this disseration, they are not equally common. The word *raŋ* 'CLF.PERSON' is most often attested with [j], while *ɽaʔ* 'market' with [ʐ], and *jŭt* 'friend' with [z]. More research is needed to understand the sociolinguistic variation with these forms. For the time being, the most common form will be used for each respective word in cited examples, such as *jaŋ, ʐaʔ*, and *zŭt*.

Coda *ŋ* is coarticulated with a labial nasal after round vowels (i.e. [ŋ͡m]), closely parallel with allophonic variation in Vietnamese. Baclawski Jr (2018c) reports that the coarticulated nasal occurs somewhat more frequently than the non-coarticulated form at a rate of about 60% to 40% in the sample. For example, both *hɔ̆ŋ* and *hɔ̆ŋ͡m* 'with' are attested in the sample (which also exhibits the *s ∼tʰ* variation, which in this case is reduced to *h*). However, again there is lexical variation. For example, *hɔ̆ŋ͡m* 'with' is predominant, while *ʔiŋ ʔɔŋ* 'frog' is more common in the sample without coarticulation. In the cited examples, the most common form will be used (i.e. *hɔ̆ŋ͡m* and *ʔiŋ ʔɔŋ*), though it should be noted that all variants are in principle attested.

Presyllable reduction is also subject to variation, as reported in Baclawski Jr (2018b). For example, the same presyllable can often be reduced to *m-, n-,* or *p-* when the following consonant is coronal. For example, *mtʰã, ntʰã,* and *ptʰã* 'Sambhur deer' are all attested. When applicable, the *m-* form is used in cited examples here, as it appears to be generally the most common.

The broad phonological transcription used here also abstracts over additional phonetic characteristics that have yet to receive full phonetic study. Baclawski Jr (2018b,c) provides some evidence that initial sonorants are geminated after certain presyllables are deleted. For example, *mlɛ̆ʔ* in (1.9) above varies with *l:ɛ̆ʔ* 'drop', with a geminate about twice the length of a normal *l* (compare *lɛ̆ʔ* 'fall'). Additionally, there appear to be prosodic contours that may or may not interact with grammatical tone. For example, negative predicates have a falling prosodic contour on the structurally highest verbal element, which could be the main verb (1.12a), the existential marker *hu* (1.12b), or an auxiliary verb at the left or right edge of the predicate (1.12c). Note that the root modal *hu* is the highest verbal element in the predicate even though it surfaces on the right. Baclawski Jr (2017) argues this is due to predicate raising.

(1.12) a. kăw **tʰằw** ʔo
           1SG know NEG
         'I don't know.'

   b. mĕj   năn **hù** puj pɛ   ʔo
      female that ∃ ɓe.happy NEG
      'That woman is not happy.'

   c. kăw ɗiʔ mt̪ɛh   **hù**   ʔo
      1SG ride motorbike ROOT NEG
      'I cannot ride a motorbike.'

In these examples, a falling tone mark is used to indicate the negation contour. In the examples throughout this dissertation, though, that contour will not be marked, as its status remains unclear. More phonetic and prosodic research is required to ascertain the status of the contour.

There are two differences between the variety of Eastern Cham presented in this dissertation and that reported in prior literature on the language. First, Chapter 2, Section 2.2.2 reports on the existence of a complementizer *po*, as in the relative clause in (1.13). Prior literature claims that there is a topicalizer *năn* or *ăn*, derived from the distal demonstrative, but explicitly argues against the existence of a general complementizer. In the author's fieldwork, the topicalizer *năn/ăn* is not found, while the complementizer *po* is observed and almost always accepted in the complementizer position by consultants.

(1.13) kăw mɔŋ ɓ̥ăŋ po   hɨ cĭh
       1SG see door COMP 2SG paint
     'I saw the door that you painted.'

A second difference involves the falling, breathy register on vowels. As mentioned above, the falling, breathy register arises from a devoicing process, followed by rightward spreading onto subsequent vowels. For example, the voiced stop *d* devoiced to *t* in **daʔa* 'invite'. The falling, breathy register spread rightward to both vowels. According to Thurgood (1996: 19), the falling, breathy register spread rightward across sonorants and fricatives, but stops blocked the register from spreading further. In the author's fieldwork, however, the voiced bilabial implosive *ɓ* does not block register spreading in the word *ɓ̥ăŋ* 'door', where the register originates from the initial consonant in **biɓăŋ*. Among the consultants tested, *ɓ̥ăŋ* 'door' forms a register minimal pair with *ɓăŋ* 'eat'. More research is needed to understand this apparent variation.

As a final note on glossed examples, the following formatting is used throughout this dissertation for emphasis and clarity. Certain words and phrases are bolded in the Eastern Cham examples for emphasis, generally DC-marked phrases. Language names are right-justified and in small capital letters. Unless otherwise noted, interlinearized examples are from Eastern Cham. Constructions of note are right-justified and noted by parentheses. Discourse relations are right-justified and marked with italics.

1.5 Outline

This introduction has sketched the syntactic, semantic, and pragmatic analysis of discourse connectedness pursued in this dissertation. The remainder of this dissertation argues in more detail that DC is an Ā-feature that marks a hierarchical discourse constraint.

Chapter 2 examines the basic case of DC-marking in Eastern Cham, movement or base generation of a phrase in Spec-CP (on the surface, topicalization). A variety of Ā-movement characteristics are found for DC-movement: sensitivity to syntactic islands, reconstruction, weak crossover, and locality effects. Locality effects in particular support the idea that DC-moved phrases are syntactically distinct from in situ phrases in the same context. Additionally, two distinct DC-probes are found in Eastern Cham: one with an EPP-feature that triggers DC-movement, and one with an Op-feature that binds a phrase in Spec-CP to a lower pronoun. The latter results in base generation of a DC-marked phrase in Spec-CP.

Chapter 3 turns to DC-marking of *wh*-phrases (on the surface, optional *wh*-movement). DC is found to be independent from *wh*, as predicted from the former's status as a hierarchical discourse constraint. Otherwise, the same syntactic probes and Ā-movement characteristics are found. A semantic account is presented in which the DC-particle and Q-particle, which marks *wh*-phrases (Cable 2010), can mark the same phrase.

Chapter 4 investigates DC-marking inside a DP (on the surface, inventory forms and partitives). A parallelism between CP and DP is found, as has been argued for other Ā-features such as *wh*. Phrases can be DC-marked at the left edge of DP just as well as CP. Both syntactic probes are found to exist on D as well: one that triggers DC-movement of a phrase to Spec-DP and one that binds a DC-phrase base generated in Spec-DP to a lower pronoun.

Chapter 5 concludes. A brief comparison to Catalan, which also exhibits DC-movement, as first argued by López (2009), is presented. Minor differences between Catalan and Eastern Cham lead to a small typology of DC-marking. Additionally, contrastive topic, which would be predicted to involve a kind of DC-marking, but is blocked in Eastern Cham by a separate morpheme, is discussed.

Chapter 2

Topicalization and DC-movement

This chapter argues that DC-movement is an instance of Ā-movement, driven by a feature that indexes a relation between the constituent it attaches to and the discourse relations between the sentences in which that constituent is mentioned. This chapter focuses on the basic case of DC-movement, which appears on the surface to be a kind of topicalization (e.g. 2.1).

(2.1)   a.   **ʔɔ́ʔ**   **ni**   kăw   t̪ɔʔ   ɓăŋ   ~~ʔɔ́ʔ ni~~
           mango   this   1SG   PROG   eat
           'This mango, I am eating.'

       b.   [$_{DCP}$ DC-particle [$_{DP}$ ʔɔ́ʔ ni]] C$_{uDC}$ kăw t̪ɔʔ ɓăŋ ~~DCP~~   (DC-movement)

DC-movement is argued to be an instance of Ā-movement driven by an Agree relation between a DC-probe on C and a DC-particle (cf. Chomsky 2000; Miyagawa 2009 on Agree). This analysis is schematized in (2.1b). A DC-particle merges with the DP *ʔɔ́ʔ ni* 'this mango', creating a DC-phrase, much like focus particles and focus phrases, Q-particles and *wh*-phrases (cf. Cable 2010). This phrase enters an Agree relation with a probe on C. Finally, as a result of an EPP-feature, the DC-phrase is moved to Spec-CP, the surface left edge of the clause, as indicated by the crossed-out ~~DCP~~.

This movement operation shares a number of characteristics with other instances of Ā-movement, such as *wh*-movement cross-linguistically. It is unbounded, it is sensitive to syntactic islands, and it exhibits weak crossover and locality effects. Each of these characteristics is expanded upon in this chapter, with a particular emphasis on locality effects via a Path Containment effect seen when multiple phrases are DC-moved to the same clause edge.

The syntactic properties of DC-movement contrast with a similar construction involving a DCP in Spec-CP and a resumptive pronoun in the base position (2.2a). It is argued that the DCP is base generated in Spec-CP and binds a resumptive pronoun in the base position (2.2b). An operator probe (Op) enters into an Agree relation with the pronoun

ɲu, ensuring it is coindexed with the DP in Spec-CP (Section 2.3.2). Again, a DC-particle merges with the phrase in Spec-CP.

(2.2)   a. **nĭʔ sɛh nǎn** kǎw ʔ̥a   ɲu
           student that 1SG invite 3.ANIM
           'That student, I invited him/her.'

       b. [$_{DCP}$ DC [$_{DP}$ nĭʔ sɛh nǎn]]$_i$ C$_{uOp}$ kǎw ʔ̥a [ɲu$_{Op}$]$_i$ (Base generation + Agree)

Throughout this chapter, Agree-driven syntactic movement will be compared both to base generation and EPP-driven movement, which lacks Agree. As laid out in Chapter 1, Section 1.2, in the absence of an Agree relation, an EPP-feature will attract the most local phrases (DCP$_1$ in (2.3a); cf. Chomsky 2008). While DC-movement requires Agree, long distance DC-movement is found to involve intermediate steps of EPP-driven movement (2.3b; Section 2.3.3).

(2.3)   a. EPP-driven movement: [$_{DCP1}$ DC] C$_{EPP}$ …D̶C̶P̶$_T$ …DCP$_2$

       b. Long distance DC-movement: [$_{DCP1}$ DC] C$_{DC}$ …C$_{EPP}$ …C$_{EPP}$ …D̶C̶P̶$_T$ …DCP$_2$

From this analysis, DC-marking can occur in two ways. A DC-feature is merged with a DP in its base position, and the resulting DCP must move to Spec-CP (DC-movement). Or, a DC-feature is merged with a DP, and that DCP is base generated in Spec-CP. That DP binds a resumptive pronoun lower in the derivation (base generation + Agree).

Before we proceed, the optionality of DC-marking should be mentioned. In all of the examples reported here, if DC-marking is felicitous, the corresponding sentence without DC-marking (i.e. an in situ phrase) is also felicitous. This kind of optionality has been cited as evidence against feature-based approaches to syntactic operations that mark pragmatics in general (cf. Chapter 1, Section 1.3). There are several possible analyses for the in situ cases in Eastern Cham, however, that accord with the syntax of DC proposed so far. First, perhaps a DC-particle is not merged at all, and as a result, neither is the DC-probe on C (2.4b). In this case, a phrase like *nĭʔ sɛh nǎn* 'that student' may satisfy the DC conditions, but it is not marked so in the syntax. Section 2.1.2 provides evidence against an analysis where the DC-particle is missing entirely, based on Maximize Presuppositon.

Second, perhaps DC-marking always occurs when possible (i.e. when a phrase satisfies the DC conditions), but sometimes it involves covert movement, leaving the phrase in situ on the surface. In (2.4a), *nĭʔ sɛh nǎn* 'that student' being interpreted in Spec-CP, but pronounced in its base position, as indicated by the strikethrough below.

(2.4)   a. n̶ĭ̶ʔ̶ ̶s̶ɛ̶h̶ ̶n̶ǎ̶n̶$_{DC}$ kǎw ʔ̥a   **nĭʔ sɛh nǎn**
                            1SG invite student that
           'I invited that student.'

       b. [$_{DCP}$ DC [$_{DP}$ ∅]] C$_{uDC}$ kǎw ʔ̥a nĭʔ sɛh nǎn

Third, perhaps a null index is merged with a DC-particle, and the resulting DC-phrase is merged in Spec-CP. Section 2.1.3 argues that null referential indices can be DC-marked within a DP. In this case, there exists a phrase that is DC-marked, but it is phonologically null. Additionally, the phrase *nǐʔ sɛh năn* 'that student' only accidentally corefers with the DC-marked index. The data presented in this dissertation will not distinguish between the second and third analyses. If a language is found with an overt DC-particle, such evidence may be found, as the second analysis would predict the DC-particle to be pronounced in the base object position in (2.4a), but in Spec-CP in (2.4b).

This chapter proceeds as follows. First, Section 2.1 demonstrates that the pragmatics of DC-marking is best understood in terms of previous mention and subordinating discourse relations. Evidence is presented that the semantics of the DC-particle must make reference to two events in a discourse that stand in a subordinating discourse relation, which raises the need for a dynamic event semantics. Section 2.2 provides some brief background on the syntactic distribution of DC-movement. Then, Section 2.3 compares the syntactic properties of DC-movement (Section 2.3.1) to base generation + Agree (Section 2.3.2). Long distance DC-movement is also examined and found to involve intermediate steps of EPP-driven movement (Section 2.3.3). This analysis leads to the account of the apparent optionality of DC-movement as sketched above (Section 2.4).

2.1 Pragmatics of DC-movement

This section introduces the pragmatics of DC-marking, *discourse connectedness*, which the remainder of the chapter will argue that the syntax makes reference to. Discourse connectedness is found to be a relation between a constituent and the discourse relations between the sentences that mention the individual that constituent refers to.

Discourse connectedness has three requisite components: a subordinating discourse relation between two sentences in a discourse, a relation between two events introduced in those two sentences, and previous mention of the phrase being DC-marked. Section 2.1.1 describes the subordinating discourse relation component. Section 2.1.2 shows that DC-marking further requires a relation between two events introduced in the two sentences in that subordinating discourse relation.[1] Then, Section 2.1.3 expands upon the previous mention component, refining the notion of previous mention as inclusion of a phrase in the sets of participants in the two relevant semantic events. Section 2.1.4 puts these components together and argues they are enforced in the semantics by a presupposition introduced by the DC-particle.

To illustrate the three components of DC-marking, consider (2.5). Sentence (2.5b) is interpreted as elaborating upon (2.5a), providing a more detailed description of Thuận's cooking of that frog. Elaboration is a subordinating discourse relation, according to theo-

---

[1]In the literature on discourse or rhetorical relations, a more appropriate descriptor of discourse unit may be 'logical form' (e.g. Asher & Lascarides 2003), though 'sentence' is often used as a shorthand (cf. Webber 1988).

ries of discourse structure (e.g. Asher & Lascarides 2003), satisfying the first component of DC. Second, elaboration infers a relation between two events, cooking that frog and cooking that frog well (a 'subtype' relation, according to Asher & Lascarides 2003). Finally, the DC-marked phrase *ʔiŋ ʔɔŋ năn* 'that frog' is a participant in both of the aforementioned events. For the remainder of this dissertation, these three components of DC-marking will be referred to as the DC conditions.

(2.5) a. tʰuːŋ͡m³¹² ʈɔʔ ŋăʔ **ʔiŋ ʔɔŋ năn**
          Thuận   PROG make frog   that
          'Thuận is cooking that frog.'

  b. **ʔiŋ ʔɔŋ năn**$_{DC}$ ɲu ʈɔʔ ŋăʔ ~~ʔiŋ ʔɔŋ năn~~ ŋi lo
      frog   that   3.ANIM PROG make   be.delicious very
      'That frog, he is cooking very well [Lit: deliciously].'

Based on the three DC conditions above, the pronoun *ɲu* is also anaphoric, in addition to the event of cooking, though Eastern Cham has no means of DC-marking matrix subjects, pronouns, or predicates because of syntactic and semantic restrictions (cf. Section 2.2 and Section 2.1.3 on pronouns).

Section 2.1.1 expands upon the subordinating discourse relation condition. Section 2.1.2 provides evidence in favor of framing the DC conditions in terms of events, not sentences. Finally, Section 2.1.3 gives more detail on the previous mention condition.

2.1.1 Subordinating discourse relations

DC-marking requires a subordinating discourse relation between two events in a discourse. According to theories of the structure of discourse, 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).[2,3] In a coordinating discourse relation, a sentence is added to the discourse and supplants the prior as the focus of attention (Figure 2.1a). By contrast, in a subordinating discourse relation, a sentence is interpreted as contributing to a prior sentence, such that both remain active points of attention. Subordinating discourse relations result in hierarchical relations, reflected in Figure 2.1b, such that the prior sentence is superordinate and the current sentence subordinate.

---

[2]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).

[3]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.

Figure 2.1: Models of sentence relations in a discourse

(a) Coordinating discourse relation

$$\boxed{S_1} \quad + \quad \boxed{S_2} \quad = \quad \boxed{S_1, S_2}$$
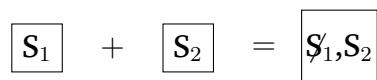
(b) Subordinating discourse relation

$$\boxed{S_1} \quad + \quad \boxed{S_2} \quad = \quad \boxed{\boxed{S_1} \\ \boxed{S_2}}$$
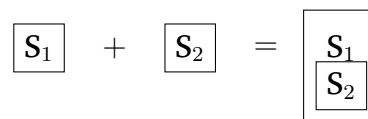
A full background on discourse relations and theories of discourse structure is beyond the scope of this chapter. Of relevance for this section, these theories have identified a set of subordinating discourse relations that provide diagnostics for discourse connectedness (cf. Asher & Lascarides 2003 and Asher & Vieu 2005 on discourse subordination; Mann & Thompson 1988 on nucleus-satellite relations). Prototypical subtypes of discourse subordination include any kind of elaboration or explanation. In other words, if a sentence is interpreted as an elaboration or explanation of another, those two sentences are in a subordinating discourse relation (2.6). In the remainder of this chapter, a down arrow ⇓ will be used to indicate a subordinating discourse relation between a superordinate and subordinate sentence (and ⇓̸ to the absence of discourse subordination). It will also be said equivalently that a superordinate sentence *discourse subordinates* the subordinate.

(2.6) DISCOURSE SUBORDINATION (⇓): Sentence ϕ ⇓ sentence ψ if ψ is interpreted as an elaboration or explanation of ϕ and ϕ remains open for further discussion after ψ

Another way to conceptualize discourse subordination is by means of what implicit questions under discussion are being answered. Subordinating discourse relations answer questions that expand upon prior sentences, such as *Why exactly did that happen* and *What exactly happened*. Coordinating discourse relations answer questions such as *What happened next*. Riester, Brunetti & Kuthy (2018) propose an implicit question under discussion test for diagnosing the discourse subordination/coordination split for the purposes of corpus annotations.

Returning to the Eastern Cham example repeated below, (2.7b) is interpreted as an elaboration on (a). 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 ⇓ b).

(2.7) a. tʰuːŋ͡m³¹² tɔʔ ŋăʔ **ʔiŋ ʔɔŋ năn**
Thuận PROG make frog that
'Thuận is cooking that frog.'

b. **ʔiŋ ʔɔŋ năn**_{DC} ɲu   ʈɔʔ   ŋăʔ   ~~ʔiŋ ʔɔŋ năn~~ ɲi   lo
   frog   that   3.ANIM PROG make   be.delicious very

   'That frog, he is cooking very well [Lit: deliciously].'   (a ⇓ b)

This pragmatic account of DC is inspired by López's (2009) description of clitic dislocation in Catalan. Clitic left- and right-dislocation also require previous mention and a subordinating discourse relation between the two sentences, formalized by López (2009) as +Anaphor, or strong anaphora. Chapter 5, Section 5.1 goes into more detail on how DC-marking in Eastern Cham compares with DC-marking in Catalan as described by López (2009), resulting in a small typology of DC.

A negative example is given in (2.8). Here, DC-marking of *ʔiŋ ʔɔŋ năn* 'that frog' is infelicitous, even though it is previously mentioned in (2.8a). What makes this example infelicitous is that the event of eating the frog is not interpreted as an explanation or elaboration of the prior sentence. Here and throughout, ⇓ refers to the absence of discourse subordination; the precise relation is left unspecified. Instead, *ʔiŋ ʔɔŋ năn* 'that frog' must be pronounced in situ.

(2.8) a. tʰuːŋ͡m³¹²   ʈɔʔ   ŋăʔ   **ʔiŋ ʔɔŋ năn**
         Thuận   PROG make frog   that

         'Thuận is cooking that frog.'

    b. #jăʔ ni **ʔiŋ ʔɔŋ năn** ɲu   ʈɔʔ   ɓăŋ ~~ʔiŋ ʔɔŋ năn~~
        now   frog   that 3.ANIM PROG eat

        INTENDED: 'Now, that frog, he is eating.'   (a ⇓ b)

When presented with this context, consultants accept the sequence of (2.8a–b) only if there is additional prior discourse. For example, prior discourse could mention a broader event of what Thuận did this evening involving that frog. In this case, superordinate material is inserted into the discourse that (2.8b) could be interpreted as elaborating upon. I take this as instructive that discourse subordination is a necessary component of DC-marking.

Based just on the examples above, it could be hypothesized that entailment is needed for DC-marking, as the event of cooking that frog well (2.8b) entails the event of cooking that frog (2.8a), but the same cannot be said of the event of eating that frog and cooking that frog (2.8a–b). However, entailment is not necessary. The two sentences involved in DC-marking can be loosely related, as long as one can be interpreted as an explanation or elaboration on the other, as in (2.9). In this context, consultants accepted (2.9b) either if it elaborates on an aspect of that door, or if it is construed as an explanation (i.e. my painting it made the door beautiful).[4]

---

[4]With regard to Section 2.1.2, it should be noted that the predicate *k̥ʰəh* 'be beautiful' in (2.9a functions as a stage-level predicate in Eastern Cham. When negated, this predicate must use the existential marker *hu*, which is used as a stage-level predicate test.

(2.9) a. **ɓǎŋ nǎn** k̥ʰəh   lo
           door that be.beautiful very
           'That door is very beautiful.'

       b. **ɓǎŋ** kǎw cǐh ɓḁ̌ŋ pan çaw
           door 1SG paint   color blue
           'The door, I painted blue.'   (a ⇓ b)

For an example involving explanation, (2.10b) gives a reason for the event of cook-ing in (2.10a). This is a subordinating discourse relation that leaves both (2.10a) and (2.10b) open for discussion: subsequent discourse could further elaborate on the reason for cooking, or it could return to the event of cooking itself. Additionally, *lɔ nǔʔ* 'chicken meat' is mentioned in both the explanans and explanandum sentences. Therefore, the DC conditions are met, and *lɔ nǔʔ* can be DC-marked. Note that Section 2.1.3 finds that apparent generics like the DC-phrase in this example denote kinds with type e.

(2.10) a. ʔaj   mɛ̆j   kǎw tǔʔ **nǔʔ**
            older.sibling female 1SG cook chicken
            'My older sister cooks chicken.'

        b. jwa **lɔ   nǔʔ**$_{DC}$ ʔaj   mɛ̆j   kǎw cəh ɓǎŋ l̶ɔ̶ ̶n̶ǔ̶ʔ̶
            because meat chicken older.sibling female 1SG like eat
            'Because chicken meat, my older sister likes to eat.'   (a ⇓ b)

        b′. #plɔh nǎn **lɔ   nǔʔ**   ʔaj   mɛ̆j   kǎw ɓǎŋ l̶ɔ̶ ̶n̶ǔ̶ʔ̶
            after that meat chicken older.sibling female 1SG eat
            'After that, she [will] eat the chicken meat.'   (a ⇓̸ b)

(2.10b′) has the same previous mention status of *lɔ nǔʔ* 'chicken meat', but the sentence is interpreted as a separate event from (2.10a), an event of eating that requires the event of cooking to be completed. This is not a subordinating discourse relation, so the DC conditions are not met and DC-movement of *lɔ nǔʔ* is impossible.

Discourse subordination can also arise in question-answer pairs is given in (2.11). If an answer elaborates upon a question by giving more or less information than requested, the answer is classified as being in a subordinating discourse relation with the question (cf. Asher & Lascarides 2003 on 'q-elaboration' and 'partial answers'). Accordingly, DC-movement is felicitous in (2.11b), which elaborates on the question. Additionally, the generic *ʔɔ̆ʔ* 'mango' is previously mentioned in (2.11a). Why should elaborating answers be subordinating? For one, the discourse in (2.11a–b) can be paraphrased with an ex-plaining relation: "Do you want to eat mango? [No. Why?] I already ate mango." Thus, (2.11b) could be covertly answering a question seeking an explanation.

(2.11) a. hɨ hu ʔĭŋ   ɓǎŋ **ʔɔ̆ʔ**   lɛ̆j
            2SG ∃ want eat mango Y/N.Q
            'Do you want to eat mango?'

b. **ʔɔ̌ʔ**$_{\text{DC}}$ kăw ɓăŋ ~~ʔɔ̌ʔ~~ jɘ̥
    mango 1SG eat already

‘I already ate mango.’   (a ⇓ b)

b′. #**ʔɔ̌ʔ** kăw ʔɨ̈ŋ ɓăŋ ~~ʔɔ̌ʔ~~
     mango 1SG want eat

‘I want to eat mango.’   (a ⇓̸ b′)

By contrast, direct answers, such as (2.11b′), are not treated as being in subordinating discourse relations with their questions. Therefore, despite the previous mention between (2.11a) and (2.11b′), DC-movement of *ʔɔ̌ʔ* is infelicitous.

DC-movement is impossible in contrastive topic contexts, which involve partial answers. Contrastive topics have been analyzed as requiring multiple embedded levels of questions under discussion (Büring 2003; Constant 2014). For example, (2.12a) prompts a paired list response: for each of you, which person did you invite? In (2.12b), the speaker answers a subpart of that question: as for myself, who did I invite? The ellipsis make it clear that the speaker intends to continue by addressing other people and their invitees.

(2.12)   a. jŭt ʔa **jaŋ** **hlɛ̌j** maj pă̰ʔ ni
           friend invite which CLF.PERSON come in this

‘Which person did you[friend] invite to come here?’

[Directed at multiple people]

b. #**tʰuːŋ͡m**^{312}$_{\text{CT}}$ çɛ̌j hu ʔa ~~tʰuːŋ͡m~~^{312} maj pă̰ʔ ni…
     Thuận self ∃ invite come here

‘I[myself] invited THUẬN to come here…’   (a ⇓? b)

Contrastive topics are different from elaborating answers in that there remains an open question under discussion; contrastive topics only partially answer a question. By contrast, elaborating answers close a question under discussion and proceed to provide further information. Chapter 5, Section 5.2 argues that the openness of a question under discussion is a crucial distinction. Contrastive topic is argued to involve a different hierarchical discourse relation than DC, based on data from Eastern Cham and clitic left-dislocation in Catalan.

Finally, there is one additional context that licenses DC-movement: gestural deixis. In an out-of-the-blue context, if consultants are presented with DC-movement, they sometimes accept it if it is preceded by a marked pointing gesture to a physical entity. This is exemplified in (2.13). In elicitation contexts, these examples are quite marked; consultants typically go out of their way to detail pointing gestures and the location of the referent. However, a more systematic study of natural speech is needed to fully ascertain the role of gesture in DC-movement.

(2.13)   a. CONTEXT: Speaker pointing at mango in their hand.

   b. **ʔɔ̆ʔ**   **ni**$_{\text{DC?}}$ kăw t̥ɔʔ ɓăŋ ~~ʔɔ̆ʔ ni~~
      mango this   1SG PROG eat
      'This mango, I am eating.'   (a ⇓? b)

This instance of DC-movement follows if gestural content can participate in discourse structure, as has been argued by Lascarides & Stone (2009). The deictic gesture itself in this example could be seen as constituting a discourse move elaborated upon by (2.13b) (cf. van Kuppevelt 1995 on metalinguistic events feeding discourse). Likewise, the DC-phrase *ʔɔ̆ʔ ni* 'this mango' could be seen as anaphoric to the deictically indicated referent.

To summarize this section, DC-movement is licensed in contexts independently argued to represent subordinating discourse relations. These include sentences interpreted as elaborations or explanations. Contrastive topic presents one exception, where DC-movement is blocked. In general, discourse subordination is sufficient, but not necessary to license pronouns, anaphoric definites, and other anaphora according to accounts of discourse structure (cf. Webber 1988, Asher 1993 on the Right-Frontier Constraint). In other notions of accessibility, if prior discourse is necessary, any prior mention is sufficient regardless of discourse structure (e.g. Prince 1992 on hearer-old; Lambrecht 1994 on the accessibility scale). For example, pronominal anaphora would be licensed in English and Eastern Cham in all of the (b) and (b′) examples throughout this section, regardless of the presence of discourse subordination. Therefore, DC imposes an additional discourse structural restriction.

2.1.2   Semantic event requirement

So far, discourse connectedness has been characterized in terms of rhetorical relations between sentences. This section finds evidence that discourse connectedness must in fact be defined in terms of *events* introduced in different sentences in a discourse, based on evidence from Kratzer's (1995) account of stage- and individual-level predicates. This will lead to the event-based semantic account of DC presented in Chapter 1 and the introduction to this chapter. Conversely, the evidence in this chapter supports event semantics and Kratzer's (1995) account of stage- and individual-level predicates in general. Furthermore, it establishes the need for a dynamic event semantics.

Discourse relations are relations between sentences that involve a constellation of inferences, of which one is a relation between two events. According to Asher & Lascarides (2003: 204), part of the explanation relation is a causal relation between two events, such that one sentence provides evidence for that causal event relation.[5] To illustrate, sentence (2.14a) and (2.14b) are in an explanation relation. Part of that explanation relation is a causal relation between the pushing event and the falling event. In this case,

---

[5]According to Asher & Lascarides (2003: 206), part of the elaboration relation is a subtype relation between two events.

the respective sentences are coterminous with the events, but this need not always be the case; the evidence for the causal relation could be indirect. Note that *Explanation*(a,b) indicates that (b) is interpreted as an explanation of (a).

(2.14)   a. Max fell.

      b. John pushed him.   *Explanation*(a,b)

In theory, discourse connectedness could be a condition on sentences in a discourse or events. If DC is a condition on sentences, DC-marking would be predicted to be licit so long as there is any subordinating discourse relation. Any two sentences in a subordinating discourse relation should be able to license DC-marking. If DC is a condition on events, by contrast, there must be two semantic events in the discourse in a causal relation inferred by a subordinating discourse relation. The notion of 'event' used here is not synonymous with 'sentence'. A single sentence can refer to multiple events, and some sentences do not introduce event variables at all. An event-based account of DC-marking makes a different prediction: the DC conditions can be satisfied if there are two events such that one event is interpreted as providing evidence for (i.e. explaining) or elaborating upon the other.

Sentence- and event-based accounts of DC-marking make different predictions regarding stage- and individual-level predicates, as per Kratzer (1995) and Fernald (2000). Stage-level predicates refer to temporary events that hold for only a certain time from the perspective of the individuals involved, while individual-level predicates refer to permanent characteristics of those individuals. A wide variety of linguistic phenomena track this distinction, such as the contrast in (2.15) from Kratzer 1995: 125 (cf. Carlson (1977) for the original distinction).

(2.15)   a. There are firemen available.   (Stage-level predicate)

      b. *There are firemen altruistic.   (Individual-level predicate)

According to Kratzer (1995), all stage-level predicates introduce event variables, but individual-level predicates do not. Fernald (2000) supports this view with a variety of evidence. If DC is a condition on events, then individual-level predicates would be predicted not be sufficient to license DC-marking either as the current or prior sentence. This prediction is borne out in Eastern Cham. DC-marked phrases must be participants in events in the current and prior sentences; neither the current nor the prior sentence can exclusively contain individual-level predicates.

To test this prediction, first a reliable test for stage- and individual-level predicates must be identified in the language. Baclawski Jr (2018a) identifies one such test based on the existential marker *hu* and negation. In Eastern Cham, *hu* has a variety of uses, from a verb meaning 'have', to an existential copula, a marker of presentational clefts, and often a marker of negation (cf. Thurgood & Li 2007). Baclawski Jr (2018a) argues

that *hu* has an identical distribution to the word *adi* in Bura (Chadic: Nigeria), which Zimmermann (2007) argues generally marks existential closure.

Of relevance for this section, Zimmermann (2007) predicts that existential closure markers such as Bura *adi* and Eastern Cham *hu* are necessary to mark negation, as negation in these languages requires overt closure over events, which can only be supplied by those markers. This is borne out in Eastern Cham. Negation generally requires *hu* in addition to the clause-final negation marker *o*. In (2.16a), the verb *ʈɔʔ* introduces a stage-level predicate, as the event of living is taken to be temporary, at least in the context elicited.

(2.16)   a. kăw hu ʈɔʔ păʔ mlĕj   năn ʔo
           1SG ∃   live in   village that NEG

          'I do not live in that village [temporarily].'   (Stage-level)

       b. kăw (*hu) nĭʔ   păʔ mlĕj   năn ʔo
           1SG   ∃   be.born in   village that NEG

          'I was not born in that village.'   (Individual-level)

       c. kăw ʄăwʔ   nĭʔ   păʔ mlĕj   năn ʔo
           1SG correct be.born in   village that NEG

          'I was not born in that village.'   (Individual-level)

Crucially, existential markers are not needed in negative individual-level predicates, as there is no event variable that needs to be existentially closed in the first place. Zimmermann (2007: 340) provides some evidence to support this prediction in Bura. Example (2.16b–c) provides that evidence for Eastern Cham. Individual-level predicates such as *nĭʔ* 'be born' in fact cannot be marked by *hu*. They can either be marked simply by *o* or by *o* and a different form *ʄăwʔ*, which means 'correct' or 'so' elsewhere in the language.

The example below provides another near-minimal pair. The negated stage-level predicate *puj pɛ* 'be happy' is typically marked by *hu*, while the individual-level predicate 'be a happy person' is more accepted with *ʄăwʔ* 'correct' (2.17a–b). One additional example is the verb *tʰăw* 'know', which very frequently is marked only by *o* when negated (2.17c).

(2.17)   a. mĕj   năn hu puj pɛ   ʔo
           female that ∃   be.happy NEG

          'That woman is not happy.'   (Stage-level)

       b. mĕj   năn ʄăwʔ   tʰa jaŋ   puj pɛ ʔo
           female that correct one CLF.PERSON happy NEG

          'That woman is not a happy person.'   (Individual-level)

       c. kăw tʰăw   ʔo
           1SG know NEG

          'I don't know.'   (Individual-level)

With the *hu*-test in mind, stage- and individual-level predicates are now diagnosable in Eastern Cham by language-internal evidence. From here, first I will show that it is ungrammatical for a phrase to be DC-marked in a clause that contains an individual-level predicate, even when the DC conditions should otherwise be satisfied. Consider the discourses in (2.18a–b). Both contain a context sentence that mentions 'that village'. DC-movement of *mlĕj năn* 'that village' is grammatical in (2.18a), which contains a stage-level predicate, as diagnosed by *hu* above. However, DC-movement of the same phrase is ungrammatical in (2.18b), which contains an individual-level predicate.

(2.18)   a. CONTEXT: I spend a lot of time in that village.

      mlĕj   năn_{DC} kăw ʈɔʔ ~~păʔ~~ ~~mlĕj năn~~
      village that   1SG live in
      'That village, I live in [temporarily].'   (Stage-level)

    b. CONTEXT: I have lived in that village for many years.

      *mlĕj   năn_{DC} kăw nĭʔ   ~~păʔ~~ ~~mlĕj năn~~
      village that   1SG be.born in
      INTENDED: 'That village, I was born in.'   (Individual-level)

It is worth noting that this ungrammaticality judgment in (2.18b) is different from typical judgments from elicitation on DC-movement, which are at worst infelicitous. This sentence is ungrammatical regardless of context. Under an event semantic account of DC, this follows because an individual-level predicate never introduces an event variable, and the DC-particle carries a presupposition that requires the current sentence to have an event variable. Therefore, the resulting sentence has no interpretation; the presupposition must fail. Note also that the prior context sentences are different between (2.18a) and (2.18b) above, in order to create felicitous explanation contexts.

As corroborating evidence, the same sentences are felicitous without DC-movement and with the connective *jwa* 'because', which enforces an explanation relation. This demonstrates that the individual-level predicate in (2.19b) can satisfy the discourse relational aspect of DC; it can be interpreted in a subordinating discourse relation with the prior sentence. Additionally, the same phrase can be mentioned in both. Nevertheless, DC-movement is ungrammatical.

(2.19)   a. CONTEXT: I spend a lot of time in that village.

      jwa   kăw ʈɔʔ păʔ mlĕj   năn
      because 1SG live in   village that
      'Because I live in that village [temporarily].'   (Stage-level)

    b. CONTEXT: I have lived in that village for many years.

      jwa   kăw nĭʔ   păʔ mlĕj   năn
      because 1SG be.born in   village that
      'Because I was born in that village.'   (Individual-level)

More evidence comes from the semantics of the prior sentence. Consider the two discourses in (2.20a–b). Here, the second sentence is the same, but the prior sentence contains either a stage- or individual-level predicate. In (2.20a), the prior sentence is a stage-level predicate, the same as above, but with a different location. If this is the prior sentence, then DC-movement of *paj ǩɔl* 'Saigon' is felicitous. Note that preposition-drop occurs, resulting in the preposition *pǎʔ* 'in' not being pronounced, a general property of DC-movement (cf. Section 2.2).

(2.20)   a. CONTEXT: I live in Saigon [temporarily].   (Stage-level)

    paj ǩɔl_{DC} mɛʔ   mɨ   kǎw naw ŋǎʔ   pjŭʔ   p̥ǎʔ p̥aj ǩɔl
    Šaigon   mother father 1SG   go   make work   in
    'My parents went to work in Saigon.'

    b. CONTEXT: I was born in Saigon.   (Individual-level)

    #paj ǩɔl_{DC} mɛʔ   mɨ   kǎw naw ŋǎʔ   pjŭʔ   p̥ǎʔ p̥aj ǩɔl
    Šaigon   mother father 1SG   go   make work   in
    INTENDED: 'My parents went to work in Saigon.'

Despite the fact that DC-movement is possible in this sentence, it is infelicitous in (2.20b). The only difference is that the prior sentence contains an individual-level predicate. This follows from the event semantic account of DC, as the DC-particle introduces a presupposition referring to a prior event variable. In the absence of such a variable, the presupposition fails and the sentence receives no interpretation again.

Finally, there is nothing blocking these sentences from being in the same rhetorical relations. In (2.21), the sentences are felicitous in both contexts without DC-movement and with the connective *jwa* 'because', again enforcing an explanation relation.

(2.21)   a. CONTEXT: I live in Saigon [temporarily].   (Stage-level)

    jwa   mɛʔ   mɨ   kǎw naw ŋǎʔ   pjŭʔ   pǎʔ paj ǩɔl
    because mother father 1SG   go   make work in   Šaigon
    'Because my parents went to work in Saigon.'

    b. CONTEXT: I was born in Saigon.   (Individual-level)

    jwa mɛʔ   mɨ   kǎw naw ŋǎʔ   pjŭʔ   pǎʔ paj ǩɔl
    jwa mother father 1SG   go   make work in   Šaigon
    'Because my parents went to work in Saigon.'

Based on this evidence, I conclude that DC must be a condition on events in a discourse as construed by Kratzer (1995). DC-marked phrases must be participants in the cause or subtype relation inferred by an explanation or elaboration (i.e. subordinating) discourse relation. Note that this event-based account of DC is not specific about where in a given sentence an event is introduced. For example, if the relevant event could be introduced

in a matrix clause, then the embedded clause would be free to contain an individual-level predicate, as the matrix clause provides the event for the purposes of DC-marking.

Conversely, DC-marking provides evidence for Kratzer's (1995) account of stage- and individual-level predicates. The event-based account of DC also suggests a need for a dynamic event semantics. For the DC conditions to be assessed, the participants of two events must be known, one of which is introduced in a prior sentence in the discourse. Under a static event semantics, event variables are introduced and then existentially closed over the course of the derivation of a sentence (cf. Champollion 2015 for a comparison and evaluation of different models of event semantics in modern compositional semantics). It is unclear how events introduced in prior sentences could be assessed for the purposes of DC. Under a dynamic event semantics, one aspect of discourse could be a set of events. Each sentence could update the discourse, adding its events to that set. Then, the DC conditions could be assessed by making reference to that set of events in the discourse. Section 2.1.4 proposes a semantics of DC along these lines.

2.1.3 Previous mention

Next, the previous mention condition of discourse connectedness is detailed. For ease of exposition, the term 'previous mention' will be used in this section, but note that the following section will refine this notion in terms of participants in semantic events and set inclusion. A broad range of phrases may be DC-moved in Eastern Cham, as long as the phrase or its referential index have been previously mentioned in a superordinate sentence. First, many of the DC-phrases above denoted individuals. Generics can also be DC-moved in Eastern Cham (2.22). The previous mention of a generic can either denote an individual (2.22a) or another generic (2.22a′).[6]

(2.22)   a. tʰuːɲm͡³¹² ʈɔʔ ŋăʔ **ʔiŋ ʔɔŋ năn**
              Thuận   PROG make frog   that
              'Thuận is cooking that frog.'

      a′. tʰuːɲm͡³¹² ʈɔʔ ŋăʔ **ʔiŋ ʔɔŋ**
              Thuận   PROG make frog
              'Thuận is cooking frog.'

      b. **ʔiŋ ʔɔŋ**₍DC₎ ɲu ŋăʔ ~~ʔiŋ ʔɔŋ~~ ɲi   lo
          frog   3.ANIM make   be.delicious very
          'Frog, he cooks very well [Lit: deliciously].'   (a ⇓ b)

I posit that DC-marked generics are in fact interpreted as kinds of type e. In (2.22b) above, *ʔiŋ ʔɔŋ* 'frog' is interpreted as a kind. The advantage of appealing to kinds is that all DC-marked phrases in Eastern Cham are then of type e, such as *ʔiŋ ʔɔŋ năn* 'that frog'

---

[6]López (2009: 100) observes the same pattern in Catalan. In both languages, there is a gap such that a generic previous mention does not license an individual DC-phrase.

in examples above and *ʔiŋ ʔɔŋ* 'frog' in (2.22b).[7] In terms of semantic implementation, Chierchia's (1997: 77) 'down' operator transforms a property of type $<e,t>$ to a corresponding kind of type e (2.23).

(2.23) For property P, $^\cap$P denotes the corresponding kind

In many of the examples above, the previous mention was identical to the DC-phrase. The previous mention relation can also be a set-subset relation, such that the DC-phrase represents a subset of its antecedent, as in (2.24b). That subset can also take the form of quantifiers or focus markers like *tʰa sĭt* 'only' (2.24b′).[8]

(2.24)   a. hu **mɨ pɔ̥h jăw** pă? ni
         ∃   5   fruit   here
        'There are 5 [kinds of] fruits here.'

     b. **pɔ̥h**   **ʔɔ̆?**   **ni**$_{DC}$ kăw ʔĭŋ   ɓăŋ   ~~pɔ̥h ʔɔ̆? ni~~
        CLF.ROUND mango this 1SG want eat
        'This mango, I want to eat.'   (a ⇓ b)

     b′. **tʰa sĭt pɔ̥h**   **ʔɔ̆?**   **ni min**$_{DC}$ kăw ʔĭŋ   ɓăŋ
        only   CLF.ROUND mango this EMPH 1SG want eat
        ~~tʰa sĭt pɔ̥h ʔɔ̆? ni min~~

        'Only this mango, I want to eat.'   (a ⇓ b)

Additionally, DC-movement is licensed in a variety of bridging contexts. For example, DC-phrases can be in part-whole and producer-product relations, which are known to license weak and strong definites, respectively, in languages like German (Schwarz 2009: 49) and Mandarin (Jenks 2018). In (2.25), the DC-phrase *ɓ̥ăŋ* 'door' is a part of the whole, *tʰaŋ* 'house'.

(2.25)   a. **tʰaŋ   năn** k̥ʰəh   lo
        house that be.beautiful very
        'That house is very beautiful.'

---

[7]In Catalan, phrases that denote individuals, kinds, and properties such as adjective phrases, prepositional phrases, and others can undergo the equivalent of DC-movement, clitic dislocation (López 2009: 4). To model Catalan, the set $\mathcal{P}_{e'}$ must be expanded to include individuals and properties. One option would be to appeal to the notion of the Universe from Discourse Representation Theory (DRT; Kamp & Reyle 1993; Kamp, van Genabith & Reyle 2011), which models the denotation of a sentence as a tuple of referents (i.e. individuals) and predicates (i.e. properties). See Chapter 5, Section 5.1.2 for further discussion.

[8]Note that the classifier *pɔ̥h* in these examples and elsewhere in this dissertation has a usage unlike other Eastern Cham numeral classifiers. It is generally optional in contexts without numerals, such as (2.24b), and it is optional in the equivalent of (2.24b) without the demonstrative.

    b. **ɓ̥ăŋ**_{DC} kăw cĭh ɓ̥ăŋ pan çaw
       door 1SG paint color blue

    'The door, I painted blue.' (a ⇓ b)

In (2.26), the DC-phrase *ták jă* 'author' is the implied producer of the product *ʈankan* 'story'.

(2.26)   a. **ʈankan năn** k̥ʰəh lo
       story that be.beautiful very

    'That story is very beautiful.'

    b. **tak^{35} ja^{214}**_{DC} kăw thăw tak^{35} ja^{214} ʔo
       author(VN) 1SG know NEG

    'The author, I do not know.' (a ⇓ b)

How can these bridging examples, and the set-subset examples above be amenable with the previous mention condition proposed for DC? For one, this is a general problem for anaphoric phenomena, as definites are known to be licensed in these bridging contexts. According to Schwarz (2009: 267), producer-product bridging is a kind of relational anaphora, where the relatum is actually merged as an index to the antecedent DP. In other words, *tak ja* 'author' would be covertly previously mentioned in (2.26a) as an index on the DP *ʈankan* 'story'. DC-marking of referential indices may also help explain the gestural deixis example from Section 2.1.1, repeated below (2.27). Here, there is no explicit prior mention, but the reference is demonstrated by pointing.

(2.27)   a. CONTEXT: Speaker pointing at mango in their hand.

    b. **ʔɔ̆ʔ ni**_{DC?} kăw ʈɔʔ ɓăŋ ʔɔ̆ʔ ni
       mango this 1SG PROG eat

    'This mango, I am eating.' (a ⇓? b)

As for part-whole bridging, perhaps subset DC-phrases covertly contain a restriction, such as 'the door (of that house)' in (2.25b), where it is the restriction that is DC-marked. Chapter 3, Section 3.2.1 gives a comparable analysis for *wh*-phrases, which may contain a covert restriction that is DC-marked. Chapter 4, Sections 4.2.1 and 4.3.1 establishes that DC-marking can also occur when the restriction is overt, such as in partitives. In these cases, however, it is that overt restriction that is DC-marked, to the exclusion of the remainder of the DP.

    There are three contexts to be noted where DC-movement is impossible. Universally unique referents such as 'the sun' and 'the president' can only be DC-moved if the referents have been discussed in prior discourse such that the phrase satisfies the DC conditions. Thus, not every context that licenses definiteness licenses DC-marking. DC-marking is also illicit in corrective contexts, where a speaker corrects the identity of a referent, as

in (2.28). This can be explained in several ways: there is no previous mention, and corrections are not intuitively interpreted as in subordinating discourse relations with the corrected sentence (see Appendix A, Section A.3.4). Therefore, there is not sufficient previous mention and not sufficient discoure structure to license DC-marking of *lo* 'meat' in (2.28b).

(2.28)   a. ʔaj   mĕj   kăw țɔʔ tŭʔ **ʄăm**
             older.sibling female 1SG PROG cook vegetable
             'My older sister is cooking vegetables.'

        b. ʄăwʔ   ʔo   // #**lo**   ʔaj   mĕj   hɨ țɔʔ tŭʔ ~~lo~~
          correct NEG   meat older.sibling female 2SG PROG cook
          'No, she is cooking meat.'   (a ⇃ b)

Finally, DC-movement requires that the DC-phrase have an overt, non-pronominal nominal restriction. Pronouns, regardless of stress and deixis, cannot be DC-moved, as reflected in (2.29a). Neither can DPs with nouns elided due to some kind of NP-ellipsis. For example, in (2.29b), the noun *mțɛh* 'motorbike' is elided, leaving a bare numeral and classifier. Note that the in situ versions of these phrases are felicitous in DC contexts.

(2.29)   a. *__**ɲu**__   kăw hu ḳɔʔ ɲɨ mĭŋ pjoj
            3.ANIM 1SG ∃ meet   in yesterday
            'I did met him yesterday.'

        b. *__**tʰa**__ __**pɔh**__   kɛn ni ɲĭm   ~~tʰa pɔh~~ jə
          one CLF.ROUND Kenny borrow   already
          'Kenny borrowed one (wheel) already.'

A full examination of anaphora in Eastern Cham is needed to ascertain why pronouns and bare classifiers cannot be DC-moved. Perhaps there is an intrinsic incompatibility between pronouns and ellipsis on one hand, and DC-marked phrases on the other. Such an incompatibility is not yet evident based on the account of DC described here, and at least some pronouns in Catalan can be DC-marked (Chapter 5, Section 5.1.2). Or perhaps, this specific pronoun and bare classifier form in Eastern Cham have different discourse or prosodic constraints than DC.

2.1.4 Semantic value of the DC-particle

To put the previous sections together, discourse connectedness has three components: a subordinating discourse relation between two sentences, two events inferred by that discourse relation, and previous mention of a phrase in those two events. This section proposes a semantic value of the DC-particle that captures the three components of DC as a presupposition.

First, in order to account for the discourse subordination and event semantic components of DC, I propose two sets: $\mathcal{E}_c$ and $\mathcal{E}_e$ (2.30a–c). $\mathcal{E}_c$ refers to the set of all events so far added to the discourse. This set is meant to capture the history of the discourse in terms of events, i.e. what has been explicitly contributed to the conversation. A full dynamic account of how $\mathcal{E}_c$ is updated is needed, but not spelled out here. For the purposes of this section, I assume that each sentence containing an event variable adds that event to $\mathcal{E}_c$.

(2.30)   a. Let $\mathcal{E}_c$ be the set of events live in a discourse at context c

   b. Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

   c. Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

The second set, $\mathcal{E}_e$ comprises the events inferred by subordinating discourse relations that the current sentence is a part of. This set includes the cause events inferred by elaboration relations and subtype events inferred by elaboration relations (see Section 2.1.2). In all cases, the current sentence is meant to be interpreted as subordinate to a superordinate sentence. In other words, the current sentence is interpreted as an explanation or elaboration.

Next, I introduce $\mathcal{P}_e$ in order to encode previous mention (2.31a). The set $\mathcal{P}_e$ is defined as the set of all participants in event e, i.e. all the theta role-bearing referents. Section 2.1.3 argued that DC-marked phrases can either be previously mentioned themselves, or their referential index can be. These facts indicate that the DC-particle can combine either with a DP or its referential index and check if it is in certain $\mathcal{P}_e$ sets.

(2.31)   a. Let $\mathcal{P}_e$ be the set of participants in event e

   b. $[\![\text{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

Putting these three sets together, I propose that the DC-particle contributes the meaning in (2.31b). At its core, the DC-particle is an identity function that combines with an individual (λx) and returns that individual (x). The DC-particle also introduces a presupposition that makes reference to two events, e and e', and the individual denoted by the phrase it combines with. The first two pieces of the presupposition enforce that there must be a prior event in the discourse inferred by a subordinating discourse relation. Specifically, that event, e' must be in the discourse history set $\mathcal{E}_c$ and the superordinate event set $\mathcal{E}_e$. The third piece of the presupposition enforces previous mention, by requiring that the individual x be a participant of both the current event e and the superordinate event e'. If all these pieces of the presupposition are satisfied, then the sentence can be interpreted. If one piece fails, the sentence receives no interpretation.

To illustrate, consider the positive and negative example of DC-marking, repeated below in (2.32). In (2.32b), the DC-particle combines with the kind *lɔ nǔʔ* 'chicken meat'.

The presupposition introduced by the DC-particle is satisfied, because *lɔ nŭʔ* 'chicken meat' is a participant in the cooking event in (2.32a), which is an event in both the set $\mathcal{E}_c$ and the set $\mathcal{E}_e$, as it is explicitly mentioned in a superordinate sentence. As a result, DC-marking of *lɔ nŭʔ* 'chicken meat' is felicitous.

(2.32) a. ʔaj   mɛ̌j   kăw tŭʔ   **nŭʔ**
             older.sibling female 1SG cook chicken
             'My older sister cooks chicken.'

     b. jwa   **lɔ**   **nŭʔ**$_{DC}$ ʔaj   mɛ̌j   kăw cəh ɓăŋ l̶ɔ̶ ̶n̶ŭ̶ʔ̶
          because meat chicken   older.sibling female 1SG like eat
          'Because chicken meat, she likes to eat.'   (a $\Downarrow$ b)

     b′. #plɔh năn **lɔ**   **nŭʔ**   ʔaj   mɛ̌j   kăw ɓăŋ l̶ɔ̶ ̶n̶ŭ̶ʔ̶
          ̥ after that meat chicken older.sibling female 1SG eat
          'After that, she [will] eat the chicken meat.'   (a $\not\Downarrow$ b)

By contrast, the DC-particle's presupposition is not satisfied in (2.32b′). The kind *lɔ nŭʔ* 'chicken meat' is mentioned in the cooking event, which is in the set $\mathcal{E}_c$. However, that event is not in the set $\mathcal{E}_e$, as (2.32a) and (2.32b′) are not in a subordinating discourse relation. Therefore, the presupposition fails, and the sentence is infelicitous because it cannot receive an interpretation.

Why should the DC conditions be modeled as a presupposition in the semantics? It appears to be the case that DC functions as a hard presupposition, i.e. one that is encoded in the semantics. Chapter 3, Section 3.2.1 will find that DC projects in questions. They also project from underneath quantifiers such as *kiʔ həŋ*^{33} 'less than', which is known to be a property of hard, not soft presuppositions, which have been argued not to be encoded directly in the semantics (e.g. Abrusán 2016).[9] For example, the phrase *mɨ jaŋ* 'five people' is DC-marked inside the quantifier phrase in (2.33b). Nevertheless, the DC conditions apply; DC-marking of *kiʔ həŋ*^{33} *mɨ jaŋ* 'less than five people' is only felicitous in subordinating discourse relations, such as the elaborating answer in this example.

(2.33) a. hɨ ʔa   **lo**   **nujh** lɛ̌j
             ̥ 2SG invite many person Y/N.Q
             'Did you invite many people?'

     b. **kiʔ həŋ**^{33} **mɨ** **[jaŋ]**$_{DC}$ kăw ʔa   k̶i̶ʔ̶ ̶h̶ə̶ŋ̶^{33} ̶m̶ɨ̶ ̶j̶a̶ŋ̶ maj
          few exceed(VN) five CLF.PERSON 1SG invite   come
          păʔ ni
          here
          'I invited less than five people to come here.'   (a $\Downarrow$ b)

---

[9]Another test distinguishing hard and soft presuppositions is cancellability. Contexts for testing cancellability for DC have not yet been found. They would be predicted to involve metalinguistic commentary on the structure of the discourse, such as 'I don't know if we are talking about X, but…'

As a presupposition, DC might be expected to be able to be accommodated. For example, a speaker could DC-mark a phrase in the absence of an appropriate subordinating discourse relation, yet the sentence may be interpretable if the hearer is able to accommodate a discourse structure that permits DC-marking. I do not yet have conclusive naturalistic data on the accommodatability of DC, but consultants have reported metalinguistic comments indicative of accommodation. For example, in Section 2.1.1, DC-marking in the absence of subordinating discourse relations is accepted by some consultants only if they are able to revise the prior discourse to add prior sentences.

Because DC is a discourse structural presupposition, there are different predictions on the accommodatability of DC in monologues and dialogues. In monologues or otherwise the speech of one speaker, accommodation should be difficult if impossible. Presuppositions involving the history of the discourse itself have been noted to be difficult to accommodate (von Fintel 2008: 25). In dialogues, accommodation should be possible, as individual speakers are known to enter conversations at different points, revise the discourse structure itself (e.g. Hunter et al. 2017), or be more willing to enter conversations *in medias res* (von Fintel 2008: fn. 17).

With the presupposition introduced by the DC-particle in mind, there are four aspects of DC-marking that should be mentioned: the apparent verb-centrism of DC, the apparent unremarkability of DC-marked phrases, Maximize Presupposition effects, and the lack of interpretive difference between long and short distance DC-movement. First, the DC-particle has been taken to combine with a DP so far. Yet, much of the presupposition it introduces refers to events. The DP is only involved insofar as it is a participant in the relevant events. For this dissertation, it will be maintained that the DC-particle combines with DPs in Eastern Cham, because it is DPs that are explicitly DC-marked in the syntax, not verbs. In Catalan (Chapter 5, Section 5.1), DC-marking could be seen as taking place both on the DP and the verb, as DPs undergo movement and clitic doubling on the verb.

A related aspect of DC-marking is that the DP the DC-particle combines with is not necessarily special. The presupposition will be satisfied as long as the DC-particle combines with *any* DP that meets the DC conditions. For example, consider the context in (2.34). Two phrases in (2.34b) are previously mentioned in the superordinate sentence (2.34a): *jŭt kăw* 'my friend' and *tʰăw* 'dog'. DC-marking of both phrases should be licit. This is borne out in (2.34b–b′). Note that DC-marking of *tʰăw năn* 'that dog' involves DC-movement and DC-marking of *jŭt kăw* 'my friend' involves base generation + Agree, as they were deemed the most natural by consultants in this context.

(2.34) a. **jŭt   kăw** hu   **tʰăw** m̥jăw
         friend 1SG have dog   new

     'My friend has a new dog.'

  b. **jŭt   kăw**$_{DC,i}$ kăw plĕj **tʰăw năn** ka ɲu$_i$
     friend 1SG   1SG s̥ell dog   that to 3.ANIM

    'I sold that dog to my friend.'   (a ⇓ b)

b′. **tʰǎw nǎn**$_{\text{DC}}$ kǎw plĕ̥j ~~tʰǎw nǎn~~ ka **jŭt** **kǎw**
   dog   that   1SG sell   to friend 1SG

   'I sold that dog to my friend.'   (a ⇓ b)

According to consultants, both (2.34b–b′) are felicitous in this context, without an obvious meaning difference. The lack of a meaning difference is predicted by the DC conditions, which do not distinguish between the two phrases in this context. Why DC-phrases should be unremarkable and the DC conditions appear to focus on the verb, while DC-marking occurs on DPs is unclear. However, as mentioned above, it is DPs that are DC-marked in Eastern Cham.

Given that the DC-particle introduces a presupposition, it is worth investigating if DC upholds the principle of Maximize Presupposition (e.g. Schlenker 2012). Maximize Presupposition holds that the strongest possible presuppositions be chosen, provided a set of possible competing forms. The introduction to this chapter raised the issue of DC-marking (2.35b) compared with the absence of DC-marking (2.35b′). As described later in this chapter, multiple phrases can be DC-marked (2.35b″). Perhaps DC-marking of two phrases is stronger than DC-marking of one.

(2.35)   a. **jŭt** **kǎw** hu **tʰǎw** m̥jǎw
         friend 1SG have dog   new

         'My friend has a new dog.'

    b. kǎw plĕ̥j **tʰǎw nǎn** ka **jŭt** **kǎw**
        1SG sell dog   that to friend 1SG

        'I sold that dog to my friend.'   (a ⇓ b)

    b′. **jŭt** **kǎw**$_{\text{DC},i}$ kǎw plĕ̥j **tʰǎw nǎn** ka ɲu$_i$
         friend 1SG   1SG sell dog   that to 3.ANIM

        'I sold that dog to my friend.'   (a ⇓ b′)

    b″. **jŭt** **kǎw**$_{\text{DC},i}$ **tʰǎw nǎn** kǎw plĕ̥j ~~tʰǎw nǎn~~ ka ɲu$_i$
         friend 1SG   dog   that 1SG sell   to 3.ANIM

        'I sold that dog to my friend.'   (a ⇓ b″)

Section 2.4 returns to the question of how in situ phrases compare with DC-movement in more detail. It is possible that (2.35b) contains an instance of covert DC-movement or base generation of a DC-marked index that is phonologically null. If either is true, then (2.35b–b′) could be equivalent in terms of Maximize Presupposition. In this context, one phrase could be DC-marked whether overtly or covertly.

As for the difference between one and two instances of DC-marking, it is unclear if two presuppositions are actually stronger than one in the context in (2.35). Recall that the presupposition introduced by the DC-particle, repeated below as (2.36), is largely organized around an event e'. In (2.35b″) above, if two phrases are DC-marked, the presupposition is nearly identical, except for a different phrase being previously mentioned.

(2.36) $[\![\text{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

It is possibly telling that sentences with multiple DC-marked phrases appear to be preferred when there are multiple subordinating discourse relations. Multiple DC-marking was readily accepted in the context in (2.37), while it is generally somewhat degraded in two sentence contexts. In (2.37), there is a chain of discourse subordination: (2.37c) is interpreted as explaining (2.37b), which in turn is interpreted as explaining (2.37a).

(2.37)   a. **jŭt**   **kăw**   puj p̥ɛ   lo
           friend   1SG   be.happy   very

           'My friend is very happy.'

     b. **jŭt**   **kăw**   hu   **tʰăw**   m̥jăw
           friend   1SG   have   dog   new

           'My friend has a new dog.'   (a $\Downarrow$ b)

     c. **jŭt**   **kăw**$_{\text{DC},i}$ **tʰăw**   **năn**$_{\text{DC}}$ kăw   plĕj   ~~tʰăw năn~~   ka   ɲu$_i$
           friend   1SG   dog   that   1SG   s̥ell   to   3.ANIM

           'I sold that dog to my friend.'   (b $\Downarrow$ c)

The existence of two separate subordinating discourse relations may license multiple DC-movement, as now the two presuppositions in (2.37c) are different. The presupposition introduced by the DC-particle that combines with *jŭt kăw* 'my friend' could be satisfied by the event in (2.37a), while that with *tʰăw năn* 'that dog' could be satisfied by the event in (2.37b).

Finally, there does not appear to be an interpretive difference between long and short distance DC-movement. For example, the phrase *ʔɔ̌ʔ* 'mango' can be DC-moved short distance to the embedded clause edge in (2.38b), or it can be DC-moved long distance to the matrix clause edge, as indicated by the bracket notation. In this context, the semantic event in (2.38b) responsible for DC-marking is the event introduced in the embedded clause. Nevertheless, *ʔɔ̌ʔ* 'mango' can be DC-moved beyond the embedded clause.

(2.38)   a. pu   hu   ʔĭŋ   ɓăŋ   **ʔɔ̌ʔ**   lĕj
           Phú   ∃   want   eat   mango   Y/N.Q

           'Does Phú want to eat mango?'

     b. {**ʔɔ̌ʔ**$_{\text{DC}}$}   ɲu   ɗom   {**ʔɔ̌ʔ**$_{\text{DC}}$}   ɲu   ɓăŋ   ~~ʔɔ̌ʔ~~   jə
           mango   3.ANIM   say   mango   3.ANIM   eat   already

           'He said he already ate mango.'   (a $\Downarrow$ b)

More data is needed to fully assess the differences between long and short distance DC-movement. Based on this initial data, it seems to be the case that the presupposition introduced by the DC-particle only requires there to be some event in the current sentence that contributes to the DC conditions.

To sum up this section, there are three components of DC-marking: a subordinating discourse relation between two sentences, two events inferred by that discourse relation, and previous mention of a phrase in those two events. These components can all be enforced by a hard (i.e. semantic) presupposition, (2.36) above.

2.2 Background on Ā-movement in Eastern Cham

This section presents syntactic distributional properties of Ā-movement in Eastern Cham that provide a backdrop for the argumentation in the next sections and chapters. Prosody distinguishes a hanging topic construction from true Ā-movement operations. Then, the optional presence of a complementizer, preposition drop, and an argument/adjunct asymmetry are all shared by DC-movement, base generation + Agree, presentational clefts, and relativization. Together, these facts present language-internal justification for treating DC-movement as a kind of Ā-movement operation.

2.2.1 Prosody

First, the prosody of DC-movement distinguishes it from a hanging topic construction. Hanging topics are present in many languages and are typically characterized by a pause after the topic and a potentially loose syntactic connection between the topic and the rest of the sentence (e.g. Aissen 1992 on hanging topics and left-dislocation). It is important to identify hanging topics, because they are known to have different syntactic properties from Ā-movement operations. Eastern Cham does have a hanging topic construction (2.39a), which contains an audibly longer pause than DC-movement (2.39b). In (2.39a), there is a null pronoun that refers to the hanging topic. Note that a comparable prosodic difference exists between base generation + Agree and a hanging topic with a pronoun in the base object position.

(2.39)   a. **pu**$_i$ // hɨ   ʔa   ∅$_i$ hu
                Phú   2SG invite *pro* ROOT
             'Phú. You can invite him.'   (Hanging topic)

          b. **pu** hɨ   ʔa   ~~pu~~ hu
                Phú 2SG invite   ROOT
             'Phú, you can invite.'   (DC-movement)

Figures 2.2a–b present pitch tracks for (2.39a) and (2.39b), respectively, which highlight the length of the pause and prosodic contours. The hanging topic in Figure 2.2a has a much longer pause (0.55s), and the prosodic contour indicates a stress followed by fall, reminiscent of sentence-final drops. DC-movement (Figure 2.2b) has a much shorter

pause (0.07s) and a more level prosodic contour. This follows if a hanging topic constitutes an independent prosodic unit, while DC-moved phrases constitute a prominent phrase within a larger sentence.[10]

Figure 2.2: Prosody of hanging topic and DC-movement

(a) Hanging topic prosody

(b) DC-movement prosody
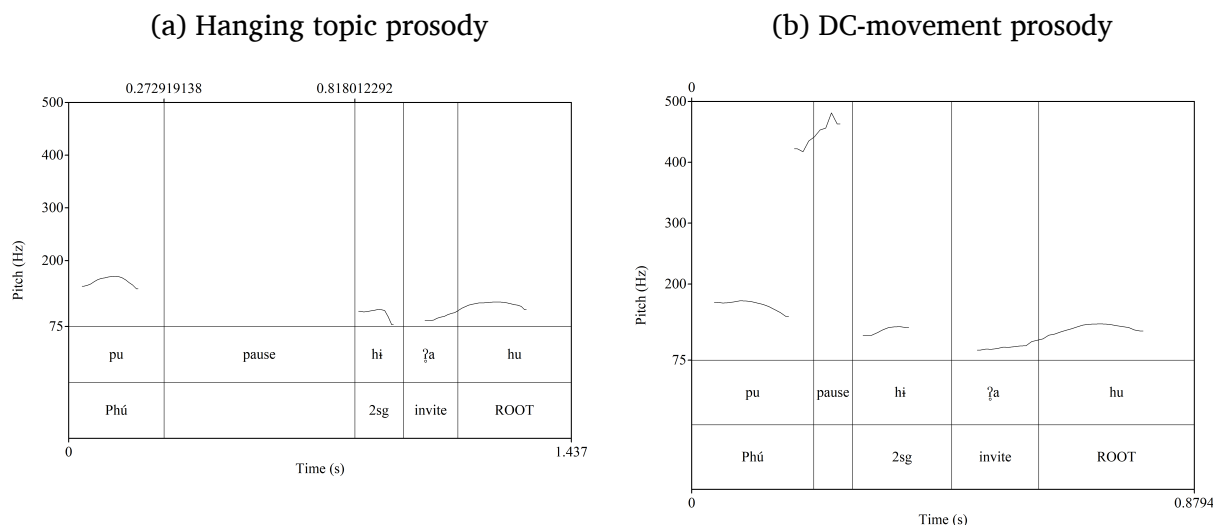
Figure 2.3 compares the prosodic contour of moved and in situ DC-phrases. Both of these sentences were offered in the same context, where *paʔ raŋ* 'four people' satisfied the conditions of being DC. While there is an increase in pitch associated with the left edge of the sentence, there is no obvious difference in the prosodic contours of the in situ and moved DC-phrases. In English topicalization, there is a characteristic rise-fall-rise prosodic contour indicative of contrast (e.g. Jackendoff 1972: 258). Unlike English, there is no contrastive or other characteristic prosodic contour in Eastern Cham. This being said, more systematic research is needed to fully ascertain the prosody of DC-marking. Perhaps there is a characteristic prosodic contour associated with DC-phrases or DC-marking that has not yet been detected.

(2.40) a. kăw ʔḁ   paʔ raŋ
        1SG invite four CLF.PERSON

    'I invited four people.'

    b. **paʔ raŋ**   kăw ʔḁ ~~paʔ raŋ~~
        four CLF.PERSON 1SG invite

    'I invited four people.'   (DC-movement)

---

[10]As a methodological note, consultants were trained to identify hanging topics as involving two separate sentences. This often proved to be a reliable meta-linguistic marker of hanging topics, at least in Eastern Cham communities.

Figure 2.3: Prosody of moved and in situ DC-phrases

(a) In situ DC-phrase

(b) Moved DC-phrase
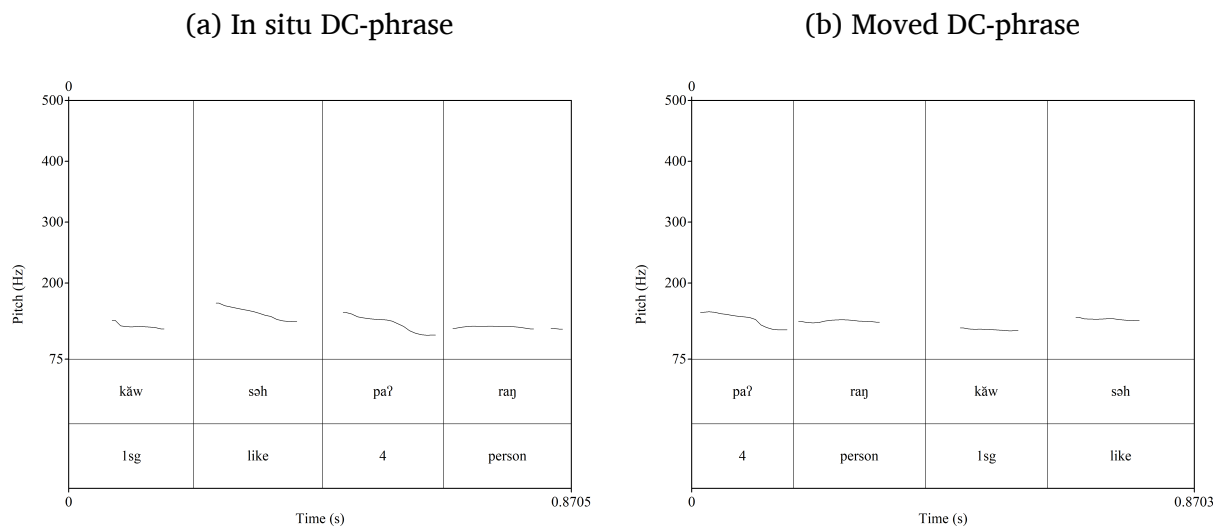
2.2.2 Complementizer

Second, there is a complementizer that optionally occurs in Eastern Cham, *po* or *plɔh*, as seen with DC-movement (2.41a), relative clauses (2.41b), and presentational clefts (2.41c). This complementizer marks Ā-extraction. It reliably occurs in instances of DC-movement, but not in hanging topics (2.41d). In other words, if the complementizer cannot appear, the relevant construction will not be considered DC-movement. Examples of DC-movement will be generally be cited with or without the complementizer, depending on the consultant's naturalistic response in the given context.

(2.41)   a. **pu**   po   hɨ   ʔa   ~~pu~~ hu
             Phú   COMP   2SG   invite   ROOT
             'Phú, you can invite.'   (DC-movement)

         b. **pu**   po   hɨ   ʔa   ~~pu~~ tɔʔ   pǎʔ ṭeh
             Phú   COMP   2SG   invite   COP   there
             'Phú, who you invited, is over there.'   (Relative clause)

         c. hu   **tʰa   jaŋ   nujh**   po   kǎw   ʔa   ~~tʰa jaŋ nujh~~
             ∃   one   CLF.PERSON   person   COMP   1SG   invite
             'There is a person who I invited.'   (Presentational cleft)

         d. *$\textbf{pu}_i$ //   po   hɨ   ʔa $\varnothing_i$ hu
             Phú   COMP   2SG   invite   *pro*   ROOT
             INTENDED: 'Phú. You can invite him.'   (Hanging topic)

This use of the form *po* is an empirical claim of this dissertation. It is unattested in previous literature on Eastern Cham. In fact, Thurgood (2005: 508) suggests that no such complementizers exist in the language. In the author's fieldwork, the form also appears as a connective and is used as an affirmative discourse particle. It appears to derive from *pō*, which Aymonier & Cabaton (1906: 309) report as an affirmative particle.

In general, *po* alternates with *plɔh*, which is attested in the literature as a connective meaning 'after'. Among the six major consultants on whom this dissertation is based, four used the form *po* in DC-movement and relative clauses, one used *plɔh*, and one used both. I conclude that these are true complementizers, as they were frequently used when the target sentence offered in English, Vietnamese, or Eastern Cham lacked any complementizer. In other words, they are not calques of a complementizer or other form in one of the contact languages.

David Blood (1977: 63) and others report that the demonstrative *năn* or a reduced form like *ăn* acts as a topicalizer. In the author's fieldwork, *năn* is consistently rejected in favor of *po* or *plɔh*. Further work is needed to understand the grammatical and sociolinguistic distribution of these forms.

2.2.3 Subject/object asymmetry

DC-movement and relativization exhibit an asymmetry such that matrix subjects cannot be moved to the immediately dominating CP. This is one instance where the complementizer plays an important role, as it is otherwise unclear if DC-movement has occurred. In (2.42a), the subject cannot appear to the left of the complementizer, indicating that it must remain in subject position; it cannot be DC-moved. The complementizer also cannot be present in the relative clause in (2.42b). Presentational clefts, by contrast, do not have a matrix subject restriction on complementizers, as shown in (2.42c). As for non-matrix subjects, embedded subjects (2.42d), or any non-subject argument can be DC-moved and accompanied by the complementizer.

(2.42) a. **nujh ni** (*po) ~~nujh ni~~ ɓăŋ pɔh ʔɔ̃ʔ năn
           person this COMP   eat CLF.ROUND mango that
           'This person ate that mango.'

       b. **nujh năn** (*po) ~~nujh năn~~ ɓăŋ ʔɔ̃ʔ tɔʔ păʔ t̥eh
           person that COMP   eat mango COP there
           'That person who ate mango is over there.'   (Relative clause)

       c. hu **tʰa jaŋ** **nujh** (po) ~~tʰa jaŋ nujh~~ ɓăŋ ʔɔ̃ʔ
           ∃ one CLF.PERSON person COMP   eat mango
           'There is a person who ate mango.'   (Presentational cleft)

       d. **nujh ni** (po) kăw hnĭŋ ~~nujh ni~~ ɓăŋ pɔh ʔɔ̃ʔ năn
           person this COMP 1SG think   eat CLF.ROUND mango that
           'This person, I think ate that mango.'

The asymmetry between matrix subjects and non-matrix subjects has a variety of possible explanations: perhaps movement does occur, but there is an independent restriction on the pronunciation of the complementizer; or perhaps the subject cannot move to Spec-CP due to a specifier-to-specifier anti-locality restriction (Erlewine 2016). See Section 2.3.1 for further discussion on specifier-to-specifier anti-locality restrictions in Eastern Cham, as opposed to other anti-locality restrictions that have been proposed. For the purposes of this dissertation, it will be assumed that subjects do not move to the immediately dominating CP, as similar anti-locality restrictions are active throughout Eastern Cham (Section 2.3.1, Section 2.3.3).

2.2.4 Preposition drop & argument/adjunct asymmetry

Finally, DC-movement triggers preposition drop or p-drop, a characteristic shared by other Ā-movement operations in Eastern Cham. When DP complements of prepositions are pronounced in their base position, the preposition must be pronounced. For example, the preposition *ka* 'to' is obligatory in (2.43a). The reverse is true under DC-movement: *ka* cannot be pronounced either in the derived or base position in (2.43b). The preposition also cannot be pronounced in other instances of Ā-movement, such as relative clauses (2.43c) and presentational clefts (2.43d).

(2.43)   a. kăw plĕj han ni   *(ka) **nĭʔ**   **năn**
      1SG g̊ive cake this   to   child that
      'I [will] give this cake to that child.'

    b. (*ka) **nĭʔ**   **năn** kăw plĕj han ni   (*ka) ~~nĭʔ năn~~
       to   child that 1SG g̊ive cake this   to
       'That child, I [will] give this cake to.'   (DC-movement)

    c. (*ka) **nĭʔ**   **năn** po   kăw plĕj han ni   (*ka) ~~nĭʔ năn~~ ʈɔʔ
       to   child that COMP 1SG g̊ive cake this   to   COP
       păʔ ʈeh
       there
       'That child who I [will] give this cake to is over there.'   (Relative clause)

    d. hu **tʰa jaŋ**   **nĭʔ** kăw plĕj han ni   (*ka) ~~tʰa jaŋ nĭʔ~~
       ∃ one CLF.PERSON child 1SG g̊ive cake this   to
       'There is a child who I [will] give this cake to.'   (Presentational cleft)

Preposition-drop (or 'p-drop') is attested in areally or genetically close languages like Mandarin Chinese (Wang 2007) and Indonesian (e.g. Sato 2011), and it is usually given a prosodic explanation, such as prepositions being too weak to be pronounced in the relevant position. Adjuncts, however, display different properties. When an adjunct preposition phrase is moved to the left periphery the preposition must be pronouced (2.44b).

(2.44) a. jŭt   ɲum ʔja   cɛ   ni   *(**mǐŋ**) **ǒŋ͡m**[45]
          friend drink water tea this with   straw(VN)

      'You[friend] drink this tea with [a] straw.'

   b. *(**mǐŋ**) **ǒŋ͡m**[45] jŭt   ɲum   ʔja   cɛ   ni   ~~mǐŋ ǒŋ͡m~~[45]
        with   straw   friend drink water tea this

      'With [a] straw, you[friend] drink this tea.'

   c. **mǐŋ** **ǒŋ͡m**[45] (*po) jŭt   ɲum   ʔja   cɛ   ni   ~~mǐŋ ǒŋ͡m~~[45]
        with   straw   C̥OMP   friend drink water tea this

      'With [a] straw, you[friend] drink this tea.'

Adjunct movement will be set aside for the purposes of this dissertation, as it does not exhibit the syntactic and pragmatic properties of DC-movement. For example, adjunct movement cannot be marked by the complementizer (2.44c), and the pragmatic conditions of DC have not been observed, but more research is needed. This is not unexpected, as adjuncts are known to have different movement properties in Southeast Asian languages (cf. Tsai 2009 on a *wh*-argument/adjunct asymmetry).

2.3   Ā-properties of DC-marking

This section turns to cross-linguistic properties associated with Ā-movement. DC-movement, which leaves a gap in the base position, is shown to be an instance of Ā-movement, involving an Agree relation between C and a DC-feature (2.45a–b). This contrasts with DC-marking that leaves a resumptive pronoun, which still involves an Agree relation (shown here with an Op feature), but base generation of the DC-phrase in Spec-CP (2.45c–d).

(2.45) a. **ʔɔ̌ʔ**   **ni** kǎw t̥ɔʔ   ɓǎŋ   ~~ʔɔ̌ʔ ni~~
            mango this 1SG PROG eat

      'This mango, I am eating.'

   b. [$_\text{DCP}$ DC [$_\text{DP}$ ʔɔ̌ʔ ni]] C$_\text{uDC}$ kǎw t̥ɔʔ ɓǎŋ ~~DCP~~   (DC-movement)

   c. **nǐʔ sɛh nǎn** kǎw ʔa   ɲu
      student that   1SG invite 3.ANIM

      'That student, I invited him/her.'

   d. [$_\text{DCP}$ DC [$_\text{DP}$ nǐʔ sɛh nǎn]]$_i$ C$_\text{uOp}$ kǎw ʔa [ɲu$_\text{Op}$]$_i$ (Base generation + Agree)

Evidence for the distinction between DC-movement and base generation + Agree comes from Ā-movement diagnostics: island constraints, weak crossover, and locality effects. Table 2.1 presents the results of these diagnostics. Both DC-marking strategies are sensitive to island constraints, implying some Agree relation with C. Only DC-movement

presents weak crossover effects, indicative of Ā-movement. Finally, when two phrases are DC-marked in the same sentence, DC-moved phrases display locality effects with each other. Base generated DC-phrases display locality effects with each other. However, these effects disappear with a mix of the two. As a result, it is clear that the two DC-marking strategies involve distinct movement operations.

Table 2.1: Movement diagnostics and DC-marking

	Island sensitivity	Weak crossover	Locality w/DC-movement	Locality w/Base generation +Agree
DC-movement	✓	✓	✓	✗
Base generation +Agree	✓	✗	✗	✓

Sections 2.3.1–2.3.2 compare the movement diagnostics above for DC-movement and base generation+Agree, respectively. Section 2.3.3 presents a complication: long distance DC-movement necessitates the existence of both DC-probes and EPP-features.

2.3.1 DC-movement

This section examines the Ā-movement properties of DC-marking that leaves a gap in the base position, through unboundedness, island sensitivity, weak crossover, and locality effects (cf. Chomsky 1977; Adger & Ramchand 2005; and many others). This DC-marking strategy consistently patterns like known instances of Ā-movement, such as *wh*-movement in English.

First, DC-movement is unbounded, as is *wh*-movement and other instances of Ā-movement (Chomsky 1973). DC-marked phrases can undergo successive cyclic movement across multiple CP boundaries. In (2.46), the DC-marked phrase *han ni* 'this cake' can be pronounced in the specifier of the embedded CP or the specifier of the matrix CP, as indicated by the bracket notation.

(2.46) {**han ni**$_{DC}$} kăw hnĭŋ {**han ni**$_{DC}$} t^{h}uːŋ͡m^{312} ʔ̥a nĭʔ mĕj sĭt
    cake this 1SG think   Thuận   invite child female small
  năn maj ɓăŋ ~~han ni~~
  that come eat
  'This cake, I think Thuận invited that little girl to come eat.'

Second, sensitivity to island constraints is widely known to be a property of syntactic movement operations (e.g. Szabolcsi 2006). Minimally, island sensitivity indicates that there exists some Agree relation between a syntactic head and a phrase inside the island.

DC-movement in Eastern Cham is consistently sensitive to strong islands, such as DC-movement out of adjunct clauses (2.47a), subjects (2.47b), and relative clauses (2.47c). In each of these examples, the relevant island is indicated by square brackets. Note that the optional complementizer *po* is used in (2.47b–c) in order to guarantee the existence of a relative clause and not a different construction, such as a serial verb construction (cf. Section 2.2).

(2.47) a. ***jaŋ**   **năn** kăw tɨk⁴⁵   mb̥ăŋ [jwa   hɨ ʔa
      CLF.PERSON that 1SG be.angry(VN) very   because 2SG invite
      ~~jaŋ năn~~ maj   ɲum ɓiə³³   ni   ]
              come drink beer(VN) this
      INTENDED: 'I am very angry because you invited that person to come drink this beer.'   (Adjunct clause island)

     b. ***mɛʔ**   **kăw** [ɗo²¹   ɓăŋ po   ~~mɛʔ kăw~~ ŋăʔ   ] tɔʔ çɛʔ
      mother 1SG   stuff(VN) eat COMP   make   COP near
      păʔ ʈɛh
      there
      INTENDED: 'The food that my mother makes is over there.'   (Subject relative clause island)

     c. ***mɛʔ**   **kăw** kăw plĕj [ɗo²¹   ɓăŋ po   ~~mɛʔ kăw~~ ŋăʔ   ]
      mother 1SG 1SG ɓuy   stuff(VN) eat COMP   make
      INTENDED: 'I buy the food that my mother makes.'   (Object relative clause island)

The corresponding sentences without DC-movement are grammatical (2.48a–c). Therefore, there is something about DC-movement that causes the ungrammaticality above. This could be due to syntactic movement, or at least an Agree relation between C and the gap inside the island.

(2.48) a. kăw tɨk⁴⁵   mb̥ăŋ [jwa   hɨ ʔa ʔaj   maj
      1SG be.angry(VN) very   because 2SG invite older.sibling come
      ɲum **ɓiə**³³   **ni**   ]
      drink beer(VN) this
      'I am very angry because you invited them[older sibling] to come drink this beer.'   (Adjunct clause)

     b. [ɗo²¹   ɓăŋ po **mɛʔ**   **kăw** ŋăʔ   ] tɔʔ çɛʔ păʔ ʈɛh
      stuff(VN) eat COMP mother 1SG   make   COP near there
      'The food that my mother makes is over there.'   (Subject)

     c. kăw plĕj [ɗo²¹   ɓăŋ po **mɛʔ**   **kăw** ŋăʔ   ]
      1SG ɓuy   stuff(VN) eat COMP mother 1SG   make
      'I buy the food that my mother makes.'   (Relative clause)

It should be noted that these judgments 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, the other two consultants were often permissive with judgments in general, and I will set aside their judgments in what follows and focus on the majority pattern reported above.[11] This is not unexpected, as it is well known that there can be widespread inter-speaker variation with regard to grammaticality judgments of island constraints (e.g. Szabolcsi 2006).

Third, weak crossover indicates that DC-movement involves Ā-movement, or minimally the lack of A-movement (cf. Ruys 2004). Weak crossover occurs when a DP cannot be moved over a coreferential pronoun, even though that pronoun does not c-command the base position of the DP (Lasnik & Stowell 1991; Ruys 2000, 2004). The context for the diagnostic used in this section comes from ditransitive verbs. The base order of arguments in ditransitive predicates is direct object–indirect object, as in (2.49a).[12] The direct object can bind a pronoun in the indirect object, but not vice versa (2.49b). The binding pattern in (2.49) implies that the direct object c-commands the indirect object in this structure, if c-command is the relevant factor in determining binding here.

(2.49) a. kăw m̥jan l̥ăjʔ [nǐʔ   mjaw năn]$_i$ ka po   ɲu$_i$
            1SG return   child cat   that   to owner 3.ANIM

    'I returned that kitten to its owner.'

   b. *kăw m̥jan l̥ăjʔ ɲu$_i$ ka po   [nǐʔ   mjaw năn]$_i$
            1SG return   3.ANIM to owner child cat   that

    INTENDED: 'I returned it to the kitten's owner.'

If an indirect object is DC-moved over a direct object, a potential crossover context arises. For example, the DP *tʰa sǐt pu mǐn* 'only Phú' crosses over the direct object in (2.50a), which contains a pronoun. Note that the pronoun *ka* 'to' is deleted when its DP complement is DC-moved (cf. Section 2.2). If DC-movement were an instance of Ā-movement, the pronoun should not be able to refer to *Phú*. This prediction is borne out, as demonstrated by the ungrammaticality of the *i* index on the pronoun. Instead, the pronoun can only refer to someone else in the context. By contrast, there is no crossover in (2.50b), as *tʰa sǐt mjaw mǐn* 'only the cat' always c-commands the pronoun; it never crosses over. Thus, no weak crossover effect obtains; the pronoun may corefer with the DC-moved DP.

(2.50) a. **tʰa sǐt pu$_i$ mǐn** kăw m̥jan l̥ăjʔ bɔp^{45} ɲu$_{*i/j}$ ~~ka~~ ~~tʰa sǐt pu mǐn~~
            only   Phú EMPH 1SG return   wallet(VN) 3.ANIM to

    'I only returned Phú his wallet.'

----

[11]Contra claims made in Baclawski Jr (2016), which were based on judgments from one of the other two consultants.

[12]Objects can be shifted to result in other relative orders, which can be diagnosed by clause-final modals and aspect markers, along with focus interpretation (Baclawski Jr 2017).

b. **tʰa sĭt mjaw**$_i$ **mĭn** kăw m̥jan l̥ăjʔ? ~~tʰa sĭt mjaw mĭn~~ ka po ɲu$_{i/j}$
only cat EMPH 1SG return to owner 3.ANIM
'I only returned the cat to its owner.'

Note that the in situ versions of (2.50a–b) exhibit the same grammaticality pattern. The in situ indirect object cannot bind the direct object in (2.50a), whereas the in situ direct object can bind the indirect object in (2.50b). This grammaticality pattern is the same as that shown above in (2.49).

The existence of weak crossover effects indicates that Eastern Cham DC-movement is an instance of Ā-movement, given certain theories of weak crossover, or is minimally not an instance of A-movement, given others (Ruys 2004 on configurational and licensing approaches to weak crossover, respectively). Other Ā-movement operations known to exhibit weak crossover include *wh*-movement in English, topicalization and relativization in Hebrew (e.g. Sells 1984) and Irish (e.g. McCloskey 1990), among many others. Eastern Cham DC-movement patterns with these instances of Ā-movement.[13]

The presence of weak crossover effects in the DC-movement examples above is surprising, given apparent exceptions in weak crossover effects in English topicalization. Weak crossover effects are reported to be absent when referential DPs are topicalized (2.51a), but not quantificational DPs (2.51b). Ruys (2004) analyzes cases like (2.51a) as accidental coreference, where pronouns like *it* are not true bound variables.

---

[13]Baclawski Jr (2016) argues that there are no weak crossover effects in Eastern Cham, based on data points like (ia). Here, the R-expression *jŭt kăw* 'my friend' appears to be DC-moved over a coreferential pronoun. Upon closer inspection, however, this appears to be a different movement operation entirely, a passivization perhaps calqued from Vietnamese. For one, the Vietnamese adversative passive marker *bị* can be added in the same position as in Vietnamese (ib). Consultants consider this sentence equivalent to (ia). Passivization in Vietnamese is known not to give rise to weak crossover due to mixed A- and Ā-movement properties (cf. Bruening & Tran 2015). Another piece of evidence is that the complementizer *po* cannot be added to this sentence (ic). This is unexpected, because the complementizer is generally optionally available in DC-movement and relativization. Its impossibility would follow if the complementizer marks Ā-extraction, not passivization.

(i)   a. [**jŭt kăw**]$_i$ tʰăw ɲu$_i$ kĕʔ ~~jŭt kăw~~
friend 1SG dog 3.ANIM bite
'My friend was bitten by his dog.'

b. [**jŭt kăw**]$_i$ ɓi^{21} tʰăw ɲu$_i$ kĕʔ ~~jŭt kăw~~
friend 1SG PASS(VN) dog 3.ANIM bite
'My friend was bitten by his dog.'

c. *[**jŭt kăw**]$_i$ po tʰăw ɲu$_i$ kĕʔ ~~jŭt kăw~~
friend 1SG COMP dog 3.ANIM bite
INTENDED: 'My friend, his dog bit.'

(2.51)   a. This book$_i$, I expect its$_i$ author to buy ~~this book~~.

   b. *Everybody else$_i$, I told his$_i$ wife that I called ~~everybody else~~.

   c. **pu**$_i$ kăw m̥jan l̥ăjʔ bɔp^{45} ɲu$_{*i/j}$ ~~ka pu~~
      Phú 1SG return   wallet(VN) 3.ANIM to
      'I returned Phú his wallet.'

The pronoun *ɲu* in Eastern Cham does not appear to have the ability to avoid weak crossover through accidental coreference, as referential DPs give rise to weak crossover effects. Whether *tʰa sĭt pu mĭn* 'only Phú' is quantificational or not in the examples above, *pu* 'Phú' certainly is in (2.51c). I hypothesize that this is a fact about the pronominal system in Eastern Cham, as one of the repairs for weak crossover effects replaces the overt pronoun *ɲu* with a null pronoun (2.52). Note that the possessor of *bɔp*^{45} 'wallet' must be *pu* 'Phú' in this example, as indicated by the ungrammaticality of the j index.

(2.52) **pu**$_i$ kăw m̥jan l̥ăjʔ bɔp^{45} ∅$_{i/*j}$ ~~ka pu~~
       Phú 1SG return   wallet(VN) *pro* to
       'I returned Phú his wallet.'

This hypothesis predicts that null pronouns in Eastern Cham may be interpreted as bound variables or as free variables for the purposes of weak crossover, while the overt pronoun *ɲu* may only be interpreted in these contexts as bound variables. By contrast, English pronouns do not exhibit a distinction, and may accidentally corefer in weak crossover contexts. A full exploration of Eastern Cham pronouns is needed to assess this hypothesis.

Locality effects provide additional evidence that DC-movement involves Ā-movement through path containment effects. This paradigm will also be crucial in future sections and chapters for determining which phenomena pattern with DC-movement. Locality effects occur when structural closeness determines which phrases can interact with syntactic probes (e.g. Chomsky 2000). In Eastern Cham, there are such locality effects, but only when multiple phrases are DC-moved to the same left periphery (i.e. the left edge of the matrix clause).

The basic pattern is demonstrated in (2.53a–b). Descriptively, this pattern can be explained in terms of path containment (cf. Pesetsky 1982). Here, two phrases are DC-moved, *han ni* 'this cake', bolded throughout this chapter, and *nĭʔ mĕj sĭt năn* '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 (2.53a), the movement path of the underlined phrase is contained within that of the bolded phrase. If the paths are crossed, however, as in (2.53b), the resulting sentence is strongly and consistently ungrammatical.

(2.53) a. **han** **ni** nĭʔ mɛ̌j sĭt năn tʰuːŋ͡m³¹² ʔḁ ~~nĭʔ mɛ̌j sĭt năn~~
             cake this child female small that Thuận invite

             maj ɓăŋ **han ni**
             come eat

             'This cake, Thuận invited that little girl to come eat.'

    b. *nĭʔ mɛ̌j sĭt năn **han** **ni** tʰuːŋ͡m³¹² ʔḁ ~~nĭʔ mɛ̌j sĭt năn~~
             child female small that cake this Thuận invite

             maj ɓăŋ **han ni**
             come eat

             INTENDED: 'This cake, Thuận invited that little girl to come eat.'

Note that the discourse context is not given in the examples in this section, as the relevant contrast is one of grammaticality. No context has been found that licenses ungrammatical sentences such as (2.53b), and consultants consistently strongly reject the sentences in a way not observed with infelicitous DC-marking in an otherwise grammatical sentence. DC-marking of multiple phrases is most felicitous if there are two superordinate sentences, each of which introduces one of the DC-marked phrases (see Section 2.1.4).

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 on English *wh*-movement). Baclawski Jr & Jenks (2016) analyze a similar phenomenon in Moken (Austronesian: Thailand) with two CPs with two separate C-probes. An analysis of (2.53a) is depicted in Figure 2.4. This analysis is elaborated upon in the following figures. For ease of exposition, the bolded DCPs correspond with the bolded phrases in the interlinearized Eastern Cham examples, and the underlined DCPs with the underlined phrases. Also note that the trees have been abbreviated to just CP projections and the relative structural hierarchy of the base positions of the two DCPs.

Figure 2.4: Path containment derivation (cf. 2.53a)
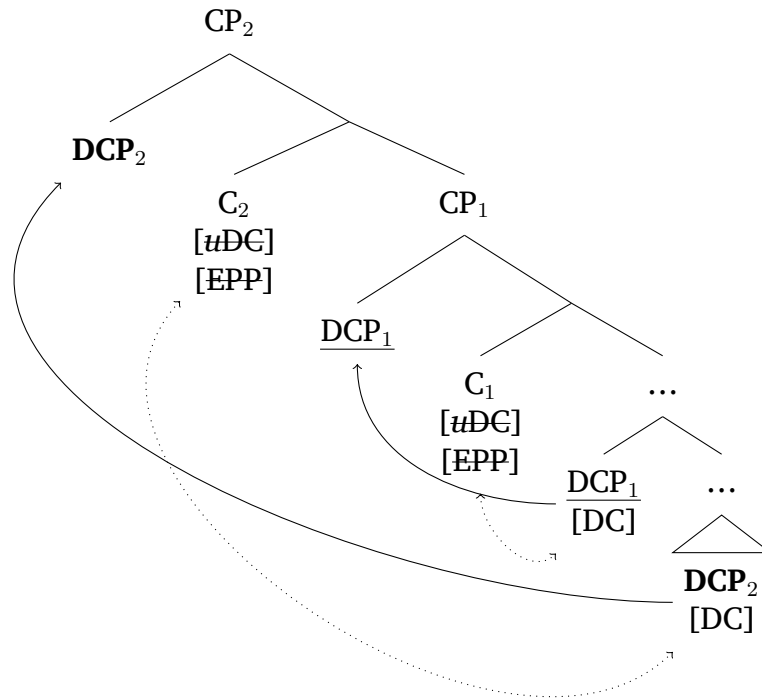
The first relevant step of this derivation is the projection of CP_1, assuming that each CP is projected in turn. C_1 has a probe that searches for a phrase bearing the DC-feature. As shown in Figure 2.5a, C_1 Agrees with the structurally closest such phrase, DCP_1. It is that DCP that moves to Spec-CP_1 (Figure 2.5b), assuming that DC-probes attract phrases to their specifiers.

Figure 2.5: Projection of CP_1

(a) C_1 Agrees with DCP_1 (b) DCP_1 moves to Spec-CP_1
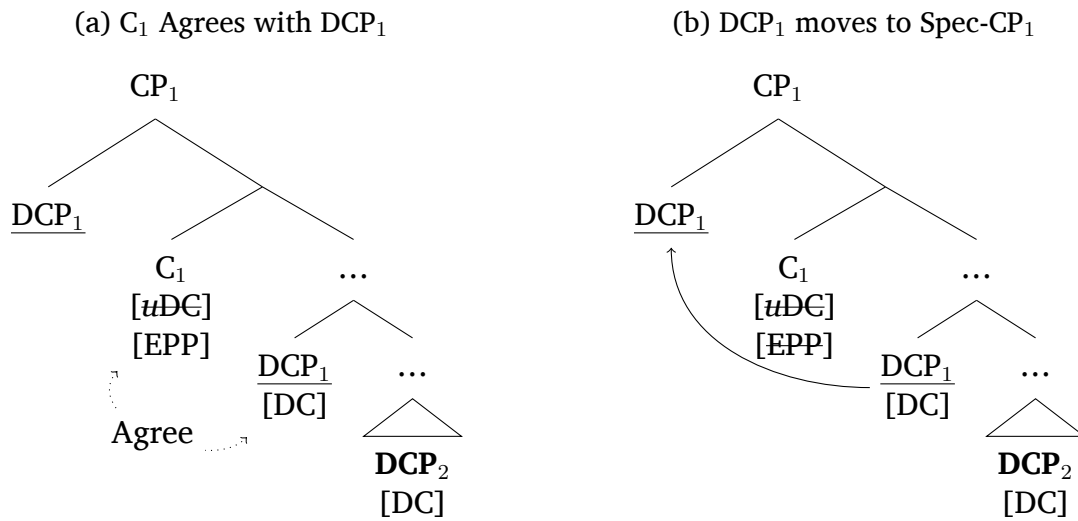
With the DC-probe satisfied and DCP_1 moved, CP_2 is projected, along with another DC-probe on C_2. At this point, C_2 agrees with the next closest DCP, DCP_2 (Figure 2.6a). It is that phrase that is moved to Spec-CP_2 (Figure 2.6b). This scenario guarantees a path containment effect, because the innermost probe must agree with the structurally highest DCP, and the outermost probe with the lowest DCP.

Figure 2.6: Projection of CP_2

(a) C_2 Agrees with DCP_2

(b) DCP_2 moves to Spec-CP_2
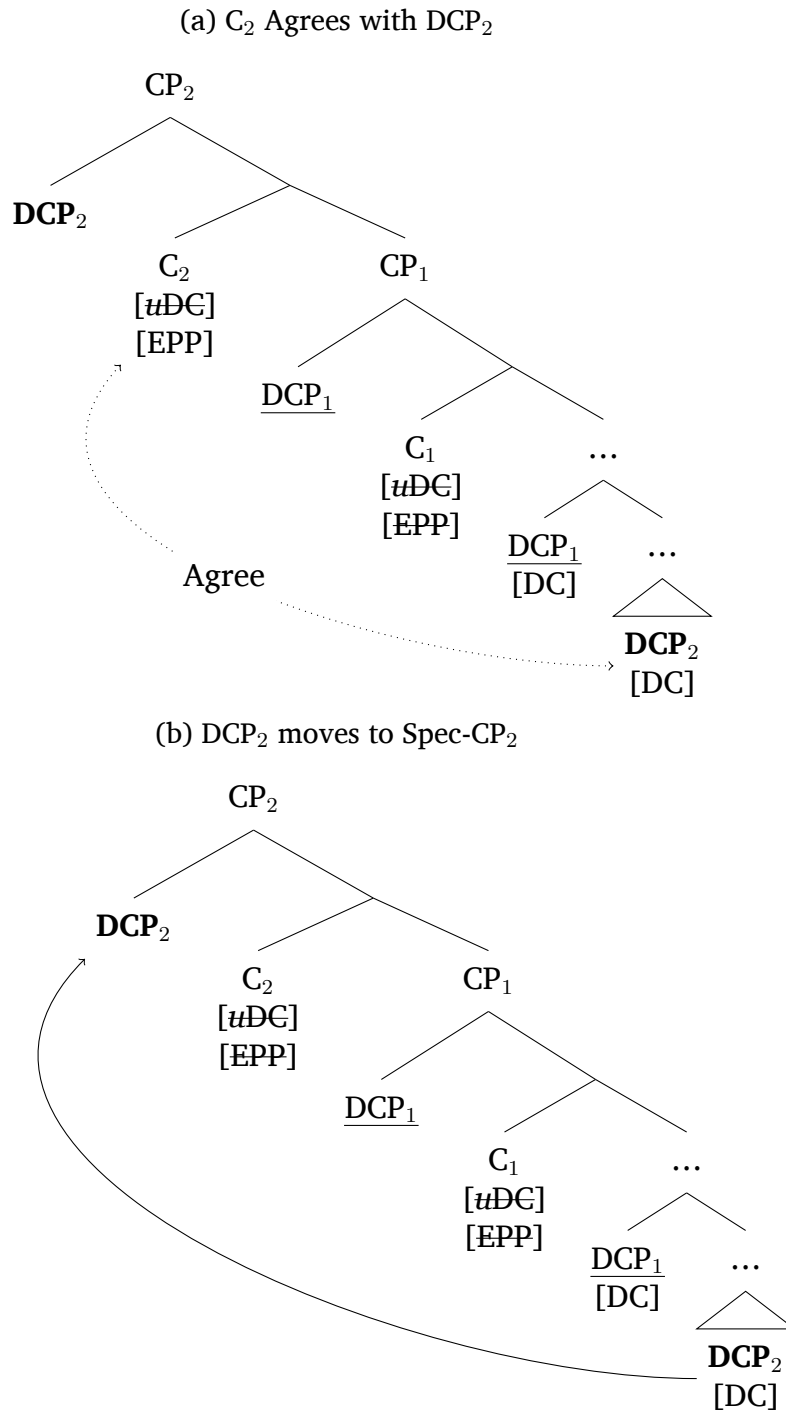
A crucial explanandum in this analysis is why the probe in C_2 cannot agree with DCP_1. There are at least two motivations for this restriction, both independently proposed in syntax. First is criterial freezing. Rizzi (2010) and others have proposed that certain Ā-movement operations disallow phrases from participating in subsequent movement operations. Criterial freezing in particular has been proposed for topicalization and other left peripheral movement operations. Topicalization is a relevant comparison, as multiple phrases can be topicalized in many languages (e.g. Rizzi 1997). If Eastern Cham DC-movement also results in criterial freezing, then we would not predict a DC-moved DCP to be movable beyond Spec-CP. This is depicted in Figure 2.7a.

Figure 2.7: Projection of CP_1

(a) Criterial freezing in Spec-CP_{DC} (b) Anti-locality
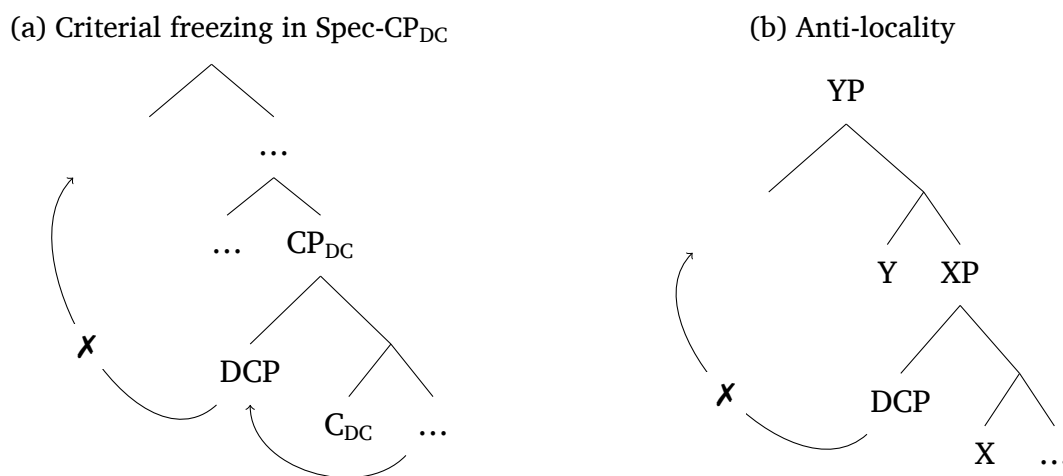
Second is specifier-to-specifier anti-locality, which has been proposed to be a general constraint on syntactic movement (e.g. Erlewine 2016). According to specifier-to-specifier anti-locality, a phrase in the specifier of an XP must cross at least one other phrasal projection (i.e. not just XP) if moved. Movement from Spec-XP to the specifier of the immediately dominating YP, as in Figure 2.7b, is impossible. Other restrictions on short movement have been proposed, but they do not capture the present effect. For example, Abels (2003) argues for a ban on movement from the complement of XP to the specifier of that XP. Grohmann (2003) argues that anti-locality bans movement within a domain such as *v*P. The Eastern Cham effect in the figure above involves movement from a specifier position in one domain (TP) to a specifier position in a different domain (CP). Both criterial freezing and specifier-to-specifier anti-locality predict that the phrase in Spec-CP_1 cannot move to Spec-CP_2 in the path containment derivation above. Data from long distance DC-movement will suggest that both of these motivations are independently needed in Eastern Cham (Section 2.3.3).

Another explanandum is why DCP_1 does not intervene on the Agree relation between C_2 and DCP_2. Criterial freezing holds that a criterial feature can only be checked once

in a derivation. It will be seen throughout this dissertation that the DC-feature can only be checked once in a derivation for a specific phrase (Section 2.3.3 on long-distance DC-movement, Chapter 4, Sections 4.2.3 and 4.3.3 on the interaction between inventory forms and partitives, respectively, and the left periphery).

To summarize the path containment effect data, a derivation with crossed paths can never occur, because it would have to violate a more general syntactic constraint like locality or criterial freezing. On its own, this path containment effect does not diagnose Ā-movement; the same effect could be derived through multiple syntactic operator and variable paths, as suggested for base generation + Agree in the next section. But, given the island and weak crossover data, it is clear that DC-movement is an instance of Ā-movement, just like *wh*-movement in English, as summarized in Table 2.2, modified from above.

Table 2.2: Movement diagnostics and DC-movement

	Unbounded	Island sensitivity	Weak crossover	Locality w/ DC-moved phrase
Eastern Cham DC-movement	✓	✓	✓	✓
English *wh*-movement	✓	✓	✓	✓

Finally, it is worth briefly noting other movement diagnostics not used in this section. The absence of agreement morphology in Eastern Cham renders a variety of movement diagnostics inapplicable (e.g. Adger & Ramchand 2005 on identity effects). Idioms are also frequently cited as diagnostics distinguishing between Ā-movement and base generation; specifically, the idiomatic meaning is preserved under movement, but not under base generation (e.g. Chomsky 1993). However, idioms that can be used to test movement in Eastern Cham have not yet been identified.[14] Lastly, anti-pronominal contexts could be used to test whether the gap in the base position of the DC-marked phrase is a true pronoun. If the phrase is DC-moved, the gap should not be a pronoun, but if it is base generated, there would be a pronoun (cf. Postal 1994). However, a reliable anti-pronominal context in Eastern Cham that could test this has not yet been identified, in part due to the requirement that personal pronouns be animate.

2.3.2 Base generation + Agree

This section turns to the resumptive pronoun strategy for DC-marking. This section finds that this strategy involves base generation of the DC-phrase in Spec-CP and an Agree

---

[14]Though, cf. Baclawski Jr (2016) for the possible idiom *naw çwă? ḳlaj* 'go to the restroom (LIT.: go step [into] the forest'.

relation between C and the resumptive pronoun in the base position, again, using evidence from syntactic islands, weak crossover, and locality effects. In the remainder of this section, the term 'base generation + Agree' will be used to refer to this DC-marking strategy.

Like DC-movement, base generation + Agree is sensitive to strong syntactic islands. The presence of a pronoun in the base position does not alleviate adjunct clause or subject islands (2.54a–b). For some consultants, a pronoun inconsistently alleviates the object relative clause island (2.54c).

(2.54) a. *__jaŋ__   __năn__$_i$ kăw tɨ́k   m̥ə̆ŋ [jwa   hɨ   ʔa
           CLF.PERSON that 1SG be.angry(VN) very   because 2SG invite
        ɲu$_i$ maj   ɲum ɓiə³³   ni   ]
        3.ANIM come drink beer(VN) this
        INTENDED: 'I am very angry because you invited that person to come drink this beer.'   (Adjunct clause)

     b. *__mɛʔ__ __kăw__$_i$ [ɗo²¹   ɓḁ̆ŋ po   ɲu$_i$ ŋă̆ʔ   ] t̥ɔʔ çɛʔ
         mother 1SG   stuff(VN) eat COMP 3.ANIM make   COP near
        pă̆ʔ t̥ɛh
        there
        INTENDED: 'The food that my mother makes is over there.'   (Subject)

     c. ??__mɛʔ__ __kăw__$_i$ kăw plɛ̆j [ɗo²¹   ɓḁ̆ŋ po   ɲu$_i$ ŋă̆ʔ   ]
         mother 1SG   1SG buy   stuff(VN) eat COMP 3.ANIM make
        INTENDED: 'I buy the food that my mother makes.'   (Relative clause)

Eastern Cham, thus, patterns with Vata (Kru: Côte d'Ivoire; cf. Koopman & Sportiche 1986) and other languages in that resumptive pronouns do not alleviate island effects. This would make sense in at least two scenarios. First, there could be syntactic movement in (2.54a–c) that leaves a pronoun in the base position. Second, the pronoun in the base position could independently enter into an Agree relation with C (cf. Adger & Ramchand 2005 on Ā-constraints on Agree relations with in situ phrases). Evidence from weak crossover favors the latter analysis.

Unlike with islands, weak crossover effects are alleviated if there is a resumptive pronoun in the base position of DC-marking. (2.55a–b) repeats the weak crossover context from the previous section, adding pronouns in each respective base position. Note that the preposition *ka* 'to' is restored in (a), as the preposition is not dropped when it has overt material in its complement position. When the pronoun is added to the base position in (a), the possessor pronoun in *bɔp*⁴⁵ *ɲu* 'his wallet' is free to corefer to *pu* 'Phú'.[15,16]

---

[15]A pronoun in the base position also alleviates the weak crossover in (ic) reported in footnote 13.

[16]Another grammatical variant of (2.55a) replaces the overt possessor pronoun with a null pronoun.

(2.55)   a. **tʰa sĭt pu**$_i$ **mĭn**   kăw m̥jan l̥ăjʔ bɔp^{45} ɲu$_{i/j}$ ka ɲu$_i$
          only   Phú EMPH 1SG return   wallet(VN) 3.ANIM to 3.ANIM

      'I only returned Phú his wallet.'

    b. **tʰa sĭt mjaw**$_i$ **mĭn**   kăw m̥jan l̥ăjʔ ɲu$_i$ ka po   ɲu$_{i/j}$
          only   cat   EMPH 1SG return   3.ANIM to owner 3.ANIM

      'I only returned the cat to its owner.'

If Ā-movement is the relevant trigger for weak crossover effects, not Agree, then the absence of weak crossover effects in (2.55a) is predicted under a base generation + Agree account. In (2.55a), there is no Ā-movement, and the pronoun in the base position is allowed to be Ā-bound because of an Agree relation with C. The pattern in Eastern Cham is also seen in other languages like Irish (McCloskey 2011: 110): weak crossover effects with Ā-movement, but no effects with base generation + Agree.

Locality effects further evidence the syntactic distinction between DC-movement and base generation + Agree. Recall that path containment effects arise when multiple phrases are DC-moved. At first glance, the same holds for base generation. When multiple phrases in the left periphery are coindexed with corresponding pronouns in their base positions, the same path containment effect obtains: nested paths lead to grammaticality (2.56a), and crossed paths to ungrammaticality (2.56b).

(2.56)   a. **jŭt**   **ni**   nĭʔ   mĕj   sĭt   năn tʰuːŋ͡m^{312} ʔa   ɲu   maj
          friend this child female small that Thuận   invite 3.ANIM come

      k̥ɔʔ **ɲu**
          meet 3.ANIM

      'This friend, Thuận invited that little girl to come meet.'

    b. *nĭʔ   mĕj   sĭt   năn **jŭt**   **ni** tʰuːŋ͡m^{312} ʔa   ɲu   maj
          child female small that friend this Thuận   invite 3.ANIM come

      k̥ɔʔ **ɲu**
          meet 3.ANIM

      INTENDED: 'This friend, Thuận invited that little girl to come meet.'

However, a different effect arises if there is a mix of DC-movement and base generation + Agree. If only one of the phrases in the left periphery is DC-moved and the other is coindexed with a pronoun, the result is grammatical, even if the paths are crossed. In (2.57a), it is *nĭʔ mĕj sĭt năn* 'that little girl' that has a pronoun in its base position, and in (2.57b), it is *jŭt ni* 'this friend' that does. It is worth noting that the base position pronouns seem to serve a purely syntactic role. They do not perform a disambiguating function, as the third person pronoun could in theory refer to either phrase in the left periphery.

(2.57) a. <u>nĭ̠ʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> **jŭt** **ni** tʰuːŋ͡m³¹² ʔa ɲu maj
     child female small that friend this Thuận invite 3.ANIM come

    ḳɔʔ ~~**jŭt ni**~~
    meet

    'This friend, Thuận invited that little girl to come meet.'

    b. <u>nĭ̠ʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> **jŭt** **ni** tʰuːŋ͡m³¹² ʔa ~~nĭ̠ʔ mĕj sĭt năn~~
      child female small that friend this Thuận invite

     maj ḳɔʔ **ɲu**
     come meet 3.ANIM

     'This friend, Thuận invited that little girl to come meet.'

These facts follow if base generation + Agree involves some Agree relation with C, but crucially not the same Agree relation as DC-movement. Figure 2.8 presents two possible analyses of this pattern, using (2.57a) as a test case. In Figure 2.8a, there is a DC-movement chain and a base generation + Agree chain, in line with the DC-marking mechanisms proposed in this chapter. Here, the pronoun in the base position corefers with the phrase in Spec-CP_2 by means of an Op-feature (cf. also Adger & Ramchand's (2005) ID-feature). The pronoun bears an Op-feature, which enters an Agree relation with the Op-probe on C_2. Then, the pronoun becomes coindexed with the DCP in Spec-CP_2 thanks to a predicate abstraction operation typically associated with Op-features.

Figure 2.8: Mixed DC-movement and base generation + Agree for (2.57a)

(a) Base generation + Agree
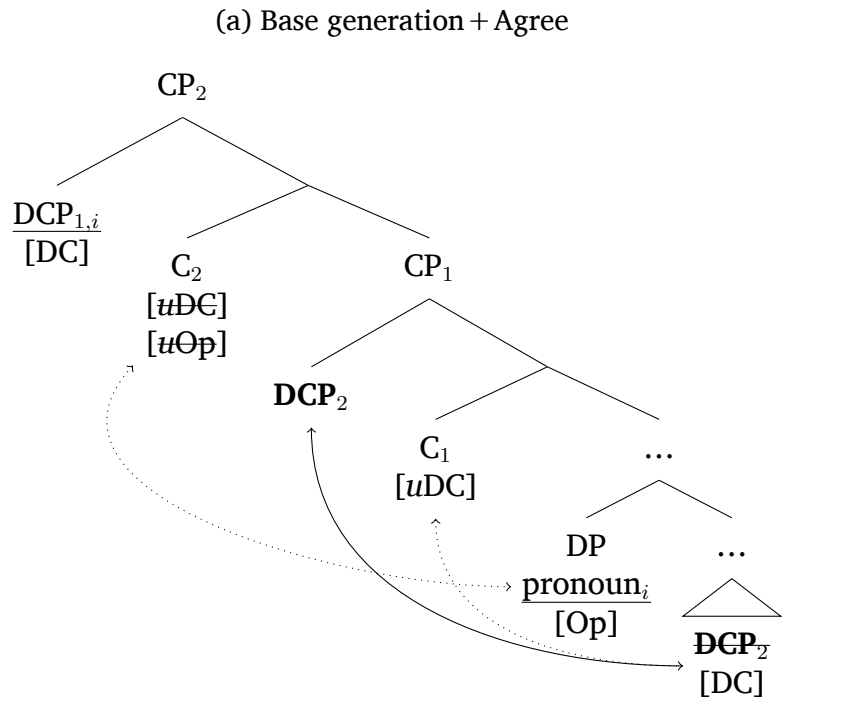
(b) Movement driven by some feature (F)
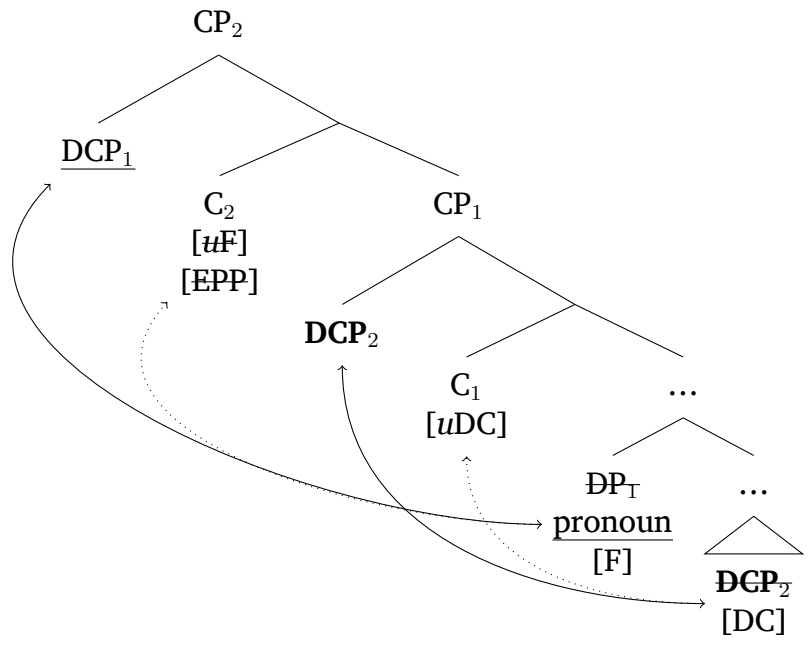
In Figure 2.8b, there is a DC-movement chain and a chain involving a hypothetical other feature F. In this case, the pronoun in the base position spells out a lower copy of the moved phrase. Both of these analyses account for the island and locality effect data presented in this section. However, the movement analysis in Figure 2.8b fails to predict weak crossover effects to be alleviated, as Ā-movement should give rise to weak crossover effects. Additionally, it would have to be coincidental that movement chains driven by the feature F always have the pragmatic interpretation of DC (cf. Section 2.1).

According to a base generation + Agree account, in situ pronouns are sensitive to islands and locality, because they enter Agree relations with C. Weak crossover is correctly predicted not to arise, as there is no Ā-movement. Phrases base generated in Spec-CP are also correctly predicted to bear DC-features, as the base generation + Agree C-head also bears a DC specification. Given that the syntactic and pragmatic facts fall out from a base generation + Agree account, the analysis in Figure 2.8a is favored here.

2.3.3   Long distance DC-movement

This section presents a complication to DC-movement involving long distance movement. This complication can be accounted for with the ingredients of DC-movement independently needed above: locality, criterial freezing, and anti-locality. Additionally, a new ingredient in the analysis of DC is needed: a different kind of syntactic probe mediating long distance movement.

The basic complication is that Eastern Cham path containment effects do not arise with DC-movement to different peripheries. In English, however, these are the circumstances under which path containment effects do arise. Consider (2.58a–b): the path of the *wh*-movement of *who*, underlined, is nested inside that of the topicalization of *this problem* in (a). When the movement paths are crossed, as in (2.58), the resulting sentence is judged ungrammatical. In both cases, topics are moved to the matrix left periphery, while *wh*-phrases are moved to the left periphery of the embedded CP.

(2.58)   a. **This problem**, Mary knows <u>who</u> to consult ~~who~~ about ~~**this problem**~~.

   b. ****This specialist**, Mary knows <u>what problems</u> to consult ~~**this specialist**~~ about ~~what problems~~.   (Pesetsky 1982: 268)

In the previous sections, path containment effects arose in Eastern Cham when multiple phrases were moved to the matrix left periphery. Path containment effects disappear when one is moved to the matrix and one to the embedded periphery (2.59). The addition of matrix material in (2.59b) is sufficient to alleviate the crossed paths (compare this example with 2.53b above). Eastern Cham, thus, does not show the kind of path containment effects that English shows.

(2.59) a. **han ni** kăw hnĭŋ <u>nĭʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> tʰuːɲ͡m³¹² ʔa
cake this 1SG think child female small that Thuận invite
~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
come eat
'This cake, I think Thuận invited that little girl to come eat.'

b. <u>nĭʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> kăw hnĭŋ **han ni** tʰuːɲ͡m³¹² ʔa
child female small that 1SG think cake this Thuận invite
~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
come eat
'This cake, I think Thuận invited that little girl to come eat.'

The reason for this pattern comes from the characteristics of long distance DC-movement. First, it is worth noting that embedded DC-movement does exist in Eastern Cham and has the same characteristics as matrix DC-movement. As seen in (2.60a–b), the same path containment effect obtains if two DPs are DC-moved to the left periphery embedded under *kăw hnĭŋ* 'I think'. This implies that the same DC-probe can appear on embedded C-heads.

(2.60) a. kăw hnĭŋ **han ni** <u>nĭʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> tʰuːɲ͡m³¹² ʔa
1SG think cake this child female small that Thuận invite
~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
come eat
'This cake, I think Thuận invited that little girl to come eat.'

b. *kăw hnĭŋ <u>nĭʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> **han ni** tʰuːɲ͡m³¹² ʔa
1SG think child female small that cake this Thuận invite
~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
come eat
INTENDED: 'This cake, I think Thuận invited that little girl to come eat.'

A different pattern arises with long distance movement. Multiple DPs cannot be DC-moved long distance to the matrix periphery of (2.61), regardless of their relative order. Path containment effects do not arise, because all relative orders are ungrammatical. Why should long distance movement have different path containment properties than short movement? The answer lies in which syntactic heads allow intermediate movement.

(2.61) a. ***han ni** <u>nĭʔ</u> <u>mĕj</u> <u>sĭt</u> <u>năn</u> kăw hnĭŋ tʰuːɲ͡m³¹² ʔa
cake this child female small that 1SG think Thuận invite
~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
come eat
INTENDED: 'This cake, I think Thuận invited that little girl to come eat.'

b. *nĭʔ   mĕj   sĭt   năn **han ni** kăw hnĭŋ tʰuːŋ͡m³¹² ʔa
     child female small that cake this 1SG think Thuận   invite
   ~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
                       come eat

INTENDED: 'This cake, I think Thuận invited that little girl to come eat.'

It has been proposed for a variety of languages that long distance Ā-movement is mediated by movement to each intermediate CP. For languages like Irish and Dinka, that intermediate movement is overtly reflected by forms of complementizers (e.g. McCloskey 2001), empty specifier positions (e.g. van Urk & Richards 2015), or resumptive pronouns (e.g. van Urk 2018). For Eastern Cham, there is no overt manifestation of these intermediate positions, yet it is apparent that long distance movement does not pattern like regular (short distance) DC-movement, based on the distributional evidence above.

These distributional facts can be explained if DC-movement is preceded by intermediate movement driven by a different syntactic head, $C_{embedded}$. Unlike C_{DC}, $C_{embedded}$ cannot iterate; i.e. there can only be one in a given periphery. $C_{embedded}$ only attracts DC-phrases, but it does not check DC-features in a criterial freezing sense, as DC-phrases always proceed to C_{DC}. For these reasons, $C_{embedded}$ will be assumed to have a general EPP-feature.

Table 2.3: Characteristics of C_{DC} and $C_{embedded}$ in Eastern Cham

	Probe	Iterate?	Final landing site of DC-movement?
C_{DC}	uDC	✓	✓
$C_{embedded}$	EPP	✗	✗

Long distance DC-movement, then, proceeds through as many embedded CPs as necessary before reaching its final landing site, a CP_{DC}. This is schematized in (2.62).

(2.62) Long distance DC-movement: C_{DC} …$C_{embedded}$ …$C_{embedded}$ …

This view of long distance DC-movement allows for an explanation of the facts in this section, requiring only the independently needed facts about DC-movement proposed in the sections above. First, multiple DPs cannot be moved long distance, because $C_{embedded}$ does not iterate. It only allows one DP to be moved long distance through a particular periphery. Second, $C_{embedded}$ allows path containment to be violated in the particular case of (2.63), repeated below. Here, one DP is moved long distance, and one short.

(2.63) nĭʔ   mĕj   sĭt   năn kăw hnĭŋ **han ni** tʰuːŋ͡m³¹² ʔa
     child female small that 1SG think cake this Thuận   invite
   ~~nĭʔ mĕj sĭt năn~~ maj ɓăŋ **han ni**
                       come eat

'This cake, I think Thuận invited that little girl to come eat.'

The derivation for this sentence can be modeled as follows. In the embedded clause, two CPs are projected: one $CP_{embedded}$ and one CP_{DC} (Figure 2.9a). The former attracts the closest DP, DP_1. The latter probes for the closest eligible DP bearing a DC-feature. Here, DP_1 cannot move to Spec-CP_2 because that movement would violate anti-locality (Section 2.3.1). Thus, it is DP_2 that agrees with C_2 and moves. Because CP_2 has a DC-probe, this is the final landing site for DP_2.

Figure 2.9: Mixed long and short DC-movement (2.63)

(a) Projection of embedded CP_{1-2}

CP_2

DCP$_2$
[DC]

C_2
[*u*DC]
[EPP]

CP_1

$\underline{DCP_1}$
[DC]

C_1
[EPP]

...

$\underline{DCP_1}$
[DC]

...

DCP$_2$
[DC]
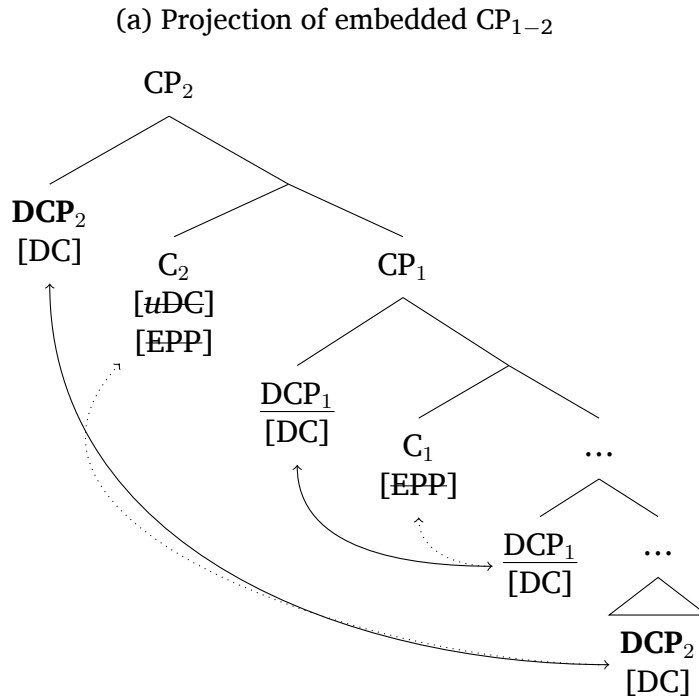
(b) Projection of matrix CP_3

CP_3

$\underline{DCP_1}$
[DC]

C_3
[*u*DC]
[EPP]

...

...

CP_2

DCP$_2$
[DC]

C_2

CP_1

$\underline{DCP_1}$
[DC]

C_1 ...
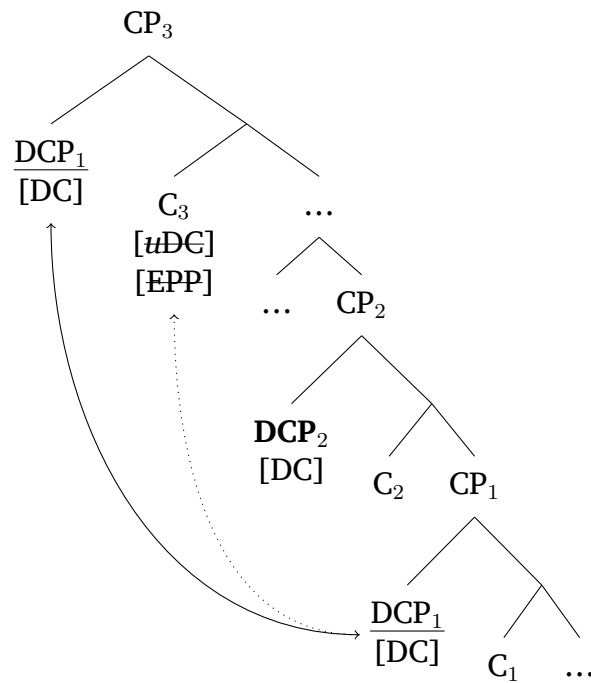
Turning to the matrix periphery, shown in Figure 2.9b, one CP is projected, CP_3. This CP now has a DC-probe. It probes for the closest eligible DP. DP_2 cannot move because of criterial freezing: it is frozen in Spec-CP_2. DP_1 is not frozen, however; it is free to agree with C_3 and move to Spec-CP_3. This derivation results in crossed paths enabled by $C_{embedded}$, criterial freezing, and anti-locality.

If the derivation is changed such that C_1 bears a DC-probe and C_2 bears just an EPP-feature, then it is DCP_2 that is predicted to be DC-moved to Spec-CP_3. Such a derivation, where DCP_2 is DC-moved to the matrix CP periphery and DCP_1 to the embedded CP periphery, is correctly predicted to be licit.

If there were only DC-probes, it is unclear what would allow crossed paths. Conversely, if there were only EPP-features, the same problem would arise. Long distance movement provides evidence for the need of both kinds of probes in Eastern Cham. In turn, this provides support for the idea that DC-movement involves Agree-driven movement, not base generation or EPP-driven movement: the latter two are independently needed in Eastern Cham, but they have different syntactic properties than DC-movement.

2.4 DC-movement is obligatory

Two means of DC-marking in Eastern Cham have been put forth: DC-movement and base generation + Agree. Both involve DC-phrases, but differ in the syntactic probe merged on C. To conclude this chapter, I turn back to the apparent optionality of DC-movement. As described in the introduction, the movement operation in (2.64a) appears to be optional, because in all derivations where a moved phrase is felicitous, so too is a corresponding in situ phrase. In other words, there do not appear to be any contexts where the moved phrase is licit, but not the in situ phrase. This kind of optionality has been argued to be a point in favor of a non-syntactic approach, as the movement operation itself appears to be optional (e.g. Erteschik-Shir 2007 on topicalization; Chapter 1, Section 1.3).

(2.64)   a. {ʔɔ̆ʔ   **ni**$_{DC}$} kăw ʈɔʔ   ɓăŋ {ʔɔ̆ʔ ni$_{DC?}$}
            mango this   1SG PROG eat
         'This mango, I am eating.'

      b. Derivation of moved phrase:
         [$_{DCP}$ DC [$_{DP}$ ʔɔ̆ʔ ni]] C_{uDC} kăw ʈɔʔ ɓăŋ ~~DCP~~

      c. Derivation of in situ phrase #1 (no DC-particle or probe):
         kăw ʈɔʔ ɓăŋ [$_{DP}$ ʔɔ̆ʔ ni]

      d. Derivation of in situ phrase #2 (covert movement):
         ~~DCP~~ C_{uDC} kăw ʈɔʔ ɓăŋ [$_{DCP}$ DC [$_{DP}$ ʔɔ̆ʔ ni]]

      e. Derivation of in situ phrase #3 (null index):
         [$_{DCP}$ DC [$_{DP}$ ∅]] C_{uDC} kăw ʈɔʔ ɓăŋ [$_{DP}$ ʔɔ̆ʔ ni]

There are a variety of possible analyses that are consistent with DC-marking so far, in which the movement operation is not optional. First, perhaps the DC-particle and corresponding CP are not merged at all (2.64c), as has been proposed for optional *wh*-movement (Denham 2000). The phrase *ʔɔ̆ʔ ni* 'this mango' stays in situ, as it is not marked by any Ā-feature.[17] This would be possible in Eastern Cham, as the DC-particle has no pronunciation, so a DC-marked phrase is pronounced identically to a non-DC-marked phrase. However, an analysis lacking a DC-particle entirely leads to a violation of Maximize Presupposition, as a derivation is chosen without the presupposition introduced by the DC-particle (see Section 2.1.2). If Maximize Presupposition is not active in this case, then an analysis lacking the DC-particle entirely is tenable.

Second, perhaps in situ DC-marked phrases undergo covert movement to Spec-CP (2.64d). The in situ DC-marked phrase would move to Spec-CP, but be pronounced in its base position. If a language is found with an overt DC-particle, a covert movement analysis would predict the DC-particle is pronounced in the base position along with the DC-marked phrase.

Third, perhaps a null index marked by a DC-particle is merged in Spec-CP (2.64e). In this case, the in situ phrase only accidentally corefers with the index. It is not the in situ phrase that is DC-marked at all. If a language is found with an overt DC-particle, the null index analysis predicts that the DC-particle can be merged in Spec-CP with a null overt or index, which corefers with an in situ phrase.

This data so far favor the second and third approaches (2.64d–e). A potential argument against the covert movement analysis, or any analysis where a DC-particle is merged low on in situ phrases, is that DC-marking is possible inside syntactic islands. For instance, (2.65b) elaborates upon or explains (2.65a), though the putative DC-marked phrase *mɛʔ kăw* 'my mother' is inside a relative clause island. If the phrase *mɛʔ kăw* 'my mother' is DC-marked, how can it enter an Agree relation with C from inside a syntactic island?

(2.65) a. kăw çuwʔ **mɛʔ** **kăw** hjĕj ni
  1SG help mother 1SG day this
  'I helped my mother today.'

  b. kăw plĕj [kan po **mɛʔ** **kăw** t̥ɔʔ ŋă̊ʔ hwă̊ʔ năn]
   1SG buy fish COMP mother 1SG PROG make eat that
   'I bought the fish that my mother is cooking.'   (a ⇓ b)

  b'. [kan po **mɛʔ** **kăw** t̥ɔʔ ŋă̊ʔ hwă̊ʔ năn] kăw plĕj
    fish COMP mother 1SG PROG make eat that 1SG buy
    ~~kan po mɛʔ kăw t̥ɔʔ ŋă̊ʔ hwă̊ʔ năn~~

    'I bought the fish that my mother is cooking.'   (a ⇓ b')

---

[17]Alternately, the DC-particle could be merged with the in situ DP, while the DC-probe is not merged on C. The lack of a corresponding DC-probe, however, would weaken the connection between DC pragmatics and DC-movement, and it would raise the issue of why DC-movement occurs at all.

It turns out that a syntactic island itself can be pied-piped to Spec-CP, as in (2.65b′). Chapter 3, Section 3.3.1 argues that the DC-feature percolates in this case to the DP headed by *kan* 'fish', allowing the DP containing the relative clause to be DC-moved. Hence, syntactic islands do not present barriers for in situ phrases to be DC-marked at a certain level.

A potential argument against the null index analysis is that pronouns in general cannot be DC-moved (Section 2.1.3). However, Chapter 5, Section 5.1 finds that certain pronouns can be DC-moved in Catalan. Additionally, Section 2.1.3 hypothesizes that the ban on overtly DC-marking pronouns in Eastern Cham reflects either a prosodic or discourse-level fact about the third person animate pronoun in Eastern Cham. Null pronouns could have different properties; there is no direct evidence yet that they cannot be merged with a DC-particle and merged in Spec-CP.

All in all, an analysis can be maintained in which Eastern Cham DC-marked phrases must be moved or base generated in Spec-CP. In other words, the DC-particle necessarily corresponds with a DC-probe on C. Chapter 4 will find that there is an additional DC-probe on D that can also license DC-marking.

It is worth asking why overt DC-marking alternates with covert DC-marking. Covert DC-marking presumably gives rise to an increase in ambiguity, as a derivation with an in situ DC-marked phrase is string-identical to a derivation with no DC-marking at all. An increase in ambiguity due to a lack of overt information is seen elsewhere in the realm of discourse markers. For instance, there is a range of sentence connectives that indicate certain discourse relations, such as *That's because* (e.g. Grosz & Sidner 1986 on 'cue phrases'; Taboada & Das 2013). The connective in (2.66b) indicates the existence of an explanation relation. However, its absence does not prevent the same explanation relation from being parsed (2.66b′). Increased ambiguity may result. For example, an elaboration relation may be posited, such that the speaker is not expressing evidence for why they are cooking chicken; they are only expanding upon their relation to it. Nevertheless, the intended discourse remains interpretable. Perhaps the optionaltiy of connectives like *That's because* can inform the analysis of the apparent optionality of DC-marking.

(2.66)   a. I'm cooking chicken for dinner.

      b. That's because I like chicken.

      b′. I like chicken.

Chapter 3 explores the interaction between DC and *wh* in more detail. DC-particles are found to be independent from Q-particles, as both can be merged with the same phrase. DC remains an Ā-feature, but one that marks a hierarchical discourse constraint along a different dimension of discourse tracking than *wh*.

Chapter 3

Wh-phrases can be DC-marked

This chapter argues that *wh*-phrases can be DC-marked either through DC-movement or base generation + Agree. Eastern Cham is a *wh*-in situ language, so *wh*-phrases are pronounced in their base position in the unmarked case (3.1a). Under certain circumstances, *wh*-phrases are pronounced at the left edge of a clause (3.1b). These cases are shown to be the result of *wh*-phrases being DC-marked. DC-marked *wh*-phrases, as in (3.1b), have the same pragmatic and syntactic properties as DC-phrases observed in the previous chapter.

(3.1) a. hɨ t̪ɔʔ ɓăŋ **k̥eʔ**
        2SG PROG eat what
        'What are you eating?'

   b. **k̥eʔ**$_{DC}$ hɨ t̪ɔʔ ɓăŋ k̥eʔ
        what 2SG PROG eat
        'What are you eating?'

This analysis must also address two puzzles that arise with the DC-movement of *wh*-phrases in a *wh*-in situ language. First, how can the semantics of *wh*-phrases interact with the semantics of DC? *Wh*-questions are analyzed as sets of possible answers, according to Hamblin question semantics (Hamblin 1973). Rooth (1992) posits that *wh*-phrases have no normal semantic value; they only have an interpretation along a focus dimension of meaning, where they function as sets of alternatives defined by the *wh*-phrase. *Wh*-phrases must combine with a focus operator introduced by C in order to be interpreted at all. As a consequence, it has been argued that *wh*-phrases cannot be marked for information statuses such as topic, as topicality requires a phrase to have a normal interpretation as some kind of referent (e.g. Comorovski 1996: 160; Cruschina 2012: 158).

This puzzle directly concerns DC-marking, as the DC-particle must combine with a phrase interpreted as an individual, not a set of individuals. For a *wh*-phrase to be marked as DC, it must be specified how DC-marking can coexist with the computation of alternative sets. In this chapter, it is proposed that DC-marking occurs in an embedded constituent of the *wh*-phrase, where the interpretation of the embedded NP or DP is of

type e, before the *wh*-D-head introduces the computation of alternatives. Then, the DC-feature percolates up to the highest DP, and the whole DP is moved as a result through pied-piping. This analysis provides a model for explaining how different information status features can be assigned to *wh*-phrases; the features enter the derivation below on a phrase embedded under the *wh*-D-head. In turn, it can be used to account for cross-linguistic data, which show that certain *wh*-phrases exhibit the syntactic properties of topicalization (cf. D-linking below; e.g. Polinsky 2001; Grewendorf 2012; Pan 2014).

Second, how can a *wh*-phrase optionally move? Cheng's (1991, 1997) Clausal Typing Hypothesis (CTH) predicts that there are no languages with both *wh*-movement and *wh*-in situ in the general case. According to the CTH, every clause must be typed either by a *wh*-particle on C (i.e. *wh*-in situ) or by movement of a *wh*-phrase to Spec-CP. In the Agree framework, the CTH has been adapted to posit that every language is predicted to have one Agree mechanism between C and *wh* in the general case (cf. Roussou & Vlachos 2011, and references therein). According to the CTH, *wh*-phrases should not optionally move.

Denham (1998, 2000) and others have proposed means of deriving optional *wh*-movement consistent with the Minimalist Program (cf. also Cheng & Rooryck 2000 and others on optional in-situness). As described in Chapter 2, Section 2.4, perhaps both the *wh*-feature and the C-head bearing the *wh*-probe are missing from the derivation. Alternately, the movement of *wh*-phrases may be the result of a different construction, such as concealed clefts or pseudoclefts (Cheng 1991 on Bahasa Indonesian; Paul 2001, Potsdam 2006 on Malagasy); or the movement may be triggered by an independent feature such as topic or focus (cf. Jensen 2013 on focus-movement in the Chamic language Jarai).

This chapter takes the latter approach: the movement of *wh*-phrases in Eastern Cham is due to an independent feature, DC. Evidence is presented that any feature appealing to *wh*, such as *wh* itself or a generalized Ā-feature (Aravind 2017, 2018), is insufficient in accounting for DC-movement. Fronted *wh*- and non-*wh*-phrases share pragmatic and syntactic properties, and they can syntactically intervene on one another.

A generally agreed-upon exception to the CTH is D-linking (Pesetsky 1987). D-linked *wh*-phrases are those with contextually salient sets, such as *which book*. They are known to exceptionally move in *wh*-in situ languages or stay in situ in *wh*-movement languages (e.g. Comorovski 1996; Pesetsky 2000), while the precise featural content of D-linking is yet unclear. This chapter further shows that the movement of *wh*-phrases in Eastern Cham cannot be ascribed to D-linking, as D-linking is a separate phenomenon in Eastern Cham. Much like DC-movement in the previous chapter, *wh*-phrases can be base generated in the left periphery with resumptive pronouns in their base position (3.2). The capacity to be resumed is shown to correlate with D-linking (in addition to DC-marking), as resumption and D-linking both require referential indices in a way DC does not.

(3.2)   a. **nujh**   **hlɛ̆j**$_{\text{DC},i}$ hɨ   ʔḁ   ɲu$_i$
          person   which   2SG   invite   3.ANIM
          'Which person did you invite?'

      b. [$_{\text{DP}}$ [$_{\text{DCP}}$ DC nujh] hlɛ̆j$_{\text{D-linked}}$]$_i$ C$_{\text{uOp}}$ hɨ ʔḁ [ɲu$_{\text{Op}}$]$_i$ (Base generation + Agree)

DC-movement, thus, is a unitary movement operation driven only by a DC-feature. This supports the CTH, in that there is only one Agree relation between C and *wh*, and it is not one that results in overt phrasal movement. The chapter proceeds as follows. Section 3.1 gives some relevant background on the DC-movement of *wh*-phrases, comparing it with Ā-movement operations in general. Section 3.2 extends my proposal of the pragmatics of *discourse connectedness* to *wh*-phrases. Additionally, Section 3.2.2 demonstrates that DC is separate from Discourse/D-linking. Section 3.3 establishes that the DC-movement of *wh-* and non-*wh*-phrases involve the same probe on C from a featural standpoint, based on distributional evidence and locality effects. Finally, Section 3.4 confirms that DC-movement is orthogonal to the *wh*-feature, based on the characteristics of in situ *wh*-phrases, and builds a complete account of how *wh*-phrases can be DC-moved.

3.1 Background and Ā-movement

Before proceeding, this section presents some background information on *wh*-phrases and Ā-movement. Eastern Cham has the basic characteristics of a *wh*-in situ language, and DC-movement of *wh*-phrases shares the Ā-movement characteristics with DC-movement describe in Chapter 2, Section 2.3.1. As with DC-movement in Chapter 2, Section 2.3.2, there is also a base generation option for *wh*-phrases with resumptive pronouns in the base position.

First, in an out-of-the-blue context, *wh*-phrases must remain in their base position (3.3a). Based just on this example, it is conceivable that the *wh*-phrase does move, but to a low position, such as the right edge of the *v*P. Such an analysis has been proposed for certain dialects of Spanish (Uribe-Etxebarria 2002) and Hindi-Urdu (Manetta 2006). Eastern Cham *wh*-phrases, however, are truly in situ, as illustrated in (3.3b). The *wh*-phrase is bounded on the right by the indirect object and the root modal *hu* (which denotes either ability or permission). Note that the modal *hu* is a phrase head that triggers predicate raising (cf. Baclawski Jr 2017), so the *wh*-phrase is not moving to the edge of or outside of the predicate.

(3.3)   CONTEXT: Out-of-the-blue.
   a. hɨ   t̥ɔʔ   ɓăŋ   k̥eʔ
      2SG PROG eat   what
      'What are you eating?'

   b. kăw [ₓ𝑝 plĕj **k̥eʔ**   ka nɨʔ   sɨt   năn] hu   v̶P̶
      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, *lĕj*, that is not found in *wh*-questions.

(3.4) hɨ   t̥ɔʔ   ɓăŋ ʋɤ̌ʔ lɛ̌j
     2SG PROG eat ITER Y/N.Q
     'Are you eating more/again?'

When *wh*-phrases are fronted to the left periphery, they are moved via Ā-movement. The previous chapter laid out a series of general Ā-movement characteristics that group DC-movement, relativization, and clefting: (i) marking by the complementizer *po*; (ii) preposition/p-drop; (iii) island sensitivity; and (iv) weak crossover effects. The DC-movement of *wh*-phrases exhibits all of these properties. First, DC-moved *wh*-phrases are optionally marked by the complementizer *po* (3.5). This is commensurate with movement to some clausal left periphery such as Spec-CP.

(3.5)   a. **tʰɛ̌j**$_{DC}$ po   hɨ   ʔa   t̶ʰ̶ɛ̶̌j̶
         who COMP 2SG invite
         'Who did you invite?'   (DC-movement)

       b. **pu** po   kăw ʔa   p̶u̶ t̥ɔʔ păʔ t̥eh
         Phú COMP 1SG invite   COP there
         'Phú, who I invited, is over there.'   (Relative clause)

Second, prepositions are dropped when argument preposition phrases are Ā-moved (cf. 'p-drop': Hoonchamlong 1991 on Thai; Wang 2007 on Mandarin Chinese; Sato 2011 on Indonesian). For example, in (3.6a), the preposition *ka* is obligatory with in situ indirect objects in ditransitive constructions. When these phrases are Ā-moved, however, *ka* cannot appear in any position (3.6b–c). I follow previous analyses of p-drop in positing that there is a pronunciation rule that deletes prepositions whose complements have been Ā-moved. P-drop also has the effect of restricting Ā-movement to DPs on the surface; adjuncts cannot be Ā-moved, as described below for hanging topics.

(3.6)   a. hɨ   plɛ̌j han năn *(ka) **tʰɛ̌j**
         2SG give cake that   to   who
         'Who did you give that cake to?'

       b. **tʰɛ̌j**$_{DC}$ hɨ   plɛ̌j han năn (*ka) t̶ʰ̶ɛ̶̌j̶
         who 2SG give cake that   to
         'Who did you give that cake to?'   (DC-movement)

       c. hu **tʰa nɨ̌ʔ mɛ̌j**   kăw plɛ̌j han ni   (*ka) t̶ʰ̶a̶ ̶n̶ɨ̶̌ʔ̶ ̶m̶ɛ̶̌j̶
         ∃ one child female 1SG give cake this   to
         'There is a girl who I [will] give this cake to.'   (Cleft)

Third, Ā-movement is consistently sensitive to syntactic islands, such as object relative clauses (3.7a–b), in line with Ā-movement cross-linguistically or minimally an Agree relation involving an Ā-probe (cf. Adger & Ramchand 2005). 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 I will set them aside and focus on the majority pattern reported here. Note also that there is nothing ungrammatical about the complex DP in the absence of movement (3.7c).

(3.7) a. ****jaŋ** **hlĕj** hɨ plĕj ɗo²¹ ɓăŋ po ~~jaŋ hlĕj~~ ŋă?
CLF.PERSON which 2SG b̥uy stuff(VN) eat C̥OMP make

INTENDED: 'Which person do you buy the food [they] make?' (DC-movement)

b. **hu **tʰa jaŋ** kăw plĕj ɗo²¹ ɓăŋ po ~~tʰa jaŋ~~ ŋă?
∃ one CLF.PERSON 1SG b̥uy stuff(VN) eat C̥OMP make

INTENDED: 'There is a person who I buy the food [they] make.' (Cleft)

c. kăw plĕj ɗo²¹ ɓăŋ po mɛʔ kăw ŋă?
1SG b̥uy stuff(VN) eat C̥OMP mother 1SG make

'I buy the food that my mother makes.'

DC-movement is also sensitive to adjunct clause islands (3.8). DC-movement of the *wh*-phrase *jaŋ hlĕj* 'which person' is ungrammatical, as is a cleft (3.8a–b). Yet, the same context is grammatical in the absence of movement (3.8c).

(3.8) a. ****jaŋ** **hlĕj** hɨ tɨk³⁵ mɓ̥ăŋ jwa kăw ʔa
CLF.PERSON which 2SG be.angry(VN) very because 1SG invite
~~jaŋ hlĕj~~ maj ɲum ɓiə³³
come drink beer(VN)

INTENDED: 'Which person are you angry because I invited [them] to come drink beer?' (DC-movement)

b. **hu **tʰa jaŋ** kăw tɨk³⁵ mɓ̥ăŋ jwa kăw ʔa
∃ one CLF.PERSON 1SG be.angry(VN) very because 1SG invite
~~tʰa jaŋ~~ maj ɲum ɓiə³³
come drink beer(VN)

INTENDED: 'There is a person who I am angry because I invited [them] to come drink beer?' (Cleft)

c. kăw tɨk³⁵ mɓ̥ăŋ jwa hɨ ʔa ʔaj maj ɲum
1SG be.angry(VN) very because 2SG invite older.sibling come drink
ɓiə³³
beer(VN)

'I am very angry because you invited him[older sibling] to come drink beer.'

Fourth, DC-movement of *wh*-phrases gives rise to weak crossover effects, again in line with Ā-movement cross-linguistically, or minimally the absence of A-movement (e.g. Ruys 2004). Weak crossover occurs when a DP cannot be Ā-moved across 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 (3.9). The direct object can bind a referent within the indirect object, but not vice versa (3.9b), as established in Chapter 2, Section 2.3.1.

(3.9)   a. hɨ   m̥jan ļăjʔ [nɨʔ   hlɛ̌j]$_i$ ka po   ɲu$_i$
            2SG return   child which to owner 3.ANIM

           'Which young animal [LIT.: animal child] did you return to its owner?'

     b. *hɨ m̥jan ļăjʔ ɲu$_i$ ka po   [nɨʔ   hlɛ̌j]$_i$
            2SG return   3.ANIM to owner child which

           INTENDED: 'Which young animal did you return to its owner?'

If an indirect object is Ā-moved over a direct object, a crossover context arises. For example, the DP *jaŋ hlɛ̌j* 'which person' crosses over the direct object in (3.10a), 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 (3.10b), as *nɨʔ hlɛ̌j* 'only the cat' always c-commands the pronoun. No weak crossover effect obtains, as the *wh*-phrase never crosses over the pronoun; the pronoun may corefer with the Ā-moved DP.

(3.10)   a. **[jaŋ**   **hlɛ̌j**]$_{\text{DC},i}$ hɨ m̥jan ļăjʔ ɓɔp^{35} ɲu$_{*i/j}$ ~~ka~~ ~~jaŋ~~ ~~hlɛ̌j~~
            CLF.PERSON which   2SG return   wallet(VN) 3.ANIM to

           'Which person did you return their wallet to?'   (DC-movement)

     b. **[nɨʔ**   **hlɛ̌j**]$_{\text{DC},i}$ hɨ m̥jan ļăjʔ ~~nɨʔ~~ ~~hlɛ̌j~~ ka po   ɲu$_{i/j}$
            animal which   2SG return   to owner 3.ANIM

           'Which animal did you return to its owner?'   (DC-movement)

Thus, DC-movement of *wh*-phrases in Eastern Cham shares Ā-movement characteristics with relativization, clefting, and DC-movement in general. In the previous chapter, these movement operations contrasted with other information structural phenomena, such as hanging topics and contrastive topics. There does not appear to be a hanging topic correlate for *wh*-phrases in Eastern Cham, but they can function as contrastive topics.

    Finally, *wh*-phrases base generated in Spec-CP with resumptive pronouns share some of these Ā-feature characteristics. Base generated phrases can be marked by the complementizer, and they are sensitive to islands (3.11). This follows if resumptive pronouns are in some Agree relation with C.

(3.11)   a. **tʰɛ̆j**$_{DC,i}$ po   hɨ   ʔ̥a   ɲu$_i$
             who   COMP 2SG invite 3.ANIM

   'Who did you invite?'

   b. *__jaŋ__   **hlɛ̆j**$_i$ hɨ   tɨk^{35} m̥ɓ̥ăŋ jwa   kăw ʔa
       CLF.PERSON which 2SG be.angry(VN) very   because 1SG invite
       ɲu$_i$ maj   ɲum   ɓiə^{33}
       3.ANIM come drink beer(VN)

   INTENDED: 'Who/which person are you angry because I invited to come drink beer?'

Resumptive pronouns, however, alleviate p-drop and weak crossover effects (3.12). As for p-drop, the pronoun provides enough phonological material in the base position to license the pronunciation of the preposition *ka*. As for weak crossover, the pronoun allows for coindexation with the phrase in the left periphery. This suggests that the phrase in the left periphery is not derived via movement. Instead, the resumptive pronoun is a true, independent pronoun, bound by the phrase base generated above through an Agree relation with C.

(3.12)   a. **tʰɛ̆j**$_{DC,i}$ hɨ   plɛ̆j han   năn *(ka) ɲu$_i$
             who   2SG g̊ive cake that   to   3.ANIM

   'Who did you give that cake to?'

   b. [__jaŋ__   **hlɛ̆j**]$_{DC,i}$ hɨ   m̥jan ḷăjʔ ɓɔp^{35} ɲu$_{i/j}$ ka ɲu$_i$
       CLF.PERSON which   2SG return   wallet(VN) 3.ANIM to 3.ANIM

   'Which person did you return their wallet to?'

This section has presented evidence that DC-movement of *wh*-phrases in Eastern Cham is an instance of Ā-movement, alongside DC-movement in general. Much like DC-movement, there is also a base generation option for *wh*-phrases as well. The following section examines the pragmatics of *discourse connectedness* and concludes that it unifies *wh*- and non-*wh*-phrases.

3.2   Pragmatics of DC-marked *wh*-phrases

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. Recall that the DC-particle combines with an individual (3.13e), and it introduces a presupposition that requires that individual to be mentioned in a prior sentence. That prior sentence must introduce an event that the current event is interpreted as a cause or subtype of. This assumes that the discourse can be updated in such a way that prior events can be tracked, along with how they relate to each other (cf. Sections 2.1.1–2.1.2 above).

(3.13)   a. Let $\mathcal{E}_c$ be the set of events live in a discourse at context c

      b. Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

      c. Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

      d. Let $\mathcal{P}_e$ be the set of participants in event e

      e. $[\![\text{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

When a DC-particle combines with a DP, the presupposition above is checked. If the presupposition is satisfied, the DC-particle otherwise acts as an identity function. If the presupposition fails, the resulting sentence has no interpretation.

Wh-phrases introduce a complication to the semantics and pragmatics of DC. Regarding pragmatics, it has been argued that *wh*-phrases can only have the information status of focus (Cable 2008 and references therein). Other information statuses are claimed to be impossible, especially those with an anaphoric character like topic. The impossibility of topic-marking is argued to be the result of the inherent interrogativity or information-seeking status of *wh*-phrases. *Wh*-phrases would presumably also not be able to be DC-marked as well, as DC requires previous mention.

In a Hamblin semantics of questions, *wh*-questions are interpreted as sets of possible answers and *wh*-phrases as sets of alternatives of a certain type. It is unclear how the DC-particle can combine with a set of individuals, as laid out so far. This problem is more pronounced under Alternative Semantics, in which *wh*-phrases have no interpretation at all along the ordinary dimension of meaning (Rooth 1992; Kotek 2019), rendering the semantics of the DC-particle inapplicable to *wh*-phrases. For example, the interpretation of *who* is the set of people such that they answer the *wh*-question. It has no ordinary semantic interpretation ((3.14a); denoted by ^{o}); they are only interpretable along a focus dimension of meaning ((3.14b); denoted by ^{f}). An operator on C is needed to convert the interpretation to the ordinary dimension of meaning (3.14c).

(3.14)   a. $[\![\text{who}]\!]^o$ = undefined

      b. $[\![\text{who}]\!]^f$ = {Alice, Ben, Carla...}

      c. $[\![\text{who } \text{C}_Q \text{ arrived}]\!]^o$ = {Alice arrived, Ben arrived, Carla arrived...}

Section 3.2.1 extends DC to *wh*-phrases and proposes that DC is marked on an embedded constituent of DPs, before the *wh*-D-head introduces the computation of alternatives. Section 3.2.2 turns to D-linking. D-linked *wh*-phrases are exceptional in both pragmatics and syntax. Of all *wh*-phrases, they are the ones argued to be able to have topic status (cf. Pan 2014). They can also uniquely obviate *wh*-movement restrictions such as superiority effects (Pesetsky 1987). It follows that the pragmatic and syntactic effects seen with the DC-movement of *wh*-phrases in Eastern Cham could be due to D-linking. This section shows that this is not the case. Instead, *wh*-resumptive pronouns correlate with D-linking, implying that it is an anaphoric phenomenon, not sensitive to discourse structure.

3.2.1   DC pragmatics

Wh-phrases can be DC-marked if they meet the DC conditions. This introduces a problem, as outlined above, if *wh*-phrases are interpreted as sets of alternatives, not individuals or properties (Rooth 1992). This section proposes a model by which DC-marking accords with the nature of *wh*-phrases. Additionally, the data indicate that DC cannot be captured by standard accounts of topicalization, as a variety of indefinites and quantifiers can be DC-moved. Before proceeding, recall the definition of discourse subordination, repeated below. The notation (a $\Downarrow$ b) will be used to indicate that a sentence (a) is superordinate to a sentence (b) that explains or elaborates on it. The symbol $\nDownarrow$ will be used to indicate the absence of a subordinating discourse relation.

(3.15) Discourse subordination ($\Downarrow$): Sentence $\phi \Downarrow$ sentence ψ if ψ is interpreted as an elaboration or explanation of ϕ

First, (3.16) illustrates the DC conditions when applied to a *wh*-phrase. In (3.16b), the *wh*-phrase *ǩɔʔ ǩeʔ* 'what pot' refers to a set of pots that constitute the possible answers to the question, assuming a Hamblin semantics for questions (Hamblin 1973; Rooth 1992).[1] The alternative set is not itself previously mentioned in (3.16a). Instead, two individuals from that set are mentioned, a pot of frog and a pot of *kiép*, a different kind of frog. It is also the case that the kind of pots is mentioned in (3.16a). As with the previous section, DC-movement here is optional. In situ *wh*-phrases are felicitous in each of these contexts.

(3.16)   a. mɔŋ mɨ   kăw tŭʔ **ʔiŋ ʔɔŋ tʰa ǩɔʔ hŏ͡ŋm kĭwʔ tʰa ǩɔʔ**
            look father 1SG boil frog   one pot with   kiép   one   pot

         'Look at my father boil one pot of frog and one of kiép.'

      b. jăʔ ni **ǩɔʔ ǩeʔ**_{DC} ʔoŋ   năn t̨ɔʔ   ŋăʔ   ~~ǩɔʔ ǩeʔ~~ năn
            now   pot   what   old.man that PROG make   that

         'Now, what pot is that old man making [working on]?'   (a $\Downarrow$ b)

In terms of discourse structure, (3.16b) is naturally interpreted as an elaborating question on (3.16a). 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 *ǩɔʔ ǩeʔ* 'what pot' is accepted, because there is some previous mention of the restriction of the *wh*-word in a superordinate sentence (3.16a).

(3.17b) 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,

---

[1]Sentence-final *năn* in these examples marks clause-level deixis and does not form a consitutent with the *wh*-phrase.

DC-movement of *ƙɔʔ ƙeʔ* 'what pot' is rejected. The moved phrase in (3.17b′) lacks an appropriate antecedent entirely. Again, DC-movement is infelicitous.

(3.17) a. mɔŋ mɨ   kăw tŭʔ **ʔiŋ ʔɔŋ tʰa ƙɔʔ hɔ̌ŋ͡m kɨwʔ tʰa ƙɔʔ**
    look father 1SG boil frog one pot with kiép one pot

    'Look at my father boil one pot of frog and one of kiép.'

   b. #jăʔ ni **ƙɔʔ ƙeʔ** ʔoŋ   năn ʈɔʔ   ɓăŋ ~~ƙɔʔ ƙeʔ~~ năn
    now pot what old.man that PROG eat   that

    INTENDED: 'Now, what pot is that old man eating?'   (a ⇊ b′)

   b′. #jăʔ ni **ƙeʔ** mu?   năn ʈɔʔ ŋăʔ ~~ƙeʔ~~ năn
    now what old.woman that PROG make that

    INTENDED: 'Now, what is that old woman making [working on]?' (a ⇊ b′)

When presented with (3.17b–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 was prior discourse about multiple people and their cooking. In these cases, there is a sentence that mentions the restriction of the *wh*-phrase that is being elaborated upon, which is lacking in the explicit discourse in (3.17).

Descriptively, *wh*-phrases can be DC-moved if they satisfy the DC conditions: they must be previously mentioned in a superordinate sentence. To model previous mention more concretely, I turn to an expanded syntax and semantics of *wh*-phrases. Bare *wh*-phrases like *who* are considered to be specified for some property or kind (i.e. $^{\cap}human$ or a more restricted kind from context; cf. Section 2.1.3 on kinds above). The structure of *who* can then be split into a *wh*-determiner and an NP-restriction (Figure 3.1a). D-linked *wh*-phrases, i.e. those with contextually salient sets, of the form *which X* have been argued to contain an embedded DP that refers to a contextual antecedent, perhaps a plural antecedent (e.g. Boeckx & Grohmann 2004). In Figure 3.1b, the contextual antecedent is indicated by a hypothetical abstract referential index 7. In both cases, it is the D-head that contains the *wh*-element. If that D-head initiates the computation of alternatives, then NP and DP_1 are interpreted as a property or kind and an individual, respectively. Note that a Q-particle is adjoined to DP in these examples, as will be shown to be necessary for Eastern Cham (cf. Cable 2010).

Figure 3.1: Syntax and semantics of *wh*-phrases

(a) Bare *wh*-phrase   (b) D-linked *wh*-phrase

DP

Q   DP

D   NP

who $\cap human$

DP$_3$

Q   DP$_2$

D   DP$_1$

which   7
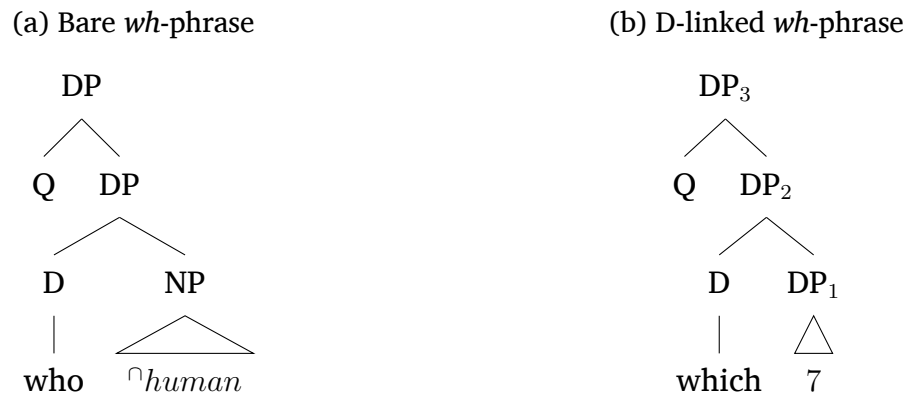
These *wh*-phrase structures present a way to mark DC without its ranging over sets of alternatives. In bare *wh*-phrases, DC can mark the NP restriction, checking if the kind referred to satisfies the DC conditions (Figure 3.2a). Note that the previous mention of a kind can be either mention the same kind or an example individual, as demonstrated in Chapter 2, Section 2.1.3. A DC-marked bare *wh*-phrase is analogous to *what kind X* in English such that *X* satisfies the DC conditions. Note also that the kind indicated in Figure 3.2a is $\cap human$; it is likely that DC-marked bare *wh*-phrases have more specific kind restrictions from context. In D-linked *wh*-phrases, DC can combine with the contextual antecedent DP (DP$_1$) checking if that individual satisfies the DC conditions (Figure 3.2b). Since D-linked *wh*-phrases have been argued to be like partitives, this is analogous to *which of the X,* such that X satisfies the DC conditions.

Figure 3.2: Syntax and semantics of DC-marked *wh*-phrases

(a) Bare *wh*-phrase

```
            DP
           /  \
          Q    DP
         / \
        D   DCP
        |   /  \
      who  DC   NP
           |     |
```
$\lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e$ $^{\cap}human$
$[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

(b) D-linked *wh*-phrase

```
            DP_3
           /   \
          Q    DP_2
         / \
        D   DCP
        |   /  \
    which  DC   DP_1
           |     |
```
$\lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e$ 7
$[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$
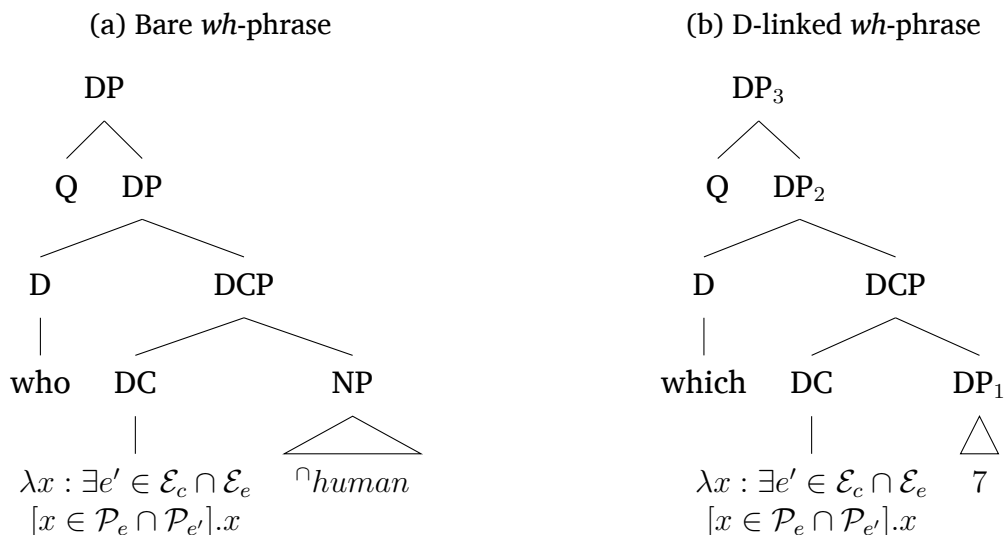
According to Alternative Semantics, no special rules are needed for DC to be interpreted in this structure. The *wh*-D-head, such as *who* or *which*, only has a focus semantic value, as will any parent node until it combines with an operator that can translate it to an ordinary semantic value (Rooth's (1992) $\sim$; Kotek's (2019) ALTSHIFT). The complement of that D-head, however, is free to have an ordinary semantic value. According to Cable's (2010: 64) implementation of Rooth (1985, 1992), a non-focussed phrase can be interpreted in the focus dimension as a set (3.18).

(3.18) The focus semantic value of an non-focussed head is simply the set containing its ordinary semantic value.
$$[\![\, X \,]\!]^f = \{\, [\![\, X \,]\!] \,\}$$ (after Cable 2010: 64)

Figure 3.3 adds the status of each node as having a meaning along the ordinary dimension of meaning (^{o}) or the focus dimension of meaning (^{f}). Note that in the case of bare *wh*-phrases in Figure 3.3a, the NP is type-shifted to be interpreted as a kind with the DC-particle. Then, the resulting DCP must be type-shifted again in order to be interpreted as a property, the original property denoted by the NP. In the case of D-linked *wh*-phrases, the embedded DP is already interpreted as an individual. The resulting DCP must be type-shifted to a property in order to be eventually interpreted as an alternative set containing that individual.

Figure 3.3: Syntax and Alternative Semantics of DC-marked *wh*-phrases
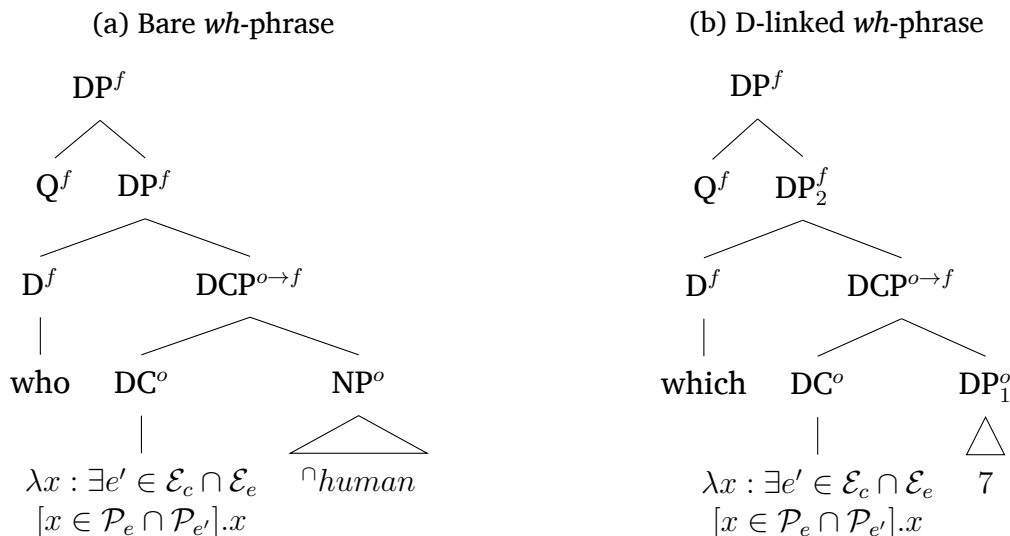
The analysis above predicts that both individuals and kinds can be marked as DC. (3.19) provides evidence that this is the case. In (3.19b), the *wh*-phrase *nǐʔ ḳeʔ* 'what animal' overtly contains the NP *nǐʔ* 'animal', and it is this NP that is previously mentioned in (3.19a). The DC-marked phrase is kind-denoting, as possible answers are kinds of animals (e.g. *paw* 'buffalo'). Note that in context the meaning of (3.19a) appears to amount to *[Look at] all the animals being raised there,* and it is understood that some of the animals are raised by the addressee. Then, (3.19b–b′) are interpreted as elaborating upon (3.19a) by asking for more information about the ownership of the animals.

(3.19)   a. p̥ih **nǐʔ tʰun** pǎʔ ni …
    all   animal   here

   '[Of] all the animals here…'

  b. **nǐʔ** **ḳeʔ**$_{DC}$ hɨ jɔŋ ~~nǐʔ ḳeʔ~~
   animal what   2SG   raise

   'What [kind of] animal do you raise?'   (a ⇓ b)

  b′. **nǐʔ** **hlɛ̌j**$_{DC}$ hɨ jɔŋ ~~nǐʔ hlɛ̌j~~
   animal which 2SG   raise

   'Which animal do you raise?'   (a ⇓ b′)

As for (3.19b′), the D-linked *wh*-phrase *nǐʔ hlɛ̌j* 'which animal', it is the contextual antecedent (i.e. the individual animals) that is previously mentioned. The DC-marked

phrase in this example is individual-denoting, as possible answers are individual animals (e.g. *paw năn* 'that buffalo'). It is worth noting that *nǐʔ tʰun* 'animals' in (3.19a) is ambiguous between a kind reading and an individual reading.

Given that constituents inside the DP can be DC-marked, it would be predicted that quantifiers and focus associators can be DC-marked in Eastern Cham. In theory, a quantifier phrase contains an NP or DP that can be DC-marked, while the quantifier itself would not be of the right semantic type to be DC-marked. Similarly, a focus associator can embed an NP or DP, but the resulting focussed phrase is interpreted as a set of alternatives. This prediction is borne out. Phrases containing the universal quantifier *pih* and focus associator *tʰa sĭt* 'only' can be DC-moved, as in (3.20b,b′), respectively. In both cases, the DC-particle combines with a constituent embedded underneath the quantifier and focus associator.

(3.20)   a. pih **nǐʔ tʰun** păʔ ni…
         all   animal   here

         '[Of] all the animals here…'

   b. **pih** [**nǐʔ   năn**]$_{DC}$ kăw jɔŋ ~~pih nǐʔ năn~~
      all   animal that   1SG raise

      'I raise all of them.'   (a ⇓ b)

   b′. **tʰa sĭt** [**nǐʔ   ni**]$_{DC}$ **mĭn**   kăw jɔŋ ~~tʰa sĭt nǐʔ ni mĭn~~
       only   animal this   EMPH 1SG raise

       'I only raise this one.'   (a ⇓ b′)

Even downward entailing quantifiers like *kiʔ həŋ*^{33} 'less than' can be DC-moved. In (3.21), it it the kind $^{\cap}person$ that is DC-marked. Here, the elaborating answer to the polar question creates a subordinate discourse relation.

(3.21)   a. hɨ ʔa   **lo**   **nujh** lĕj
         2SG invite many person Y/N.Q

         'Did you invite many people?'

   b. **kiʔ həŋ**^{33} **mɨ** [**jaŋ**]$_{DC}$ kăw ʔa   ~~kiʔ həŋ^{33} mɨ jaŋ~~ maj
      few exceed(VN)   five   CLF.PERSON 1SG invite   come
      păʔ ni
      here

      'I invited less than five people to come here.'   (a ⇓ b)

Taken together, the individual and kind readings of *wh*-phrases allow them to be DC-marked, provided DC-marking takes place below the *wh*-D-head. This explains how different types of *wh*-phrase in Eastern Cham can be DC-moved. It is worth noting that 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 di-
agnostics, DC-phrases do not align with topics, and DC-movement cannot be described
purely as topicalization. Discourse/D-linking is a related notion. It has been argued that
D-linked *wh*-phrases are *wh*-topics (e.g. Pan 2014). The following section demonstrates
that DC also cannot be captured in terms of D-linking.

3.2.2   D-linking

Because of the conceptual similarity between DC and Discourse/D-linking, it is worth
investigating to what extent they overlap. However, it is worth noting that D-linking is
often only evoked to describe *wh*-phrases. A DC-marked *wh*-phrase, as laid out in the
previous section, requires previous mention of the individuals quantified over by the *wh*-
phrase in a superordinate sentence in the discourse. D-linked *wh*-phrases are usually
characterized as denoting alternative sets saliently shared by the speaker and addressee
(Pesetsky 1987; Comorovski 1996; 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).

 Despite this conceptual similarity, my analysis of DC predicts that it is orthogonal to D-
linking, as DC does not require a contextually salient individual, and it uniquely imposes
a discourse structural requirement. Based on language-internal evidence, DC-marking
in Eastern Cham is orthogonal to D-linking. DC-marking does not follow the predicted
distribution of D-linking, but a different phenomenon does, *wh*-resumptive pronouns.

 The evidence in this section relies on Pesetsky's (1987) characterization of D-linking in
English. According to Pesetsky, the form of a *wh*-phrase determines its D-linking specifi-
cation. *Wh*-phrases of the form *which X* are taken to be obligatorily D-linked (i.e. lexically
specified as such; 3.22a). Bare *wh*-phrases are optionally D-linked, in that a D-linked read-
ing can be coerced, given an appropriate context (3.22b). *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 ((3.22c); Pesetsky 1987: fn.36; Wiltschko 1997: 113). And
finally, *wh*-phrases of the form *wh-the-hell* are described as 'aggressively non-D-linked',
never D-linked ((3.22d); cf. den Dikken & Giannakidou 2002).

(3.22)   CONTEXT: Some people$_i$ entered the room…

    a. [Which (ones)]$_i$ did Antonia talk to?   [D-linked]

    b. $^?$Who$_i$ did Antonia talk to?   [Optionally D-linked]

    c. #[What ones]$_i$ did Antonia talk to?   [Non-D-linked]

    d. *[Who the hell]$_i$ 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 can-
not be DC-moved (3.23). DC-moved *wh*-phrases are also often translated as D-linked

wh-phrases in both English and Vietnamese (e.g. 3.24). 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 (3.24), perhaps the closest translation equivalent of Eastern Cham DC-movement in English and Vietnamese is D-linking.

(3.23) {*} hɨ   tɔʔ   ŋăʔ {mɓroj k̥eʔ}
           2SG PROG do   crazy   what
    'What the hell are you doing?'

(3.24) k̥eʔ (po)   hɨ   tɔʔ   ɓăŋ k̥eʔ
          what COMP 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 (3.25a–c), 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 (3.25a) than those in (3.25b–c).

(3.25) a. **nɨʔ**   **hlĕj**$_{DC}$ hɨ   jɔŋ ~~nɨʔ hlĕj~~
            animal which 2SG raise

          'Which animal do you raise?'

       b. **nɨʔ**   **k̥eʔ**$_{DC}$ hɨ   jɔŋ ~~nɨʔ k̥eʔ~~
            animal what   2SG raise

          'What animal do you raise?'

       c. **k̥eʔ**$_{DC}$ hɨ   jɔŋ ~~k̥eʔ~~
            what   2SG raise

          'What do you raise?'

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 (3.26), consultants consistently accept resumptive pronouns with *which X* (3.26a), but not with *what X* or bare *wh*-phrases (3.26b–c).

(3.26) a. **nɨʔ**   **hlĕj**$_{DC,i}$ hɨ   jɔŋ ɲu$_i$
            animal which   2SG raise 3.ANIM

          'Which animal do you raise?'

       b. $^{??}$**nɨʔ**   **k̥eʔ**$_{DC,i}$ hɨ   jɔŋ ɲu$_i$
            animal what   2SG raise 3.ANIM

          INTENDED: 'What animal do you raise?'

c. $^{??}$**ķe?**$_{DC,i}$ hɨ   jɔŋ   ɲu$_i$
    what   2SG raise 3.ANIM
    INTENDED: 'What do you raise?'

In a subset of contexts where a D-linked reading is strongly coerced, such as the partitive-like context in (3.27), though, all three forms of *wh*-phrases are accepted with resumptives. This aligns precisely with Pesetsky's (1987) characterization of D-linking in English.

(3.27)   a. pih **nĭ? tʰun** pă? ni // **nĭ?**   **hlĕj**$_{DC,i}$ hɨ   cəh jɔŋ   ɲu$_i$
         all animal here   animal which 2SG like raise 3.ANIM
         '[Of] all the animals here, which animal do you like to raise?'

       b. pih **nĭ? tʰun** pă? ni // **nĭ?**   **ķe?**$_{DC,i}$ hɨ   cəh jɔŋ   ɲu$_i$
         all animal here   animal what   2SG like raise 3.ANIM
         '[Of] all the animals here, what animal do you like to raise?'

       c. pih **nĭ? tʰun** pă? ni // **ķe?**$_{DC,i}$ hɨ   cəh jɔŋ   ɲu$_i$
         all animal here   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. The index renders the *wh*-phrase truly anaphoric. It follows that a pronoun can refer to the same antecedent. In Eastern Cham, a moved phrase corresponding with *which X* must be both D-linked and DC-marked, as phrases with *which* involve a *wh*-determiner taking a DP complement that denotes an 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 moved *what X* and bare *wh*-phrases in Eastern Cham do not necessary license coreferential pronouns. The resulting distinction between DC and D-linking corroborates the conclusion from the previous chapter that the computation of DC is a separate process from the computation of anaphoric coherence.

3.3   Syntax of DC-marked *wh*-phrases

This section examines the syntactic properties of DC-movement and base generation + Agree applied to *wh*-phrases, along with the syntactic properties of DC-moved *wh*-phrases themselves. Section 3.3.1 examines how a whole *wh*-phrase can be DC-moved even though only a subpart is DC-marked via feature percolation and pied-piping. Section 3.3.2 turns to the syntactic properties of DC-movement and the question of what feature triggers it. The DC-movement of *wh*-phrases is shown to have identical syntactic properties to DC-movement in general. Section 3.3.4 then demonstrates that DC-movement

must be a unitary phenomenon, as a DC-marked *wh*-phrase can intervene on a DC-marked non-*wh*-phrase.

3.3.1   Feature percolation

First, this section presents a feature percolation account for how *wh*-phrases are DC-marked in Eastern Cham. The previous section argued that DC-marking occurs inside a *wh*-phrase underneath the *wh*-D-head, as in Figure 3.4. I propose that the DC-feature percolates upwards, following Cole, Hermon & Sung's (1993) account of feature percolation, resulting in a DC-feature present on the highest DPs in Figures 3.4a and 3.4b.

Figure 3.4: DC-marking in *wh*-phrases

(a) Bare *wh*-phrase

$$DP_{DC}$$

Q DP_{DC}

D   DCP

who   DC   NP

…

(b) D-linked *wh*-phrase

$$DP_{DC}$$

Q DP_{DC}

D   DCP

which   DC   DP

…
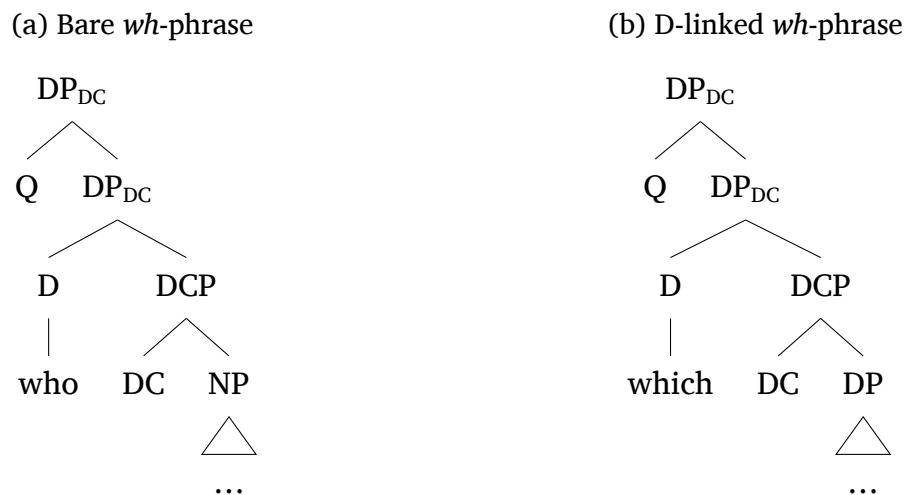
Cole, Hermon & Sung (1993) argue that a syntactic feature can percolate to higher phrases as long as it does not clash with other features. In the realm of DC, no directly related features have been proposed, so DC is free to percolate. I posit that the percolation of the DC-feature is constrained by four factors: the resulting phrase must be a DP or NP, as only DPs and NPs can be DC-marked in Eastern Cham; the smallest possible phrase is chosen to be DC-marked; island constraints; and *wh*-semantics.

In the case of *wh*-phrases, as in Figure 3.4 above, the DC-feature is free to percolate up to the highest DP level. In both cases, the highest DP is the smallest phrase that can be DC-marked. The highest DP can be DC-moved to Spec-CP, so the DC-feature does not percolate higher. Lower DPs and NPs are prevented from being DC-marked by island constraints and *wh*-semantics. If the lower DP in Figure 3.4b is DC-marked, for example, it cannot be DC-moved out of a *wh*-island to Spec-CP. Chapter 4, Section 4.2.2 posits a DC-probe on D as well, but DC semantics cannot be computed after alternatives have been

evoked by the *wh*-D-head, as argued in Section 3.2.1.[2] Therefore, the only possibility is the DC-feature percolating to the highest DP, and the whole *wh*-phrase being pied-piped to Spec-CP.

The account of feature percolation proposed here predicts that syntactic islands can be pied-piped to Spec-CP. This prediction is borne out. Consider a context such as (3.28). The DC-marked phrase is *mɛ̀ kǎw* 'my mother', so the DC-particle first merges with that DP. However, *mɛ̀ kǎw* cannot be DC-marked inside the relative clause, as it is the subject (Section 3.3.2 below; Chapter 2, Section 2.2.3). The phrase also cannot be DC-marked out of the relative clause, as the relative clause forms a syntactic island. Therefore, the DC-feature percolates to the next DP, the entire relative clause, which then itself is DC-moved to Spec-CP. Note that the presupposition is assessed for just *mɛ̀ kǎw* 'my mother', as the DC-particle itself remains low inside the relative clause.

(3.28) a. kǎw çuwʔ **mɛ̀**   **kǎw** hjɛ̌j ni
          1SG   help   mother 1SG   day   this
         'I helped my mother today.'

     b. [kan pö   **mɛ̀**   **kǎw**$_{DC}$ ʈɔʔ   ŋǎʔ   hwǎʔ nǎn] kǎw plɛ̌j
           fish COMP mother 1SG   PROG make eat   that   1SG   b̊uy
         ~~kan pö **mɛ̀ kǎw** ʈɔʔ ŋǎʔ hwǎʔ nǎn~~

         'I bought the fish that my mother is cooking.'   (a ⇓ b)

DP-internal DC-movement presents a third example of feature percolation. As argued in more detail in Chapter 4, Sections 4.2.2 and 4.3.2, NPs and DPs can be DC-moved to Spec-DP in inventory form and partitive constructions. (3.29b) gives an example of an inventory form in context. Here, the NP *tamkaj* 'watermelon' is DC-moved to Spec-DP. The DC-particle merges with that NP. Then, the feature does not percolate, because the NP is the smallest phrase that can be DC-moved, as licensed by a DC-probe on D. A corroborating fact is that inventory forms are only felicitous if the numeral (i.e. the remainder of the DP) does not satisfy the DC conditions, as evidenced in the example below by the absence of previous mention of the numeral.

(3.29) a. hɨ   naw ʐaʔ   hu   plɛ̌j **tamkaj**   lɛ̌j
         2SG go   market EXIST b̊uy   watermelon   Y/N.Q
         Q: 'Did you go to the market and buy watermelons?'

     b. kǎw plɛ̌j [$_{DP}$ **tamkaj**$_{DC}$ klǎw pɔ̊h   ~~tamkaj~~]
         1SG   b̊uy   watermelon   three CLF.ROUND
         A: 'I bought watermelon, three.'   (a ⇓ b)

---

[2]If a different element of the DP is questioned, such as the numeral in a *how many* question, DC-movement can occur to Spec-DP in a *wh*-phrase, as argued in Chapter 4, Section 4.2.2.

Feature percolation, thus, accounts for the ability of *wh*-phrases to be DC-moved, even though the DC-particle is merged lower than D. Feature percolation is constrained by general syntactic and semantic factors. Syntactic islands are active throughout Eastern Cham. DC semantics was independently argued to be incompatible with semantic alternatives. It is a general fact that only NPs and DPs can be DC-moved in Eastern Cham. The constraint that the smallest possible phrase be DC-moved has not yet been mentioned, but it could be parsed as a constraint favoring economy or disfavoring ambiguity. As for economy, movement of larger phrases presumably involves a greater cost, so movement of smaller phrases would be favored from production and processing standpoints. As for ambiguity, movement of a large phrase such as the relative clause in (3.28b) above, results in ambiguity; in principle, both *mɛʔ kăw* 'my mother' and the whole relative clause headed by *kan* 'fish' could be intended to be DC-marked. I envision this factor as a violable constraint, as the larger phrase is still moved in (3.28b).

3.3.2 DC-movement

Turning to the syntactic properties of movement, I will show that the DC-movement of *wh*-phrases has the same properties of DC-movement in general observed in the last chapter: (i) there is a (matrix) subject-object asymmetry; (ii) multiple phrases can be moved; and (iii) path containment effects arise. These facts distinguish DC-movement from clefts, which have been argued to be the source of apparently optional *wh*-movement in languages like Bahasa Indonesian (Cheng 1991) and Malagasy (Potsdam 2006).

First, there is a subject-object asymmetry, such that a subject cannot be in the specifier of the immediately dominating CP and be marked by the complementizer *po* (3.30a). This could be due to a restriction on movement such as Anti-Locality, or a restriction on the pronunciation of the complementizer.

(3.30) a. **tʰɛ̆j* po   tʰɛ̆j ɓăŋ pɔh   ʔɔ̆ʔ   năn
            who COMP   eat   CLF.ROUND mango that
         INTENDED: 'Who ate that mango?'   (DC-movement)

       b. hu **tʰɛ̆j** po   tʰɛ̆j ɓăŋ pɔh   ʔɔ̆ʔ   năn
            ∃   who COMP   eat   CLF.ROUND mango that
         'Who is it that ate that mango?'   (Cleft)

In Eastern Cham, *wh*-phrases can be used in a presentation cleft-like construction, as in (3.30b). This construction has the hallmarks of Ā-movement (cf. Section 3.1), but it does not have the same subject-object asymmetry.

Second, multiple *wh*-phrases can be DC-moved, again unlike clefts (3.31). Note that (3.31b) is ungrammatical with or without the second existential marker *hu*. If the movement in (3.30a) and (3.31a) were due to some kind of concealed *wh*-cleft, an explanation

would have to be given for why it differs in these respects from the clefts in the (b) examples. Additionally, it would have to be explained why DC-movement in general has the same characteristics as well.

(3.31)   a. **ḳeʔ**_{DC} **tʰɛ̌j**_{DC} hɨ   ʔa   tʰɛ̌j maj ɓăŋ ḳeʔ
        what who 2SG invite   come eat

      'Who did you invite to come eat what?'   (DC-movement)

    b. *hu **ḳeʔ** hu **tʰɛ̌j** hɨ   ʔa   tʰɛ̌j maj ɓăŋ ḳeʔ
        ∃   what ∃   who 2SG invite   come eat

      INTENDED: 'Who is that what is it that you invited to come eat?'   (Cleft)

Third, when multiple *wh*-phrases are DC-moved, the same path containment effect arises as with DC-movement in general. The movement path of one phrase must be completely contained within that of the other (3.32a). When paths are crossed, the resulting sentence is consistently ungrammatical (3.32b). This is unexpected for *wh*-movement, as it represents an Anti-Superiority effect (cf. Baclawski Jr & Jenks 2016 on Moken). Note that one movement path is marked in bold and the other is underlined in these examples.

(3.32)   a. **ḳeʔ**_{DC} <u>t^hɛ̌j</u>_{DC} tʰuːŋ͡m^{312} ʔa   tʰɛ̌j maj ɓăŋ **ḳeʔ**
        what who Thuận(VN) invite   come eat

      'Who did Thuận invite to come eat what?'

    b. *<u>tʰɛ̌j</u> **ḳeʔ** tʰuːŋ͡m^{312} ʔa   tʰɛ̌j maj ɓăŋ **ḳeʔ**
        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). However, the path containment effect persists in embedded clauses in Eastern Cham (3.33).'

(3.33)   a. hɨ   hnĭŋ **ḳeʔ**_{DC} <u>tʰɛ̌j</u>_{DC} tʰuːŋ͡m^{312} ʔa   tʰɛ̌j maj ɓăŋ **ḳeʔ**
        2SG think what who Thuận(VN) invite   come eat

      'Who do you think Thuận invited to come eat what?'

    b. *hɨ   hnĭŋ <u>tʰɛ̌j</u> **ḳeʔ** tʰuːŋ͡m^{312} ʔa   tʰɛ̌j maj ɓăŋ **ḳeʔ**
        2SG think who what Thuận(VN) invite   come eat

      INTENDED: 'Who do you think Thuận invited to come eat what?'

As with DC-movement, I analyze these sentences with multiple syntactic probes, which are each constrained by locality, or structural closeness (e.g. Pesetsky 1982 on English *wh*-movement; Baclawski Jr & Jenks 2016 on Moken). When the first C-head is merged, the first probe (C_1) searches for a phrase with the relevant feature, here DP_1. That DP is moved to Spec-CP_1 (Figure 3.5a). When the second C-head is merged, the second probe (C_2) searches for the same feature. It cannot Agree with DP_1 either due to criterial freezing (e.g. Rizzi 2010) or Anti-Locality (e.g. Erlewine 2016). Instead, it Agrees with DP_2, and it is that phrase that moves to Spec-CP_2 (Figure 3.5b).

Figure 3.5: Path containment derivation

(a) Derivation of CP_1 (b) Derivation of CP_2
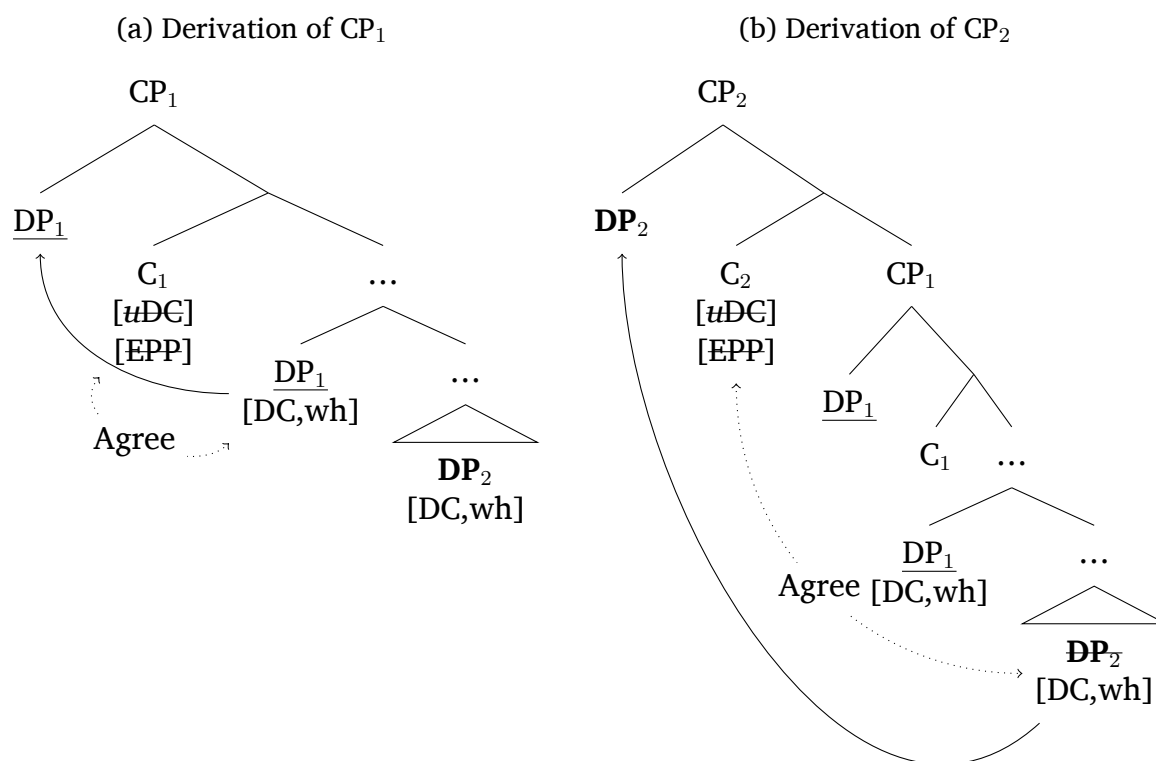
3.3.3   Base generation

In the previous chapter, base generation of DC-phrases with resumptive pronouns (RPs) was examined. It was concluded that there is a separate Agree relation between RPs and C, involving an Op feature. The Op feature ensures the RP is interpreted as coreferential

with the phrase in Spec-CP. If *wh*-phrases can also be DC-marked, I would predict that the same option is available for *wh*-phrases, represented in (3.34) below.

(3.34)   a. **nujh**   **hlĕj**$_{DC,i}$ hɨ   ʔa   ɲu$_i$
        person   which   2SG   invite   3.ANIM
        'Which person did you invite?'

      b. [$_{DP}$ [$_{DCP}$ DC nujh] hlĕj$_{D-linked}$]$_i$ C$_{uOp}$ hɨ ʔa [ɲu$_{Op}$]$_i$ (Base generation + Agree)

In fact, not only is this option available in Eastern Cham, it has the same syntactic effects. Base generation does not interact with DC-movement, but multiple base generation exhibits path containment effects. First, (3.35) demonstrates the path containment effect seen when multiple DC-phrases are base generated in Spec-CP, and there are multiple RPs. Contained paths lead to grammaticality (3.35a), while crossed paths lead to ungrammaticality (3.35b).

(3.35)   a. **jŭt**   **ni**   n̄ɨʔ   mĕj   sĭt   năn   tʰuːɲm͡³¹²   ʔa   ɲu   maj
        friend   this   child   female   small   that   Thuận   invite   3.ANIM   come
        ɓăŋ   **ɲu**
        eat   3.ANIM
        'This friend, Thuận invited that little girl to come meet.'

      b. *n̄ɨʔ   mĕj   sĭt   năn   **jŭt**   **ni**   tʰuːɲm͡³¹²   ʔa   ɲu   maj
        child   female   small   that   friend   this   Thuận   invite   3.ANIM   come
        ɓăŋ   **ɲu**
        eat   3.ANIM
        INTENDED: 'This friend, Thuận invited that little girl to come meet.'

However, base generation does not interact with DC-movement. In (3.36), there is a mix of one base generation path and one movement path. The paths are free to be nested or crossed, and the resulting sentences are grammatical.

(3.36)   a. n̄ɨʔ   mĕj   sĭt   năn   **jŭt**   **ni**   tʰuːɲm͡³¹²   ʔa   ɲu   maj
        child   female   small   that   friend   this   Thuận   invite   3.ANIM   come
        ɓăŋ   ~~**jŭt ni**~~
        eat
        'This friend, Thuận invited that little girl to come meet.'

      b. n̄ɨʔ   mĕj   sĭt   năn   **jŭt**   **ni**   tʰuːɲm͡³¹²   ʔa   ~~n̄ɨʔ mĕj sĭt năn~~
        child   female   small   that   friend   this   Thuận   invite
        maj   ɓăŋ   **ɲu**
        come   eat   3.ANIM
        'This friend, Thuận invited that little girl to come meet.'

These facts follow if base generation involves an Agree relation with some feature (Op) that is different from the feature that drives DC-movement (DC). Figure 3.6a–b represent such a crossed path derivation. When C_1 is merged, the C-probe searches for a phrase bearing an Op-feature. When C_2 is merged, the C-probe searches for a phrase bearing a DC-feature. Those probes do not interact; they Agree with different phrases entirely.

Figure 3.6: Crossed path derivation for (3.36a)

(a) Derivation of CP_1 (b) Derivation of CP_2
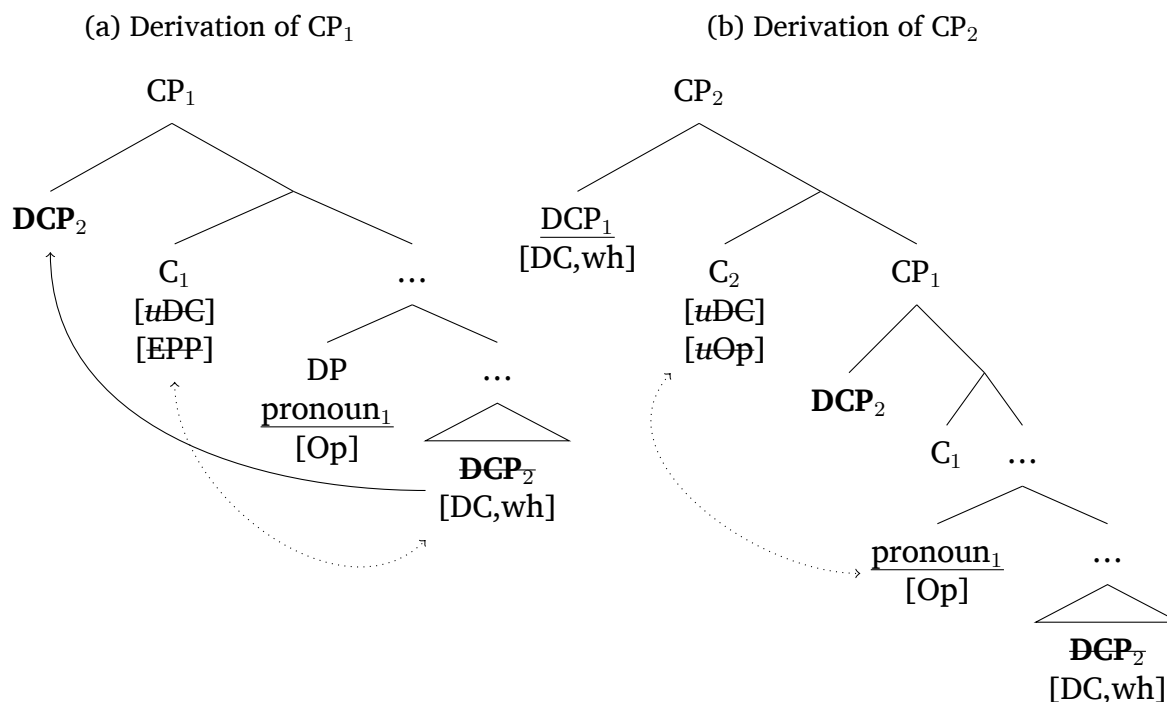
DC-movement, then, is accompanied by a base generation strategy with resumptive pronouns, regardless of the *wh*-nature of the DC-phrase. The next section probes further into the issue of what feature exactly drives DC-movement.

3.3.4 Locality effects

Thus far in this chapter, I have assumed that a DC-feature is responsible for the DC-movement of *wh*-phrases based on their pragmatic and syntactic properties of the moved phrases and the movement operation itself. However, there is an alternative analysis that can explain much of the data: a generalized Ā-feature. This section will sketch a generalized Ā-feature analysis of locality effects between moved *wh*-phrases and moved *wh*-phrases. The following section will provide support for a DC-feature analysis over a generalized Ā-feature analysis, based on the properties of in situ *wh*-phrases. To begin, consider again the basic instances of DC-movement with a *wh*- and non-*wh*-phrase below.

(3.37)   a. **k̥eʔ**   hɨ   t̥ɔʔ   ɓăŋ   ~~k̥eʔ~~
            what 2SG PROG eat

         'What are you eating?'

       b. **ʔɔ̆ʔ**   **ni**   kăw t̥ɔʔ   ɓăŋ   ~~ʔɔ̆ʔ ni~~
            mango this 1SG PROG eat

         'This mango, I am eating.'

Table 3.1 outlines three different featural analyses for the probe that drives DC-movement in both cases. The first possibility (Table 3.1a) is a hypothetical optional *wh*-movement analysis in the vein of Denham (2000).[3] According to this analysis, it is the *wh*-feature alone that is responsible for the movement of *wh*-phrases in Eastern Cham. A pure *wh*-feature analysis over-generates, as not all *wh*-phrases can in fact be moved. It would have to be stipulated that moved *wh*-phrases happen to correspond with DC-marking. A prediction of a pure *wh*-feature analysis is that *wh*-phrases should not intervene on the movement of non-*wh*-phrases, as the movement operations are driven by different syntactic features and probes.

Table 3.1: Featural analyses of DC-movement

		Wh-phrase (3.37a)			Non-*wh*-phrase (3.37b)		
a.	Optional *wh*, DC	k̥eʔ$_{wh}$	C$_{uwh}$	k̥eʔ	ʔɔ̆ʔ ni$_{DC}$	C$_{uDC}$	~~ʔɔ̆ʔ ni~~
b.	Generalized Ā-feature	k̥eʔ$_{Ā:wh}$	C$_{uĀ}$	k̥eʔ	ʔɔ̆ʔ ni$_{Ā:DC}$	C$_{uĀ}$	~~ʔɔ̆ʔ ni~~
c.	DC-feature	k̥eʔ$_{DC,wh}$	C$_{uDC}$	k̥eʔ	ʔɔ̆ʔ ni$_{DC}$	C$_{uDC}$	~~ʔɔ̆ʔ ni~~

As schematized in Table 3.1b, movement of *wh*-phrases can also in theory be analyzed with a generalized Ā-probe. Aravind (2017) proposes a feature hierarchy of Ā-features, as in Figure 3.7 (cf. also Starke 2001, Rizzi 2004, and Abels 2012). According to this hierarchy, a lower-order feature entails those above it. For instance, a generalized Ā-probe at the top level can interact with a *wh*-feature (i.e. Ā:wh) or a topic feature (i.e. Ā:topic). This hierarchy explains interactions between seemingly disparate syntactic movement operations like *wh*-movement and topicalization in English, Malayalam, and other languages (Aravind 2017, 2018). A generalized Ā-feature could hypothetically account for DC-movement and the movement of *wh*-phrases in Eastern Cham, as it would be satisfied by either *wh* (i.e. Ā:wh) or DC (i.e. Ā:DC), if it is accepted that DC is an Ā-feature.

---

[3]It should be noted that a DC-feature analysis of apparently optional *wh*-movement appears to be within the spirit of Denham (1998, 2000), who lists topic and other features as possible drivers of such movement.

Figure 3.7: Ā-feature hierarchy (after Aravind 2017: (44))
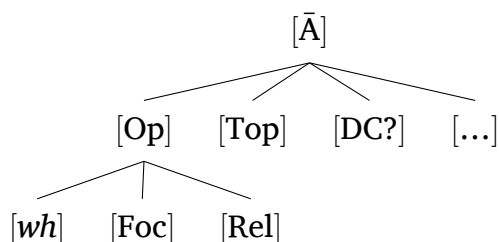
Finally, a DC-feature analysis posits the only relevant feature for DC-movement is DC itself, as in Table 3.1c. This is the analysis that has been argued for throughout this chapter and Chapter 2. Both the generalized Ā-feature and DC-feature analyses predict that moved *wh*-phrases should be able to intervene on the movement of non-*wh*-phrases, because the syntactic probes interact with both moved *wh*-phrases and moved non-*wh*-phrases.

Moved *wh*-phrases do in fact intervene on the movement of non-*wh*-phrases, affirming the generalized Ā- and DC-feature analyses, and providing further evidence against an optional *wh* analysis. In (3.38), one *wh*- and one non-*wh*-phrase are DC-moved to the left periphery. As with previous examples, the resulting sentence is grammatical if the movement paths are nested (3.38a) and ungrammatical if they are crossed (3.38b).

(3.38)   a. **han ni** nǐʔ m̥ĕj sǐt hlĕj tʰuːŋ͡m³¹² ʔḁ n̶ǐ̶ʔ̶ ̶m̶ĕ̶j̶ ̶s̶ǐ̶t̶ ̶h̶l̶ĕ̶j̶
             cake this child female little which Thuận(VN) invite
             maj ɓăŋ **han ni**
             come eat
             'Which little girl did Thuận invite to come eat this cake?'

         b. *nǐʔ m̥ĕj sǐt năn **han hlĕj** tʰuːŋ͡m³¹² ʔḁ n̶ǐ̶ʔ̶ ̶m̶ĕ̶j̶ ̶s̶ǐ̶t̶ ̶n̶ă̶n̶
             child female little which cake this Thuận(VN) invite
             maj ɓăŋ **han hlĕj**
             come eat
             INTENDED: 'Which cake did Thuận invite that little girl to come eat ?'

The manners in which the generalized Ā- and DC-feature analyses account for the grammatical path containment example above are presented in Figure 3.8. In Figure 3.8a, the generalized Ā-probe on C₁ probes and Agrees with the *wh*-phrase in DP. An identical probe on C₂ probes and Agrees with the DC-marked phrase, DCP. The ungrammatical path containment example (3.38b) cannot be derived, as C₁ would have to Agree with a non-local phrase. In Figure 3.8b, the DC-probes only probe for a DC-feature. The *wh*-DP only happens to also bear a *wh*-feature.

Figure 3.8: Featural analyses of (3.38a)

(a) Ā-probe analysis

(b) DC-probe analysis
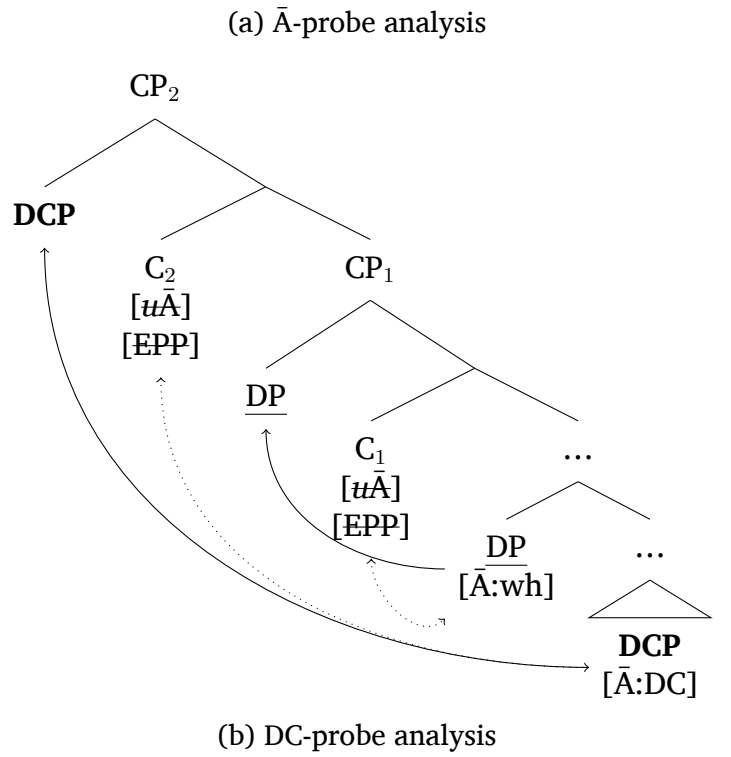
In order to differentiate between these two analyses, the properties of in situ *wh*-phrases must be examined. A generalized Ā-feature approach predicts that in situ *wh*-phrases be visible to Ā-probes, as they still have *wh*-features. This would result in syntactic interaction between in situ *wh*-phrases and movement of a non-*wh*-phrase. By contrast, a DC-feature approach does not make that prediction, as in situ *wh*-phrases could lack the relevant DC-feature. The following section presents data which favors the latter approach, as in situ *wh*-phrases do Agree with C, but show no interaction with DC-movement.

3.4   In situ *wh*-phrases

In this chapter, I have referred to *wh*-in situ as a unitary phenomenon. However, a closer look is needed to understand how *wh*-phrases interact with C. In situ *wh*-phrases in Eastern Cham are shown to Agree with C, but in a way that does not interact with DC-movement. This leads to arguments that favor a DC-feature over a generalized Ā-feature.

There are known to be multiple structures that surface as *wh*-in situ. Figure 3.9 presents three of these. First, in situ phrases can Agree with C, but C does not attract them to Spec-CP (cf. Cable 2010 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 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 3.9: Types of *wh*-in situ

(a) Agree, −Attract   (b) Covert *wh*-movement   (c) Variable interpretation
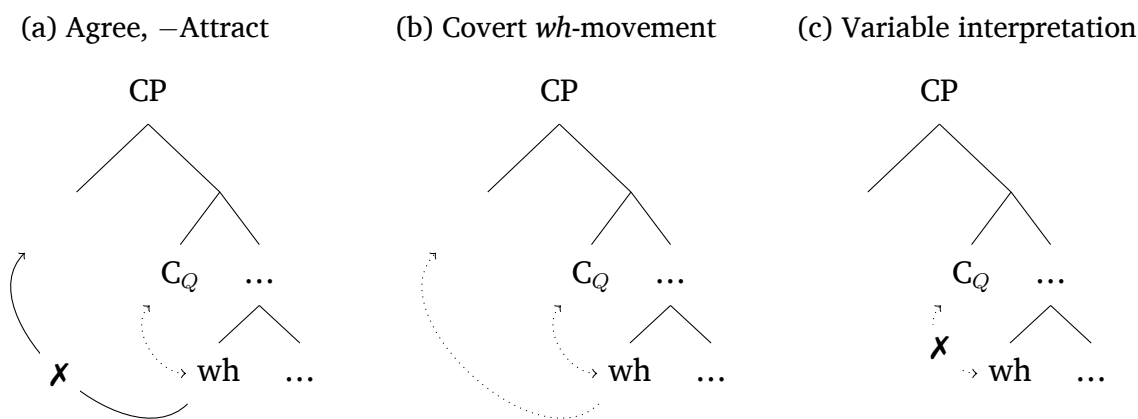
These structures can be differentiated by applying movement diagnostics to in situ *wh*-phrases. Evidence from island constraints and intervention effects in Eastern Cham point to the Agree, but not Attract structure in Figure 3.9a. First, *wh*-phrases are ungrammatical

within syntactic islands. For instance, in a complex NP, the existence of an in situ *wh*-phrase leads to ungrammaticality (3.39a), 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 (e.g. Adger & Ramchand 2005). It should be noted that this context is generally permissible in the absence of overt movement and *wh*-phrases (3.39b).

(3.39) a. *hɨ plĕj ɗo²¹ ɓăŋ po **tʰĕj** ŋă?
         2SG buy stuff(VN) eat COMP who make
         INTENDED: 'You buy the food that who makes?'

       b. kăw plĕj ɗo²¹ ɓăŋ po mɛ? kăw ŋă?
         1SG buy stuff(VN) eat COMP 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* (3.40a). In this schema, the C_Q-head introduces a $\sim$-operator that interprets the *wh*-alternative set (cf. Rooth 1992; Kotek's (2019) ALTSHIFT), while the focus operator *only* functions as the intervener. Intervention effects arise because the focus operator cannot interpret the *wh*-alternative set.

(3.40) a. ✓Intervention: $[Q \ [\sim_C \ [\text{ONLY}_C \ \text{T} \ \ldots \text{which soup}]]]$

       b. ✗Intervention: $[\text{which soup} \ [Q \ [\sim_C \ [\text{ONLY}_C \ \text{T} \ \ldots \text{which soup}]]]]$

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 (3.40b), 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 (3.41a), there is a *wh*-phrase, *ʔja paj hlĕj* 'which soup' under the scope of 'only'.[4] If covert *wh*-movement were possible, this sentence would be expected to be grammatical. Since the sentence is ungrammatical, I conclude that in situ *wh*-phrases have no means of moving out of the scope of intervening focus operators. Similarly, the focus operator *çɨŋ* 'also' gives rise to an intervention effect in (3.41b).

(3.41) a. *tʰa sĭt ʔaj tʰuːŋ͡m³¹² krɨ **ʔja paj hlĕj** mĭn
          only older.sibling Thuận(VN) like soup which EMPH
          INTENDED: 'Which soup does only Thuận like to eat?'

---

[4]Note that the emphatic particle *mĭn* coccurs with the focus operator *tʰa çĕj* 'only' and seems to indicate the right edge of its scope.

b. CONTEXT: There are some restaurants everyone can go to, and some that
Kenny cannot.

*kɛn ni çɨŋ naw ɓăŋ pă? **ɲa²¹ haŋ²¹** **hlɛ̆j** 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 (3.42), 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.

(3.42) a. **?ja paj hlɛ̆j**$_{DC}$ tʰa çɛ̆j ?aj tʰuːŋ͡m³¹² kri ~~?ja paj hlɛ̆j~~ mĭn
soup which only older.sibling Thuận(VN) like EMPH

'Which soup does only Thuận like to eat?'

b. **ɲa²¹ haŋ²¹** **hlɛ̆j**$_{DC}$ kɛn ni çɨŋ naw ɓăŋ ~~ɲa²¹ haŋ²¹ hlɛ̆j~~ 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 (3.43a) and the antecedent of conditionals (3.43b). This further demonstrates that *wh*-phrases cannot covertly move out of the scope of operators, given that interrogative readings are impossible in these contexts.

(3.43) a. hɨ hu ɓăŋ **k̥e?** ?o
2SG ∃ eat what NEG

'You didn't eat anything.' / *'What didn't you eat?' (Negation)

b. **tʰɛ̆j** ɲum ka²¹ fe³³ hl̥ă? hl̥ă? n̥ɔŋ nujh năn
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 3.10a). This contrasts with DC-movement, where the *wh*-phrase Agrees with C and moves to Spec-CP (Figure 3.10b).

Figure 3.10: Positions of Eastern Cham *wh*-phrases

(a) *Wh*-in situ
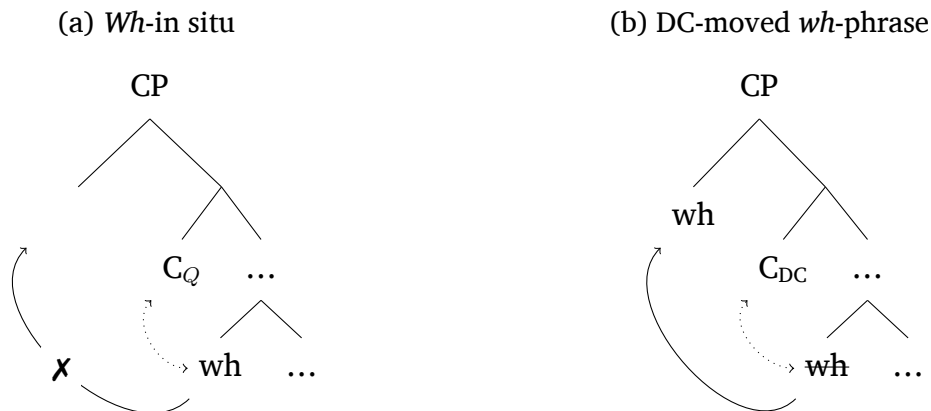
(b) DC-moved *wh*-phrase

To distinguish between the DC-feature and generalized Ā-feature analyses of moved *wh*-phrases, I turn to the interaction between in situ and DC-moved *wh*-phrases. 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 (3.44a). For example, the object *what* cannot move across the subject *who*. Here, the $*^{PL}$ notation indicates that the question loses its paired list and single answer readings. The major exception (outside echo questions) is when the *wh*-phrases are D-linked (e.g. Pesetsky 1987). Whatever the underlying explanation, D-linked *wh*-phrases are exceptional.

(3.44) a. $*^{\mathrm{PL}}$What$_{\mathrm{\bar{A}:wh}}$ C$_{\mathrm{u\bar{A}}}$ did who$_{\mathrm{\bar{A}:wh}}$ buy?

  b. Which book$_{\mathrm{\bar{A}:wh,D\text{-}linked}}$ C$_{\mathrm{u\bar{A}}}$ did which student$_{\mathrm{\bar{A}:wh,D\text{-}linked}}$ read?
  (Pesetsky 2000: 15–16)

The generalization from English is that in situ *wh*-phrases do compete for *wh*-movement, unless they are D-linked. This accords with a generalized Ā-feature analysis, as all *wh*-phrases are taken to be assigned [Ā:wh].

In situ *wh*-phrases in Eastern Cham do not have such an interaction. In situ phrases never show any signs of competing for DC-movement. An object DC-phrase (3.45a) or *wh*-phrase (3.45b) can be DC-moved over an in situ *wh*-subject. This is unexpected under a generalized Ā-feature analysis, as the in situ *wh*-subject should be visible to the probe on C and intervene on the movement of the object.

(3.45) a. **han ni**$_{\mathrm{DC}}$ C$_{\mathrm{uDC}}$ tʰuːŋm^{312} ʔa̰ tʰɛ̌j$_{\mathrm{wh}}$ maj ɓăŋ ~~**han ni**~~
         cake this   Thuận(VN) invite who come eat
    'Who did Thuận invite to come eat this cake?'

b. ḳeʔ$_{DC,wh}$ C$_{uDC}$ tʰuːŋ͡m^{312} ʔa t̪ʰĕj$_{wh}$ maj ɓăŋ ḳeʔ
   what   Thuận(VN) invite who come eat
   'Who did Thuận invite to come eat what?'

The general pattern in English and Eastern Cham is outlined in Table 3.2. In English, the default situation is for *wh*-phrases to move. D-linking exceptionally allows them to stay in situ. In Eastern Cham, the default situation involves no movement. Instead, the exceptional case results in movement. An Ā-feature analysis would have to posit that all in situ *wh*-phrases are exceptional. However, there is no obvious category like D-linking that can explain this exception. Movement falls out naturally from a DC-feature analysis, as the presence of a DC-feature entails the presence of a DC-probe.

Table 3.2: Movement and non-movement of *wh*-phrases

	Agree, −Attract	Agree, +Attract
English	+D-linked	Default (i.e. only *wh*)
Eastern Cham	Default (i.e. only *wh*)	+DC

3.5 Conclusion

This chapter has shown how *wh*-phrases in Eastern Cham can be DC-marked. The same phrase can contain a DC-particle, which marks DC, and a Q-particle, which marks *wh*, as shown by the syntactic and pragmatic properties of *wh*-phrases moved to or base generated in Spec-CP. The same two DC-probes on C proposed in Chapter 2, Section 2.3 derive the range of DC-marking of *wh*-phrases. A third C-probe is needed in order to agree with the Q-particle and interpret *wh*-phrases.

DC-movement of a *wh*-phrase proceeds as in Figure 3.11. First, the in situ *wh*-phrase enters into an Agree relation with C$_Q$ (more specifically, C Agrees with the Q-particle, as per Cable 2010). This ensures that the *wh*-phrase is interpretable. Next, C$_{DC}$ probes for a DC-feature and moves the *wh*-phrase if it bears that feature. In the absence of a DC-feature, the phrase cannot move. 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.

Figure 3.11: Eastern Cham DC-movement of a *wh*-phrase

```
                          CP
                         /  \
                      DP
                       /   \
                    C_DC    CP
                           /  \
                         C_Q   ...
                              /  \
                            DP    ...
                           /  \
                          Q    DP
                              /  \
                             D    DCP
                                 /  \
                                DC   ...
```
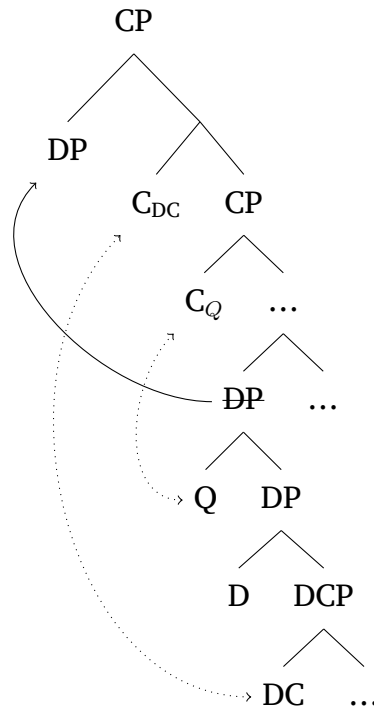
The independence of DC and *wh* reinforces that DC must be a distinct Ā-feature in Eastern Cham syntax. It also informs the apparent optionality of DC-marking. DC-marking introduces a presupposition that checks the DC conditions on the phrase it combines with. Otherwise, it acts as an identity function. As a consequence, the DC-particle can merge anywhere in a derivation as long as it combines with an individual. By contrast, the Q-particle merges with a *wh*-phrase, and subsequently the phrase has no interpretation along the ordinary dimension of meaning (in Alternative Semantics). It must enter an Agree relation with C in order to regain an interpretation.

Chapter 4

DP-internal DC-marking

The previous chapters have argued that discourse connectedness (DC) is an Ā-feature (alongside others like *wh*) based on the movement of phrases to clause edges. This chapter presents another shared characteristic between DC and other Ā-features: the CP/DP parallelism. In some cases, Ā-features can trigger movement to the left edge of DPs, not only that of CPs. *Pied piping with inversion,* seen in many Mesoamerican languages, exemplifies DP-internal movement driven by a *wh*-feature. In Tzotzil (Mayan: Mexico), possessors must follow possessees in the NP (4.1a). When the possessor is a *wh*-phrase, though, (in bold) it must be moved to the left edge of the DP, and the whole DP is moved to the left edge of the clause (4.1b).

(4.1) a. [$_{DP}$ s-p'in   li   Maruch-e   ]
           A3-pot   the   Maruch-ENC
        'Maruch's pot' (Aissen 1996: 454)   TZOTZIL

       b. [$_{DP}$ **Buch'u**   x-ch'amal   ~~buch'u~~]   i-cham   ~~DP~~?
              who   A3-child   CP-died
        'Whose child died?' (Aissen 1996: 457)

Aissen (1996) and others analyze the DP-internal movement as secondary *wh*-movement driven by a *wh*-probe on D (cf. Coon 2009 for an analysis with Q-particles; Broadwell 2001 for an Optimality Theoretic account). Aboh (2004b) argues more broadly that DP has an expanded left periphery much like CP, replete with Ā-probes and phrasal projections. If DC is also an Ā-feature, DP-internal DC-marking would be predicted to occur as well.

This prediction is borne out in Eastern Cham. When a subpart of a DP is DC-marked, it is pronounced at the left edge of the DP. This analysis unifies two phenomena not previously related to one another: inventory forms and partitives. Inventory forms occur when a noun is DC-marked, to the exclusion of a numeral or quantifier (4.2a; DC-phrases bolded throughout). Partitives occur when an embedded DP is DC-marked to the exclusion of the matrix DP (4.2b). In both of these phenomena, I posit DP-internal DC-movement.

(4.2) a. kăw plěj [$_{DP}$ **ʔɔ̌ʔ**$_{DC}$ çŭh pɔ̥h   ~~ʔɔ̌ʔ~~]
           1SG buy   mango 7   CLF.ROUND

           'I bought mangoes, seven.'   (Inventory form)

       b. kăw plěj [$_{DP}$ [$_{DP}$ **ʔɔ̌ʔ**   **ni**$_{DC}$] çŭh pɔ̥h   ~~ʔɔ̌ʔ ni~~]
           1SG buy   mango this   7   CLF.ROUND

           'I bought seven of these mangoes.'   (Partitive)

These two phenomena interact with DC-marking at the CP periphery. In the same contexts as (4.2) above, the DC-marked subpart of the DP can be base generated in Spec-CP and bind a null NP-pronoun. For the moment, the surface order of the sentences in (4.3a–b) appear derivable by DP-movement of the NP. However, evidence from locality effects will show that the NP is instead base generated in Spec-CP much like base generation + Agree in Chapter 2, Section 2.3.2. The parallelism between (4.2) above and (4.3) below confirms that DC can be marked at the left edges of both CPs and DPs.

(4.3) a. **ʔɔ̌ʔ**$_{DC,i}$ kăw plěj [$_{DP}$ çŭh pɔ̥h pro_i]
           mango 1SG buy   7   CLF.ROUND

           'Mangoes, I bought seven.'   (Inventory form)

       b. **ʔɔ̌ʔ**   **ni**$_{DC,i}$ kăw plěj [$_{DP}$ çŭh pɔ̥h pro_i]
           mango this   1SG buy   7   CLF.ROUND

           'These mangoes, I bought seven of.'   (Partitive)

The evidence in this chapter confirms the pragmatics and syntax of DC from the previous chapters, with one addition: the same DC-movement probe on C can also appear on D. There is a probe on D that triggers DC-movement. This probe is illustrated in Figure 4.1 for inventory forms and partitives.[1]

---

[1]Note that the DP syntax depicted here is simplified to account for the necessary overt components.

Figure 4.1: DP-internal DC-probe

(a) Inventory form
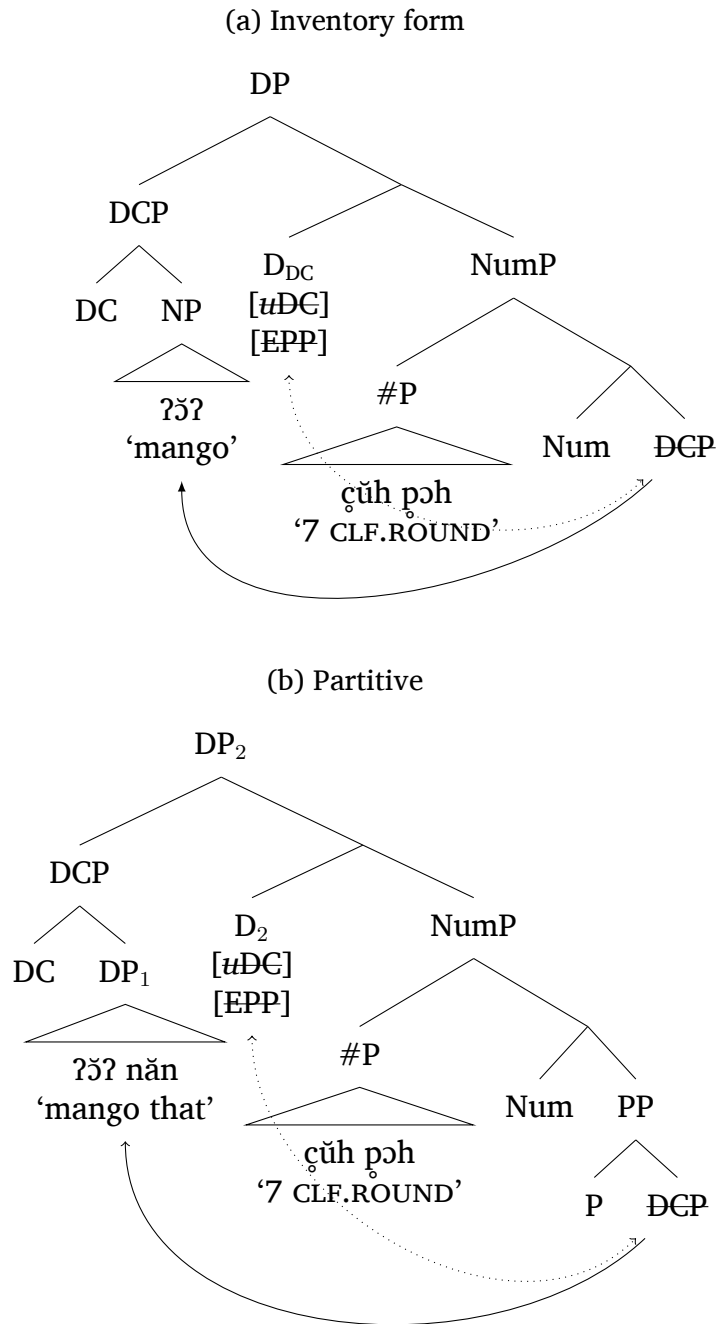
(b) Partitive

The remainder of this chapter is structured as follows. Section 4.1 gives some background on DP structure and introduces basic distributional properties of inventory forms and partitives. Section 4.2 examines the pragmatic and syntactic properties of inventory forms, concluding they provide evidence for the CP/DP parallelism for DC. Section 4.3 gives corroborating evidence from partitives. In Section 4.5, it is found that DP-internal DC-movement can feed subsequent A-movement, retaining DC pragmatics, as predicted by the syntax and pragmatics of DC-movement.

Finally, a lingering issue is to what extent DC is comparable to anaphoric definiteness and familiarity in general. A third potential instance of DP-internal DC-marking is explored in Section 4.6: domain restricting (DR)-appositives (4.4a). However, it is found that DR-appositives are licensed not by DC, but anaphoric definiteness.

(4.4) kǎw ʔa [$_{DP}$ **pu hɔ̌ŋm hwa**$_{+Def,i}$ t̥wa jaŋ ɲu$_i$]
     1SG invite   Phú with Hoa   two CLF.PERSON 3.ANIM
     'I invited those two people, Phú and Hoa.'   (DR-Appositive)

This chapter, thus, finds an additional characteristic shared by DC and other Ā-features: a CP/DP parallelism. The CP/DP parallelism opens the door to find more phenomena that can be explained in terms of DC-marking, namely instances of DP-internal movement. DC is also further distinguished from information structural notions such as definiteness and familiarity.

4.1 Background

Before we proceed to DP-internal DC-marking, some background is needed on Eastern Cham DP syntax. First, Section 4.1.1 argues that the basic DP order is derived by movement of the Numeral Phrase (NumP-movement). This movement is driven by a probe on D. Second, Section 4.1.2 introduces two deviations from basic DP order: inventory forms and partitives.

4.1.1 Basic DP structure

Following research on other Chamic languages (Jensen 2013, 2014 on Jarai) and Vietnamese (Nguyễn 2004, 2013), largely from the Cartographic enterprise, I posit that the basic DP structure in Eastern Cham is derived by Numeral Phrase (NumP) movement. The surface distribution of the Eastern Cham DP is typically described as in (4.5) (cf. Thurgood 2005; Brunelle and Phú 2018). The dashes indicate relative ordering; for example, quantifiers reliably precede numerals, and relative clauses are claimed to precede demonstratives.

(4.5) DP: Quantifier – Numeral – Classifier – Noun – Adjective – Relative clause – Demonstrative

These ordering effects are identical as those reported by Jensen (2013, 2014) on the related Chamic language Jarai. Likewise, they are identical to those reported by Nguyễn (2004, 2013) on Vietnamese, which follows given the status of Mainland Southeast Asia as a sprachbund and the two languages' long history of language contact. Two examples are given below that demonstrate that the Eastern Cham order is parallel to that in Vietnamese.

(4.6) a. h̥ǎʔ   haʔ [DP ʈwa plah   ʔan   pjŏŋ năn]
           1SG.POL rip   two CLF.FLAT paper big   that
           'I tore those two big pieces of paper.'   EASTERN CHAM

       aʹ. Tôi xé [DP hai mẩu   giấy lớn đó]
           1SG rip   two CLF.FLAT paper big that
           'I tore those two big pieces of paper.'   VIETNAMESE

       b. h̥ǎʔ   haʔ [DP ʈwa plah   ʔan   pjŏŋ po   jŭt   plĕj]
           1SG.POL rip   two CLF.FLAT paper big   COMP friend buy
           'I tore two big pieces of paper that you bought.'   EASTERN CHAM

       bʹ. Tôi xé [DP hai mẩu   giấy lớn mà bạn   đã   mua]
           1SG rip   two CLF.FLAT paper big REL friend PST buy
           'I tore the two big pieces of paper that you bought.' VIETNAMESE

Two differences between Eastern Cham and Vietnamese are worth noting. Both are numeral classifier languages, but classifiers exclusively accompany numerals in Eastern Cham, while Vietnamese classifiers also occur with definite nouns and in other contexts (e.g. Löbel 2000). The typical Eastern Cham pattern is shown in (4.7a). The numeral classifier *çĕj* is impossible in definite noun phrases without numerals, but obligatory with them.

(4.7) a. (*çĕj)   thăw năn / tʰa *(çĕj)   thăw
           CLF.ANIMAL dog that   one CLF.ANIMAL dog
           'That dog / One dog.'   (Pattern 1)

       b. (nɨʔ)   thăw năn / tʰa *(nɨʔ)   thăw
           CLF.ANIMATE dog that   one CLF.ANIMATE dog
           'That dog / One dog.'   (Pattern 2)

A second pattern is found with Vietnamese-dominant speakers of Eastern Cham (4.7b). Here, the Vietnamese classifier *con*, which also means 'child', is calqued with the Eastern Cham word *nɨʔ* meaning 'child'. For these speakers, the calqued classifier is possible in definite phrases without numerals, much like in Vietnamese. The data for this chapter are all taken from speakers who exhibit Pattern 1.

For another difference with Vietnamese, Eastern Cham determiners are DP-final. Nguyễn (2004, 2013) posits that Vietnamese determiners are DP-initial, based on overt evidence involving the two plural markers, which are analyzed as determiners in (4.8a). By contrast, demonstratives are DP-final (4.8a).

(4.8)   a. [$_{DP}$ tất cả {những/các} cuốn sách này]
                all   PL/PL   CLF   book this

           'all these books' (Nguyễn 2004: 54)   Vietnamese

      b. [$_{DP}$ pŭʔ hlĕj   (*năn)] / [$_{DP}$ tʰa jaŋ   ɲu   (*năn)]
          book which   that   one CLF.PERSON 3.ANIM   that

          'which (of those) books' / 'him/her [LIT.: one him/her]'   Eastern Cham

In Eastern Cham, there are no comparable plural markers, but other possible D-heads, such as *wh*-determiners and pronouns inside complex pronouns are DP-final (4.8b). These forms cannot cooccur with demonstratives, which implies that they occupy similar positions within the DP.[2] In the remainder of this chapter, I will assume that *wh*-determiners and pronouns inside complex pronouns are D-heads.

Both Jensen (2013, 2014) and Nguyễn (2004, 2013) propose the general DP spine below for Jarai and Vietnamese, respectively. These analyses take a Cartographic, Antisymmetric approach to syntax, wherein the same phrasal spine is assumed for all languages and all trees are right-branching (Kayne 1994; Cinque 2005; cf. also Simpson 2005 on DP structure in other Mainland Southeast Asian languages; cf. Bruening, Xuyen Dinh & Lan Kim 2018 for a contrary approach to Vietnamese).

(4.9) DP: Quantifier ≫ Determiner ≫ Demonstrative ≫ Numeral/Classifier ≫ NP

Jensen (2013, 2014) proposes that numeral–classifier sequences in Jarai are constituents in the specifier position of Numeral Phrases (Figure 4.2). This departs from Nguyễn's (2013) analysis of Vietnamese, which assumes separate Numeral and Classifier Phrase projections.

---

[2]Nevertheless, we follow cross-linguistic evidence that determiners and demonstratives head distinct phrasal projections (e.g. Bernstein 1997), as does Jensen (2013, 2014) on Jarai.

Figure 4.2: DP spine for Eastern Cham, Jarai (Jensen 2013, 2014)

```
              (QuantP)
               /    \
         (Quant)    DP
                   /  \
                  D   DemP
                      /  \
                   Dem   NumP
                         /  \
                       #P    \
                      /  \   Num   NP
                     #   Clf
```
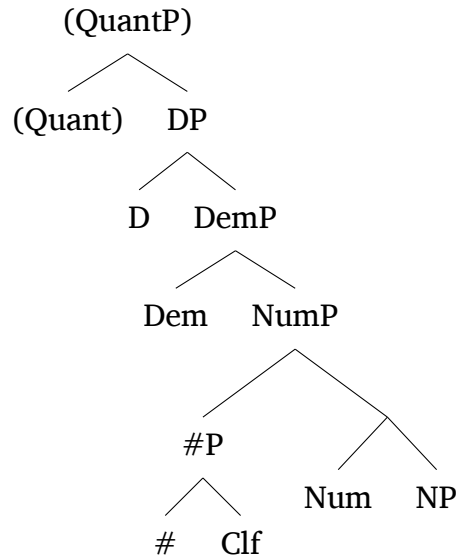
The structure in Figure 4.2 is evidenced by a separate morpheme, optionally marking plurals, that appears between numeral–classifier sequences and nouns. This separate morpheme overtly instantiates the separate Num-head. Jensen (2013: 36) argues that the Jarai form *khul* overtly spells out the Num-head in (4.10a). Eastern Cham has a likely cognate *kʰɔl* in this position (4.10b). This form is marginal in modern colloquial speech and is restricted to pronouns. However, it marks plural nouns in general in formal speech and in earlier manuscripts (cf. Aymonier & Cabaton 1906: 94 on <khu̯ɔʔl>). A more common form in this position is *nɛh*.

(4.10) a. [_DP_ rɔma ƀĕ (khul)   kɔyâo anŭn   ]
              five CLF PL.NHUM   tree DEM.MED

        'those five trees' (Jensen 2014: 36)   JARAI

     b. [_DP_ mɨ jaŋ   (%kʰɔl/nɛh) hɨ   ]
             five CLF.PERSON   PL/PL   2SG

        'the five of you'   EASTERN CHAM

With the DP spine above, the surface word order in Eastern Cham and Jarai can be derived by movement of the Numeral Phrase (NumP) to Spec-DP (Figure 4.3). This accounts for the DP-finality of demonstratives and determiners. In an Agree framework, NumP-movement can be derived through a probe on D that searches for the Num feature.

Figure 4.3: NumP-movement for Eastern Cham, Jarai

```
                      (QuantP)
                      /      \
                 (Quant)     DP
                            /   \
                        NumP      \
                       /    \      D        DemP
                   #P        \   [uNum]     /    \
                  /  \   Num   NP  [EPP]  Dem   NumP
                 #   Clf  ↑
```
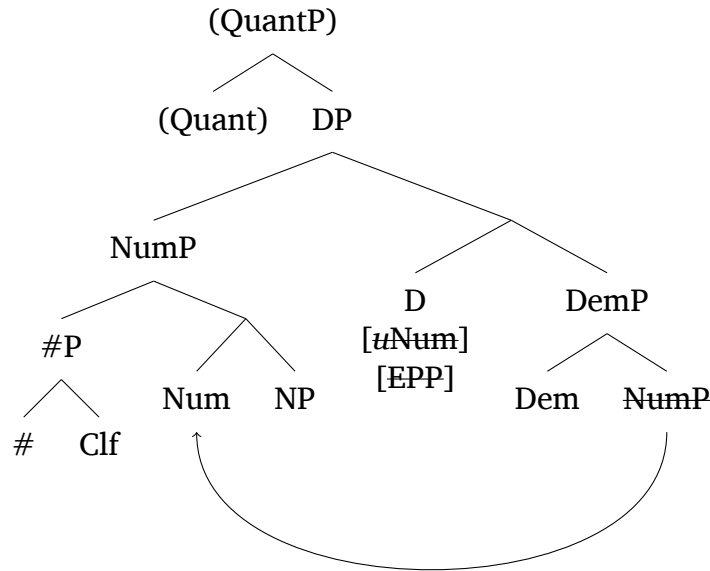
Exactly why NumP-movement is involved in DP structure is yet unclear and requires further comparative and diachronic research. Nevertheless, NumP-movement follows from the phrase structure assumed here and proposed for Jarai. For now, I posit that the most basic D-head in the lexicon of these languages contains a Num-probe and an EPP-feature. Additional evidence for NumP-movement is that there exist structures in which it does not occur, because the D-head bears a probe with a different specification. The following sections find that there exist other D-heads that do not trigger NumP-movement. The relevant constructions are introduced below.

4.1.2   Deviations from basic DP structure

This chapter focuses on two deviations from basic DP order: inventory forms and partitives. Each of these deviations involves some constituent at the left edge of the DP. Pragmatically, these constituents have been described as having topic-like or context-bound properties in prior literature on comparable constructions in other languages.

4.1.2.1   Inventory forms

The first deviation from basic DP structure is the inventory form. Inventory forms typically consist of a bare noun dislocated to the left edge of the DP and appear most often in

list contexts (Greenberg 1975). Example (4.11a–b) gives examples from Indonesian and English.

(4.11)   a. saya mau membeli [$_{DP}$ beras dua kilo ~~beras~~]
         I   want buy   rice two kilo
         'I want to buy two kilos of rice.' (Simpson 2005: (21–22))   INDONESIAN

         b. Sugar, three pounds. Bread, two loaves. Wine, four bottles.

Inventory forms are common cross-linguistically, though they appear to vary in frequency and productivity. In English, they are restricted to list contexts such as recipes and grocery lists, while in Indonesian and Eastern Cham, they are common outside of such contexts. Inventory forms have been described in terms of topic and focus, such that the dislocated noun must be topical, and the remainder of the DP is focussed (Simpson 2005: 97). In the Indonesian example, *beras* 'rice' must be old or otherwise topical information, and in the English example, the superset of ingredients or groceries must be the topic around which the list is organized.

(4.12) kăw plĕj [$_{DP}$ **ʔɔʔ**   çŭh pɔh   ~~ʔɔʔ~~]
       1SG buy   mango 7   CLF.ROUND
       'I bought mangoes, seven.'

The following section argues that the dislocated nouns in Eastern Cham inventory forms, such as *ʔɔʔ* 'mango' must be discourse connected (DC), and their position is derived by DP-internal DC-movement, as indicated in (4.12) above.

4.1.2.2 Partitives

Second are partitives, which are string-identical to inventory forms, except for the presence of a demonstrative in the constituent at the left edge of the DP. Interpretively, partitives consist of two referents: a contextually bound set, such as *ʔɔʔ năn* 'those mangoes' in (4.13), and the whole partitive, which denotes a subpart of that contextual set.

(4.13) kăw plĕj [$_{DP}$ **ʔɔʔ**   **năn** çŭh pɔh   ~~ʔɔʔ năn~~]
       1SG buy   mango that 7   CLF.ROUND
       'I bought seven of those mangoes.'

Much like inventory forms, partitives are known to contain a topic-like or context-bound constituent. In this case, it is a contextually given or salient individual or plural individual (e.g. Enç 1991). Again like inventory forms, the subpart of the DP referring to that set is at the left edge of the DP. Section 4.3.2 argues that the syntax of these partitives is derived by a DC-probe on D, the same one involved in inventory forms. In the case of partitives, it is an embedded DP that is DC-marked and moved to the left edge of the matrix DP. Section 4.3 also argues that the construction in (4.13) is the only true partitive in Eastern Cham that shares cross-linguistic characteristics associated with partitives.

4.1.2.3 Summary

In this section, we have examined two constructions: inventory forms and partitives. In both, there is a subpart of the DP which has a topic-like or context-bound pragmatics and appears at the left edge of the DP. These two observations are related. In the next sections, it is found that phrases are merged at the left edge of the DP if they have DC pragmatics. The following sections examine the pragmatic and syntactic properties of each construction in turn.

4.2 Inventory forms

First, this section examines inventory forms (IFs). IFs provide the clearest case of DP-internal movement driven by discourse connectedness (DC). They also give evidence that DC is an Ā-feature due to the CP/DP parallelism; there is a DC-probe on D, just as the DC-probe on C posited in previous chapters.

Before we proceed, more detail about the basic characteristics of IFs must be given. Cross-linguistically, IFs are described as constructions where a noun is dislocated to the edge of a DP from its base position, typically in list contexts. In Eastern Cham, IFs occur when the noun is dislocated to the left edge. They are visible when there is material to the left of the noun in the DP, such as numeral–classifier sequences (4.14a), quantifiers (4.14b), and measure phrases (4.14c).

(4.14)   a. kăw plĕj [DP **ʔɔ̌ʔ** cŭh pɔh ʔɔ̌ʔ]
          1SG b̥uy   mango 7   CLF.ROUND
        'I bought mangoes, seven.'

     b. kăw plĕj [DP **ʔɔ̌ʔ** pih ʔɔ̌ʔ]
        1SG b̥uy   mango all
       'I bought mangoes, all.'

     c. kăw plĕj [DP **ʔɔ̌ʔ** cŭh kiʔ ʔɔ̌ʔ]
        1SG b̥uy   mango 7   kg
       'I bought mangoes, seven kilograms.'

When there is no material to the left of the noun, such as NP–demonstrative sequences like *ʔɔ̌ʔ năn* 'that mango', there is no surface evidence of leftward movement of the noun. At present, it is unclear if an IF can occur in these structures.

As in other languages, IFs in Eastern Cham do occur in list contexts. For example, (4.15) gives a list of items the speaker bought at a market. In languages like English, IFs are highly restricted to shopping list or recipe contexts. In Eastern Cham, this is not the case. As the next subsection will show, IFs occur when the noun satisfies the DC conditions, which results in a broader set of possible contexts.

(4.15) kăw plěj [$_{DP}$ **ʔɔ̌ʔ** çŭh pɔh ʔɔ̌ʔ] [$_{DP}$ **hɔŋ̊m** mɨ pɔh
   1SG buy   mango 7 CLF.ROUND   papaya 5 CLF.ROUND

  ~~hɔŋ̊m~~] hɔ̌ŋ̊m [$_{DP}$ **na**   nam pɔh   ~~na~~]
      with   pineapple 6 CLF.ROUND

  'I bought mangoes, seven, papaya, five, and pineapples, six.'

The reader may notice that (4.15) is a contrastive topic (CT) context. In CT contexts, there is a complex question under discussion, such as *What, and how many of each thing, did you buy?* One set is picked out as the set around which the answer is organized. In this case, that is the set of things bought at the market. Individual mentions from this set become, in the answer, CTs. In Chapter 2, Section 2.1.1, it was observed CTs cannot be DC-marked. Yet, in this one case of DP-internal movement, they can. Chapter 5 will examine how in this case, CT- and DC-marking is allowed in the same DP.

In the remainder of this section, the pragmatics of IFs are shown to involve DC-marking of the noun to the exclusion of the DP (Section 4.2.1). IFs are also shown to involve DP-internal DC-movement driven by a DC-probe on D, analogous to the DC-probe on C. This conclusion is based on DP-internal syntax (Section 4.2.2) and the interaction of IFs and the CP periphery (Section 4.2.3).

4.2.1 Pragmatics of IFs

This section examines the pragmatics of inventory forms in Eastern Cham. Inventory forms are found to be licensed only when the noun satisfies the conditions of *discourse connectedness* (DC) to the exclusion of the remainder of the DP. Before we examine inventory forms, it is worth reexamining the DC conditions and what it would mean for them to be satisfied within a DP. As laid out in Chapters 2–3, DC is defined in terms of relations between sentences in a discourse, the events introduced in those sentences, and their participants. Specifically, (4.16a–c) define the set of live events in the context, $\mathcal{E}_c$, and the set of events inferred by subordinating discourse relations to the current sentence, $\mathcal{E}_e$. Additionally, $\mathcal{P}_e \cap \mathcal{P}_{e'}$ is defined as the intersection of the sets of participants in the current event and a prior event in $\mathcal{E}_e$. An individual, x satisfies the DC conditions if it is in that intersection.

(4.16)   a. Let $\mathcal{E}_c$ be the set of events live in a discourse at context c

    b. Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

    c. Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

    d. Let $\mathcal{P}_e$ be the set of participants in event e

    e. $[\![\text{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

A basic example of DC-movement is given in (4.17). Sentence (4.17b) elaborates on (4.17a), indicated by (a $\Downarrow$ b). Therefore, from the perspective of (4.17b), the event of cooking in (4.17a) is in $\mathcal{E}_e$. Additionally, both the events of cooking in (4.17a) and (4.17b) contain *ʔiŋ ʔɔŋ nằn* 'that frog'. Chapter 2, Section 2.1.2 argues that this is what licenses the movement operation marked in (4.17b).

(4.17)   a. tʰuːŋ̊m³¹²   t̥ɔʔ   ŋ̆ăʔ   **ʔiŋ ʔɔŋ   nằn**
         Thuận(VN)   PROG   make   frog   that
         'Thuận is cooking that frog.'

     b. **ʔiŋ ʔɔŋ   nằn**$_{DC}$ ɲu   t̥ɔʔ   ŋ̆ăʔ   ~~ʔiŋ ʔɔŋ nằn~~   ŋi   lo
        frog   that   3.ANIM   PROG   make   be.delicious   very
        'That frog, he is cooking very well [Lit: deliciously].'   (a $\Downarrow$ b)

     b'. $g(x) \in \mathcal{P}_e \cap \mathcal{P}_{e'}$

Inventory forms consist of two parts: an NP at the left edge of the DP, and a numeral (alongside a numeral classifier or measure phrase) or quantifier. They are felicitous if the NP does satisfy the DC conditions, while the numeral or quantifier phrase does not. In the examples below, the DC conditions are explicitly assessed for the NP and numeral or quantifier phrase, respectively. Noun phrases are indicated by the kinds they denote, such as $^\cap watermelon$ for the noun *watermelon*. Numeral phrases are indicated by an additional # predicate added to the semantic representation of a definite DP, such as $\iota x[watermelon(x) \wedge \#(x) = 3]$ for *three watermelons*, following work on numeral semantics (cf. Kennedy 2013 and references therein).

Consider the polar question in (4.18a). Abstracting away from the semantics of questions and the serial verb construction (i.e. *go…buy*), the polar question contains an event of buying such that the agent is *hɨ* '2SG' and the theme is *tamkaj* 'watermelon'. This event semantics in (4.18a') follows a Neo-Davidsonian event semantics (cf. Champollion 2015 for a comparison of different models of event semantics in compositional semantics).

(4.18)   a. hɨ   naw   z̥aʔ   hu   plĕj   **tamkaj**   lĕj
         2SG   go   market   EXIST   b̥uy   watermelon   Y/N.Q
        Q: 'Did you go to the market and buy watermelons?'

     a'. $[\![(4.18a)]\!]$ $=$ …$\exists e[\mathbf{agent}(e) = \mathbf{2SG} \wedge \mathbf{buying}(e) \wedge \mathbf{theme}(e) = {}^\cap watermelon]$

     b. kăw   plĕj   [$_{DP}$ **tamkaj**$_{DC}$ klăw   pɔh   ~~tamkaj~~]
        1SG   b̥uy   watermelon   three   CLF.ROUND
        A: 'I bought watermelon, three.'

     b'. (a $\Downarrow$ b)   DC ✓

     b''. $^\cap watermelon \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

     b'''. $\iota x[watermelon(x) \wedge \#(x) = 3] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

The answer (4.18b) presents a positive example of inventory form pragmatics. There are three components that must be satisfied for an inventory form to be licensed in Eastern Cham. First, the answer (4.18b) elaborates upon the polar question. A direct answer would be of the form 'Yes, I bought watermelons'. Instead, the speaker implicitly answers the question and then elaborates on the watermelon with a specific number they bought (cf. Chapter 2, Section 2.1.1 on elaborating answers). Therefore, the answer is discourse subordinate to the question (i.e. 4.18a ⇓ 4.18b). This is denoted by (a ⇓ b) in this and subsequent examples. Second, the kind *tamkaj* 'watermelon' is also mentioned in the superordinate event of buying. This is indicated by $^\cap watermelon \in \mathcal{P}_e \cap \mathcal{P}_{e'}$.

Third, the remainder of the inventory form, including the numeral phrase, is not previously mentioned in the superordinate event. Nowhere in the superordinate sentence is 'three watermelons' mentioned. This is indicated by $\iota x[watermelon(x) \wedge \#(x) = 3] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$, which is intended to refer to the denotation of the whole DP consitituent. Note that the numeral predicate $\#(x) = 3$ on its own is not assessed in these examples, as it is not interpreted as an individual and therefore cannot be DC-marked. The DC conditions here are only computed for the numeral with regard to $watermelon$.

All three of these components are necessary to license inventory forms, denoted by "DC ✓". Subsequent examples will mark each of these three components and whether they do (✓) or do not (✗) predict DC-marking in the inventory form.

Inventory forms are infelicitous if the whole DP satisfies the DC conditions. In (4.19), there is again an elaborating answer to a polar question; hence discourse subordination. However, the whole DP, *klăw pɔh tamkaj* 'three watermelons' is mentioned in the superordinate event (that of the polar question). It is not the case that the noun satisfies the DC conditions to the exclusion of the remainder of the DP. In other words, the minimal difference between (4.19) and (4.18) is the presence of *klăw* 'three' in the superordinate sentence, and that is enough to render the inventory form infelicitous.

(4.19) a. hɨ   ʔɨ̆ŋ   naw zạʔ   plɛ̆j **klăw pɔh**   **tamkaj**   lɛ̆j
           2SG want go   market buy three CLF.ROUND watermelon Y/N.Q

        Q: 'Do you want to go to the market and buy three watermelons?'

    a′. ⟦(4.19a)⟧ =
        ...$\exists e[\mathbf{agent}(e) = \mathbf{2SG} \wedge \mathbf{buying}(e) \wedge \mathbf{theme}(e) = \iota x[watermelon(x) \wedge \#(x) = 3]$

    b. #kăw plɛ̆j [$_{DP}$ **tamkaj**   klăw pɔh   ~~tamkaj~~] jə
        1SG buy   watermelon three CLF.ROUND   already

        INTENDED: A: 'I already bought watermelon, three.'

        b′. (a ⇓ b)   DC ✓

        b″. $^\cap watermelon \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

        b‴. $\iota x[watermelon(x) \wedge \#(x) = 3] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

Instead, the whole DP *klăw pɔh tamkaj* 'three CLF.ROUND watermelons' can be DC-moved to the CP periphery, as in (4.19b′). This is not an inventory form, it is DC-movement of a DP with its basic word order.

(4.20) [$_{\text{DP}}$ **klăw pɔh   tamkaj**] kăw plĕj ~~klăw pɔh tamkaj~~ jə
              one   CLF.ROUND watermelon   1SG b̊uy   already

    A: 'I already bought three watermelons.'   (a ⇓ b)

Inventory forms are also infelicitous in the absence of discourse subordination. In (4.21), the polar question in (4.21a) is directly answered by (4.21b). This is not an instance of discourse subordination, because it is a direct answer to a question, so the context is not sufficient to license DC-marking. This is despite the fact that the kind $^\cap watermelon$ is found in the question. Note that the meaning of (4.21a) is an alternative set of possible answers, only one of which is represented in (4.21a′). The answer chosen to be represented is 'two watermelons', in order to indicate that 'three watermelons' is not sufficiently previously mentioned in the question, but 'watermelon' is.

(4.21)   a. hɨ   plĕj ṭom   **pɔh**   **tamkaj**
             2SG b̊uy   how.many C̊LF.ROUND watermelon

        Q: 'How many watermelons did you buy?'

      a′. $[\![(4.21a)]\!]$ =
          $\ldots\exists e[\mathbf{agent}(e) = \mathbf{2SG} \wedge \mathbf{buying}(e) \wedge \mathbf{theme}(e) = \iota x[watermelon(x) \wedge \#(x) = 2]$

      b. #kăw plĕj [$_{\text{DP}}$ **tamkaj**   klăw pɔh   ~~tamkaj~~]
          1SG b̊uy   watermelon three C̊LF.ROUND

        INTENDED: A: 'I bought watermelon, three.'

        b′. (a ⇍ b)   DC ✗

        b″. $^\cap watermelon \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

        b‴. $\iota x[watermelon(x) \wedge \#(x) = 3] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

In Eastern Cham, thus, inventory forms are only licensed if the noun satisfies the DC conditions to the exclusion of the remainder of the DP.

4.2.2   DP-internal syntax of IFs

This section examines the DP-internal syntax of inventory forms and finds that they involve DP-internal movement driven by a syntactic probe on D. Given that inventory forms function to mark DC, it follows that this probe is a DC-probe identical to that on C proposed in the previous chapters. This DP/CP parallelism aligns DC with other Ā-features. Aboh (2004b) argues that the DP has an expanded left periphery much like CP, with regard to topicalization, focus-movement, and others. Under an Agree analysis, this would predict that the same kinds of Ā-probes can exist on D as they can on C.

The DP/CP parallelism is also found with the *wh*-feature through *pied-piping with inversion,* as laid out in the introduction to this chapter. Recall that possessors typically follow possessees in languages like Tzotzil (4.22a). *Wh*-possessors, though, move to the left edge of the DP before the whole DP undergoes *wh*-movement (4.22b). This initial movement has been analyzed via a *wh*-probe on D, which triggers short-distance movement of the *wh*-possessor (Aissen 1996). If the DC-probe can be merged on D or C, it would provide an additional piece of evidence that DC patterns with Ā-features in general.

(4.22)   a. [$_{DP}$ s-p'in   li   Maruch-e   ]
              A3-pot   the   Maruch-ENC

      'Maruch's pot' (Aissen 1996: 454)   TZOTZIL

   b. [$_{DP}$ **Buch'u** D$_{uwh}$ x-ch'amal ~~buch'u~~] C$_{uwh}$ i-cham ~~DP~~?
             who   A3-child   CP-died

      'Whose child died?' (Aissen 1996: 457)

Before we proceed, recall the Eastern Cham DP syntax proposed in Section 4.1.1, repeated in Figure 4.4 below. Here, a structure is given for the unmarked DP order, *c̥ŭh pɔh ʔɔ̆ʔ năn* 'those seven mangoes'. The unmarked word order is derived by movement of the Numeral/NumP to Spec-DP, driven by a probe on D. This results in the demonstrative *năn* 'that' appearing linearly at the right edge of the DP.

Figure 4.4: Basic DP structure
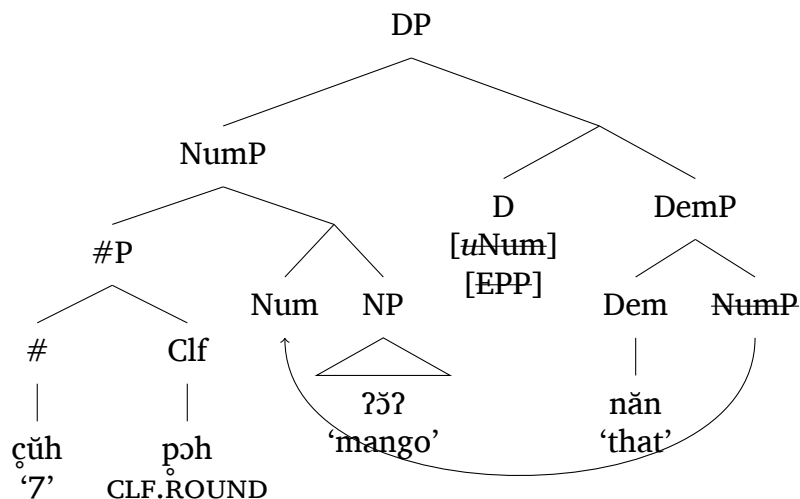
Inventory forms arise, as in the previous section, when an NP is marked as DC, and it moves to the left edge of the DP (4.23).

(4.23)   a. kăw plĕj [DP çŭh pɔ̊h   ʔɔ̆ʔ   ]
            1SG buy   7   CLF.ROUND mango

         'I bought seven mangoes.'

         b. kăw plĕj [DP **ʔɔ̆ʔ**DC   DuDC çŭh pɔ̊h   ~~ʔɔ̆ʔ~~]
            1SG buy   mango   7   CLF.ROUND

         'I bought mangoes, seven.'   (Inventory form)

Based on cross-linguistic evidence in Southeast Asian languages, Simpson (2005) analyzes inventory forms as NP-movement in the absence of NumP-movement. Figure 4.5 adapts this proposal to an Agree framework. In this derivation, there is a different D-head with a DC-probe instead of a NumP-probe. In the case of an inventory form, only the NP is assigned the DC feature. As a result, only the NP is moved to the DP left periphery.

Figure 4.5: Derivation of inventory form (4.23b)
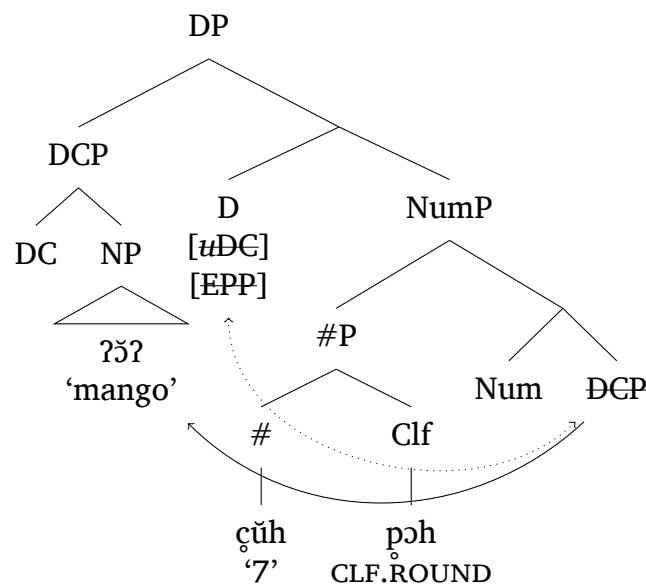
This analysis makes a variety of predictions that will be elaborated upon in this and the subsequent sections. First, demonstratives should be ungrammatical at the right edge of inventory forms, because there is no NumP-movement enabling them to appear in that position. This prediction is borne out: demonstratives are ungrammatical phrase-finally in Eastern Cham inventory forms (4.24b). Note that Eastern Cham is unlike other languages such as Indonesian in this regard, in that they do apparently allow for DP-final demonstratives (4.24c). Also note that demonstratives are ungrammatical DP-medially. They cannot be pronounced in their base position in inventory forms, as that base position is DP-medial. Section 4.3.2 goes into more detail on the DP-final requirement.

(4.24) a. kăw plĕj [ᴅᴘ çŭh pɔh   ʔɔ̌ʔ   năn]
            1sɢ buy   7   ᴄʟꜰ.ʀᴏᴜɴᴅ mango that
          'I bought those seven mangoes.'

       b. *kăw plĕj [ᴅᴘ **ʔɔ̌ʔ**ᴅᴄ çŭh pɔh   ~~ʔɔ̌ʔ~~ năn]
            1sɢ buy   mango 7   ᴄʟꜰ.ʀᴏᴜɴᴅ   that
          ɪɴᴛᴇɴᴅᴇᴅ: 'I bought those mangoes, seven.'

       c. … mengerjakan [**lobang** sa   buah itu] sampai lima enam hari
          … make   hole   one ᴄʟꜰ that took   5   6   day
          'Indeed it took 5 or 6 days just to dig that one hole.'   ɪɴᴅᴏɴᴇꜱɪᴀɴ
          (Simpson 2005: (16))

Second, the NP in an inventory form should function as a left-branch extraction island, as it has been moved to a specifier position. By contrast, the remainder of the DP would not function as an extraction island on its own. *Wh*-in situ confirms this configuration. Recall from Chapter 3, Section 3.4 that in situ *wh*-phrases are sensitive to island constraints. *Wh*-phrases cannot be embedded inside an extraction island, as they must enter into an Agree relation with C, and that Agree relation is sensitive to island constraints. Figure 4.6 illustrates this prediction. In Figure 4.6a, C cannot Agree with a *wh*-phrase embedded inside the NP in Spec-DP. But in Figure 4.6b, C can Agree with a *wh*-phrase inside the NumP.

Figure 4.6: *Wh*-probes and inventory forms

(a) *Wh*-in situ inside NP

(b) *Wh*-in situ inside NumP
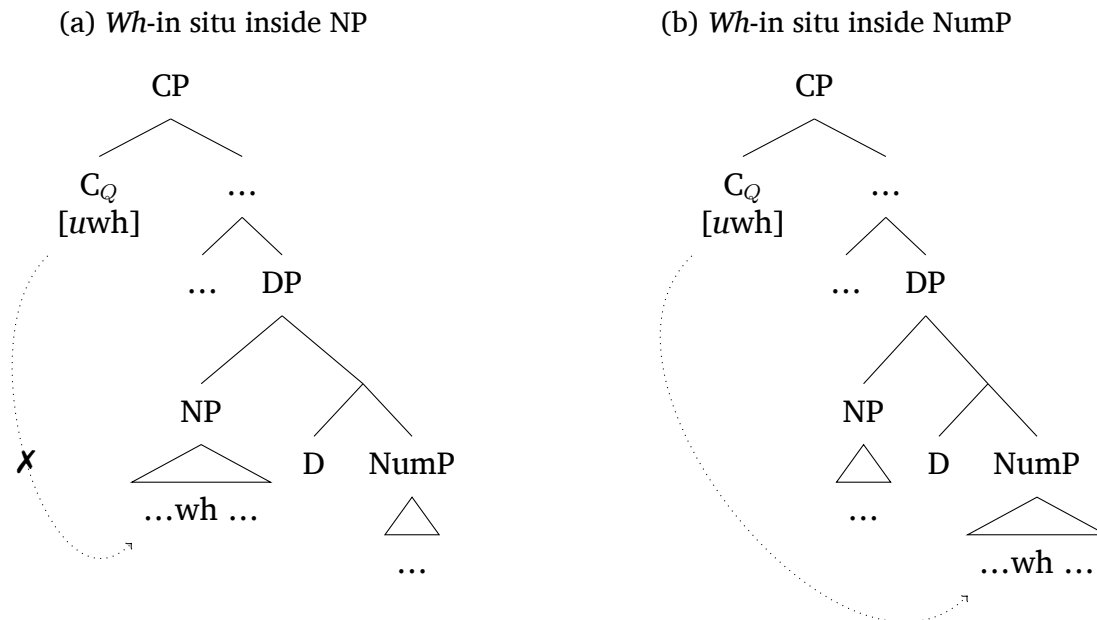
This prediction is borne out in Eastern Cham. It is ungrammatical for a *wh*-phrase to be embedded inside the NP of an inventory form, as in the possessive in (4.25a; *wh*-phrases bolded). Note that a non-*wh* possessor is grammatical. By contrast, the *wh*-word ṭom 'how many' can be merged in the Num position in an inventory form, as in (4.25b). Presumably in this case, C$_Q$ enters into an Agree relation with the in situ phrase ṭom 'how many'.

(4.25) a. *hɨ   plĕj [$_{DP}$ [$_{NP}$ pɔh zăw [**tʰɛ̆j**]] tʰa   kiʔ]
              2SG buy   fruit   who   one kg
         INTENDED: 'Whose fruit did you buy one kilogram of?'

      b. hɨ   plĕj [$_{DP}$ [$_{NP}$ ʔɔ̆ʔ   ] **ṭom**   kiʔ]
         2SG buy   mango   how.many kg
         'How many kilograms of mango did you buy?'

      c. *hɨ   plĕj [$_{DP}$ [$_{NumP}$ tʰa kiʔ pɔh zăw [**tʰɛ̆j**]] D]
              2SG buy   one kg fruit   who
         INTENDED: 'Whose fruit did you buy one kilogram of?'

It is also worth noting that a *wh*-phrase cannot be merged inside the NumP in the basic DP order, as in (4.25c). In this case, the *wh*-phrase $t^h\check{e}j$ 'who' is embedded inside an NumP that has been moved to Spec-DP and becomes an extraction island.

Based on these two sections, the CP/DP parallelism exists for DC in Eastern Cham: similar probes on C and D can drive DC-movement. This raises the question as to whether inventory forms interact with the CP periphery. Section 4.2.3 finds that this is the case. The noun in inventory forms can also be merged in the CP periphery if it satisfies the DC pragmatic conditions to the exclusion of the rest of the DP.

4.2.3 IFs and the CP periphery

In this last subsection, I examine the interaction between inventory forms and DC-marking at the CP periphery. Here, we find direct evidence that syntactic probes on C and D can both search for features that index DC pragmatics. Before proceeding, consider again the Tzotzil example, repeated below as (4.26). This example illustrates the CP/DP parallelism for the *wh*-feature, as the *wh*-phrase *buch'u* 'who' is moved to the left edge of the DP, and the whole DP is moved to the left edge of the CP.

(4.26)   a. [$_{DP}$ s-p'in   li   Maruch-e   ]
            A3-pot   the   Maruch-ENC

      'Maruch's pot' (Aissen 1996: 454)   TZOTZIL

   b. [$_{DP}$ **Buch'u**   x-ch'amal   ~~buch'u~~]   i-cham   ~~DP~~?
         who   A3-child   CP-died

      'Whose child died?' (Aissen 1996: 457)

As much as this example illustrates the CP/DP parallelism, there is a tension with another hypothesized fact about many Ā-features: they can only enter into one Agree relation (e.g. Rizzi 2010 on criterial freezing). In the example above, how can there be two instances of *wh*-movement if there is only one *wh*-feature? Cable (2010: 186) accounts for the Tzotzil example by hypothesizing two Ā-movement operations involving related, but not identical syntactic features: Q and *wh*. Under this analysis, the *wh*-phrase *buch'u* 'who' bears a *wh*-feature. That feature Agrees with D_{wh}. Then, Cable's (2010) Q-particle merges with the whole DP and agrees with C_Q.

(4.27) [$_{DP}$ **Buch'u**$_{wh}$ D_{wh} x-ch'amal   ~~buch'u~~]$_Q$ C_Q i-cham   ~~DP~~?
         who   A3-child   CP-died

   'Whose child died?'

If Eastern Cham inventory forms involve the same kind of DC-marking as phrases merged in the CP periphery, we might expect an interaction between the two. However, we have only posited one solitary DC-feature, not a dyad like Q and *wh*. Therefore, we would not expect DC-movement of the same phrase to Spec-DP and to Spec-CP in the

same derivation. This is in fact borne out: the NP in an inventory form can be merged either in Spec-DP or Spec-CP, but not both.

Turning to the data, (4.28) presents a context that licenses inventory forms in general, as in (4.28b). Here, the NP *ʔɟʔ* 'mango' moves to Spec-DP from its base position. In the same context, *ʔɟʔ* can just as well be merged in Spec-CP, as shown in (4.28b′). In this example, the base position of *ʔɟʔ* 'mango' is filled by a null NP-pronoun with an Op-feature. Later in this section, evidence from locality effects will demonstrate that this sentence does not involve Ā-movement, but base generation of *ʔɟʔ* in Spec-CP and binding of the null pronoun much like base generation + Agree in Chapter 2, Section 2.3.2.

(4.28) a. hɨ   naw z̥aʔ   hu   plɛ̆j **ʔɟʔ**   lɛ̆j
        2SG go   market EXIST   buy   mango   Y/N.Q

    Q: 'Did you go to the market and buy mangoes?'

   b. kăw plɛ̆j [$_{DP}$ **ʔɟʔ**$_{DC}$ c̥ŭh pɔh   ʔ̶ɟ̶ʔ̶]
      1SG buy   mango 7   CLF.ROUND

    A: 'I bought mangoes, seven.'

   b′. **ʔɟʔ**$_{DC,i}$ kăw plɛ̆j [$_{DP}$ c̥ŭh pɔh   *pro$_i$*]
       mango   1SG buy   7   CLF.ROUND pro

    A: 'Mangoes, I bought seven.'

In general, if a context licenses an inventory form, it also licenses merging of the NP in Spec-CP. Conversely, in contexts that fail to license inventory forms, the NP cannot be merged in Spec-CP. (4.29) presents such a context. Taken together, the association between Spec-CP and DC pragmatics from Chapters 2–3 and the association of Spec-DP with DC pragmatics from the previous sections, it follows that there is a true syntactic parallelism between the two positions.

(4.29) a. hɨ   naw z̥aʔ   hu   plɛ̆j **c̥ŭh** pɔh   **ʔɟʔ**   lɛ̆j
        2SG go   market EXIST   buy   7   CLF.ROUND mango   Y/N.Q

    Q: 'Did you go to the market and buy seven mangoes?'

   b. #kăw plɛ̆j [$_{DP}$ **ʔɟʔ**$_{DC}$ c̥ŭh pɔh   ʔ̶ɟ̶ʔ̶]
      1SG buy   mango 7   CLF.ROUND

    INTENDED: A: 'I bought mangoes, seven.'

   b′. #**ʔɟʔ**$_{DC,i}$ kăw plɛ̆j [$_{DP}$ c̥ŭh pɔh   *pro$_i$*]
       mango   1SG buy   7   CLF.ROUND pro

    INTENDED: A: 'Mangoes, I bought seven.'

Unlike *wh*-phrases in Tzotzil, though, entire inventory forms in Eastern Cham cannot be merged in Spec-CP; they cannot be pied-piped. As demonstrated in (4.30a), there is nothing syntactically illicit about a numeral phrase being merged in Spec-CP. However,

an inventory form containing an NP in a derived position (4.30b) is ungrammatical, regardless of context.

(4.30) a. [_{DP} çŭh pɔ̥h ʔɔ̌ʔ] kăw plĕj <s>çŭh pɔ̥h ʔɔ̌ʔ</s>
        7 CLF.ROUND mango 1SG buy
    'Seven mangoes, I bought.'

  b. *[_{DP} **ʔɔ̌ʔ**_{DC} çŭh pɔ̥h ʔɔ̌ʔ] kăw plĕj <s>ʔɔ̌ʔ çŭh pɔ̥h</s>
        mango 7 CLF.ROUND 1SG buy
    INTENDED: 'I bought mangoes, seven.'

This observation follows if a phrase can only be associated with one DC-related feature in a derivation, unlike *wh*, which has a Q-feature counterpart. In (4.30b), there is only one DC-feature, and it is merged with *ʔɔ̌ʔ* 'mango'. Once the DC-feature is checked by D, it can no longer enter an Agree relation with C. Note that there is no evidence for a D-head bearing only an EPP-feature, unlike what was proposed for embedded C-heads in Chapter 2, Section 2.3.3. As a result, DC-movement to Spec-CP transiting through Spec-DP would not be expected, as it would have to involve two instances of checking a DC-feature on the same phrase.
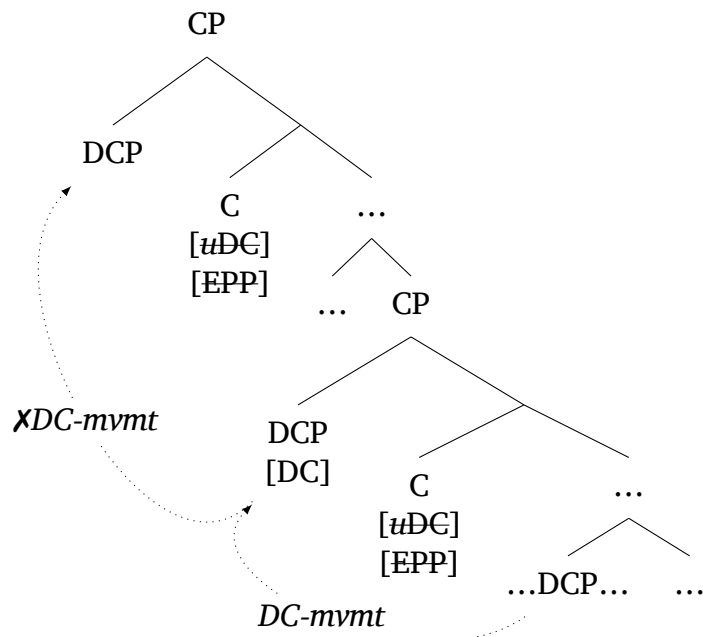
A corroborating analysis was raised in Chapter 2, Section 2.3.3: long-distance DC-movement cannot be mediated by embedded DC-probes; instead, it is mediated by intermediate probes with only EPP-features. There, only the topmost C-head was found to bear a DC-probe. Thus, for a given phrase in Eastern Cham, there can only be one DC-feature which can only enter into one Agree relation. A schematization of the avoidance of double-DC-marking involving inventory forms and embedded clauses is given in Figure 4.7 below.

Figure 4.7: Absence of double-DC-marking

(a) DC-movement to DP, then CP

CP

DCP

C
[*u*DC]
[EPP]

...

DP   ...

✗*DC-mvmt*

DCP
[DC]

D
[*u*DC]
[EPP]

...

...DCP...   ...

DC-mvmt
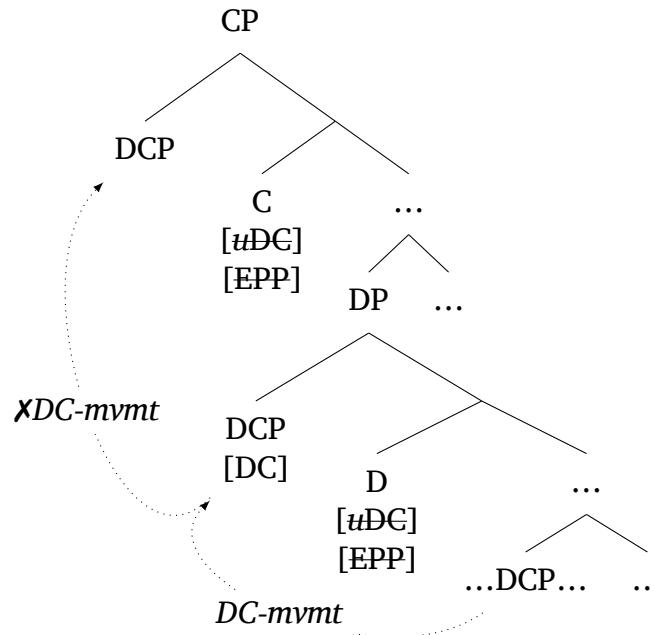
(b) DC-movement to embedded CP, then matrix CP

CP

DCP

C
[*u*DC]
[EPP]

...

...   CP

✗*DC-mvmt*

DCP
[DC]

C
[*u*DC]
[EPP]

...

...DCP...   ...

DC-mvmt

The example (4.28b) above, then, raises a question: when an NP is merged in Spec-CP from a putative inventory form, does it transit through Spec-DP? Given that there can only be one DC-probe in a derivation, we would not predict that this is the case. Indeed, evidence from path containment effects indicate that such NPs are base generated in Spec-CP; they do not transit through Spec-DP.

Recall from Chapters 2–3 that a path containment effect arises when multiple phrases are DC-moved to the same clausal periphery. In (4.31a), the movement path of *nĭʔ mĕj sĭt năn* 'that little girl' is contained inside that of *han ni* 'this cake'. The result in Eastern Cham is grammatical. But when the movement paths are crossed, as in (4.31b), the result is ungrammatical. Chapter 2, Section 2.3.1 argued that this pattern is the result of two C-heads each with a probe searching for a DC-feature.

(4.31) a. **han   ni**   nĭʔ   mĕj   sĭt   năn   tʰuːɲm͡³¹²   ʔa   n̶ĭ̶ʔ̶ ̶m̶ĕ̶j̶ ̶s̶ĭ̶t̶ ̶n̶ă̶n̶
            cake this child female small that Thuận   invite
         maj   ɓăŋ   **h̶a̶n̶ ̶n̶i̶**
         come eat
         'This cake, Thuận invited that little girl to come eat.'

        b. *nĭʔ   mĕj   sĭt   năn   **han   ni**   tʰuːɲm͡³¹²   ʔa   n̶ĭ̶ʔ̶ ̶m̶ĕ̶j̶ ̶s̶ĭ̶t̶ ̶n̶ă̶n̶
            child female small that cake this Thuận   invite
         maj   ɓăŋ   **h̶a̶n̶ ̶n̶i̶**
         come eat
         INTENDED: 'This cake, Thuận invited that little girl to come eat.'

By contrast, if one phrase is DC-moved and one base generated, the path containment effect is absent. In the sentences in (4.32), the DP *nĭʔ mĕj sĭt năn* 'that little girl' is DC-moved, while *jŭt ni* 'this friend' is base generated, as evidenced by the resumptive pronoun *ɲu*. The result is grammatical whether the paths are nested (4.32a) or crossed (4.32b). Chapter 2, Sections 2.3.1–2.3.2 argues that the ungrammaticality is alleviated because there are two C-probes searching for different features: one DC and one Op, which binds the pronoun to the base generated phrase.

(4.32) a. nĭʔ   mĕj   sĭt   năn   **jŭt   ni**   tʰuːɲm͡³¹²   ʔa   ɲu   maj
            child female small that friend this Thuận   invite 3.ANIM come
         ɓăŋ   **j̶ŭ̶t̶ ̶n̶i̶**
         eat
         'This friend, Thuận invited that little girl to come meet.'

        b. nĭʔ   mĕj   sĭt   năn   **jŭt   ni**   tʰuːɲm͡³¹²   ʔa   n̶ĭ̶ʔ̶ ̶m̶ĕ̶j̶ ̶s̶ĭ̶t̶ ̶n̶ă̶n̶
            child female small that friend this Thuận   invite
         maj   ɓăŋ   **ɲu**
         come eat   3.ANIM
         'This friend, Thuận invited that little girl to come meet.'

Turning back to inventory forms, if the apparent subextraction of the NP were DC-movement, a path containment effect with other instances of DC-movement would be predicted. If it were base generation, we would not predict this effect to arise. In fact, the path containment effect is absent. Nested paths lead to grammaticality (4.33a), as do crossed paths (4.33b). This can most readily be explained if the NP *han* 'cake' is base generated in Spec-CP, and it binds a null pronoun in the object base position.

(4.33) a. **han** nǐʔ měj sǐt năn tʰuːŋ͡m³¹² ʔa ~~nǐʔ měj sǐt năn~~ maj
  cake child female small that Thuận invite come
  ɓăŋ **klăw klɛh**
  eat three piece
  'Thuận invited that little girl to come eat cake, three (pieces).'

b. nǐʔ měj sǐt năn **han** tʰuːŋ͡m³¹² ʔa ~~nǐʔ měj sǐt năn~~ maj
  child female small that cake Thuận invite come
  ɓăŋ **klăw klɛh**
  eat three piece
  'Thuận invited that little girl to come eat cake, three (pieces).'

As corroborating evidence, if the NP were base generated in Spec-CP, we would also predict path containment effects to arise with another base generated DC-phrase. This is because there would be two C-probes both searching for an Op-feature. This is indeed the case. In (4.34), there are two movement paths: one with the NP *han* 'cake' binding the inventory form, and one with *nǐʔ měj sǐt năn* 'that little girl' binding a resumptive pronoun. When the movement paths are nested, the result is grammatical (4.34a), but when the paths are crossed, the result is ungrammatical (4.34b).

(4.34) a. **han** nǐʔ měj sǐt năn tʰuːŋ͡m³¹² ʔa ɲu maj ɓăŋ
  cake child female small that Thuận invite 3.ANIM come eat
  klăw klɛh
  three piece
  'Thuận invited that little girl to come eat cake, three (pieces).'

b. *nǐʔ měj sǐt năn **han** tʰuːŋ͡m³¹² ʔa ɲu maj ɓăŋ
  child female small that cake Thuận invite 3.ANIM come eat
  klăw klɛh
  three piece
  INTENDED: 'Thuận invited that little girl to come eat cake, three (pieces).'

It is worth mentioning that there is no semantic question about what is being referred to in the two movement paths in the examples above. Therefore, it does not appear to be the case that the examples involve garden path interpretations, and that the resumptive

pronouns and numerals purely serve to aid with processing (e.g. Alexopoulou & Keller 2007).

To summarize, NPs in inventory forms can appear in the CP periphery, because they are DC-marked. To do so, they must be base generated in that position, as the DC-feature cannot be checked twice. Base generation of the NP patterns with the base generation described in Chapter 2, Section 2.3.2: a DC-probe on C probes for an Op-feature and binds an in-situ pronoun with the DC-phrase in Spec-CP, as in Figure 4.8.

Figure 4.8: Base generation of NP in Spec-CP
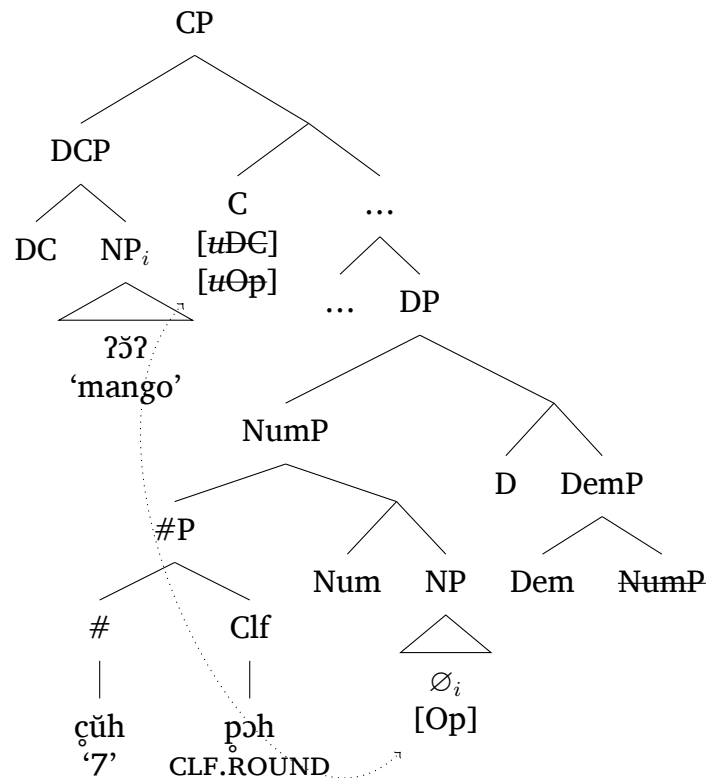
In Figure 4.8, there is a numeral–classifier sequence followed by a null pronoun. This null pronoun construction is independently attested in Eastern Cham. When the antecedent is accessible in the context, *çŭh pɔh* 'seven CLF.ROUND' can refer to '7 mangoes'. In the tree, there is no DP-internal DC-movement. Instead, there is the unmarked NumP-movement. This is the most parsimonious option, as overt pronouns cannot generally undergo DC-movement (cf. Chapter 2, Section 2.1.3).

4.2.4 Summary

This section has demonstrated that inventory forms in Eastern Cham function to DC-mark an NP to the exclusion of a numeral or quantifier. This raises a new possible analysis of inventory forms cross-linguistically. In terms of syntax, a DC-probe on D drives DP-internal DC-movement. When an NP undergoes this movement, an inventory form arises. This probe is identical to the C-probe proposed in the previous sections, supporting the CP/DP parallelism proposed for Ā-features in general. The interaction between inventory forms and the CP periphery further provides evidence for the parallelism, given the additional restriction that there is only one DC-related feature in the syntax. The next section turns to the partitive, another construction that involves DP-internal DC-marking and supports the CP/DP parallelism.

4.3 Partitives

This section examines the pragmatics of true partitives in Eastern Cham. Much like inventory forms, partitives involve DC-marking of a subpart of a DP, such as *ʔɔ̆ʔ năn* 'those mangoes' in (4.35), and they are derived by a DC-probe on D, which triggers DP-internal DC-movement. The only difference from inventory forms is that an embedded DP, not an NP is DC-marked. Partitives in Eastern Cham, then, primarily serve to distinguish a subpart of a DP as DC.

(4.35) kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̆ʔ** **năn**$_{DC}$] çŭh pɔh ~~ʔɔ̆ʔ năn~~]
    1SG buy   mango that   7   CLF.ROUND
     'I bought seven of those mangoes.'

Before proceeding, it must be established that the construction above is a true partitive. In contrast to the gloss above, the unmarked DP order cannot be interpreted as a partitive (4.36a). The numeral cannot be interpreted as referring to a subset of the set of mangoes. Measure phrases can have a pseudopartitive interpretation (4.36b), but they too cannot have a partitive interpretation with a demonstrative (4.36c).

(4.36)   a. kăw plĕj [$_{DP}$ çŭh pɔh ʔɔ̆ʔ năn]
         1SG buy   7   CLF.ROUND mango that
         'I bought those seven mangoes.' / *'I bought seven of those mangoes.'

     b. kăw plĕj [$_{DP}$ tʰa li^{35} ʔja cɛ]
         1SG buy   one cup(VN) water tea
         'I bought one cup of tea.'

     c. kăw plĕj [$_{DP}$ tʰa li^{35} ʔja cɛ năn]
         1SG buy   one cup(VN) water tea that
         'I bought that one cup of tea.' / *'I bought one cup of that tea.'

The closest construction to the English partitive purely in terms of word order is that in (4.37a). Here, the contextual set is inside a prepositional phrase headed by *mĭŋ* 'from' or *l̥am* 'in/inside'. The question then becomes which of these two constructions, (4.37a) or (b) repeated below, is analyzeable as a true partitive.

(4.37)   a. kăw plěj [DP klăw (mĭŋ/l̥am) [DP nam pɔh   nŭʔ   ]]
             1SG buy   three from/in   6   CLF.ROUND chicken
          'I bought three of the six eggs.'

         b. kăw plěj [DP [DP **ʔɔ̌ʔ   năn**DC] çŭh pɔh   ~~ʔɔ̌ʔ năn~~]
             1SG buy   mango that   7   CLF.ROUND
          'I bought seven of those mangoes.'

Turning to the cross-linguistic properties of partitives, we find that the construction with the prepositional phrase (4.37a) does not act as a true partitive, but that in (4.37b) does. In terms of the pragmatics of partitives, one set acts as a contextual restriction and must have some kind of discourse status such as specificity (Enç 1991). In terms of the distributional properties of partitives, the 'partitive constraint' holds that the contextual set must be definite or demonstrative (Jackendoff 1977). Additionally, many partitives display an 'empty noun' restriction such that the other set cannot repeat the head noun (cf. Martí i Girbau 2011).

In Eastern Cham, the partitive-like construction with a prepositional phrase shares none of these characteristics. It is felicitous regardless of the discourse status of the contextual set (4.38a). Therefore, there is no special discourse status driving the pragmatics of this construction. The partitive constraint, which requires a demonstrative, and the empty noun constraint, which bars repetition of the head noun, are also not obeyed (4.38b). Instead of a partitive, this construction functions as a DP with a prepositional phrase adjunct.

(4.38)   ✓CONTEXT: Did you buy three eggs?

            ✓CONTEXT: Did you buy six eggs?

         a. kăw plěj [DP klăw mĭŋ nam pɔh   nŭʔ   ]
             1SG buy   three from 6   CLF.ROUND chicken
          'I bought three of the six eggs.'

         b. kăw plěj [DP klăw pɔh   nŭʔ   ni mĭŋ nam pɔh
             1SG buy   three CLF.ROUND chicken this from 6   CLF.ROUND
             nŭʔ   năn]
             chicken that
          'I bought these three eggs of those six eggs.'

By contrast, the construction without a preposition patterns in nearly all ways with partitives cross-linguistically. The contextual set has a special discourse status, as demonstrated in (4.39a). In the next subsection, this discourse status will be argued to be DC. The partitive constraint is active; there cannot be two demonstratives (4.39b). The empty noun constraint is active as well; the head noun cannot be repeated, even if the nouns are not identical as in (4.39c).

(4.39)   ✗CONTEXT: Did you buy seven mangoes?

           ✓CONTEXT: Did you buy those mangoes?

   a. kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̃ʔ**   **năn**$_{DC}$] çŭh pɔh   ~~ʔɔ̃ʔ năn~~]
      1SG buy   mango that   7   CLF.ROUND
      'I bought seven of those mangoes.'

   b. *kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̃ʔ**   **năn**] çŭh pɔh   ~~ʔɔ̃ʔ năn~~ ni]
      1SG buy   mango that   7   CLF.ROUND   this
      INTENDED: 'I bought these seven of those mangoes.'

   c. *kăw plĕj [$_{DP}$ [$_{DP}$ **pɔh jăw năn**] çŭh pɔh   ʔɔ̃ʔ   ]
      1SG buy   fruit   that   7   CLF.ROUND mango
      INTENDED: 'I bought seven mangoes of those fruits.'

There is one difference between Eastern Cham partitives and those in other languages. In Eastern Cham, there is a stricter ban on repeating material than the empty noun constraint. The numeral as well cannot be repeated, even the two DPs have different numerals. Contrast the ungrammaticality of the Eastern Cham (4.40) with the grammaticality of the English gloss. This also distinguishes Eastern Cham from the closely related Jarai, which allows numeral doubling (Jensen 2013: 69).

(4.40) *kăw plĕj [$_{DP}$ **tʰa plŭh ʔɔ̃ʔ**   **năn** çŭh pɔh   ~~tʰa plŭh ʔɔ̃ʔ năn~~]
       1SG buy   10   mango that 7   CLF.ROUND
       INTENDED: 'I bought seven of those ten mangoes.'

What motivates this difference is unclear. Perhaps it is a stronger version of the constraint that manifests itself as the empty noun constraint in other languages. Nevertheless, this construction shares pragmatic and distributional properties of partitives cross-linguistically. Therefore, it will be referred to as a partitive in the remainder of this chapter.

Much like inventory forms, partitives consist of a contextually bound phrase at the left edge of the DP. The following subsection will find that for a phrase to be at that left edge, it must be discourse connected (DC). The subsequent subsections will find that partitives are also derived by DC-movement driven by a DC-probe on D.

4.3.1 Pragmatics of partitives

Partitives have a similar pragmatic distribution to inventory forms. In partitives, the contextually bound referent must satisfy the DC conditions to the exclusion of the remainder of the DP. The polar question (4.41a) introduces an event of buying, the theme of which is the complex demonstrative *tamkaj nǎn* 'those watermelons'. The semantic representation of complex demonstratives in (4.41a′) follows Nowak's (2016) analysis of deictic demonstratives, which includes the definite article, a nominal property $watermelon(x)$, and a deictic predicate $x = g(5)$. The deictic predicate contains an assignment function from the variable x to an individual, here and throughout using a hypothetical abstract index 5. As in the previous section, the numeral phrase will be represented by adding the numeral predicate $\#(x) = 7$.[3]

(4.41)   a. hɨ   hu   plěj **tamkaj**   **nǎn** lěj
            2SG EXIST buy   watermelon that Y/N.Q

           Q: 'Have you bought those watermelons?'

     a′. ⟦(4.41a)⟧ =
         $...\exists e[\textbf{agent}(e) = \textbf{2SG} \wedge \textbf{buying}(e) \wedge \textbf{theme}(e) =$
         $\iota x.[watermelon(x) \wedge x = g(5)]$

     b. kǎw plěj [$_{DP}$ **tamkaj**   **nǎn**$_{DC}$ çǔh pɔ̥h   ~~tamkaj nǎn~~]
        1SG buy   watermelon that   7   CLF.ROUND

        A: 'I bought seven of those watermelons.'

      b′. (a ⇓ b)   DC ✓

      b″. $\iota x.[watermelon(x) \wedge x = g(5)] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

      b‴. $\iota x.[watermelon(x) \wedge x = g(5) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

Again, (4.41a) discourse subordinates the elaborating answer (4.41b), by virtue of the subordinate properties of elaborating answers to questions. The complex demonstrative *tamkaj nǎn* 'those watermelons' satisfies the DC conditions in (4.41b), because the plurality is a participant in an event in the superordinate sentence. However, the whole DP, 'those seven watermelons' does not satisfy the DC conditions, as it is not a participant in any event in the discourse. In other words, the contextually bound referent *tamkaj nǎn* 'those watermelons' does satisfy the DC conditions, but the subset *çǔh pɔ̥h* 'seven' does not.

This account predicts that partitives are infelicitous if the entire partitive satisfies the DC conditions. This is borne out. In (4.42), there is an elaborating answer to a polar question; hence a subordinating discourse relation. However, the demonstrative numeral phrase, *çǔh pɔ̥h tamkaj nǎn* 'those seven watermelons' is a participant in the

---

[3]Note that a more complete semantics of partitives contains a set–subset relation denoting seven of the set of those watermelons.

event introduced in the question. Thus, it is not the case that the demonstrative phrase satisfies the DC conditions to the exclusion of the remainder of the DP. Therefore, DP-internal DC-marking is infelicitous. Instead, the whole DP * çŭh pɔh tamkaj năn* 'those seven watermelons' would be DC-marked (4.42b′).

(4.42)   a. hɨ   ʔĭŋ   plĕj   **çŭh pɔh**   **tamkaj**   **năn** lĕj
        2SG want buy 7   CLF.ROUND watermelon that Y/N.Q
        Q: 'Do you want to buy those seven watermelons?'

    a′. $[\![(4.42a)]\!] = ...\exists e[\textbf{agent}(e) = \textbf{2SG} \wedge \textbf{buying}(e) \wedge \textbf{theme}(e) = \iota x.[watermelon(x) \wedge x = g(5) \wedge \#(x) = 7]$

    b. #kăw   plĕj   [$_{DP}$ **tamkaj**   **năn** çŭh pɔh   ~~tamkaj năn~~] jə
        1SG buy   watermelon that 7   CLF.ROUND   already
        INTENDED: A: 'I already bought seven of those watermelons.'

     b′. (a ⇓ b)   DC ✓
     b″. $\iota x.[watermelon(x) \wedge x = g(5)] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓
     b‴. $\iota x.[watermelon(x) \wedge x = g(5) \wedge \#(x) = 7] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

Finally, partitives are infelicitous in the absence of discourse subordination. In (4.43), there is no discourse subordination, as the polar question in (4.43a) is directly answered (not elaborated upon) by (4.43b). Note that the meaning of the question in (4.43a) is a set of possible answers. The semantic representation in (4.43a′) presents one possible answer to the question: 'two of those watermelons', in order to signify that 'seven of those watermelons' is not sufficiently previously mentioned in the question, but 'those watermelons' is.

(4.43)   a. hɨ   plĕj   ʈom   **pɔh**   **tamkaj**   **năn**
        2SG buy how.many CLF.ROUND watermelon that
        Q: 'How many of those watermelons did you buy?'

    a′. $[\![(4.43a)]\!] = ...\exists e[\textbf{agent}(e) = \textbf{2SG} \wedge \textbf{buying}(e) \wedge \textbf{theme}(e) = \iota x.[watermelon(x) \wedge x = g(5) \wedge \#(x) = 2]$

    b. #kăw   plĕj   [$_{DP}$ **tamkaj**   **năn** çŭh pɔh   ~~tamkaj năn~~]
        1SG buy   watermelon that 7   CLF.ROUND
        INTENDED: A: 'I bought seven of those watermelons.'

     b′. (a ⇏ b)   DC ✗
     b″. $\iota x.[watermelon(x) \wedge x = g(5)] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓
     b‴. $\iota x.[watermelon(x) \wedge x = g(5) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

Like inventory forms, thus, partitives are licensed when the contextually bound set, the complex demonstrative, satisfies the DC conditions to the exclusion of the remainder of the DP. The next subsection demonstrates that partitives are derived by DP-internal DC-movement.

4.3.2 DP-internal syntax of partitives

This section proposes a DP-internal syntax of Eastern Cham partitives consisting of a standard underlying partitive syntax with an additional DC-movement operation. Stickney (2009) analyzes partitives in languages like English with two DP layers, one embedded inside a preposition phrase headed by *of* (Figure 4.9). This analysis follows from the observation that there are two distinct referents in partitives. In English, the contextually bound referent is the one embedded inside the preposition phrase, here DP_1.

Figure 4.9: Partitive syntax (after Stickney 2009)
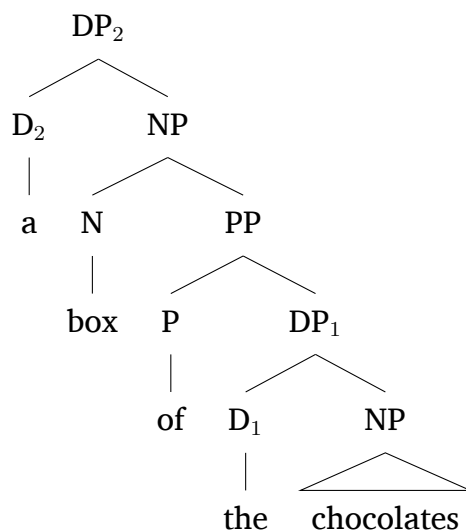
I follow Stickney's (2009) analysis of the underlying syntax of partitives and derive Eastern Cham partitives through DP-internal movement. In Eastern Cham, the contextually bound set is at the left edge of the DP. In (4.44), repeated below, that set is *ʔɔ̆ʔ năn* 'those mangoes'. As represented in the example, this is derived by movement to the left edge of the DP.

(4.44) kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̆ʔ** **năn**$_{DC}$] c̥ŭh pɔh ~~ʔɔ̆ʔ năn~~]
  1SG buy   mango that   7 CLF.ROUND
  'I bought seven of those mangoes.'

Inside the embedded DP, DP_1, the demonstrative is at the right edge. Therefore, there is a normal Num-probe on D_1, as in Figure 4.10a. In the matrix DP, by contrast, a D-head with a DC-probe is merged, the same probe found in inventory forms. This probe searches for a DC-feature, and this time it is DP_1 that bears that feature. As a result, DP_1 is moved to Spec-DP_2 (Figure 4.10b). This accounts for the DC pragmatics of DP_1 to the exclusion of the rest of DP_2, as established in the previous subsection.

Figure 4.10: Eastern Cham partitive derivation of (4.44)

(a) Derivation of DP$_1$

(b) Derivation of DP$_2$
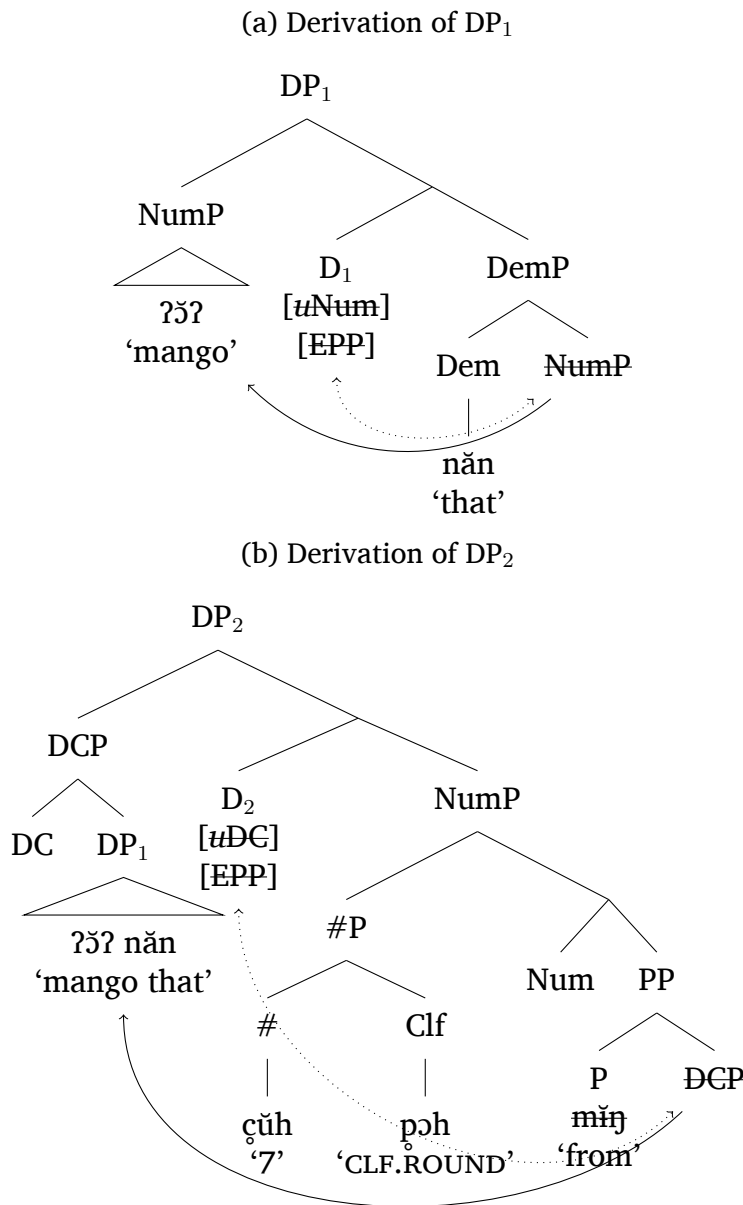
There are three additional motivations for this structure, even though there is no overt preposition equivalent to English *of*. First, as seen in the introduction to this section, there is a similar construction with an overt preposition, such as (4.45) below. This

construction could be seen as instantiating the structure in Figure 4.10b without DC-movement. Accordingly, there is no special discourse status for either of the DPs in this construction.

(4.45) kăw plěj [_DP klăw (mĭŋ/l̥am) [_DP nam pɔh   nŭʔ   ]]
        1SG buy   three from/in   6   CLF.ROUND chicken
    'I bought three of the six eggs.'

Second, the absence of prepositions in true partitives falls out from the general phenomenon of preposition- or p-drop (cf. Chapter 2, Section 2.2.4). When an argument prepositional phrase is DC-moved to the CP periphery in Eastern Cham, the preposition is dropped. For example, *ka* 'to' cannot be pronounced in any position in (4.46a). Likely, the DP is subextracted from the PP, stranding the preposition, and it is too prosodically weak to pronounce on its own. If partitives are also derived by DC-movement, p-drop would also be predicted. The preposition *mĭŋ* 'with' is used in (4.46b), as it and *l̥am* 'in' are seen in similar constructions such as (4.45) above. However, the choice of *mĭŋ* 'from' is arbitrary, as it is ungrammatical to pronounce it overtly anywhere in (4.46b).

(4.46)   a. **nĭʔ   năn** kăw plɛj han ni ~~ka nĭʔ năn~~
            child that 1SG   g̊ive cake this to
        'That child, I [will] give this cake to.'

        b. kăw plěj [_DP [_DP **ʔɔ̆ʔ   năn**_DC] c̥ŭh pɔh   ~~mĭŋ   ʔɔ̆ʔ năn~~]
            1SG buy   mango that   7   CLF.ROUND from
        'I bought seven of those mangoes.'

Third, the partitive constraint, which bans two demonstratives, falls out from the syntax of partitives. As seen with inventory forms in Section 4.2.2, phrase-final demonstratives are only licit in Eastern Cham in the presence of NumP-movement, which results in demonstratives being DP-final (Figure 4.11a). Demonstratives do not appear DP-medially in Eastern Cham, as would occur under DC-movement (Figure 4.11b). Perhaps demonstratives have a constraint that restricts their pronunciation to DP-final structures. In Eastern Cham partitives, the embedded DP has NumP-movement, while the matrix DP has DC-movement. Therefore, we would correctly predict demonstratives only to occur in the embedded DP.

Figure 4.11: Demonstratives and D-heads in Eastern Cham

(a) NumP-movement
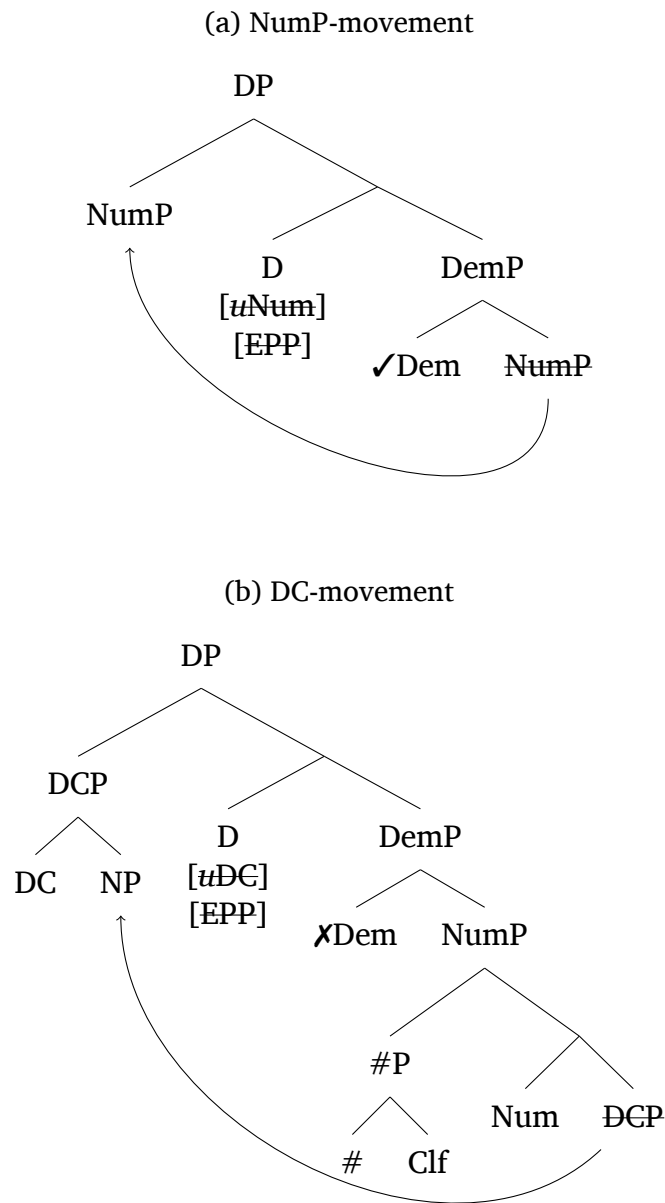
(b) DC-movement

Taken together, these strands of evidence show that a DC-movement analysis of Eastern Cham partitives, starting from a more typical partitive syntax, is possible. However,

it remains a stipulation that there exists an underlying preposition in Eastern Cham partitives, and the exact identity of that preposition is unclear. This structure makes the further prediction that the embedded DP, which consists minimally of a noun and demonstrative, is a constituent at the left edge of the DP. This prediction will be confirmed by the interaction of partitives and the CP periphery in the following section. A remaining puzzle is the empty noun and empty numeral restriction.

Before we proceed, it is worth noting a potential alternative analysis of partitive word order. Recall the structure of Eastern Cham inventory forms proposed in the previous section, repeated below in Figure 4.12a. Perhaps partitives simply occur when the demonstrative position in this structure is filled, as in Figure 4.12b.

Figure 4.12: Alternative analysis of partitives

(a) Inventory form

(b) Inventory form with demonstrative
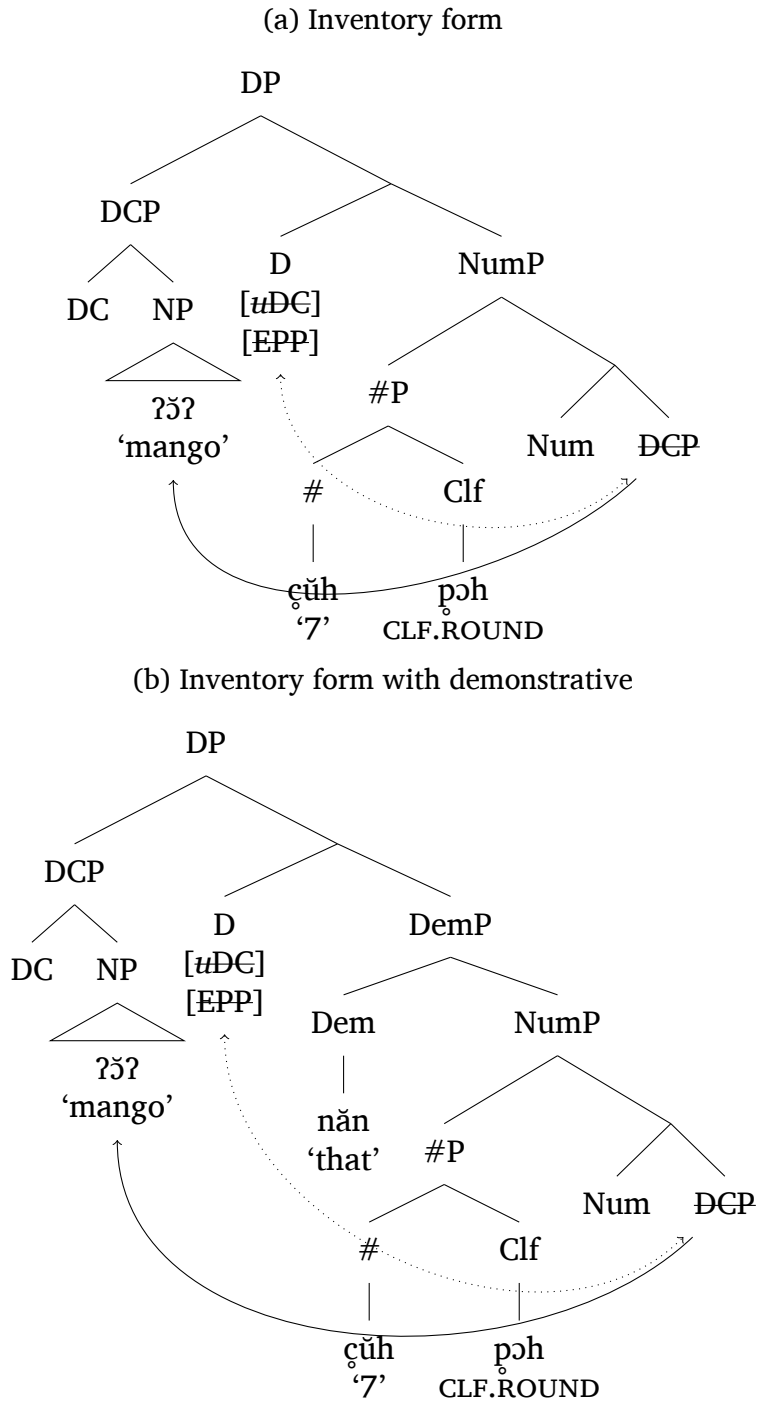
This would be a simpler syntactic structure than the partitive syntax above, and it makes some good predictions. The partitive constraint is captured, because there is only one demonstrative position. Two demonstratives are impossible, as there is no second DP at all. The empty noun and empty numeral fall out for similar reasons. There is only one DP, so the noun and numeral cannot be repeated. However, under this analysis, the demonstrative and NP do not form a constituent together. This is inconsistent with the pragmatics in the previous subsection, under which the NP and demonstrative are DC-marked together. Additionally, it violates the DP-final requirement posited for Eastern Cham demonstratives.

4.3.3 Partitives and the CP periphery

Finally, this section examines the interaction between partitives and the CP periphery. The same set of constraints is observed for partitives as was observed for inventory forms (Section 4.2.3). Again, as with inventory forms, only the DC-marked phrase can appear in the CP periphery. Furthermore, it must be base generated in Spec-CP, not derived by movement. This result supports the CP/DP parallelism for DC and the ban on double-DC-marking within a derivation. First, (4.47b–b′) illustrates that the same context that licenses partitives also licenses merging DP_1 in Spec-CP, with the subset DP remaining in its base position.

(4.47)   a. hɨ   naw   ẓaʔ   hu   plĕj   **ʔɔ̆ʔ**   **năn** lĕj
          2SG   go   market EXIST buy   mango   that   Y/N.Q
          Q: 'Did you go to the market and buy those mangoes?'

    b. kăw   plĕj   [$_{DP2}$ [$_{DP1}$ **ʔɔ̆ʔ**   **năn**$_{DC}$] cŭh pɔh   ~~ʔɔ̆ʔ năn~~   ]
       1SG   buy   mango   that   7   CLF.ROUND
       'I bought seven of those mangoes.'

    b′. [$_{DP1}$ **ʔɔ̆ʔ**   **năn**$_{DC,i}$] kăw   plĕj   [$_{DP2}$ cŭh pɔh pro_i]
         mango   that   1SG   buy   7   CLF.ROUND
        'I bought seven of those mangoes.'

By contrast, the whole partitive cannot appear in the left periphery, as in (4.48b), due to a ban on double DC-marking. In the previous section on inventory forms, it was posited that a phrase can only be associated with one DC-feature in a derivation. In the case of partitives, the DC-feature associated with *ʔɔ̆ʔ năn* 'those mangoes' enters into an Agree relation with D_2, resulting in DP-internal movement. That same feature is no longer capable of Agreeing with C. Pied-piping of the whole partitive is impossible.

(4.48)   a. hɨ   naw   ẓaʔ   hu   plĕj   **ʔɔ̆ʔ**   **năn** lĕj
          2SG   go   market EXIST buy   mango   that   Y/N.Q
          Q: 'Did you go to the market and buy those mangoes?'

b. *[$_{DP2}$ [$_{DP1}$ **ʔɔ̆ʔ** **năn**$_{DC}$] ç̊ŭh pɔ̊h] kăw plĕj
         mango that   7   CLF.ROUND   1SG buy
~~ʔɔ̆ʔ năn ç̊ŭh pɔ̊h~~

INTENDED: 'I bought seven of those mangoes.'

If DP-internal movement had not occurred, then DP$_2$ in (4.48b) may have been able to be DC-moved, as the DC-feature would only have been checked once, by C. However, Chapter 3, Section 3.3.1 posited a restriction that only the smallest possible DC-phrase is moved, so feature percolation from DP$_1$ to DP$_2$ would not be predicted.

When the contextual set DP in a partitive is merged in Spec-CP, we would predict that it is base generated there, not moved through the partitive base position, because of the impossibility of double-DC-marking and feature percolation. This prediction is borne out through evidence from path containment effects. Recall that path containment effects arise when there are two C-probes searching for the same feature. Nested paths result in grammaticality (4.49a), while crossed paths result in ungrammaticality (4.49b).

(4.49) a. **han** **ni** nĭʔ mĕj sĭt năn tʰuːŋ͡m³¹² ʔa ~~nĭʔ mĕj sĭt năn~~
     cake this child female small that Thuận invite
     maj ɓăŋ **han ni**
     come eat
     'This cake, Thuận invited that little girl to come eat.'

b. *nĭʔ mĕj sĭt năn **han** **ni** tʰuːŋ͡m³¹² ʔa ~~nĭʔ mĕj sĭt năn~~
     child female small that cake this Thuận invite
     maj ɓăŋ **han ni**
     come eat
     INTENDED: 'This cake, Thuận invited that little girl to come eat.'

If there are multiple phrases in the CP periphery, one DC-moved and one subextracted from a partitive, path containment effects go away. Nested paths lead to grammaticality (4.50a), as do crossed paths (4.50b). This is explained if the DP *han ni* 'this cake' is base generated in Spec-CP, and the C-probe searches for an Op-feature, not a DC-feature. Therefore, the two C-probes do not interact: there is one DC-probe and one Op-probe.

(4.50) a. **han** **ni** nĭʔ mĕj sĭt năn tʰuːŋ͡m³¹² ʔa ~~nĭʔ mĕj sĭt năn~~
     cake this child female small that Thuận invite
     maj ɓăŋ **klăw klɛh**
     come eat three piece
     'Thuận invited that little girl to come eat three pieces of this cake.'

b. <u>nĭʔ</u>   <u>mĕj</u>   <u>sĭt</u>   <u>năn</u>   **han**   **ni**   tʰuːŋ͡m³¹²   ʔḁ   ~~nĭʔ mĕj sĭt năn~~
child   female   small   that   cake   this   Thuận   invite

maj   ɓăŋ   **klăw**   **klɛh**
come   eat   three   piece

'Thuận invited that little girl to come eat three pieces of this cake.'

If there are two phrases base generated in Spec-CP, however, path containment effects reappear. In (4.51), one phrase is linked to a partitive, *han ni* 'this cake', while another, *nĭʔ mĕj sĭt năn* 'that little girl', is base generated and binds a resumptive pronoun *ɲu*. Here, nested paths result in grammaticality (4.51a), but crossed paths result in ungrammaticality (4.51b). This occurs because there are two C-probes both searching for an Op-feature.

(4.51)   a. **han**   **ni**   <u>nĭʔ</u>   <u>mĕj</u>   <u>sĭt</u>   <u>năn</u>   tʰuːŋ͡m³¹²   ʔḁ   ɲu   maj
cake   this   child   female   small   that   Thuận   invite   3.ANIM   come

ɓăŋ   **klăw**   **klɛh**
eat   three   piece

'Thuận invited that little girl to come eat three pieces of this cake.'

b. *<u>nĭʔ</u>   <u>mĕj</u>   <u>sĭt</u>   <u>năn</u>   **han**   **ni**   tʰuːŋ͡m³¹²   ʔḁ   ɲu   maj   ɓăŋ
child   female   small   that   cake   this   Thuận   invite   3.ANIM   come   eat

klăw   **klɛh**
three   piece

INTENDED: 'Thuận invited that little girl to come eat three pieces of this cake.'

To summarize, DP$_1$ in inventory forms can appear in the CP periphery, because it is DC-marked. To do so, it must be base generated in that position. Base generation of the NP patterns with the base generation described in Chapter 2, Section 2.3.2: a DC-probe on C probes for an Op-feature and binds an in-situ pronoun with the DC-phrase in Spec-CP, as in Figure 4.13. It should be noted that the null NP-pronoun in the examples in this section can bear an Op-feature, while null pronouns elsewhere cannot. If null pronouns in general can bear Op-features, then it would be expected that base generation + Agree in the general case can be string-identical with DC-movement. It must be stipulated that the kind of null pronoun that accompanies numerals and classifiers has different properties (i.e. an Op-feature) from null pronouns that stand alone.[4]

---

[4]It is also possible that the NP pronouns in this section in fact reflect NP-ellipsis, and that these instances of NP-ellipsis can bear Op-features as well.

Figure 4.13: Base generation of DP$_1$ in Spec-CP
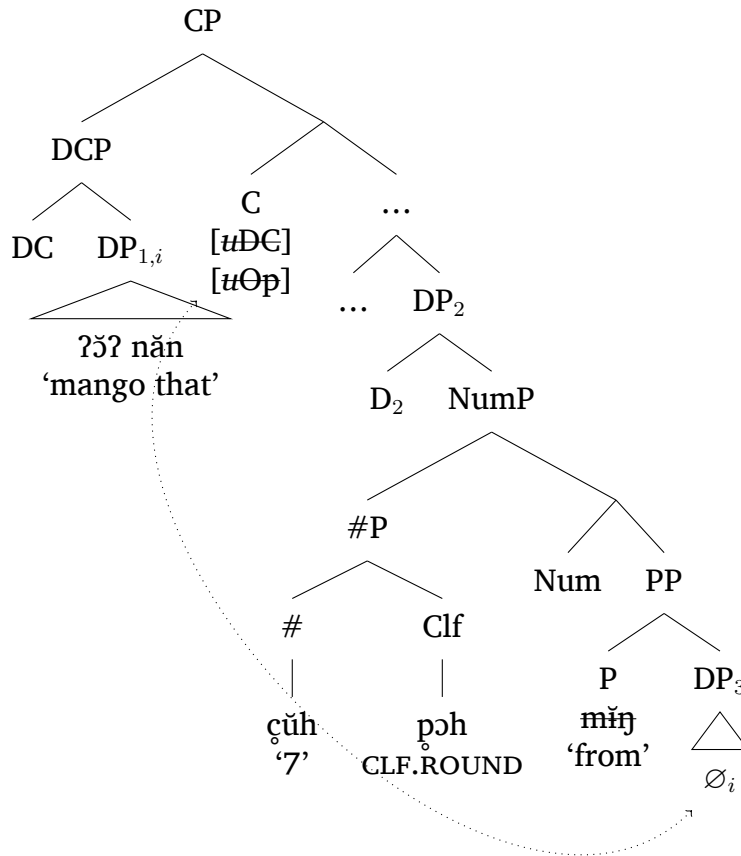
In Figure 4.13, the phrase in Spec-CP is labeled DP$_1$ to preserve the descriptive equivalence with DP$_1$ in the partitives described so far. However, here it binds a null pronoun in another embedded DP, labeled DP$_3$.

4.3.4 Summary

This section has demonstrated that true partitives in Eastern Cham function to DC-mark the embedded DP to the exclusion of the rest of the partitive. This raises a new possible analysis of partitives cross-linguistically. In terms of syntax, a DC-probe on D drives DP-

internal DC-movement. This probe is identical to the C-probe proposed in the previous sections, supporting the CP/DP parallelism proposed for Ā-features in general.

4.4 Interim summary

Taken together, the evidence in these sections points to a connection between the left edge of the DP and DC. Inventory forms and partitives are derived via DP-internal movement of a DC-marked NP or DP, respectively, to Spec-DP. Therefore, there are two DC-probes that trigger movement: one merged on C (Figure 4.14a) and one merged on D (Figure 4.14b).

Figure 4.14: Eastern Cham DC-probes on D and C
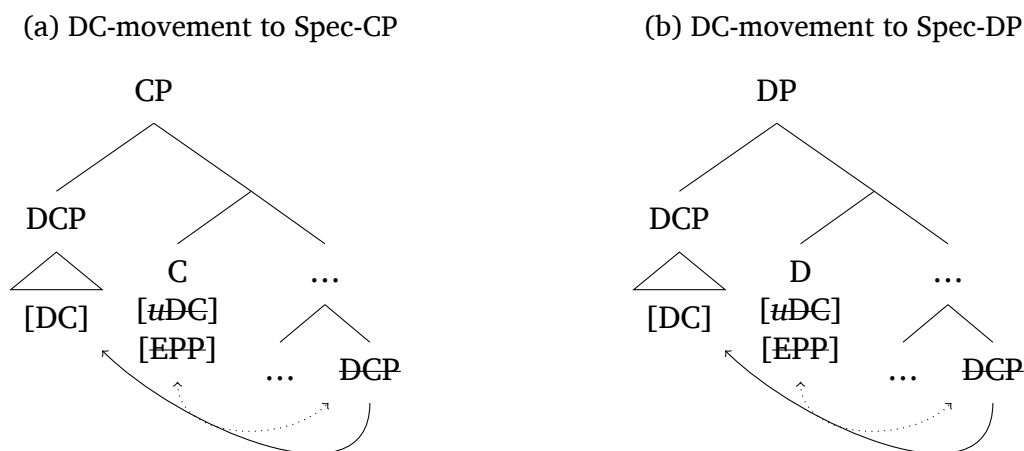
Much like other Ā-features, DC exhibits a CP/DP parallelism: the same syntactic probe can appear on C and D. Also like *wh*, DC can pied-pipe a phrase if the feature percolates, as argued in Chapter 3, Section 3.3.1. Unlike certain features like *wh*, DC cannot undergo short DC-movement inside a DP, which is then itself DC-moved, as there is only one type of DC-feature in the syntax, and it must merge with the phrase that satisfies DC pragmatics.

The next section provides corroborating evidence for the analysis put forth above. Section 4.5 demonstrates that DP-internal DC-marking can feed A-movement. Even after A-movement, the phrases retain their DC interpretation, which is predicted if they first Agree with a DC-probe.

Finally, it should be noted that there are also C-heads and D-heads that lack a DC specification. Embedded C-heads in long-distance DC-movement only bear an EPP-feature, as argued in Chapter 2, Section 2.3.3. Anaphoric definite D-heads will also be found not to be specified for DC in Section 4.6. This section finds that domain restricting- or DR-appositives are best explained in terms of anaphoric definiteness, not DC.

## 4.5   DP-internal DC-marking can feed A-movement

This section finds that DP-internal DC-marking can feed A-movement, specifically move-
ment to subject position, as in (4.52). When an inventory form feeds A-movement, the
DC-marked phrase retains its DC interpretation, even if it is not pronounced in a posi-
tion associated with DC in general, because a DC-feature is necessarily associated with a
DC-probe. This accords with the syntactic analysis of DP-internal DC-marking proposed
above.

(4.52) [_{DP} **hɔŋ̂m**_{DC} ç̥ŭh p̥ɔh] lɛ̌ʔ tʃŭn [_{DP} ~~hɔŋ̂m~~ ç̥ŭh p̥ɔh   ]
           papaya   fall down   7   CLF.ROUND
    'Papaya, seven fell down.'

This section will focus on inventory forms, but identical facts hold for partitives. There
are two ways in which IFs can interact with A-movement. First, the IF itself can be A-
moved and pronounced in subject position (Figure 4.15a). Second, the same movement
can occur, followed by distributed deletion (e.g. Fanselow & Ćavar 2002), where a sub-
part of a phrase is deleted in one copy, while the rest of the phrase is deleted in the
other copy, resulting in apparent movement such as subextraction. Distributed deletion
results in only the DC-phrase being pronounced in Spec-TP (Figure 4.15b). IFs will be
said to feed A-movement in the distributed deletion case, because the movement cannot
occur without the first step of DP-internal movement. In both cases, the NP must satisfy
the DC conditions, even though subject position is not independently associated with DC
pragmatics.

Figure 4.15: A-movement and DP-internal DC-marking in inventory form (IF)

(a) A-movement of whole IF
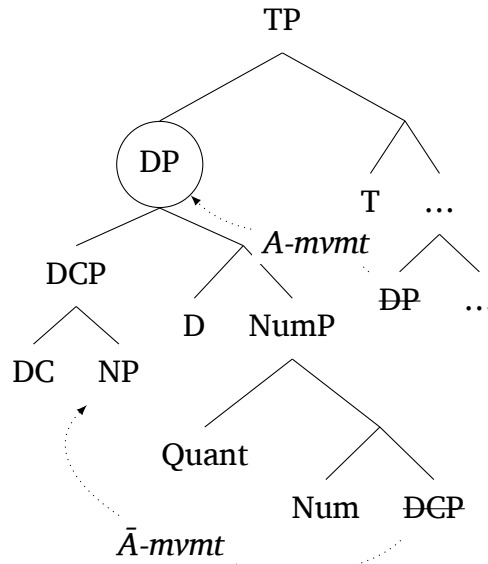
(b) A-movement and distributed deletion
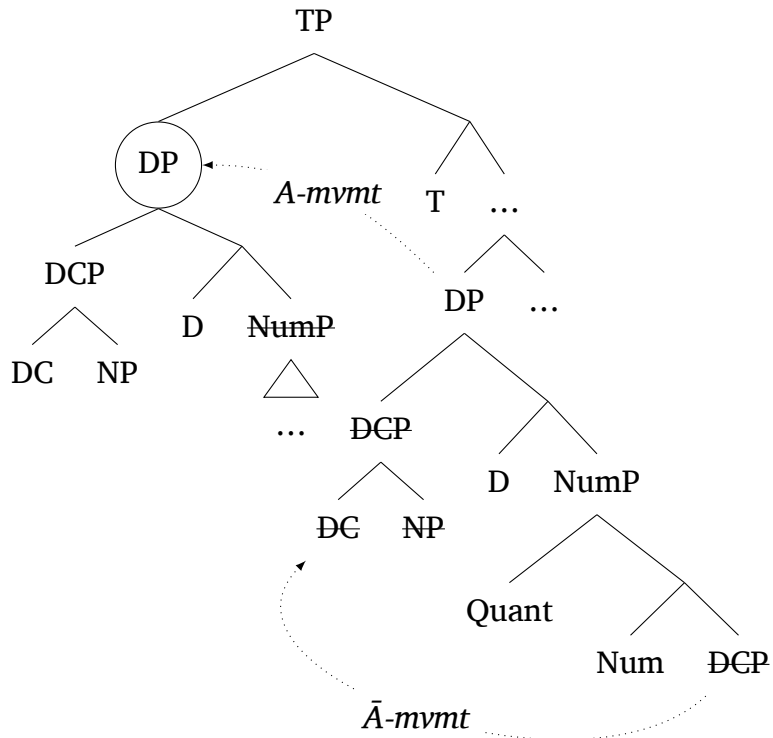
There are three components to this analysis. First, movement to subject position targets DPs, resulting in a quantifier float-like phenomenon. Second, subject position is not on its own a DC-marking position. Third, DC-marked NPs pronounced in subject position retain their DC interpretation. The remainder of this section gives evidence for each of these components in turn.

First, movement to subject position targets DPs. Eastern Cham patterns like many other languages in that the subject position must be filled (4.59a), and it must be filled by a DP, which can be a bare noun (4.59b).[5] Other phrases, such as prepositional phrases are illicit in that position (4.59c). This can be explained if there is an EPP feature on T, which is specified to attract D.

(4.53) a. *lɛ̌ʔ tʃŭn çŭh pɔh   hɔŋ͡m
            fall down 7   CLF.ROUND papaya
        INTENDED: 'Fell down seven papayas.'

     b. [NP **hɔŋ͡m**] lɛ̌ʔ tʃŭn ~~hɔŋ͡m~~
            papaya   fall down
        'Papaya fell down.'

     c. *[PP ʈi hlŭʔ   ] lɛ̌ʔ tʃŭn çŭh pɔh   hɔŋ͡m
            on ground   fall down 7   CLF.ROUND papaya
        INTENDED: 'On the ground fell seven papayas.'

For inventory forms, the whole DP is moved to subject position and then either pronounced there or partially deleted. This is most clearly observed in unaccusative predicates, where the subject DP is externally merged in object position. The relevant pattern for inventory forms is given in (4.54). The whole inventory form can be pronounced in subject position, as in (4.54a). Or, only the DC-marked NP can be pronounced there, here *hɔŋ͡m* 'papaya' (4.54b). This results in the numeral or quantifier being pronounced in its base position. The numeral or quantifier cannot be pronounced in subject position to the exclusion of the NP (4.54c).

(4.54) a. [DP **hɔŋ͡m** çŭh pɔh   ] lɛ̌ʔ tʃŭn ~~hɔŋ͡m çŭh pɔh~~
            papaya   7   CLF.ROUND   fall down
        'Papaya, seven fell down.'

     b. [DP **hɔŋ͡m** ~~çŭh pɔh~~] lɛ̌ʔ tʃŭn [DP ~~hɔŋ͡m~~ çŭh pɔh   ]
            papaya   fall down   7   CLF.ROUND
        'Papaya, seven fell down.'

     c. *çŭh pɔh   lɛ̌ʔ tʃŭn [DP **hɔŋ͡m** ~~çŭh pɔh~~]
            7   CLF.ROUND fall down   papaya
        INTENDED: 'Papaya, seven fell down.'

---

[5]Bare nouns could also be analyzed as NPs, and subject position restricted to DPs and NPs.

The pattern above is reminiscent of quantifier float. Quantifier float, as in languages like English, exhibits a pattern where the quantifier can be pronounced in a lower position than the noun, as in (4.55b–c). Conversely, the lower copy of a noun cannot be pronounced (4.55d). This example depicts quantifier float in terms of one proposed analysis: subextraction of a DP out of a quantifier phrase, stranding the quantifier (Miyagawa 1989; Sportiche 1988).

(4.55)   a. [All the papayas] are about to fall down.

b. The papayas are [all ~~the papayas~~] about to fall down.

c. The papayas are about to [all ~~the papayas~~] fall down.

d. *All are about to [~~all~~ the papayas] fall down.

In Eastern Cham, quantifiers can be pronounced in their base position, as in (4.54b) above. However, unlike English, they cannot be pronounced in an intermediate position. In (4.56a), the numeral cannot be pronounced after the prospective aspect marker *çɛ̆ʔ ṭi*. In (4.56b), it cannot be pronounced after the predicate-final iterative aspect marker *vɨ̆ʔ*. Baclawski Jr (2017) finds that this position is generally available for focussed objects, and it is structurally higher than the initial merge position of the *v*P. Otherwise, the form and interpretation of the Eastern Cham phenomenon parallels that of quantifier float.

(4.56)   a. **hɔɲ͡m**   çɛ̆ʔ ṭi   {*}   lɛ̆ʔ   tʃŭn   {çŭh   pɔh}
            papaya   PROSP   fall   down   7   CLF.ROUND
            'Papayas, seven are about to fall down.'

b. **hɔɲ͡m**   lɛ̆ʔ   tʃŭn   {çŭh   pɔh}   vɔ̆ʔ   {*}
    papaya   fall   down   7   CLF.ROUND   ITER
    'Papayas, seven fell down again.'

Much like quantifier float in languages like Thai (Jenks 2013), the function of the Eastern Cham phenomenon seems to be to pronounce the numeral or quantifier low for the purposes of scope. Pronouncing the numeral or quantifier low ensures that it has a low scope interpretation. To illustrate, there is a clear interaction between the universal quantifier *pih* 'all' and negation depending on whether it is pronounced in subject position or low. Note that a partitive construction is used here, as the judgments are the most clear.

(4.57)   a. **hɔɲ͡m**   **năn pih**   hu   lɛ̆ʔ   tʃŭn   ~~hɔɲ͡m năn pih~~   ʔo
            papaya   that all   EXIST   fall   down   NEG
            'All of those papayas did not fall down.' $\forall > \neg, *\neg > \forall$

b. **hɔɲ͡m**   **năn**   hu   lɛ̆ʔ   tʃŭn   [~~hɔɲ͡m năn~~   pih]   ʔo
    papaya   that   EXIST   fall   down   all   NEG
    'Those papayas did not all fall down.' $*\forall > \neg, \neg > \forall$

In the remainder of this section, I will treat the Eastern Cham phenomenon as a restricted kind of quantifier float. I follow Jenks (2013) in analyzing quantifier float through distributed deletion (cf. also Fanselow & Ćavar 2002), though nothing hinges on this choice for purposes of this section. Other analyses like subextraction/stranding (Sportiche 1988; Miyagawa 1989; Shlonsky 1991) would have similar implications for the syntax and pragmatics of DC. According to distributed deletion, quantifier float arises from movement and deletion such that the noun is pronounced high and the quantifier low. This is demonstrated for English in (4.58a) and Eastern Cham in (4.58b).

(4.58)   a. [~~all~~ The papayas] are [all ~~the papayas~~] about to fall down.

   b. [**hɔŋ͡m**   ~~çŭh pɔh~~] lɛ̆ʔ tʃŭn   [~~hɔŋ͡m~~ çŭh pɔh   ]
      papaya   fall down   7   CLF.ROUND
      'Papaya, seven fell down.'

Movement to subject position, thus, targets DPs and results in a quantifier float-like phenomenon. In the remainder of this section, I will refer to the Eastern Cham phenomenon as quantifier float.

Second, subject position does not inherently have an association with DC pragmatics. One example is given below. In this narrative, sentence (4.59b) is interpreted as a separate event from sentence (4.59a), as indicated by the temporal adverbials m̥jăw mɨh 'first', and plɔh năn 'then'. Therefore, there is no subordinating discourse relation between the two. This context is sufficient to allow ʔiŋ ʔɔŋ năn 'the frog' to be in subject position in (4.59b), but it is not sufficient to license DC-movement of ʔiŋ ʔɔŋ năn 'the frog' in (4.59b′).

(4.59)   a. m̥jăw mɨh kăw mɨ̆ʔ   ʔiŋ ʔɔŋ
            first   1SG catch frog
         'First, I caught (the) frog.'

   b. plɔh năn ʔiŋ ʔɔŋ năn plɔŋ͡m l̥am ʔja   ʋɛ̆ʔ
      after that frog   that jump   into water ITER
      'After that, the frog jumped back in the water.'   (a ∦ b)

   b′. plɔh năn {#} kăw ŋă̆ʔ   {ʔiŋ ʔɔŋ năn}
       after that   1SG make frog   that
       'After that, I cooked that frog.'   (a ∦ b)

It is true that Eastern Cham has a tendency to have topical or old information subjects. When a subject is focussed, it typically must be in a cleft construction with the existential marker *hu* (Baclawski Jr 2018a). However, there is no evidence that subjects must be specifically DC-marked. In the examples that follow in the remainder of this section, the unmarked DP order *çŭh pɔh hɔŋ͡m* 'seven papayas' can be a felicitous subject in each of the (b) sentences, even when it does not meet the DC conditions.

Third, NPs pronounced in subject position retain their DC pragmatics. This is expected if the NP also undergoes DP-internal DC-movement. First, (4.60b) is felicitous in a context where the DC conditions are satisfied for the underlying inventory form. In (4.60), the NP *hɔŋ͡m* 'papaya' satisfies the DC conditions. The whole DP including the numeral, however, does not satisfy the DC conditions. See Section 4.2.1 for more detail on the DC conditions that license inventory forms.

(4.60)   a. **pɔh**   **hɔŋ͡m**   hu   lɛ̆ʔ   tʃŭn   lɛ̆j
          CLF.ROUND   papaya   EXIST   fall   down   Y/N.Q
        Q: 'Did papaya fall down?'

    a′. $[\![(4.60a)]\!] = \ldots\exists e[\textbf{falling}(e) \wedge \textbf{theme}(e) = {}^{\cap}papaya]$

    b. **hɔŋ͡m**   lɛ̆ʔ   tʃŭn   [$_{DP}$ ~~hɔŋ͡m~~ çŭh pɔh   ]
        papaya   fall   down   7   CLF.ROUND
        A: 'Seven papayas fell down.'

      b′. (a $\Downarrow$ b)   DC ✓

      b″. ${}^{\cap}papaya \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

      b‴. $\lambda x.[papaya(x) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

In the absence of previous mention in a superordinate sentence, quantifier float is infelicitous. This is exemplified in (4.61), where the *wh*-question makes no reference to papayas. Note that the semantic representation in (4.60a′) contains one possible answer to the question: the kind 'mango', in order to signify that 'papaya' is not sufficiently mentioned in the question. The complete meaning of the question is a set of all possible answers.

(4.61)   a. ķeʔ   lɛ̆ʔ   tʃŭn   năn
         what   fall   down   that
        Q: 'What fell down?'

    a′. $[\![(4.60a)]\!] = \ldots\exists e[\textbf{falling}(e) \wedge \textbf{theme}(e) = {}^{\cap}mango]$

    b. #**hɔŋ͡m**   lɛ̆ʔ   tʃŭn   [$_{DP}$ ~~hɔŋ͡m~~ çŭh pɔh   ]
         papaya   fall   down   7   CLF.ROUND
        INTENDED: A: 'Seven papayas fell down.'

      b′. (a $\Downarrow$ b)   DC ✓

      b″. $\lambda x.papaya(x) \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

      b‴. $\lambda x.[papaya(x) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

The absence of discourse subordination also results in infelicity. The question in (4.62a) is directly answered by (4.62b), which does not constitute a subordinating discourse relation. Even though papayas are mentioned in (4.62a) to the exclusion of the

numeral, quantifier float is degraded. Note that, as above, the semantic representation in (4.62a′) contains one possible answer to the question: 'two papayas'. The answer 'two papayas' is chosen in order to demonstrate that 'seven papayas' is not sufficiently previously mentioned in the question, but 'papaya' is.

(4.62)   a. ţom   **pɔh**   **hɔ̆ŋ̊m**   lĕʔ tʃŭn
         how.many CLF.ROUND papaya fall down

       Q: 'How many papayas fell down?'

    a′. ⟦(4.60a)⟧ = …∃e[**falling**$(e) \wedge$ **theme**$(e) = \iota x.[papaya(x) \wedge \#(x) = 2]$

    b. #**hɔŋ̊m**   lĕʔ tʃŭn   [$_{DP}$ ~~hɔ̆ŋ̊m~~ çŭh pɔh   ]
       papaya fall down   7   CLF.ROUND

       INTENDED: A: 'Seven papayas fell down.'

      b′. (a ⇓ b)   DC ✗

      b″. $\lambda x.papaya(x) \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

      b‴. $\iota x.[papaya(x) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

Lastly, if both the NP and the whole DP are participants in the event introduced in the question, quantifier float is again degraded. This is predicted, as inventory forms are in general infelicitous if the whole DP including the numeral also satisfies the DC conditions.

(4.63)   a. **çŭh pɔh**   **hɔ̆ŋ̊m** çĕʔ ţi lĕʔ tʃŭn ʃăwʔ lĕj
        7   CLF.ROUND papaya PROSP fall down correct Y/N.Q

       Q: 'Seven papayas are about to fall down, is that right?'

    a′. ⟦(4.60a)⟧ = …∃e[**falling**$(e) \wedge$ **theme**$(e) = \lambda x.[papaya(x) \wedge \#(x) = 7]]$

    b. #**hɔŋ̊m**   lĕʔ tʃŭn   [$_{DP}$ ~~hɔ̆ŋ̊m~~ çŭh pɔh   ] jə
       papaya fall down   7   CLF.ROUND   already

       A: 'Seven papayas fell down already.'

      b′. (a ⇓ b)   DC ✓

      b″. $\lambda x.papaya(x) \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

      b‴. $\lambda x.[papaya(x) \wedge \#(x) = 7] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

All in all, given the obligatoriness of DC-marking in quantifier float, or movement to subject position, I conclude that this movement is fed by an initial step of DC-movement inside the DP. Without this first step, we would expect optionality in DC-marking at most. Why exactly distributed deletion requires a first step of DP-internal movement is not clear. Perhaps in Eastern Cham, distributed deletion can only apply such that the structurally higher constituent is pronounced in the higher copy. In other words, the NP can only be pronounced in the higher copy if it is the highest constituent. The NP can only become structurally high via DP-internal movement.

4.6 Domain restricting appositives

Throughout this dissertation, DC has been compared to and found to be independent from information structural notions such as topicality and D-linking. A lingering question is whether DC-effects can be explained in terms of definiteness or familiarity. This section explores a third DP-internal construction, the domain restricting (DR-) appositive. DR-appositives require anaphoric definiteness, not DC-marking. DR-appositives have a different pragmatic distribution from inventory forms and partitives. This difference in pragmatic distributions indicates that DC is independent from anaphoric definiteness. More specifically, definiteness does not impose the discourse structural requirement of DC.

The first part of this section examines the structural properties of DR-appositives and how they are distinct from close-appositives (Section 4.6.1). The second part examines the pragmatics of DR-appositives and finds that they cannot be characterized in terms of DC (Section 4.6.2).

4.6.1 Structure of DR-appositives

DR-appositives in Eastern Cham consist of a complex pronoun and a name that specifies the reference of the pronoun. For example, the plural name *pu hɔ̌ŋm hwa* 'Phú and Hoa' specifies the reference of the complex pronoun *țwa jaŋ ɲu* 'those two people' [LIT.: 'two them'] in (4.64a).

(4.64)   a. hɨ   ʔa   [$_{DP}$ **pu** **hɔ̌ŋm** **hwa**$_{+\text{Def},i}$ țwa jaŋ   ɲu$_i$]
            2SG invite   Phú with   Hoa   two CLF.PERSON 3.ANIM
         hu
         ROOT
         'You can invite Phú and Hoa, those two people .'

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

Huang, Li & Li (2009: 298) identify a similar construction in Mandarin Chinese, where a name and/or pronoun specify the reference of a demonstrative phrase. In (4.64b), the plural name *Zhangsan, Lisi* 'Zhangsan and Lisi' specifies the reference of *na ji-ge guai haizi* 'those good kids'. Huang, Li & Li (2009) provide arguments that this construction functions as a single DP and has a different distribution from close appositives. Jenks (2018) further analyzes this construction as an anaphoric definite whose reference is overly specified by a pronoun or name. That pronoun or name is generated in Spec-DP of the anaphoric definite DP.

According to Huang, Li & Li's (2009) and Jenks's (2018) analyses, the name in a DR-appositive is generated within the definite DP. These analyses predict that the name and definite DP should appear in one specific order, and that no other phrase can intervene in between them. By contrast, close appositives typically have variable order and can be stacked, as they involve two DPs, one of which is attached at a high or low position such that their relative order is not fixed.

In Eastern Cham, DR-appositives confirm the predictions from Mandarin Chinese. The name and complex pronoun must be in a specific order, and they cannot be interrupted by another phrases. First, close appositives are similar to DR-appositives (and distinct from non-restrictive appositives) in that they lack a prosodic break or pause between the two appositive elements. However, the relative ordering of elements in a close appositive is not strict, as shown in (4.65a–b).

(4.65)   a. hḷăʔ   cəh [$_{DP}$ ʔoŋ   pu ǩjuʔ tɔ hḷăʔ   ]
            1SG.POL like   grandfather Phú teacher 1SG.POL
            'I like Mr. [LIT: Grandfather] Phú, my teacher.'

        b. hḷăʔ   cəh [$_{DP}$ ǩjuʔ tɔ hḷăʔ   ʔoŋ   pu]
            1SG.POL like   teacher 1SG.POL grandfather Phú
            'I like my teacher, Mr. [LIT: Grandfather] Phú.'

A lack of relative ordering is also found with a name and a demonstrative phrase in Eastern Cham (4.66a–b). While this configuration results in a DR-appositive in Mandarin Chinese, as in (4.64b) above, it does not appear to in Eastern Cham. It is not yet clear what conditions this variable ordering, but it is reflective of close appositives cross-linguistically (e.g. Lekakou & Szendrői 2012 on polydefinites and close appositives).

(4.66)   a. hɨ   ʔa   [$_{DP}$ **pu hɔ͡ŋm hwa** ṭwa jaŋ   năn] hu
            2SG invite   Phú with Hoa two CLF.PERSON that   ROOT
            'You can invite Phú and Hoa, those two people .'

        b. hɨ   ʔa   [$_{DP}$ ṭwa jaŋ   năn **pu hɔ͡ŋm hwa**] hu
            2SG invite   two CLF.PERSON that Phú with Hoa   ROOT
            'You can invite Phú and Hoa, those two people.'

Instead, there is only one configuration that results in a DR-appositive: a name followed by a complex pronoun. On analogy with complex demonstratives, a complex pronoun is a pronoun of the form *pro F,* such as *we linguists.* In Eastern Cham, complex pronouns are common on their own, such as *ṭwa jaŋ ɲu* 'those two people' (LIT: 'two them'). Here, the name specifies the reference of the complex pronoun. The order of the two is strict; the complex pronoun cannot precede the name (4.67b).

(4.67) a. hɨ   ʔa   [DP **pu hɔ̆ŋ͡m hwa**ᵢ ṭwa jaŋ   ɲu*ᵢ*   ] hu
           2SG invite   Phú with   Hoa   two CLF.PERSON 3.ANIM   ROOT
        'You can invite Phú and Hoa, those two people .'

     b. *hɨ   ʔa   [DP ṭwa jaŋ   ɲu*ᵢ*   **pu hɔ̆ŋ͡m hwa**ᵢ] hu
           2SG invite   two CLF.PERSON 3.ANIM Phú with   Hoa   ROOT
        INTENDED: 'You can invite Phú and Hoa, those two people .'

At first glance, this ordering restriction might seem to be a fact about binding. Perhaps the pronoun *ɲu* cannot precede the name that binds it, so the order in (4.67b) must be ungrammatical. In general, within *v*P in Eastern Cham, binding relations follow linear order (left to right), not c-command. For example, the indirect object can bind the direct object if the indirect object precedes the direct object in the predicate in terms of linear order, even though neither phrase c-commands the other. If the direct object is pronounced to the right of the indirect object, as through object shift, binding is possible, even though neither DP c-commands the other (4.68a–b). Note that object shift involves a phrase being pronounced to the right of a predicate-final aspect marker or modal, such as the root modal *hu* (cf. Baclawski Jr 2017 for details and diagnostics for object shift; cf. also Simpson 2001 on related phenomena in other Southeast Asian languages).

(4.68) a. *kăw m̥jan   li⁴⁵   ʔja   cɛ   ɲu*ᵢ*   hu   ka pu*ᵢ*
           1SG   give.back glass(VN) water tea 3ANIM ROOT to Phú
        INTENDED: 'I can give Phú back his tea.'

     b. kăw m̥jan   ka pu*ᵢ* hu   li⁴⁵   ʔja   cɛ   ɲu*ᵢ*
           1SG give.back to Phú ROOT glass(VN) water tea 3ANIM
        'I can give Phú back his tea.'

DR-appositives, however, do not follow this rule. The name can undergo object shift out of a DR-appositive and still corefer with the complex pronoun (4.69a). Interestingly, the complex pronoun cannot be shifted (4.69b). Perhaps this can be explained via locality, as the name constituent is more local within the DP. Alternately, object shift could involve distributed deletion akin to the quantifier float construction in Section 4.5 above. However, more research is needed on the specifics of object shift and its relation to DR-appositives.[6]

(4.69) a. hɨ   ʔa   [DP ~~pu hɔ̆ŋ͡m hwa~~ᵢ ṭwa jaŋ   ɲu*ᵢ*   ] hu   **pu**
           2SG invite   two CLF.PERSON 3.ANIM   ROOT Phú
        hɔ̆ŋ͡m hwaᵢ
        with   Hoa
        'You can invite Phú and Hoa, those two people.'

---

[6]A further unexplained complication is that inventory forms and partitives cannot be split by object shift in this way.

b. *hɨ ʔa [DP **pu hɔ̌ŋ̂m hwa**$_i$ t̥wa jaŋ ɲu$_i$] hu t̥wa
2SG invite Phú with Hoa ROOT two

jaŋ ɲu$_i$
CLF.PERSON 3.ANIM

INTENDED: 'You can invite Phú and Hoa, those two people.'

DR-appositives also cannot be interrupted by a close appositive (4.70a), even though a DR-appositive can combine with a close appositive in general (4.70b). Huang, Li & Li (2009: 304) cite similar contrasts in Mandarin Chinese.

(4.70) a. *hɨ ʔa [DP **pu hɔ̌ŋ̂m hwa** nɨʔ sɛh hɨ t̥wa jaŋ
2SG invite Phú with Hoa child student 2SG two CLF.PERSON

ɲu] hu
3.ANIM ROOT

INTENDED: 'You can invite Phú and Hoa, your students, those two people.'

b. hɨ ʔa [DP nɨʔ sɛh hɨ **pu hɔ̌ŋ̂m hwa** t̥wa jaŋ
2SG invite child student 2SG Phú with Hoa two CLF.PERSON

ɲu] hu
3.ANIM ROOT

'You can invite your students, Phú and Hoa, those two people.'

These facts are explained if DR-appositives involve a name generated as a DP in Spec-DP, which specifies the reference of a complex pronoun. It is worth mentioning that DR-appositives cannot be analyzed as nominal appositives attached either high in the CP layer as in right-dislocation or low in the DP, as neither predicts that a DP can bind into its appositive (cf. Potts 2005; Onea 2013, 2016). This alternate analysis is discussed in more detail in Section 4.6.2.

4.6.2 Pragmatics of DR-appositives

Domain restricting (DR-) appositives are felicitous not only DC contexts, but also any context where the phrase can be interpreted as definite. DR-appositives have a broader pragmatic distribution than the partitives and inventory forms described in the previous sections, with regard to what kinds of questions they can follow. First, (4.71) gives a context where the name alone satisfies the DC conditions. This is comparable to the contexts where partitives and inventory forms would be licensed. The answer is interpretable as an elaboration on the question, as it provides more information about who Phú and Hoa are (i.e. the people with those names over there). The DC conditions would be satisfied for the plural individual Phú and Hoa, as they are mentioned in the superordinate sentence (4.71a). Note that Phú and Hoa are represented semantically as the plural individual $\mathbf{p} \oplus \mathbf{h}$, while the complex pronoun *t̥wa jaŋ ɲu* 'those two people' is represented by a

complex demonstrative expression containing the definite article ιx, a nominal property $person(x)$, an assignment function from x to the plural individual Phú and Hoa, and the numeral predicate $\#(x) = 2$, as per Nowak's (2016) analysis of complex demonstratives. This assumes that the semantics of complex pronouns is comparable to that of complex demonstratives.

(4.71) a. kăw ʔḁ **pu hɔ̆ŋ͡m hwa** maj pă?ni hu lɛ̆j
1SG invite Phú with Hoa come here ROOT Y/N.Q
'Can I invite Phú and Hoa to come here?'

    a'. $[\![(4.71a)]\!] = ...\exists e[\textbf{agent}(e) = \textbf{1SG} \wedge \textbf{inviting}(e) \wedge \textbf{theme}(e) = \textbf{p} \oplus \textbf{h}]$

  b. hɨ ʔḁ [$_{DP}$ **pu hɔ̆ŋ͡m hwa**$_i$ ʈwa jaŋ ɲu$_i$] hu
2SG invite Phú with Hoa two CLF.PERSON 3.ANIM ROOT
'You can invite Phú and Hoa, those two people.'

    b'. (a ⇓ b)   DC ✓

    b''. $\textbf{p} \oplus \textbf{h} \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

    b'''. $\iota x : [person(x) \wedge x = \textbf{p} \oplus \textbf{h} \wedge \#(x) = 2] \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

    DR-appositives are also felicitous in cases where the complex pronoun satisfies the DC conditions. (4.72) gives an identical context to the one above, except the complex pronoun is mentioned in the superordinate sentence. Note that in (4.72a'), the complex pronoun contains the deictic predicate $x = g(5)$, because it has not yet been overtly linked to Phú and Hoa in the discourse. Despite the lack of previous mention, *pu hɔ̆ŋ͡m hwa* 'Phú and Hoa' can appear at the left edge of the DR-appositive. This is not expected if only DC-marked phrases can appear at the left edge of DR-appositives. Note that appositives in general are known to be able to introduce new information (e.g. Del Gobbo 2003, Constant 2011 on appositives in Mandarin Chinese).

(4.72) a. kăw ʔḁ **ʈwa jaŋ ɲu** maj pă?ni hu lɛ̆j
1SG invite two CLF.PERSON 3.ANIM come here ROOT Y/N.Q
'Can I invite those two people to come here?'

    a'. $[\![(4.72a)]\!] = ...\exists e[\textbf{agent}(e) = \textbf{1SG} \wedge \textbf{inviting}(e) \wedge \textbf{theme}(e) = $
$\iota x : [person(x) \wedge x = g(5) \wedge \#(x) = 2]]$

  b. hɨ ʔḁ [$_{DP}$ pu hɔ̆ŋ͡m hwa$_i$ **ʈwa jaŋ** ɲu$_i$] hu
2SG invite Phú with Hoa two CLF.PERSON 3.ANIM ROOT
'You can invite Phú and Hoa, those two people.'

    b'. (a ⇓ b)   DC ✓

    b''. $\textbf{p} \oplus \textbf{h} \notin \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

    b'''. $\iota x : [person(x) \wedge x = g(5) \wedge \#(x) = 2] \in \mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

DR-appositives are also felicitous if no element of the appositive is mentioned in a superordinate sentence, as in (4.73). The pragmatics of DR-appositives cannot be explained by the DC conditions between the (a) and (b) sentences in these examples. Note that the semantic representation in (4.73a′) presents one possible answer. The complete meaning of the question is a set of all possible answers. The answer chosen to be represented is a different individual Thuận, **t**, in order to signify that Phú and Hoa are not sufficiently previously mentioned in the question.

(4.73)   a. kăw ʔa   **tʰɛ̆j** maj   pă? ni hu   lɛ̆j
             1SG invite who come here   ROOT Y/N.Q

         'Who can I invite to come here?'

   a′. ⟦(4.73a)⟧ = ...∃e[**agent**(e) = **1SG** ∧ **inviting**(e) ∧ **theme**(e) = **t**]

   b. hɨ   ʔa   [$_{DP}$ pu   hɔ̆ŋ̂m hwa$_i$ t̥wa jaŋ   ɲu$_i$] hu
        2SG invite   Phú with Hoa   two CLF.PERSON 3.ANIM   ROOT

      'You can invite Phú and Hoa, those two people.'

   b′. (a ⇓ b)   DC ✓

   b″. **p** ⊕ **h** ∉ $\mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✗

   b‴. ιx : [person(x) ∧ x = **p** ⊕ **h** ∧ #(x) = 2] ∉ $\mathcal{P}_e \cap \mathcal{P}_{e'}$ DC ✓

The data so far establish that DR-appositives do not require DC. Instead, if Jenks's (2018) analysis of Mandarin Chinese DR-appositives is applicable to Eastern Cham, they require anaphoric definiteness. In (4.73b) above, the DR-appositive is licensed essentially if the complex pronoun is licensed, either through prior discourse, salience, or familiarity.

An alternate analysis of appositive-like constructions should be mentioned, wherein the appositive contains a reduced sentence, which itself is in discourse relations. This alternate analysis provides a possible avenue for analyzing DR-appositives in terms of DC. However, this analysis fails to account for a variety of facts in Eastern Cham.

Onea (2013, 2016) argues that nominal appositives (NAPs) involve two elements: the first a part of the full, overt sentence at hand, and the second a part of a reduced sentence computed in a separate derivation. For example, an NAP like *Rishabh, an old friend* in (4.74a) involves two derivations: the derivation of *I met Rishabh at the pub* and *Rishabh is an old friend*. Onea (2013, 2016) argues that the reduced sentences in NAPs answer potential questions raised by the full sentence, as spelled out in (4.74b)

(4.74)   a. I met Rishabh, **an old friend,** at the pub.

          b. I met Rishabh at the pub. Who is Rishabh? ~~Rishabh is~~ an old friend.

The resulting picture of NAPs is sketched in (4.74b). As evidence of this analysis, Onea (2013, 2016) argues that NAPs exhibit the same kind of ellipsis as fragment answers. The fragment answer *An old friend* can answer the question *Who is Rishabh?* which is identical

to the NAP *an old friend*. According to Onea, this is not a coincidence; NAPs are fragment answers to a covert question (and syntactically independent derivations).

This analysis of NAPs is relevant for DC, as NAPs involve two sentences in a certain discourse relation.[7] Perhaps *Rishabh* could be seen as satisfying the DC conditions within those two sentences, schematized as (4.75) below. The superordinate sentence (4.75a), corresponds with the full, overt sentence above, while the subordinate sentence (4.75b), corresponds with the reduced sentence. That reduced sentence elaborates upon the rest of the sentence, providing more information about *Rishabh*; hence (a ⇓ b). Furthermore, *Rishabh* is a participant in the event introduced in the superordinate sentence. Hence, the elided copy of *Rishabh* in the NAP does satisfy the DC conditions.

(4.75)   a. I met Rishabh at the pub.

      b. Rishabh$_{DC}$ is an old friend.   (a ⇓ b)

At first glance, it is possible that Eastern Cham DR-appositives could have the discourse properties of Onea's (2013) NAPs. DR-appositives appear to be interpreted as elaborating upon the identity of a referent. For example, in (4.76), the names *pu hɔ̆ŋ̂m hwa* elaborate on the identity of the complex pronoun.

(4.76)   a. kăw ʔa **ṭwa jaŋ**   **ɲu**   maj pă̆ʔ ni hu   lɛ̆j
             1SG invite two CLF.PERSON 3.ANIM come here   ROOT Y/N.Q
             'Can I invite those two people to come here?'

      b. hɨ ʔa [$_{DP}$ pu hɔ̆ŋ̂m hwa$_i$ **ṭwa jaŋ**   **ɲu$_i$**] hu
           2SG invite   Phú with Hoa two CLF.PERSON 3.ANIM   ROOT
           'You can invite Phú and Hoa, those two people.'   (a ⇓ b)

However, there are two reasons to doubt the claim that Eastern Cham DR-appositives involve a structure like NAPs. First, if DR-appositives are comparable to NAPs, it is unclear which element elaborates upon what. In (4.76) above, the names elaborate upon the complex pronoun. However, in (4.77), it would seem to be the case that the complex pronoun elaborates upon the identity of the names.

(4.77)   a. kăw ʔa **pu hɔ̆ŋ̂m hwa** maj pă̆ʔ ni hu   lɛ̆j
             1SG invite Phú with   Hoa come here   ROOT Y/N.Q
             'Can I invite Phú and Hoa to come here?'

      b. hɨ ʔa [$_{DP}$ **pu hɔ̆ŋ̂m hwa$_i$** ṭwa jaŋ   ɲu$_i$] hu
           2SG invite   Phú with Hoa two CLF.PERSON 3.ANIM   ROOT
           'You can invite Phú and Hoa, those two people.'   (a ⇓ b)

---

[7]Onea (2016) incorporates discourse relations from Segmented Discourse Representation Theory. Discourse subordination could be argued to be one type of potential question in this system.

Second, leading analyses of appositives predict that the matrix sentence cannot bind into an appositive. This includes Onea's (2013, 2016) analysis of NAPs, Potts's (2005) Comma operator account, and high attachment analyses of appositives (e.g. Del Gobbo 2003; Schlenker 2010). This is not borne out in DR-appositives, as the names *pu hŏŋm hwa* clearly bind the pronoun *ɲu* in the examples throughout this section.

Based on this evidence, I conclude that Eastern Cham DR-appositives are not licensed by DC either overtly, as with inventory forms and partitives, or covertly, by means of relations between reduced sentences. Instead, DR-appositives are licensed by definiteness or familiarity.

4.6.3 Summary

DR-appositives, as seen in this section, do not require DC contexts. Instead, they are licensed in more general definite or familiar contexts. The fact that inventory forms and partitives have different pragmatic distributions from DR-appositives shows that DC is independent from definiteness and familiarity more broadly. DC uniquely imposes a discourse structural requirement.

An additional pattern of facts is worth noting: the interaction between appositives and the left periphery. This interaction can be explained in terms of base generation and Agree (Chapter 2, Section 2.3.2) and an anaphoric definite picture of DR-appositives. First, close appositives in any order can be DC-moved to the CP periphery (4.78a–b).

(4.78)   a. [**pu   hŏŋm   hwa** ʈwa jaŋ   năn] hɨ   ʔa
              Phú with   Hoa   two CLF.PERSON that 2SG invite
              ~~pu hŏŋm hwa ʈwa jaŋ năn~~ hu
                                           ROOT
         'You can invite Phú and Hoa, those two people.'

         b. [ʈwa jaŋ   năn **pu   hŏŋm   hwa**] hɨ   ʔa
              two CLF.PERSON that Phú with   Hoa   2SG invite
              ~~ʈwa jaŋ năn pu hŏŋm hwa~~ hu
                                           ROOT
         'You can invite those two people, Phú and Hoa.'

Second, DR-appositives cannot be DC-moved to the left periphery, as in (4.79a). This fact falls out from the general ban on the DC-movement of pronouns (Chapter 2, Section 2.1.3).

(4.79)   a. *[**pu   hŏŋm   hwa**$_i$ ʈwa jaŋ   ɲu$_i$] hɨ   ʔa
               Phú with   Hoa   two CLF.PERSON 3.ANIM   2SG invite
               ~~pu hŏŋm hwa$_i$ ʈwa jaŋ ɲu$_i$~~ hu
                                                   ROOT
   INTENDED: 'You can invite Phú and Hoa, those two people.'

b. **pu hɔ̌ŋm̂ hwa**$_i$ hɨ ʔa [t̪wa jaŋ   ɲu$_i$] hu
   Phú with Hoa 2SG invite   two CLF.PERSON 3.ANIM   ROOT
   'You can invite Phú and Hoa, those two people.'

The one configuration that approximates DC-marking of a DR-appositive is that in (4.79b). Here, the name is pronounced in Spec-CP and the complex pronoun in the base position. Given that complex pronouns can be pronounced on their own, this sentence can be analyzed as base generation of the name in Spec-CP, which binds the complex pronoun due to an Op-probe on C (see the analysis of base generation + Agree in Chapter 2, Section 2.3.3).

4.7 Conclusion

To conclude, DP-internal DC-marking presents another way in which DC patterns with other Ā-features. Ā-features such as *wh* are known to be specified on probes not only on C, but also on D. This CP/DP parallelism extends to DC. DC-movement is triggered by a probe either on C or D (Figure 4.16a). Inventory forms and partitives arise from this DP-internal DC-movement.

Figure 4.16: Eastern Cham DC-probes on D

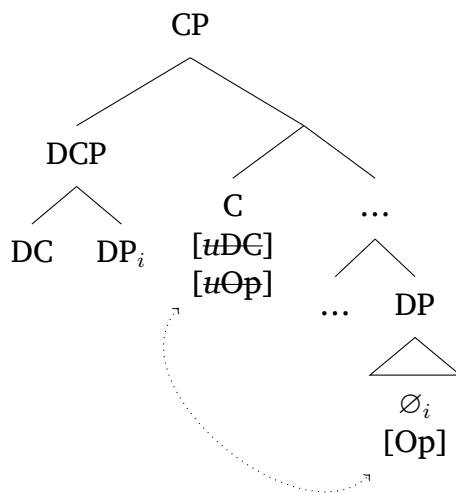(a) DC-movement to Spec-CP   (b) DC-movement to Spec-DP
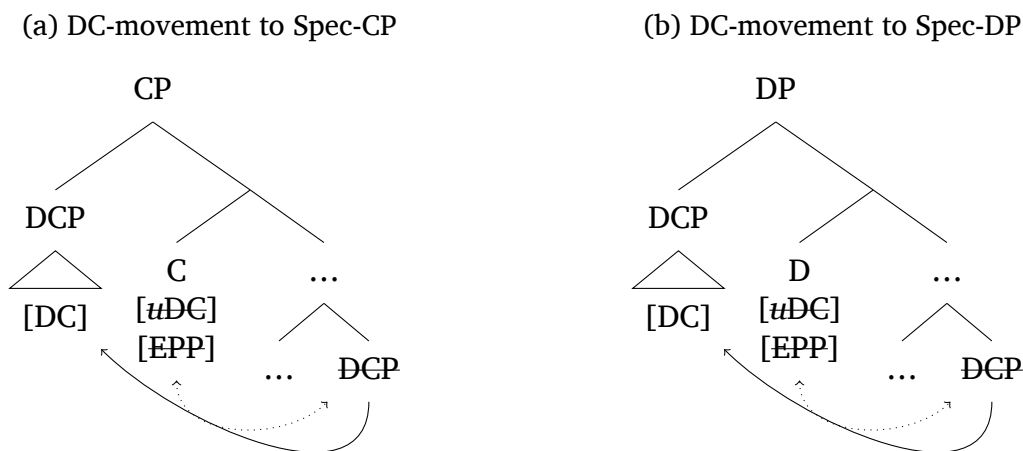
Additionally, this chapter has found that there can only be one DC-related feature in the syntax. Hence, a phrase can only be DC-marked once in a derivation. This precludes pied-piping of an inventory form or partitive, or secondary DC-movement (cf. the Tzotzil example in the introduction). DC-marking can feed A-movement, however, due to other syntactic features and probes. In this case, the DC-phrase retains its DC pragmatics.

Finally, this chapter explored domain restricting- or DR-appositives and found that they are licensed not by DC but by definiteness or familiarity. Based on the different pragmatic distributions between inventory forms and partitives on the one hand and DR-appositives on the other, DC is concluded to be an independent grammatical notion from definiteness. One complication is that, while DC-marked phrases can be base generated in Spec-DP (Figure 4.17a), there does not appear to be a corresponding probe on D (Figure 4.17b). Instead, names base generated in Spec-DP, at least in DR-appositives, need only be definite. The DC/Op-probe, thus, only appears to be merged on C, not D in Eastern Cham.

Figure 4.17: Eastern Cham DC-probes on D

(a) Base generation in Spec-CP and Agree   (b) Base generation in Spec-DP
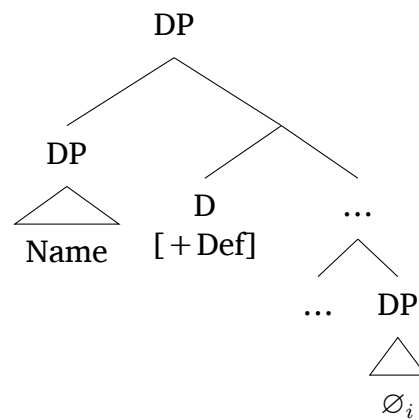
This chapter has unified two phenomena, inventory forms and partitives, as DP-internal DC-movement. A further prediction is that any instance of DP-internal movement be an instance of DC-movement in Eastern Cham, unless another probe is found. By contrast, DR-appositives demonstrate that base generation in Spec-DP reflects anaphoric definiteness, not DC. Another further prediction is that other base generation phenomena not be instances of DC-movement in Eastern Cham. Further research is needed on phenomena such as possessives, which may present opportunities to test these predictions.

Chapter 5

Conclusion

This dissertation has argued for the existence of a new Ā-feature, *discourse connected*, or DC, which exists at the interface between syntax and discourse structure. On one hand, DC is an Ā-feature on par with *wh*. DC-movement exhibits hallmarks of Ā-movement, such as unboundedness, sensitivity to syntactic islands, weak crossover effects, and locality effects when multiple phrases are moved. DC-marking also exhibits a parallelism between CP and DP, which has been noted to occur with other Ā-features such as *wh*.

On the other hand, DC marks a hierarchical discourse constraint on par with sentence connectives like *For example* and *That's because*. In a sense, DC-marking has a similar purpose to those connectives, but it does so via Ā-movement of a phrase that is a participant in two semantic events evoked by a subordinating discourse relation. In this sense, there must be an interface between syntax and discourse structure. Minimally, this interface must include DC, but there may well be other hierarchical discourse constraints marked by Ā-features.

DC adds to the typology of Ā-features as well. Ā-features include inherent features of lexical items such as *wh*, information structural features such as topic and focus, and now features that mark hierarchical discourse constraints. It is important to search for additional possible Ā-features. For one, a complete understanding of Ā-movement phenomena requires a complete understanding of the possible features. For another, the scope of possible Ā-features must be understood in order to fully examine generalized Ā-feature effects or Ā-feature hierarchy effects, which apply to all or a subset of the possible Ā-features.

In this dissertation, a range of different syntactic constructions in Eastern Cham were found to be instances of DC-marking. Topicalization (5.1a) involves either DC-movement of a phrase or base generation of the phrase in Spec-CP, which then binds a resumptive pronoun. Optional *wh*-movement (5.1b) involves DC-movement of a bare or D-linked *wh*-phrase, or likewise base generation of the phrase in Spec-CP, which then binds a resumptive pronoun.

(5.1)   *Manifestations of DC-marking in Eastern Cham:*

    a. **ʔiŋ ʔɔŋ năn**_{DC} ɲu   ʈɔʔ   ŋăʔ   ~~ʔiŋ ʔɔŋ năn~~ ɲi   lo

       frog   that   3.ANIM PROG make   be.delicious very

       'That frog, he is cooking very well [Lit: deliciously].'   (Topicalization)

    b. **ḳeʔ**_{DC} hɨ   ʈɔʔ   ɓăŋ ~~ḳeʔ~~

       what 2SG PROG eat

       'What are you eating?'   (Optional *wh*-movement)

    c. kăw plĕj [_{DP} **ʔɔ̆ʔ**_{DC} çŭh pɔh   ~~ʔɔ̆ʔ~~]

       1SG buy   mango 7   CLF.ROUND

       'I bought mangoes, seven.'   (Inventory form)

    d. kăw plĕj [_{DP} [_{DP} **ʔɔ̆ʔ**   **ni**_{DC}] çŭh pɔh   ~~ʔɔ̆ʔ ni~~]

       1SG buy   mango this   7   CLF.ROUND

       'I bought seven of these mangoes.'   (Partitive)

    e. **hɔŋm͡**_{DC} lĕʔ tʃŭn [_{DP} ~~hɔŋm͡~~ çŭh pɔh   ]

       papaya fall down   7   CLF.ROUND

       'Papaya, seven fell down.'   (Inventory form feeding quantifier float)

Inventory forms (5.1c) arise when an NP is DC-moved to the left edge of a DP to the exclusion of the rest of the phrase. Partitives (5.1d) are similar, where the embedded DP which represents the contextually bound set is DC-moved to the left edge of a DP to the exclusion of the subset phrase. Finally, DC-marking underlies other phenomena such as quantifier float (5.1e). All of these constructions were analyzed with a simple architecture involving one feature, DC, and a small number of corresponding syntactic probes.

This dissertation has argued against some competing hypotheses for the Eastern Cham effects, from edge feature-movement to definiteness, D-linking, and topicalization. If the effects observed in this dissertation are hypothesized to be due to edge feature-movement (i.e. syntactic movement driven by an EPP-feature, but not a specific feature of the moved phrase), it would have to be explained why the particular semantics and pragmatics of DC arise, as opposed to any other pragmatic interpretation. The DC-particle is argued to introduce a presupposition that makes reference to semantic events, encoded in the semantics. An edge feature analysis would only predict general discourse or pragmatic effects. DC-marking could also potentially be analyzed as prosodic movement in the sense of Zubizarreta (1998). To assess a prosodic movement hypothesis, a full picture of prosody in Eastern Cham is needed. If it can be shown that subordinating discourse relations in particular correspond with a certain prosody, and that prosody is responsible for DC-marking, then a prosodic movement analysis could be tenable.

As for feature-based analyses, DC could be argued to be a type of topicality or another existing pragmatic notion. In other words, perhaps there is no need for a DC-feature, only a topic feature. A topic-based analysis would have to explain why DC effects specifically

require subordinating discourse relations and semantic events in Eastern Cham, while such effects have not been observed in many other languages. Conversely, some predictions proposed for topicality are not borne out in Eastern Cham, such as the ability of *wh*-phrases and quantifiers to be DC-marked. However, the extent to which topicality and DC overlap is an empirical question; more languages need to be tested for DC-marking.

Taken all together, any analysis of the effects observed in this dissertation would have to account for the syntactic similarities between DC and other Ā-features such as *wh* and the specific semantics and pragmatics introduced by DC. The analysis of DC proposed in this dissertation accounts for these observations in a straightfoward manner and unifies a range of syntactic constructions under the umbrella of DC-marking, using a simple syntactic architecture of one feature and a small number of corresponding syntactic probes.

Going forward, this disseration has laid out a general framework for diagnosing discourse connectedness, using discourse-pragmatic elicitation and related syntactic facts. Appendix A gives a short questionnaire and guide for testing for DC-marking. In order to more fully understand DC, more languages and varieties remain to be tested, along with more constructions that could be explained in terms of DC-marking. A greater range of hierarchical discourse constraints should also be investigated, such as those involving coordinating discourse relations or individual discourse relations such as narration. Additional hierarchical discourse constraints could contribute to the Ā-feature hierarchy and our understanding of the syntax–discourse structure interface. Finally, this dissertation has established that multi-sentence discourses controlled for discourse structure are necessary when testing syntactic phenomena.

The remainder of this chapter begins to explore two of these directions. First, Section 5.1 sketches basic facts about DC-movement in Catalan, much of which were originally laid out by López (2009). A variety of differences between Catalan and Eastern Cham are presented, ranging from the kinds of phrases that can be DC-marked, to the pragmatic effects of multiple DC-movement. A small typology of DC-marking is proposed. Second, Section 5.2 proposes that contrastive topic represents another hierarchical discourse constraint, which at first appears to be a subtype of DC. In the end, contrastive topic will be analyzed as a separate hierarchical discourse constraint from DC, based on both Eastern Cham and Catalan.

5.1 DC-marking in Catalan

This section examines DC-marking in Catalan in more detail, in order to test the analysis proposed for Eastern Cham and to expand the typology of DC-marking. Overall, the analysis from Eastern Cham is affirmed: DC must be a syntactic feature alongside *wh*. Syntactic differences between Catalan and Eastern Cham, such as the ability of pronouns and *wh*-phrases to be DC-marked are ascribed to independent syntactic properties of Catalan. In terms of semantics, Catalan necessitates a second type of DC-particle that combines with properties in addition to individuals (5.2a–b). Eastern Cham only has one DC-particle,

which combines with individuals. This difference between Catalan and Eastern Cham is due to the semantic interpretation of DC-marked nouns as kinds in Eastern Cham and predicates in Catalan.

(5.2) a. $[\![\text{DC}_1]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$ (Individual DC-marking)

  b. $[\![\text{DC}_2]\!] = \lambda P : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].P$ (Property DC-marking)

  c. DISCOURSE LOCALITY: For $x \in \mathcal{P}_e \cap \mathcal{P}_{e'} \gg y \in \mathcal{P}_e \cap \mathcal{P}_{e''}$, $e' \Downarrow e'' \Downarrow e$

In terms of pragmatics, an implicature is found when the order of DC-moved phrases is free: *discourse locality* (5.2c). In short, if a DC-marked phrase is moved to a higher syntactic position than another, it is interpreted as being previously mentioned in a higher superordinate event in the discourse. Discourse locality arises in Catalan, but not Eastern Cham in the general case, because the relative order in Catalan is not fixed, perhaps due to the existence of doubled clitics.

Throughout this section, uncited examples from Catalan were elicited from one native speaker consultant. Examples from López (2009) were checked with that consultant. When cited, the forms and glossing from López (2009) are used, though the precise dialect and register of Catalan is not the same as that spoken by the consultant.

The following sections present the basic generalization of DC-marking in Catalan, then three differences between DC-marking in Eastern Cham and Catalan: the presence or absence of a restriction to individuals; the ability of *wh*-phrases to be DC-marked; and the presence or absence of discourse locality effects. Finally, one apparent difference in contrastive topic-marking will be found that will then be explored in Section 5.2.

5.1.1   CLRD is DC-marking

The very notion of discourse connectedness is rooted in López's (2009) account of clitic right- and left-dislocation in Catalan. López (2009) first pointed out the connection between discourse subordination and syntactic movement. Clitic-right dislocation (CLRD) in particular shares many characteristics with DC-movement in Eastern Cham.

According to López (2009: 90), CLRD involves A-movement of a phrase to a rightward specifier of *v*P, though there are other analyses that have been put forth (López 2009: 247). Both the type of movement and its landing site are different from DC-movement in Eastern Cham, which involves Ā-movement to Spec-CP or Spec-DP. Nevertheless, the parallelism between DC and *wh* remains. It has been argued in a variety of languages that *wh*-phrases are moved to Spec-*v*P, in addition to other Ā-movement operations such as focus-movement and topicalization (e.g. Belletti 2004). For example, *wh*-phrases in Hindi-Urdu (Manetta 2006) and Spanish (5.3) have been argued to move to Spec-*v*P. In (5.3a), the unmarked position of a temporal phrase *el martes* 'on Tuesday' appears in the middle of the predicate. However, the corresponding *wh*-phrase *cuándo* 'when' must be pronounced at the right edge of the predicate (5.3b–c).

(5.3)   a. ¿Sergio llegó **el martes** en bicicleta?
            Sergio arrived on Tuesday by bicycle

        'Sergio arrived by bicycle on Tuesday.'   SPANISH

       b. ?? ¿Sergio llegó **cuándo** en bicicleta?
            Sergio arrived when   by bicycle

        INTENDED: 'When did Sergio arrive by bicycle?'

        (Uribe-Etxebarria 2002: 221, citing Jiménez 1997)

       c. ¿Sergio llegó ~~cuándo~~ en bicicleta **cuándo**?
            Sergio arrived   by bicycle when

        'When did Sergio arrive by bicycle?'

CLRD in Catalan is exemplified in (5.4). In (5.4a), the object *la carn* 'the meat' is doubled by the clitic *la* and dislocated to the right. In (5.4b), the subject is doubled by a null clitic, as subjects cannot have overt clitic doubling, and dislocated. Note that Catalan does not have a matrix subject restriction like Eastern Cham; the matrix subject can be CLRD'd. The restriction in Eastern Cham is likely a more general restriction on Ā-movement (cf. Section 2.2.3).

(5.4)   a. El Joan la$_i$ va   cuinar, **la carn**$_i$.
            the Joan CL.ACC PAST.3SG cook.INF the meat

        'Joan cooked the meat.'   CATALAN

       b. $\varnothing_i$ Va   cuinar   la   carn, **el Joan**$_i$.
            pro PAST.3SG cook.INF the meat   the Joan

        'Joan cooked the meat.'

Turning to the pragmatics, CLRD is a case of DC-movement. To illustrate, consider the positive and negative examples of CLRD in (5.5b–b′). CLRD of the phrase *la carn* 'the meat' is felicitous in (5.5b). As expected from Eastern Cham DC-marking, CLRD is felicitous here if *la carn* is previously mentioned in a sentence that the current one explains or elaborates upon. In this context, (5.5b) either explains why Joan is cooking the meat, or elaborates upon Joan's relationship to it. In either case, *la carn* is previously mentioned in a superordinate sentence. Note that the relevant referents are marked with indices in these examples in order to clarify the clitic doubling. In this example, *la carn* is doubled by a null clitic, as typically occurs with subjects in Catalan.

(5.5)   a. El Joan$_i$ va   cuinar **la carn**$_j$.
            the Joan PAST.3SG cook.INF the meat

        'Joan cooked the meat.'   CATALAN

b. Ø*ⱼ* Li*ᵢ*   agrada   molt   ~~la carn~~, **la**   **carn***ⱼ*.
     pro   CL.DAT   like.3SG   much   the   meat
     'He likes the meat very much.'   (a ⇓ b)
     [LIT: The meat pleases him very much.]

b′. #Després   se*ᵢ*   la*ⱼ*   va   menjar ~~la carn~~, **la**   **carn***ⱼ*.
     afterwards   CL.REFL   CL.ACC   PAST.3SG   eat.INF   the   meat
     INTENDED: 'Afterwards he ate it/the meat.' (López 2009: (2.56))   (a ⇗̸ b′)

By contrast, there is no discourse subordination between (5.5a) and (5.5b′), as the latter transitions to a new event of eating, despite the previous mention of *la carn* 'the meat'. This minimal difference renders CLRD infelicitous. Note that there is a felicitous version of both (5.5b,b′) in which *la carn* is omitted and only the clitic is pronounced.

López (2009) presents a range of monologic examples where CLRD is felicitous in subordinating discourse relation contexts and infelicitous in non-subordinating ones. Catalan also exhibits the difference between direct and elaborating answers to questions (cf. Chapter 2, Section 2.1.1). Answers that elaborate upon questions, by providing more information than asked, are licit contexts for CLRD. For example, (5.6b) gives more information than the polar question asks. In fact, it provides either an elaboration upon or an explanation for a negative answer (i.e. 'No. Why? I already cooked the meat.').

(5.6)   a. Vols   que cuini   **la**   **carn***ᵢ*?
         want.2SG   that   cook.1SG.SBJV   the   meat
         A: 'Do you want me to cook the meat?'

     b. Ja   l*ᵢ*'he   feta   jo, **la**   **carn***ᵢ*.
         already   CL.ACC'have.1SG   made   I   the   meat
         B: 'I already cooked the meat.' (López 2009: (2.67))   (a ⇓ b)

     b′. #No, no   la*ᵢ*   vull   que cuinis   ~~la carn~~, **la**   **carn***ᵢ*.
         NEG   NEG   CL.ACC   want.1SG   that   cook.2SG.SBJV   the   meat
         B: 'No, I don't want you to cook the meat.'   (a ⇗̸ b)

The direct answer in (5.6b′) does not license CLRD, however. This is because direct answers only answer the question raised and do not on their own provide an explanation or elaboration.

The only way for a direct answer to exhibit CLRD is if the whole question-answer pair itself serves to elaborate or explain upon prior discourse. This argumentation is a focus of Baclawski Jr's (2018d) analysis of the discourse effects of the question *What happened* and will play a role in the analysis of contrastive topic put forth in Section 5.2. To illustrate, consider (5.7). The question-answer pair composed of (5.7b–c) serves to explain why Paolo no longer works at the café. Even though (5.7c) is a direct answer to the previous question, it contributes to the explanation of (5.7a), where *el cafè* 'the café' is previously mentioned.

(5.7)   a. A: Paolo stopped working at **the café**$_i$.

     b. ¿Què va   passar?
         what PAST.3SG happen.INF
         B: '[Why] What happened?'   (a $\Downarrow$ b/c)

     c. Es$_i$ va   tancar, **el** **cafè**$_i$.
         CL.REFL PAST.3SG close.PART the café
         A: 'The café closed.' (Baclawski Jr 2018d: 90)   (a $\Downarrow$ b/c)

If the question-answer pair does not serve to explain or elaborate upon prior discourse, though, CLRD becomes infelicitous again. A minimal pair is presented in (5.8), with the same question-answer pair as above. Here, however, the question-answer pair is interpreted as filling in a sequence of events. The same judgments will be found in Eastern Cham in Section 5.2.

(5.8)   a. A: First, **the café**$_i$ was renamed. Then it$_i$ was moved. Then…

     b. ¿Què va   passar?
         what PAST.3SG happen.INF
         B: '[Then] What happened?'   (a $\nDownarrow$ b/c)

     c. #Es$_i$ va   tancar, **el** **cafè**$_i$.
         CL.REFL PAST.3SG close.PART the café
         INTENDED: B: 'The café closed.' (Baclawski Jr 2018d: 89) (a $\nDownarrow$ b/c)

CLRD, then, is a kind of DC-movement. It is licensed only if the moved phrase satisfies the DC conditions.

5.1.2   Individuals and properties

Recall from Chapter 2, Section 2.1.3 that DC-marking in Eastern Cham is restricted syntactically to DPs and semantically to individuals (and properties vis-à-vis kind operators). In the semantics below, the DC-particle combines with an individual x and introduces a presupposition that makes reference to a set of participants in two events ($\mathcal{P}_e \cap \mathcal{P}_{e'}$).

(5.9)   a. Let $\mathcal{E}_c$ be the set of events live in a discourse at context c

     b. Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

     c. Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

     d. Let $\mathcal{P}_e$ be the set of participants in event e

     e. $[\![\text{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$ EASTERN CHAM

Catalan does not exhibit these syntactic and semantic restrictions. Adjective phrases and prepositional phrases can be CLRD'd. For instance, an adjective phrase is CLRD'd in (5.10b). Here, the current sentence is interpreted as providing an explanation for Maria's cleverness, or an elaboration upon it. Furthermore, *llesta* 'clever' serves as a sufficient previous mention of *d'intel·ligent* 'intelligent', because, according to consultants, intelligence is interpreted as one of the components or prerequisites of cleverness in this context.

(5.10)   a. La   Maria   és   molt   **llesta**,   oi?
         the   Maria   is   very   clever,   isn't.she

   A: 'Maria is very clever, isn't she?'

   b. Sí   que   ho   és,   **d'intel·ligent**.
      yes   that   CL   is   of.intelligent

   B: 'Indeed, she is intelligent.' (López 2009: (2.42))   (a ⇓ b)

To account for examples like this, a broader semantics of DC must be possible. While a full analysis will not be presented here, one avenue would be to introduce a second DC-particle that combines with a property (5.11b). In cases where adjective phrases and prepositional phrases are CLRD'd in Catalan, perhaps the AP or PP combines with DC_2. This DC-particle combines with a property, checks if that property is previously mentioned in the appropriate event, and returns that property. This DC-particle would require an articulated notion of $\mathcal{P}$ that tracks both individuals and predicates mentioned in events.

(5.11)   a. $[\![DC_1]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$ CATALAN

   b. $[\![DC_2]\!] = \lambda P : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].P$

Another potential avenue, not explored here, would be to use the notion of Universe from Discourse Representation Theory (Kamp & Reyle 1993; Kamp, van Genabith & Reyle 2011). Universes are tuples of (at least) individual referents and predicates that a sentence is composed of. If the meaning of the DC-particle makes reference to universes, the two DC-particles could be reduced to one. This option is explored in prior work, at least as a heuristic for the semantics of DC (Baclawski Jr forthcoming(b)). However, it is unclear if it could account for Eastern Cham, especially with regard to the event semantic data found in Chapter 2, Section 2.1.2.

A second difference between the kinds of phrases that can be DC-marked in Eastern Cham and Catalan is found in pronouns. Chapter 2, Section 2.1.3 showed that pronouns cannot be DC-marked in Eastern Cham. By contrast, Catalan pronouns can be CLRD'd, as in (5.12).

(5.12) a. A: Did you see **Joan**$_i$?

    b. Sí, el$_i$ vaig   veure, **a ell**$_i$.
       yes CL.ACC PAST.1SG see.INF A him
       B: 'Yes, I saw him.' López 2009: (2.115)   (a ⇓ b)

To briefly summarize, the ability of adjective and prepositional phrases to be CLRD'd in Catalan requires an expansion of the semantics of DC: there must be a version of the DC-particle that allows for DC-marking of properties. The ability of pronouns in be CLRD'd can be explained by the independent syntax of pronouns.

5.1.3 *Wh*-phrases

One major difference between Eastern Cham and Catalan arises with *wh*-phrases. Chapter 3, Sections 3.2.1–3.2.2 argued that *wh*-phrases of all kinds can be DC-marked in Eastern Cham. In Catalan, however, none can be. Specifically, *wh*-phrases cannot be clitic dislocated, as in (5.13a). Instead, they are *wh*-moved without clitic doubling (5.13b). Note that López (2009: 120) acknowledges that D-linked *wh*-phrases of the form *which X* should in theory be anaphoric on the same level as clitic dislocated phrases (López's (2009) [+a]). Nevertheless, they cannot be clitic dislocated.

(5.13) a. $^{??}$Quina pel·lícula la   vas   veure?
       which movie   CL.ACC PAST.2SG see.INF
     INTENDED: 'Which movie did you see?' López 2009: (3.91)

    b. Quina pel·lícula vas   veure?
       which movie   PAST.2SG see.INF
   'Which movie did you see?'

Why should *wh*-phrases be unable to be being DC-marked? One option would be to stipulate that *wh*- and DC-phrases are incompatible in Catalan. Using Cable's (2010) Q-particle, perhaps the Q-particle in Catalan cannot select a phrase headed by the DC-particle (Figure 5.1a). Conversely, no such restriction would exist in Eastern Cham. This account stipulates a narrow incompatibility between *wh* and DC and makes the prediction that some languages have this incompatibility, while others do not.

Figure 5.1: Possible explanations for the lack of DC-marking of *wh*-phrases

(a) Q- and DC-particle incompatibility
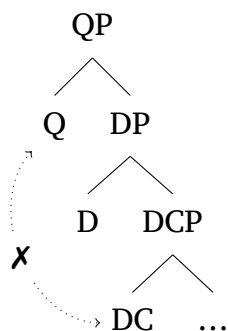
(b) Antilocality or criterial freezing
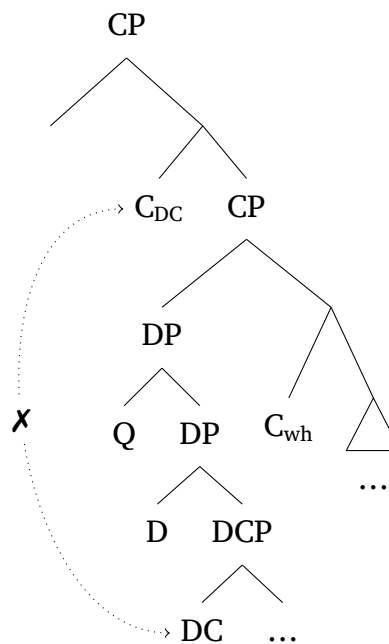
A second option would be to posit that *wh*-movement is incompatible with DC-marking. Perhaps phrases that move to Spec-CP$_{wh}$ cannot enter an Agree relation with the DC-probe due to specifier-to-specfier anti-locality or criterial freezing as in Figure 5.1b (cf. Section 2.3.1 for a similar discussion for DC-movement in Eastern Cham). Or, perhaps DC-phrases and Q-phrases are interveners to each other in a meaningful sense in Catalan. This restriction would not arise in Eastern Cham, as there is no phrasal *wh*-movement (cf. Section 3.4). Note that the DC-feature would be predicted to percolate up to the highest DP (Chapter 3, Section 3.3.1). This account makes a stronger prediction than the prior account: there should not be any *wh*-movement languages that also exhibit DC-movement.

This dissertation will not provide evidence to distinguish between these two accounts. There are languages where the data needed could be found. For example, in situ *wh*-phrases in certain varieties of French (e.g. Cheng & Rooryck 2000) and Italian (e.g. Munaro 2003 on Bellunese) are reported to have properties similar to D-linking. Additionally, in Romanian, the relative ordering of *wh*-phrases in the left periphery seems to be sensitive to the topical or D-linked properties, suggesting there is a phrasal projection above the landing site for *wh*-movement (Comorovski 1996: 2). The existence of a

D-linking-like property reported for *wh*-phrases implies that DC-marking is a possibility, given the surface similarity between DC and D-linking (Section 3.2.2). However, these languages have not been investigated in terms of DC.

The two accounts proposed here for the lack of DC-marking on *wh*-phrases in Catalan focus on syntactic restrictions. It is possible that Q- and DC-particles fail to mark the same phrase for semantic reasons. However, it is unclear why an NP such as *pel·lícula* 'film' would fail to combine with the DC-particle based on the analysis proposed for Eastern Cham in Chapter 3, Section 3.2.1, especially given that the DC-particle has a broader distribution in Catalan (Section 5.1.2 above). Syntactic restrictions are more parsimonious in this case, as they ascribe the difference between Eastern Cham and Catalan to differences in the respective lexicons.

5.1.4 Discourse locality

Next, Catalan exhibits an effect López (2009: 39) calls a kind of discourse-level locality. In Catalan, the hierarchy of DC-marked phrases must follow the hierarchy of superordinate sentences in the discourse. Recall that multiple DC-phrases have a strict relative order in Eastern Cham due to syntactic locality restrictions (cf. Chapter 2, Section 2.3.1 on Path Containment effects). Multiple DC-phrases in Catalan do not exhibit strict ordering effects. Instead, the relative order of multiple DC-phrases reflects different discourse structures.

To illustrate, the context in (5.14) contains a chain of two discourse subordination relations, which are reflected in the relative order of DC-moved phrases in the final sentence. The first sentence (5.14a) is elaborated by the question (5.14b), which asks an elaborating question about the event of going to the Mass. Then, the answer in (5.14c) elaborates upon the question in (5.14b) by providing more information than asked by that polar question.

(5.14)   a. *Llavors, ahir vas anar **a missa**$_i$, oi?*
             A: 'So then, yesterday, you went to the Mass, right?'

         b. *Què hi vas veure **el Joan**$_j$?* A: 'Did you see Joan?'   (a $\Downarrow$ b)

         c. Si, l$_j$'hi$_i$ vaig   veure, **el Joan**$_j$, **a missa**$_i$.
            yes CL.ACC'CL.LOC PAST.1SG see.INF the   Joan   at   Mass
            B: 'Yes, I saw Joan at the Mass.'   (b $\Downarrow$ c)

         c′. $^{??}$*Si, l$_j$'hi$_i$ vaig veure, **a missa**$_i$, **el Joan**$_j$.* (After López 2009: (2.40))   (b $\Downarrow$ c)

According to López (2009) and confirmed by my consultant, in this context, *a Missa* 'at Mass' is preferred to be clitic right-dislocated to the right of *el Joan* 'Joan', as in (5.14c). The opposite order in (5.14c′) is dispreferred. López (2009: 40) calls this a slight preference, but my consultant reports a robust distinction. Contrast (5.14) with (5.15). In (5.15), *el Joan* 'Joan' is introduced higher in the discourse than *a Missa* 'at Mass', and the respective order of CLRD'd phrases in (5.15c) must be flipped from the example above.

(5.15)   a. *Vas veure **en Joan**$_j$ ahir?* A: 'Did you see Joan yesterday?'

b. *Era **a Missa**$_i$.* A: 'He was at the Mass.'   (a ⇓ b)

c. $^{??}$*Si, l$_j$'hi$_i$ vaig veure, **el Joan**$_j$, **a missa**$_i$.*   (b ⇓ c)

c′. Si,   l$_j$'hi$_i$ vaig   veure, **a missa**$_i$, **el   Joan**$_j$.
     yes   CL.ACC'CL.LOC PAST.1SG see.INF at Mass   the Joan
     B: 'Yes, I saw Joan at the Mass.'   (b ⇓ c)

Syntactically, López (2009) analyzes CLRD as movement to a rightward specifier of *v*P and hypothesizes that the ordering restriction is due to some kind of pragmatic accessibility in the sense of Ariel (1990) or Lambrecht (1994). Depending on the discourse, *el Joan* 'Joan' is either more or less accessible than *a Missa* 'at Mass'. The underlying syntax involves two specifiers of *v*P that have no syntactic ordering restriction. This does not raise an issue for López (2009), as CLRD is not driven by Agree relations.

If CLRD is due to an Agree relation with a DC-probe as in Eastern Cham, however, we would expect a syntactic ordering restriction. In other words, both orders of 'Joan' and 'at Mass' should not be possible. If there is one DC-probe, we would expect a tucking-in order, or if there are multiple probes searching for the same feature, we would expect a Path Containment effect, both reflecting the phrases' base positions (cf. Chapters 2–3).

Instead, Catalan exhibits free ordering of multiple CLRD'd phrases. In the examples above and throughout López (2009), one DP and one PP are CLRD'd. However, free order also obtains for multiple DPs (5.16a) and multiple PPs (5.16b).

(5.16)   a. $\emptyset_i$ La$_j$ va   cuinar, **el   Joan**$_i$, **la   carn**$_j$.
           pro CL.ACC PAST.3SG cook.INF the Joan   the meat
           'Joan cooked the meat.'

a′. $\emptyset_i$ La$_j$ va   cuinar, **la   carn**$_j$, **el   Joan**$_i$.
     pro CL.ACC PAST.3SG cook.INF the meat   the Joan
     'Joan cooked the meat.'

b. N$_i$'hi$_j$ vaig   donar   una, de joguina, **al   nen**$_i$, **al**
     CL.DAT'CL.LOC PAST.1SG give.INF a   of toy   to.the child in.the
     parc$_j$.
     park
     'I gave the child a toy in the park.'

b′. N$_i$'hi$_j$ vaig   donar   una, de joguina, **al   parc**$_j$,
     CL.DAT'CL.LOC PAST.1SG give.INF a   of toy   in.the park
     al   nen$_i$.
     to.the child
     'I gave the child a toy in the park.'

There is a different syntactic analysis, however, that preserves CLRD as DC-marking: clitic binding. CLRD is different from DC-marking in Eastern Cham in that the moved phrases are clitic doubled. López (2009) analyzes doubled clitics as adjoining to *v*, as depicted in Figure 5.2 below. If the syntactic heads associated with CLRD, here XP and YP, have Op-probes, they enter Agree relations with lower pronouns and bind them with the phrases in their specifier positions (cf. Chapters 2–3 on base generation + Agree).

Figure 5.2: Discourse locality in (5.14)

(a) Derivation of (5.14c) (b) Derivation of (5.15c′)
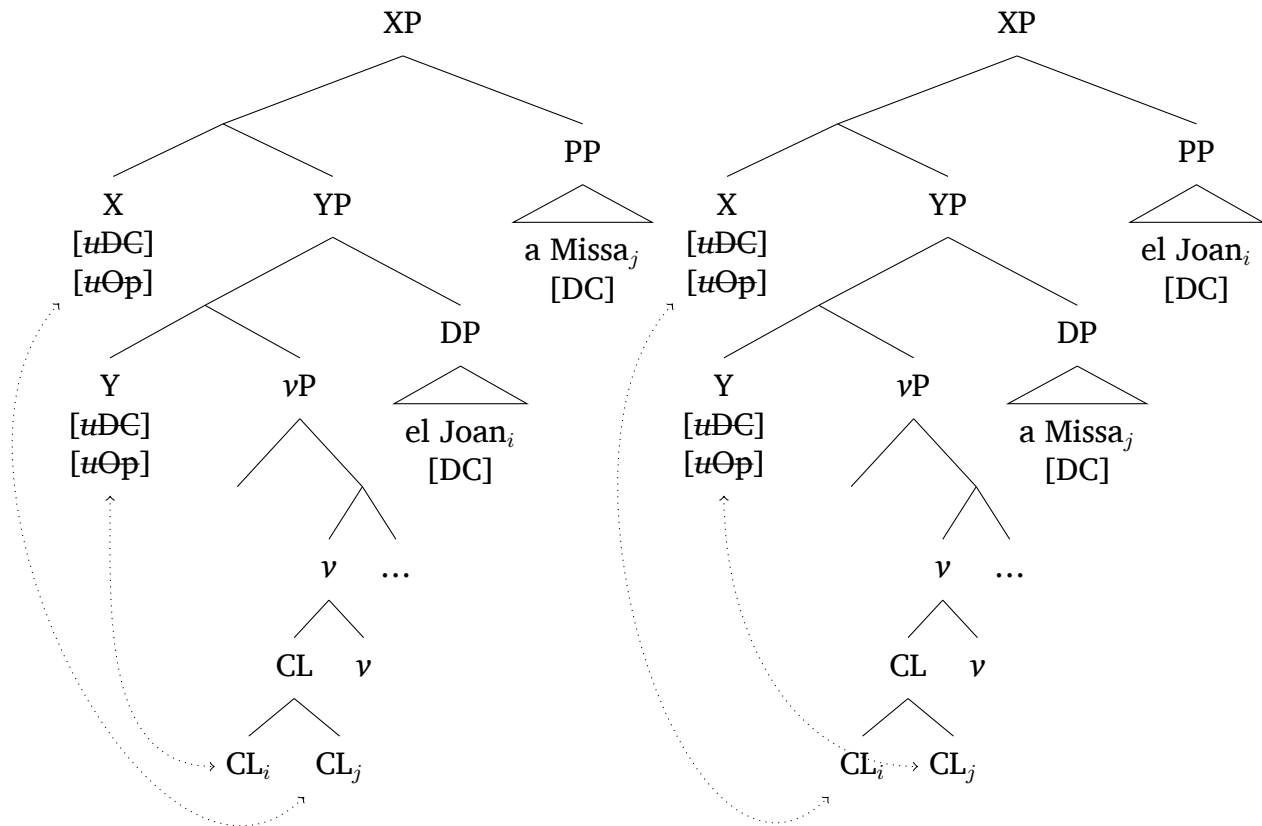
If both clitics in a multiple CLRD derivation are adjoined to *v*, they would be equally local. Therefore, a probe such as the Op-probe on Y can be free to enter an Agree relation with CL_i, as in Figure 5.2a or CL_j, as in Figure 5.2b. Eastern Cham base generation + Agree differs in that the resumptive pronouns are in their base positions and are liable to locality differences.

What then results in the discourse locality interpretation? Perhaps DC-marking introduces an implicature along the following lines. DC-marking of a phrase presupposes that the phrase is previously mentioned in a superordinate sentence. An implicature arises that the material it c-commands is subordinate to that sentence. Consider the case of (5.17c), repeated below and represented in Figure 5.2b above.

(5.17)   a. *Vas veure **en Joan**$_j$ ahir?* A: 'Did you see Joan yesterday?'

      b. *Era **a Missa**$_i$.* A: 'He was at the Mass.'   (a $\Downarrow$ b)

      c′. Si, l$_j$'hi$_i$ vaig   veure, **a missa**$_i$, **el Joan**$_j$.
          yes CL.ACC'CL.LOC PAST.1SG see.INF at Mass   the Joan
          B: 'Yes, I saw Joan at the Mass.'   (b $\Downarrow$ c)

In (5.17c), *el Joan* 'Joan' is the structurally highest CLRD'd phrase. According to the implicature sketched above, DC-marking of *el Joan* implicates that YP in Figure 5.2b contains material in subordinate sentences to the one that mentions *el Joan*. Then, DC-marking of *a Missa* 'at Mass' implicates that the material in *v*P is subordinate to that sentence. The only preceding discourse that conforms to these implicatures is one in which *el Joan* is mentioned in a sentence superordinate to the sentence that mentions *a Missa*, which in turn is superordinate to the current sentence.

Discourse locality is schematized in (5.18) below. For a DC-marked phrase x mentioned in an event in one previous sentence and y previously mentioned in an event in another, if x is structurally higher than y, an implicature arises that there is a discourse subordination chain such that the event containing x (here, e') is superordinate to the event containing y (here, e'')

(5.18) DISCOURSE LOCALITY: For $x \in \mathcal{P}_e \cap \mathcal{P}_{e'} \gg y \in \mathcal{P}_e \cap \mathcal{P}_{e''}$, $e' \Downarrow e'' \Downarrow e$

A full account of discourse locality will require further research, but the hypothesis above makes the following prediction. Discourse locality effects should arise if there are multiple possible orders licensed by the syntax. They do not arise in Eastern Cham, as multiple DC-marking is constrained by syntactic locality of phrases in their base positions.

5.1.5 Contrastive topic

Finally, Catalan differs from Eastern Cham in that contrastive topic (CT) appears to be marked in a similar way to DC-marking. Clitic right-dislocation corresponds with DC-movement. Clitic left-dislocation (CLLD) is claimed to mark CT (cf. Arregi 2003). López (2009) analyzes CLLD syntactically as CLRD with an additional movement step to Spec-CP and pragmatically as essentially DC-marking (López's (2009) [+a]) with an added contrast feature ([+c]).

To illustrate, *les taules* 'the tables' and *les cadires* 'the chairs' are clitic left-dislocated in (5.19b). López (2009) analyzes examples like these as involving essentially DC-marking

with an added contrastive element (i.e. that the tables contrast with the chairs, which are both subsets of the furniture). Section 5.2 explores in detail whether CT should be thought of as a kind of DC-marking; it will find in fact that it is not.

(5.19)   a. A: What did you do with **the furniture**?

b. **Les taules**$_i$ les$_i$ hi   vaig   portar al   matí,   però **les**
the   tables   CL.ACC   CL.LOC   PAST.1SG   bring   in-the   morning   but   the
cadires les$_i$ hi   vaig   portar al   vespre.
chairs   CL.ACC   CL.LOC   PAST.1SG   leave   in-the   evening
B: 'The tables I brought in the morning, but the chairs I brought in the
evening.' (López 2009: (2.48))

One datum that supports the notion that CLLD transits through the CLRD position is that a subpart of a CLRD'd phrase can be CLLD'd, leaving a remnant below. For example, *del seu avi* 'of her grandfather' is subextracted out of the CLRD'd phrase *les històries del seu avi* 'the stories of her grandfather'.

(5.20) **Del**   **seu** **avi,**   la$_i$ Maria   les   coneix totes,   les   històries$_i$
of-the   her   grandfather   the   Maria   CL.ACC   knows   all   the   stories
~~del seu avi~~

'Maria knows all of her grandfather's stories.' (López 2009: (4.40))

Eastern Cham, by contrast, does not generally use DC-marking in contrastive topic contexts. Section 5.2 shows that CT is marked instead with the existential marker *hu* and without syntactic movement. While this appears to be a difference between Eastern Cham and Catalan on the surface, it will be found that CLLD is not necessarily a type of DC-marking on its own.

5.1.6   Towards a typology of DC

Putting this section together, there is one fundamental difference between DC-marking in Eastern Cham and that in Catalan. In the former, the DC-particle can only combine with individuals. In the latter, DC-particles can combine with individuals or properties (5.21).

(5.21)   a. $[\![\text{DC}_1]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$ EASTERN CHAM + CATALAN

b. $[\![\text{DC}_2]\!] = \lambda P : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].P$ CATALAN

Other differences between Catalan and Eastern Cham have been accounted for in terms of differences in each language's respective lexicons. The Catalan Q-particle cannot select a DC-phrase, resulting in a ban on *wh*-phrases not being DC-markable. Eastern Cham pronouns cannot be DC-marked owing to their syntactic properties. Discourse locality

effects arise when there are multiple possible relative orders of DC-phrases, which occurs in Catalan more readily than Eastern Cham, possibly due to the clitic doubling aspect of DC-marking in Catalan.

More research is needed on other languages that exhibit DC-marking. Perhaps there are other major syntactic and semantic differences in DC-marking yet to be found. The following section turns to contrastive topic as a construction, which has arisen as a possible point of contrast between Eastern Cham and Catalan, but upon closer investigation marks a separate hierarchical discourse constraint from DC.

5.2 Contrastive topic

Finally, this section examines contrastive topic (CT) and finds that it involves a hierarchical discourse constraint, much like DC. However, despite the conceptual similarities between DC and CT, the two hierarchical discourse constraints are found not to overlap, based on evidence from Eastern Cham and a reappraisal of the evidence from Catalan. This finding in turn reinforces the event semantic approach to DC over a sentence-based rhetorical relation approach.

Contrastive topic is worth examining in the context of discourse connectedness, because both involve what appears on the surface to be topicalization. For example, CT-marking is likely the closest analog of DC-marking in languages like English (5.22). Note that throughout, CTs can either be topicalized (i.e. Ā-moved to the left edge of the clause) or marked in situ by CT prosody (cf. Jackendoff's (1972) B-accent, or the 'rise-fall-rise' contour).

(5.22) The GAZPACHO$_{CT}$, PERSEPHONE$_{FOC}$ brought…(after Constant 2014)

CT is also worth examining from a theoretical perspective, as it has been analyzed in terms of anaphora within a complex Question Under Discussion (Büring 2003; Constant 2014), which involves the QUD analog of discourse subordination. Section 5.2.1 lays out this QUD analysis of CT-marking, along with conceptual similarities and differences with DC. Section 5.2.2 finds that CT-marking is inconsistent with DC-marking in Eastern Cham; the former is marked not by movement, but the existential marker *hu* (Baclawski Jr 2018a). Then, Section 5.2.3 reevaluates Catalan CT-marking and finds that it too is inconsistent with DC-marking. In all, CT and DC represent two different hierarchical discourse constraints that are at odds with each other, as they involve complementary discourse structures.

5.2.1 DC and sub-QUDs

Contrastive topic, as analyzed by Büring (2003), Constant (2014), and others, is similar to discourse connectedness, in that it involves anaphora within a sequence of discourse moves in a conversation. Based purely on sentence-based rhetorical relations, CT appears

to be a subtype of DC. However, the event semantic approach for DC proposed in this dissertation calls the similarity into question; CT instead involves a separate hierarchical discourse constraint from DC.

Consider an informal description of discourse connectedness, based on rhetorical relations (5.23a) and the event semantics-based meaning of the DC-particle used throughout the dissertation (5.23b).

(5.23)   a. Individual x is DC iff x is mentioned in sentence ϕ and sentence ψ and ψ is interpreted as an explanation or elaboration of ϕ ($\phi \Downarrow \psi$)   (Informal)

   b. Let $\mathcal{E}_c$ be the set of events live in a discourse at context c

   Let $\mathcal{R}$ be a relation between two events, e and e', such that $e'\mathcal{R}e$ iff e is interpreted as a cause or subtype of e' (e being an event introduced in a sentence that explains or elaborates upon another)

   Let $\mathcal{E}_e$ be the set of all e' such that $e'\mathcal{R}e$

   Let $\mathcal{P}_e$ be the set of participants in event e

   $[\![\mathrm{DC}]\!] = \lambda x : \exists e' \in \mathcal{E}_c \cap \mathcal{E}_e[x \in \mathcal{P}_e \cap \mathcal{P}_{e'}].x$

Under the rhetorical relation account in (5.23a), DC requires a specific kind of relation between two sentences (what Asher & Lascarides (2003) and others call *discourse subordination*). This states that any anaphoric relation between two sentences mediated by discourse subordination should be a potential instance of DC-marking. Under the event semantic account in (5.23b), DC-marking presupposes a relation between two events, introduced in different sentences.

Turning to contrastive topic, CT-marking requires an anaphoric relation inside a complex answer to a Question Under Discussion, according to Büring (2003), Constant (2014), and others. Büring (2003) uses Roberts's (1998) Question Under Discussion (QUD) framework to account for topicalization in languages like English. In the QUD framework, discourse is modeled in terms of questions, which may or may not be overtly expressed and when overt are expressed either as questions or as corresponding answers. To illustrate, consider (5.24) from Constant (2014). The answer (5.24c) answers not the question in (5.24a), but the sub-question (5.24b).

(5.24)   a. Who ate what at the potluck?
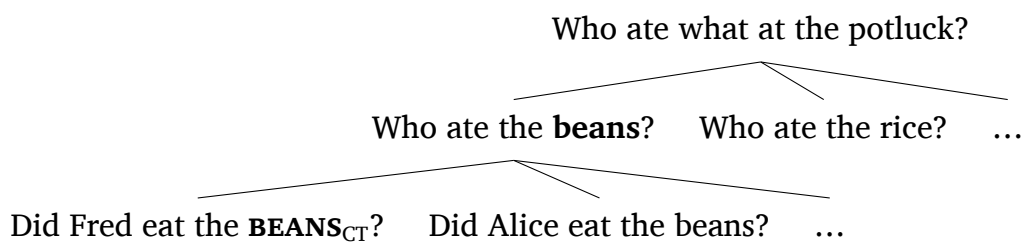
   b. (What about the beans? Who ate them?)

   c. FRED_FOC ate the BEANS_CT...(after Jackendoff 1972)

CT-marking requires the following QUD structure. First, there must be a major QUD such as (5.24a). Second, that QUD must be answered via a complex, or multiple-move answer. As a paired list question, (5.24a) must be answered in stages. The implicit sub-QUD (5.24b) first addresses each possible answer of the subject *wh*-phrase. Then, the

answer in (5.24c) answers the object *wh*-phrase with regard to *Persephone*. A complete answer would involve additional subject–object pairs, as indicated by the ellipsis.

Büring (2003) introduces a QUD-tree visualization for CT-marking, as shown in Figure 5.3. In this tree, each node represents a QUD. The question in (5.24a) is represented at the top of the tree. An intermediate node exists for each possible answer of the object *wh*-phrase, though none of these QUDs are overtly pronounced. Finally, a polar question asks about each subject–object pair. The sentence (5.24c) is represented by the focus pronunciation of the underlined node. Here, *Fred* is focussed, as it answers the QUD directly above.

Figure 5.3: QUD-tree for (5.24)

Who ate what at the potluck?

Who ate the **beans**?   Who ate the rice?   …

Did Fred eat the **BEANS**$_{\text{CT}}$?   Did Alice eat the beans?   …

Beans is a CT in (5.24c), as it is previously mentioned in the implicit sub-QUD above in the QUD-tree. Büring (2003) and Constant (2014) argue that this specific QUD structure is necessary for CT-marking: previous mention inside a complex answer to a QUD. A consequence of this analysis is that CTs must also be *anti-exhaustive* (Constant 2014: 48). In other words *beans* cannot exhaustively answer the question *Who ate what*. Instead, it only partially answers the QUD, by contributing to an answer to a sub-QUD.

With this analysis in mind, how does CT-marking compare to DC-marking? Under a sentence-based rhetorical relation account of DC, there is reason to think that CT is a subtype of DC. In the discourse above, *Fred ate the beans* is interpreted as a partial answer to *Who ate what*; it only contributes one answer pair. If the implicit sub-QUD is accepted in the example above, then CT-marking requires previous mention in a superordinate sentence in the discourse, just like DC. CT-marking would be a subtype of DC, though, in that other instances of DC-marking do not require open QUDs and sub-QUDs in this restricted manner (though, cf. Riester, Brunetti & Kuthy 2018 for a corpus implementation of the QUD framework that includes elaboration and explanation as examples of sub-QUDs). This account, then, predicts that CTs can be DC-marked.

Under an event semantic approach to DC, however, CT would not amount to a subtype of DC. Recall that DC-marking requires two events in a discourse, e and e'. CT-marking, despite the articulated QUD structure, only involves one event, in (5.24) the event of eating. CT-marking instead reflects a particular organization of the answers to constituent questions about the participants in that one event. This account, by contrast, predicts that CTs cannot be DC-marked.

Throughout this dissertation, it has been suggested that contrastive topics are not DC-marked in Eastern Cham. However, the previous section suggested that CTs are DC-marked in Catalan. The next sections go into Eastern Cham and Catalan in more detail and find that CT in fact marks a separate hierarchical discourse constraint from DC.

5.2.2 Eastern Cham CT-marking

This section examines contrastive topic-marking in Eastern Cham. CT is found to be marked not by DC-movement, but by the existential marker *hu*. Two cases are found where CT appears to be marked alongside DC, but these are found to involve additional pragmatic conditions.

First, contrastive topics are marked by *hu*, as argued by Baclawski Jr (2018a). The form *hu* has a range of uses, from a verb meaning 'have' to an existential copula, presentational cleft marker, and negation marker in stage-level predicates (cf. Section 2.1.2). Baclawski Jr (2018a) argues that *hu* is a general marker of existential closure, based on Zimmermann's (2007) conclusion about a form with a similar range of uses in Bura (Chadic: Nigeria).

In cases where subjects are CTs, *hu* precedes the subject in a presentational cleft-like construction. To illustrate, consider (5.25). The question in (5.25a) represents a QUD seeking a paired list response, as the question is directed at a group of different people and asks each of them which person they invited. Crucial to the argument is that D-linked *wh*-phrases (e.g. *which person*) indicate the *wh*-phrase around which the answer is organized in multiple *wh*-questions. This is explicitly argued by Comorovski (1996) and others, using Kuno's (1987) notion of the 'sorting key'. With this in mind, the D-linked *wh*-phrase *jaŋ hlĕj* 'which person' indicates that the answer is organized around inviters, not invitees, as indicated by the implicit sub-QUD in (5.25b).

(5.25)   a. **jaŋ**   **hlĕj** ʔa   jŭt   maj   păʔ ni
            CLF.PERSON which invite friend come here
           A: 'Which person invited you [friends] to come here?'   [Directed at group]

      b. (Who did **Thuận** invite?)

      c. hu **tʰuːŋ͡m**^{312} ʔa   çĕj   maj   păʔ ni…
          ∃   Thuận   invite self come here
          B: 'THUẬN$_{CT}$ invited ME$_{FOC}$ to come here…'

      c′. #**tʰuːŋ͡m**^{312} hu ʔa   çĕj   maj   păʔ ni…
          Thuận   ∃   invite self come here
          INTENDED: B: 'THUẬN$_{CT}$ invited ME$_{FOC}$ to come here…'

In this context, where *which person* represents the phrase around which the answer is organized, the subject is expected to be CT-marked in the partial answer (5.25c). Here, the existential marker *hu* must precede the subject; it cannot appear predicate-initially.

By contrast, consider the minimal difference in (5.26), in which the object is now CT-marked. In the question (5.26a), the D-linked *wh*-phrase now indicates that the object is the sorting key; the answer is organized around the invitees. In this case, the existential marker *hu* must appear predicate-initially (5.26c′); it cannot precede the subject (5.26c).

(5.26) a. jŭt ʔa **jaŋ** **hlĕj** maj pă? ni
           friend invite CLF.PERSON which come here

        A: 'Which person did you [friends] invite to come here?' [Directed at group]

    b. (Who invited **Thuận**?)

    c. #hu c̥ĕj ʔa **tʰuːŋ͡m**³¹² maj pă? ni…
           ∃ self invite Thuận come here

        INTENDED: B: 'I_FOC invited Thuận_CT to come here…'

    c′. c̥ĕj hu ʔa **tʰuːŋ͡m**³¹² maj pă? ni…
           self ∃ invite Thuận come here

        B: 'I_FOC invited Thuận_CT to come here…'

According to Constant's (2014) analysis of CT, CT-marking is predicted to be infelicitous with exhaustive answers to the main QUD. This prediction is borne out in Eastern Cham. Exhaustive answers with *hu* are dispreferred. The answer in (5.26b′) is interpreted as exhaustively answering the question above. It is no longer a partial answer or part of a complex QUD.

(5.27) b′. #hl̥ă? hŏŋ͡m tʰuːŋ͡m³¹² hu ʔa kɛn ni maj pă? ni mĭn
             1SG.POL with Thuận ∃ invite Kenny come here EMPH

        INTENDED: B: 'Thuận and I only invited Kenny.'

    b″. hl̥ă? hŏŋ͡m tʰuːŋ͡m³¹² ʔa kɛn ni maj pă? ni mĭn
             1SG.POL with Thuận invite Kenny come here EMPH

        B: 'Thuận and I only invited Kenny.'

The existential marker, thus, tracks the presence of the CT in a sentence. Why exactly an existential marker would perform this function is not clear (cf. Baclawski Jr 2018a for a hypothesis). Additionally, it is worth noting that *hu* does not uniquely identify the contrastive topic, when it appears predicate-initially. If there are multiple arguments in the predicate as in ditransitives, *hu* is ambiguous between marking the direct and indirect objects. Perhaps CT prosody provides the relevant disambiguation. A similar distribution has been noted for the CT-marker in Paraguayan Guaraní (Tonhauser 2012).

In the cases so far, CT is marked exclusively by the existential *hu*, not by DC-movement. In fact, when presented with DC-movement, as in (5.28c), consultants report that the result is degraded. This evidence supports the event semantics approach to DC outlined above; CT-marking is not clearly an instance of DC-marking in Eastern Cham. Note that

the position of *hu* in (5.28c) will become the clear as the appropriate combination of DC- and CT-marking.

(5.28) a. jŭt ʔa **jaŋ hlɛ̌j** maj pǎʔ ni
       friend invite which CLF.PERSON come here

       A: 'Which person did you [friends] invite to come here?' [Directed at group]

   b. (Who invited **Thuận**?)

   c. $^{??}$**t^{h}uːŋ͡m**^{312} çɛ̌j hu ʔa t^{h}uːŋ͡m^{312} maj pǎʔ ni…
      Thuận self ∃ invite come here

      INTENDED: B: 'Thuận$_{CT}$, I$_{FOC}$ invited to come here…'

There are two ways in which DC-marking appears to cooccur with CT-marking. First, a CT-marked object can be DC-moved if DC-movement of the corresponding *wh*-phrase is moved in the question. This context is seen in (5.29), which is different from the context above only in the presence of DC-movement in (5.29a). In other words, the presence of DC-movement in a question licenses DC-movement in the answer.

(5.29) a. **jaŋ hlɛ̌j** jŭt ʔa ~~jaŋ hlɛ̌j~~ maj pǎʔ ni
       which CLF.PERSON friend invite come here

       A: 'Which person did you [friends] invite to come here?' [Directed at group]

   b. (Who invited **Thuận**?)

   c. **t^{h}uːŋ͡m**^{312} çɛ̌j hu ʔa ~~t^{h}uːŋ͡m^{312}~~ maj pǎʔ ni…
      Thuận self ∃ invite come here

      B: 'Thuận$_{CT}$, I$_{FOC}$ invited to come here…'

Based on the analysis of DC-movement in questions from Chapter 3, Section 3.2.1, the parallelism between questions and answers makes sense (cf. also the discussion on question-answer pairs in Section 5.1). The *wh*-phrase in (5.29a) can only be DC-moved if the kind $^\cap person$ is previously mentioned in a sentence prior (for example, *Lots of people came here tonight* preceding (5.29a)). In this case, the phrase *t^{h}uːŋ͡m^{312}* 'Thuận' is DC-moved because of that relation to the prior discourse, not because of its CT status within the question-answer pair.

Second, a CT-like construction arose in Chapter 4, Section 4.2 on inventory forms. In a list context, a series of inventory forms may be used, as in (5.30b). This construction has some similarity to CT-marking, as there is a series of paired list responses consisting of a kind of fruit and a respective number of each kind bought.

(5.30) a. hɨ plɛ̌j **pɔh jǎw ḳeʔ** pǎʔ ẓaʔ
      2SG buy fruit what at market

      'What [kinds of] fruit did you buy at the market?'

b. kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̃ʔ**$_{DC}$ çŭh pɔh ʔɔ̃ʔ] [$_{DP}$ **hɔ̃ŋm͡**$_{DC}$ mɨ
 1SG b̥uy mango 7 CLF.ROUND papaya 5

pɔh h̶ɔ̶ŋ̶m̶͡] hɔ̃ŋm͡ [$_{DP}$ **na**$_{DC}$ nam pɔh n̶a̶]]
CLF.ROUND with pineapple 6 CLF.ROUND

'I bought mangoes, seven, papaya, five, and pineapple, six.' (a ⇓ b)

Two factors differentiate this context from a true CT context, however. First, (5.30b) is an exhaustive answer to (5.30a), not a partial answer. Second, (5.30b) is in fact an elaborating answer to the question in (5.30a). The question only asks for a list of kinds, not a paired list of kinds and amounts of fruits. Therefore, the answer in (5.30b) not only answers the question, but also an elaborating question along the lines of *How many of each fruit did you buy?* It is worth noting that it is possible to ask a version of (5.30a) as a *how many* question. Here, consultants offer a question with DC-marking of *pɔh jăw ni* 'these fruits' in the left periphery of the clause (cf. Section 4.2.3 on inventory forms and the left periphery).

(5.31) a. **pɔh jăw ni**$_{DC,i}$ hɨ plĕj ʈom pɔh ∅$_i$
 fruit this 2SG b̥uy how.many CLF.ROUND *pro*

'[Of] these [kinds of] fruits, how many did you buy?'

b. kăw plĕj [$_{DP}$ [$_{DP}$ **ʔɔ̃ʔ**$_{DC}$ çŭh pɔh ʔɔ̃ʔ] [$_{DP}$ **hɔ̃ŋm͡**$_{DC}$ mɨ
 1SG b̥uy mango 7 CLF.ROUND papaya 5

pɔh h̶ɔ̶ŋ̶m̶͡] hɔ̃ŋm͡ [$_{DP}$ **na**$_{DC}$ nam pɔh n̶a̶]]
CLF.ROUND with pineapple 6 CLF.ROUND

'I bought mangoes, seven, papaya, five, and pineapple, six.'

As with the question-answer pair above in (5.29), the inventory forms in (5.31b) inherit their DC status from the DC-marking of *pɔh jăw ni* 'these fruits' in (5.31a). This is because the whole question-answer pair must be in a subordinating discourse relation with discourse prior to the question.

Contrastive topic in Eastern Cham, thus, is marked by the existential marker *hu*, not by DC-marking. DC-marking can occur in CT contexts, but due to different discourse conditions. Nevertheless, CT evidently marks a particular discourse structural configuration. I hypothesize that CT marks a different hierarchical discourse constraint (HDC) from DC. DC marks anaphora between two sentences in a subordinating discourse relation that introduce *different events*. CT could be said to mark anaphora between two sentences in a subordinating discourse relation that refer to the *same event*. While CT is perhaps best explained in the QUD framework, at least as a heuristic, the CT conditions can be stated as an HDC distinct from DC.

5.2.3 Catalan CT-marking, revisited

Finally, this section turns back to Catalan CT-marking. It is found that CT-marking in Catalan is plausibly distinct from DC-marking, due to the properties of *wh*-phrases. Recall that clitic right-dislocation (CLRD) in Catalan corresponds with DC-movement in Eastern Cham. Clitic left-dislocation (CLLD) corresponds with contrastive topicalization.

López (2009) argues that CLLD proceeds through a stage of CLRD, on the logic that contrastive topics share a feature with CLRD'd phrases (López's (2009) [+a(naphor)]). This argument is based on data such as (5.32), where a CT, *del seu avi* 'of her grandfather' appears to be subextracted from the CLRD'd phrase (i.e. DC-marked phrase) *les històries* 'the stories'.

(5.32) **Del seu avi**, la Maria les coneix totes, les històries
      of-the her grandfather the Maria CL.ACC knows all the stories
     ~~del seu avi~~

     'Maria knows all of her grandfather's stories.' (López 2009: (4.40))

The Eastern Cham data from the previous section presents one possible explanation of these data. Perhaps the CT-marked phrase satisfies the CT conditions within the relevant QUD. For example, the grandfather may be contrasted with another relative in a complex QUD about relatives and their stories. The DC-marked phrase could independently satisfy the DC conditions based on prior discourse. For example, the whole QUD implicit above could elaborate upon stories in some way.

In Eastern Cham, the DC status of the whole QUD was diagnosable based on DC-marking of a *wh*-phrase in the question. Catalan, however, bans DC-marking on *wh*-phrases entirely, possibly obscuring the interplay between DC and CT. For example, Arregi (2003) argues that CLLD in Spanish marks contrastive topic, based on examples like (5.33). In the question, though, neither *a quién* 'to whom' nor *qué regalo* 'which gift' can be overtly DC-marked. Thus, it is unclear if there is possible DC-marking in (5.33) without sufficient investigation of broader discourse.

(5.33)   a. ¿A quién le diste qué regalo?
        to who him you-gave which gift
      'Who did you give which gift [to]?'   SPANISH

    b. A Juan, le di la moto, y a Pedro, le di el libro.
       to Juan him I-gave the bike and to Pedro him I-gave the book
      'Juan, I gave the bike [to] and Pedro, I gave the book [to].' (Arregi 2003: 32)

Catalan, thus, may support the conclusion that CT marks a different hierarchical discourse constraint than DC. They may only accidentally occur in the same sentence due to the appropriate discourse conditions being satisfied for both at once. Further research

is needed to fully address if CLRD and CLLD are truly independent in terms of discourse structure.

5.3 Summary

These sections have found that there are at least two languages with DC-marking (Eastern Cham and Catalan) and at least two hierarchical discourse constraints (DC and CT). Variation in DC-marking was found in semantic type restrictions with regard to individuals and properties; the ability of pronouns and *wh*-phrases to be DC-marked; and the existence of discourse locality effects. Variation in hierarchical discourse constraints was found in the presence or absence of an open question under discussion.

Appendix A

DC-marking questionnaire

This questionnaire lays out methods for testing if a linguistic phenomenon involves *discourse connected-* or DC-marking. Testing for DC-marking requires simultaneously controlling for discourse structure and previous mention of a particular phrase. Section A.1 presents basic contexts in monologues or narratives and provides guidelines for controlling discourse structure and previous mention. Section A.2 does the same for dialogues, specifically question-answer pairs. Finally, Section A.3 raises a variety of additional factors to consider, which should be kept in mind when testing for DC-marking: the distinction between stage- and individual-level predicates; metalinguistic comments and repairs in the elicitation setting; out-of-the-blue contexts; easily accommodated contexts; and other discourse relations.

A.1   DC-marking in monologues

Monologues or narratives present likely the most straightforward mode for testing DC-marking. That is because there is less uncertainty about a speaker's intended discourse structure, as there is only one speaker. In an elicitation setting, those intended discourse structures can be explicitly probed. In dialogues, there is much more uncertainty about each speaker's interpretation of the discourse and their intention for how each new sentence be added to the discourse (Hunter et al. 2017). For these reasons, it is perhaps best to begin testing for DC-marking in single-speaker monologues.

Section A.1.1 lays out basic contexts to elicit or find in monologues or narratives. To test for DC-marking, one must be able to reliably identify discourses with clear discourse structures (Section A.1.2). Section A.1.3 presents variables to keep in mind when testing how the linguistic phenomenon at hand interacts with those discourse structures. Finally, Section A.1.4 gives guidelines for examining what kinds of previous mention relations can be involved.

A.1.1 Basic DC-marking contexts

Explanation contexts are predicted to license DC-marking, because explanation is a subordinating discourse relation. Note that English is used as a guide. However, this dissertation makes no claim about the presence or absence of DC-marking in English. Throughout these examples, discourse relations are marked following Asher & Lascarides (2003) in a format such as *Explanation*(a,b), which denotes "sentence (b) is interpreted as explaining sentence (a)". Then, the prediction from DC-marking is indicated by ✓, ✗, or ? if the prediction is unclear.

(A.1)   a. Maria is cooking **tofu**.

     b. She wants to eat **tofu** today.   *Explanation*(a,b) ✓DC

       (b) explains (a) if the desire to eat is interpreted as an explanation of the cooking event

Elaboration contexts are also predicted to license DC-marking, as elaboration is also a subordinating discourse relation. Note that (A.1–A.2) could each be interpreted as explanation or elaboration, depending on the speaker interpretation. See Section A.1.2 for methods to diagnose which relation is evoked. However, for the purposes of DC-marking, it is only relevant to distinguish explanation and elaboration from the relations that follow.

(A.2)   a. Maria is cooking **tofu**.

     b. She is frying **tofu** in oil.   *Elaboration*(a,b) ✓DC

       (b) elaborates upon (a) if the frying event is a subpart of the cooking event

Narration contexts are predicted not to license DC-marking, as narration is a coordinating discourse relation. Note that two sentences in a narration relation may have a shared topic, such as *What is Maria doing in the kitchen?* (Asher & Lascarides 2003: 219). When testing narration, be sure that the tested phrase, here *tofu*, is not a part of that shared topic, avoiding topics such as *What is Maria doing with the tofu?*

(A.3)   a. Maria is cooking **tofu**.

     b. Now, she is eating **tofu**.   *Narration*(a,b) ✗DC

       (b) is in a narration relation with (a) if the eating event is temporally separate from the cooking event

Continuation contexts are also predicted not to license DC-marking, as continuation represents a coordinating discourse relation. In the sense of Asher & Lascarides 2003: 461, continuation refers to separate events that are temporally overlapping (i.e. not instances of narration). As with narration, sentences in a continuation relation may have a shared

topic, such as *What are Maria and Toshiko doing in the kitchen?* When testing continuation, avoid shared topics that mention the tested phrase, such as *What are Maria and Toshiko doing with the tofu?*

(A.4) a. Maria is cooking **tofu**.

   b. Toshiko is eating **tofu**.   *Continuation*(a,b) ✗DC

      (b) is in a continuation relation with (a) if the eating event is temporally overlapping with the cooking event

Continuation contexts are also liable to create contrastive topic contexts. Chapter 5, Section 5.2 argued that contrastive topic is distinct from DC-marking, in line with continuation being a coordinating discourse relation. However, one must be aware that contrastive topicalization may involve its own construction or marking that renders testing continuation difficult.

There are numerous other discourse relations outside of these four. Section A.3.4 discusses background, result, and others, along with why they are less clear in terms of the subordination/coordination dichotomy important for testing DC-marking.

A.1.2   Discourse relations & discourse structure

In terms of discourse relations and discourse structure, the main distinction tracked by DC-marking appears to be that between subordinating discourse relations and coordinating discourse relations. DC-marking is only licensed in the presence of a subordinating discourse relation. See, Chapter 1, Section 1.1 and Chapter 2, Section 2.1.1 for more discussion in this distinction.

The universality of discourse relations remains an open question (e.g. Bunt, Prasad & Joshi 2012). In other words, are all discourse relations found in all languages, and how comparable are they to each other? Therefore, it is best to use language-internal diagnostics to ensure that it is clear what discourse relations are intended by the speaker in a given discourse. A non-exhaustive list of three possible diagnostics are given below: cue phrases, implicit questions under discussion, and propositional anaphora. Additionally, in an elicitation setting, direct and indirect metalinguistic judgments may provide the most straightforward evidence for a speaker's intended interpretation of a discourse. As for direct judgments, one could ask if a sentence provides an explanation for another one or if it continues on to a new matter entirely. As for indirect judgments, speakers may attempt to add additional discourse before or after the target discourse. Chapter 2, Section 2.1.1 and Chapter 3, Section 3.2.1 noted supporting metalinguistic evidence in the form of added discourse. In Eastern Cham, if DC-marking is attempted in the absence of a subordinating discourse relation, consultants sometimes added prior discourse that added an appropriate subordinating discourse relation.

A.1.2.1 Cue phrases

Cue phrases, such as *For example* and *After that,* are prototypically adverbials that explicitly indicate relations between sentences or are at least correlated with certain relations (see Taboada & Das 2013 on finding markers for discourse relations). Cue phrases can provide positive evidence that a pair of sentences can be in a certain discourse relation. In (A.5), the felicity of *That's because* suggests that an explanation relation is licit. Other cue phrases are comparative less licit, at least on one reading of the sentences.

(A.5)   a. Maria is cooking tofu.

        b. That's because she wants to eat tofu today.   *Explanation*(a,b)

        b′. $^{??}$For example, she wants to eat tofu today.   *Elaboration*(a,b′)

        b″. $^{??}$After that, she wants to eat tofu today.   *Narration*(a,b″)

It should be noted that there is not a one-to-one correspondence between cue phrases and discourse relations (see Taboada & Das 2013). Not all instances of elaboration, for example, can be paraphrased with *For example*. Additionally, the same pair of sentences can have different discourse relations, depending on the reading. *She wants to eat tofu today* could be interpreted as a simple elaboration or result of (A.5a), among others.

A.1.2.2 Implicit questions under discussion

In an elicitation setting, questions under discussion can be added between sentences to assess their intended discourse relations (see Riester, Brunetti & Kuthy 2018 for more on this technique as a corpus method for encoding QUD structure). The felicity of *Why?* as a question inserted between (A.6a) and (A.6b) further informs that those sentences are in an explanation relation.

(A.6)   a. Maria is cooking tofu.

           Why? / $^{??}$What about the tofu? / $^{??}$What happened next?

        b. That's because she wants to eat tofu today.   *Explanation*(a,b)

DC-marking is predicted to be felicitous when the questions *Why?* (explanation) and *What about X?* (elaboration) can be inserted. It is predicted not to be felicitous when the question *What happened next?,* or something similar, can be inserted (narration), or *What about X?* in the sense of continuation. Note that elaboration and continuation involve similar inserted questions. The difference is that the X in elaboration refers to something in the respective (a) sentence, while the X in continuation refers to something in the shared topic between the (a) and (b) sentences.

A.1.2.3 Propositional anaphora

A further test for the subordination/coordination dichotomy is propositional anaphora in subsequent sentences. Subordinating discourse relations result in both the subordinate and superordinate sentences remaining open for subsequent anaphoric reference, according to theories of anaphora and discourse structure (e.g. Webber 1988, Asher 1993: 312 on the Right-Frontier Constraint; see Chapter 1, Section 1.1). For example, given the elaboration relation in (A.7a–b), the propositional anaphor *that* in (A.7c) can refer either to the proposition in (A.7a) or that in (A.7b).

(A.7)   a. Maria is cooking tofu.

      b. She is frying tofu in oil.   *Elaboration*(a,b)

      c. That$_{(a)/(b)}$'s really cool!

By contrast, coordinating discourse relations result in only the most recent sentence remaining open for subsequent anaphoric reference; the prior sentence becomes closed off. To illustrate, the propositional anaphor *that* in (A.8c) can only refer to the proposition in (A.8b), not (A.8a).

(A.8)   a. Maria is cooking **tofu**.

      b. Now, she is eating **tofu**.   *Narration*(a,b)

      c. That$_{\#(a)/(b)}$'s really cool!

This test diagnoses the broad distinction between subordinating and coordinating discourse relations. It should be noted that propositional anaphora show the contrast most clearly, as other referents are more susceptible to accommodation, rendering anaphora licit regardless of discourse structure. It should also be noted that the underlying theory has been disputed by corpus and experimental data, with exceptions noted in a variety of text genres and circumstances (e.g. Zeldes 2018: 166).

A.1.3 Properties of DC-marking

When testing a syntactic phenomenon for DC-marking, at least the following questions should be asked. As for (A.9a), Chapters 2–3 largely focussed on the Ā-dependency characteristics of DC-marking in Eastern Cham. If DC is an Ā-feature, DC effects are predicted to primarily surface in Ā-dependencies. As for (A.9b), a prediction made in this dissertation is that it is possible for there to be an overt DC-marker in the position of the null DC-particle posited for Eastern Cham. As for (A.9c), Chapter 5, Section 5.1.2 noted that Eastern Cham has category restrictions on what phrases can be DC-marked (only NPs and DPs), while Catalan does not. This informs whether predicates can be DC-marked in addition to individuals.

(A.9)   a. Does the phenomenon involve an Ā-dependency?

        b Is there an overt DC-marking morpheme (i.e. a DC-particle)?

        c. What phrasal categories can be DC-marked?

        d. Under DC-movement, where are DC-marked phrases moved to?

        e. Do in situ phrases ever show DC-marking effects?

        f. Can there be secondary DC-movement or pied-piping?

As for (A.9d), Chapters 2–3 described DC-movement to Spec-CP, Chapter 4 focussed on DC-movement to Spec-DP in Eastern Cham, while Chapter 5, Section 5.1.1 described possible DC-movement to Spec-vP in Catalan. At least these and possibly other landing sites should be possible. As for (A.9e), Chapter 2, Section 3.4 left it an open question whether in situ phrases ever show DC-marking effects. Such effects would be detectable only under certain circumstances, such as if there is an overt DC-marker. Finally, as for (A.9f), secondary DC-movement (i.e. movement first to Spec-DP, then movement of the whole DP to Spec-CP) was not found in Eastern Cham in Chapter 5. Secondary DC-movement would only be predicted if there were two kinds of DC-features (see Chapter 4, Section 4.2.3 on Cable's (2010) account of Q- and *wh*-features). An example of putative secondary DC-movement and pied-piping is sketched in English in (A.10) below.

(A.10)   a. Maria is cooking **tofu**.

        b. [**Tofu**$_{DC}$ two pieces ~~tofu~~]$_{DC}$ she is cooking in oil ~~tofu two pieces~~.

More generally, the syntactic properties of DC-marking are predicted to mirror the syntactic properties of *wh*-movement and other Ā-phenomena.

A.1.4 Previous mention

Next, when testing DC-marking a range of possible previous mention relations should be tested. In Eastern Cham, a wide variety of relations were found to be possible in Chapter 2, largely because an anaphoric index can be DC-marked, as can an NP inside a DP. Examples predicted to be licit include set-subset relations (A.11b), part-whole bridging (A.12), producer-product bridging (A.13), and focus associators and quantifiers (A.14). Downward-entailing quantifiers (A.14b′) are of particular importance, as they differentiate between topicality and DC-marking. It is not clear if set-superset relations (A.11b′) would be predicted to be licit, as they may inherently involve accommodation of additional prior discourse (see Chapter 2, Section 2.1.3). Note that the discourse relation that licenses each of the (b) sentences below is intended to be elaboration.

(A.11)   a. Maria is cooking **tofu**.

        b. She is frying **one piece**$_{DC}$ in oil.   (Set-subset)

b′. She is frying **vegetarian food**_{DC?} in oil. (Set-superset)

(A.12) a. Toshiko has a beautiful **house**.

  b. She painted **the door**_{DC} blue. (Part-whole)

(A.13) a. Toshiko enjoys **that book**.

  b. She knew **the author**_{DC} growing up. (Producer-product)

(A.14) a. Maria is cooking **tofu**.

  b. She is frying **only one piece**_{DC} in oil. (Focus associator)

  b′. She is frying **fewer than three pieces**_{DC} in oil. (Quantifier)

Pronouns were found not be able to be DC-marked in Eastern Cham (Chapter 2, Section 2.1.3), while strong pronouns may be in Catalan (Chapter 5, Section 5.1.2). While not explored in this dissertation, a more complete examination of the discourse properties of pronouns should be considered alongside DC-marking phenomena.

(A.15) a. **Toshiko** is coming to the party tonight.

  b. Maria invited **her**_{DC?}. (Weak pronoun)

  b′. Maria invited **HER**_{DC?}. (Strong pronoun)

Finally, deictic demonstration (i.e. pointing and a demonstrative, as in (A.16)) was found to license DC-marking in Eastern Cham, at least in one case. More examination of the interaction between DC-marking, demonstratives, and gesture would likely inform the kind of previous mention needed for DC-marking.

(A.16) I want to eat **that mango**_{DC?}. [Pointing at a mango]   (Deictic demonstration)

This section has outlined the pieces needed to examine DC-marking in monologues, as laid out in this dissertation. The following section lays out additional tests for dialogues. Finally, Section A.3 discusses additional factors to keep in mind when testing for DC-marking.

A.2 DC-marking in dialogues

This section turns to dialogues, specifically question-answer pairs. There are two main ways to test DC-marking in question-answer pairs discussed in Chapters 2–3: DC-marking in answers and DC-marking in questions. Dialogues longer than question-answer pairs become much more difficult to control for, as individual speakers can interpret discourse and attach new sentences in unpredictable ways (Hunter et al. 2017).

First, answers are distinguished between direct answers and elaborating answers. Direct answers provide the amount of information requested, as in (A.17b). If the discourse consists only of one question-answer pair, direct answers are not predicted to allow DC-marking.

(A.17) a. Is Maria cooking **tofu**?

  b. She is cooking **tofu**.   *Direct answer*(a,b) ✗DC

  b′. She already cooked **tofu**.   *Elaborating answer*(a,b′) ✓DC

Elaborating answers provide more information than requested, as in (A.17b′). Even if the discourse only consists of one question-answer pair, elaborating answers are predicted to allow DC-marking. The implicit question under discussion test provides some insight into this distinction. Only elaborating answers can answer an implicit *No. Why?* question.

Next, as argued in Chapter 3, *wh*-phrases can be DC-marked in Eastern Cham. Questions can have the same kinds of discourse relations as statements. Hence, question versions of narration and continuation can be compared to question versions of elaboration and explanation. One pairing is given in (A.18). Here, (A.18b) is in a narration relation with (A.18a) if the cooking event is interpreted as temporally subsequent to the event of preparing the vegetables. By contrast, (A.18b′) is in an elaboration relation with (A.18a) if the chopping event is interpreted as a subevent of the event of preparing the vegetables.

(A.18) a. Maria is preparing **onions** and **peppers**.

  b. Now, **what vegetable** is she cooking?   *Narration$_q$*(a,b) ✗DC

  b′. Now, **what vegetable** is she chopping?   *Elaboration$_q$*(a,b′) ✓DC

Chapter 3, Section 3.2.2 compares DC-marking of *wh*-phrases to D-linking. Overall, D-linking is predicted to be an independent notion from DC. Therefore, the same discourse relation pattern of DC-marking as above should be found in (A.19), even though phrases like *which vegetable* are inherently D-linked.

(A.19) a. Maria is preparing **an onion** and **a pepper**.

  b. Now, **which vegetable** is she cooking?   *Narration$_q$*(a,b) ✗DC

  b′. Now, **which vegetable** is she chopping?   *Elaboration$_q$*(a,b′) ✓DC

DC-marking in answers is predicted to be found if there is a more general phenomenon of DC-marking. However, Chapter 5, Section 5.1.3 found that DC-marking of *wh*-phrases is not found in Catalan, despite the presence of DC-marking on non-*wh*-phrases. The presence of absence of DC-marking on *wh*-phrases may be a point of cross-linguistic variation, or it may fall out from the interaction between DC-marking and *wh*-movement or other factors.

A.3 Additional factors

Finally, this section lays out for factors to keep in mind when testing for DC-marking: stage- vs. individual-level predicates, out-of-the-blue contexts, predictable sequences of events, and other discourse relations. These are all possible lurking factors that may influence DC-marking tests in unexpected ways if not controlled for.

A.3.1 Stage- vs. individual-level predicates

Chapter 2, Section 2.1.2 demonstrated that DC-marking in Eastern Cham makes reference to semantic events. Both the prior and current sentences must introduce event variables for DC-marking to be licit. The stage- vs. individual-level predicate distinction is relevant for DC-marking, because it has been argued that individual-level predicates do not introduce event variables. Stage-level predicates are temporary properties of the participants involved, while individual-level predicates are permanent ones.

There are multiple ways in which individual-level predicates can interfere on testing for DC-marking. For one, consider (A.20). DC-marking of *tofu* appears to be illicit because there is a continuation relation between the two sentences, a coordinating discourse relation. However, (A.20b) also contains an individual-level predicate (or at least it would in the equivalent Eastern Cham sentence). Therefore, it is unclear if it is the discourse relation or lack of event that renders DC-marking illicit.

(A.20) a. Maria is cooking **tofu**.

b. Toshiko does not know how to cook **tofu**. *Continuation*(a,b) ✗DC

For another example, consider (A.21). We might expect *tofu* to be DC-marked in (A.21b), as there is an explanation relation between the two sentences. However, (A.21a) contains an individual-level predicate, possibly bleeding DC-marking in this case.

(A.21) a. Maria knows how to cook **tofu**.

b. She learned about **tofu** in a cooking class. *Explanation*(a,b) ✗DC

Individual-level predicates, therefore, should be avoided entirely in DC-marking contexts. When accounting for the stage- and individual-level predicate distinction, it is important to use language-internal diagnostics (i.e. not relying only on English diagnostics when testing other languages). Chapter 2, Section 2.1.2 described a language-internal diagnostic for the stage- and individual-level predicate distinction in Eastern Cham (see Kratzer 1995, Fernald 2000 for additional tests in other languages).

A.3.2 Out-of-the-blue contexts

Next, out-of-the-blue contexts should be used with caution. Recall the basic narration test, repeated below (A.22). Ideally, these sentences should be elicited in an out-of-the-blue context, so there is no prior discourse that could result in a false positive for DC-marking. For example, there could be prior discourse that is interpreted in a subordinating discourse relation with (A.22b) and mentions *tofu*.

(A.22)   a. Maria is cooking **tofu**.

        b. Now, she is eating **tofu**.   *Narration*(a,b) ✗DC

One way to ensure that the prior discourse does not interfere with DC-marking is to ask consultants to make that prior discourse explicit. For example, what sentences could the speaker have said immediately before (A.22a)? Making prior discourse explicit is used as corroborating evidence for DC-marking in Chapter 2, Section 2.1.1 and Chapter 3, Section 3.2.1.

A.3.3 Predictable sequences of events

When testing for DC-marking, predictable sequences of events should be avoided, because they may easily give rise to accommodated prior discourse. For example, consider a sequence of events such as (A.23a–b). Because of the narration relation between the two sentences, *the tofu* should not be DC-marked in (A.23b).

(A.23)   a. First, you press **a block of tofu**.

        b. Then, you cut **the tofu** into small cubes.   *Narration*(a,b) ✗DC

However, to someone who is familiar with the process of cooking tofu, it may be obvious that the broader topic is something along the lines of *How do you cook tofu?* This presents a problem, because *the tofu* would be DC-marked if *How do you cook tofu?* precedes (A.23a) such that (A.23a–b) are discourse subordinate to it. As with out-of-the-blue contexts above, it is best to make this prior discourse explicit or to avoid sequences of events where the broader topic is clear.

A.3.4 Other discourse relations

Finally, it should be noted that there are numerous other discourse relations not mentioned here. This guide has focused on four discourse relations: elaboration, explanation, narration, and continuation, alongside the question versions of these, direct answers, and elaborating answers. Other discourse relations posited include result (A.24b), background (A.24b′), and correction (A.24b″), illustrated below (see Asher & Lascarides 2003: 459 for these and other discourse relations).

(A.24)   a. Maria is cooking **tofu**.

      b. As a result, there will be delicious **tofu** tonight.   *Result*(a,b) ? DC

      b′. It is a stormy night.   *Background*(a,b′) ? DC

      b″. No, Toshiko is cooking **tofu**.   *Correction*(a,b″) ? DC

These three discourse relations have been avoided in this questionnaire and dissertation, as there are potential issues with their status with regard to the discourse subordination vs. discourse coordination dichotomy. Result is known to be ambiguous between subordinating and coordinating (Onea 2019). While it may be possible to disentangle which instances of result involve discourse subordination and which do not, there are clearer discourse relations that can be used.

Sentences in background relations provide background information on foregrounded sentences. According to Asher & Lascarides (2003: 165), background appears to be an instance of discourse coordination, as backgrounded sentences do not expand upon foregrounded sentences; they give more general information. However, background appears to be a subordinating discourse relation with regard to anaphora. Instead, Asher & Lascarides (2003: 165) posit that background involves a special *foreground-background pair*. This special status renders it unclear if background should be considered a subordinating discourse relation for the purposes of DC-marking. Additionally, backgrounded sentences may often involve individual-level predicates, further rendering them unviable for testing DC-marking.

Corrections involve metalinguistic turns, wherein a speaker revises the discourse itself. Metalinguistic relations have been avoided in general in this dissertation, as they are not clearly coordinating or subordinating. One hallmark of subordinating discourse relations is that the superordinate (i.e. prior) sentence remains open for further discussion and anaphoric reference. Corrections effectively delete the prior sentence and thus appear not to be in a subordinating discourse relation. Yet, they are not obviously in a coordinating discourse relation, as they refer to the same event as the prior sentence.

Bibliography

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding.* University of Connecticut dissertation.

Abels, Klaus. 2012. The Italian left periphery: A view from locality. *Linguistic Inquiry* 43. 229–254. `https://doi.org/10.1162/LING_a_00084`.

Aboh, Enoch Oladé. 2004a. Left or right? A view from the Kwa periphery. In David Adger, Cécile De Cat & George Tsoulas (eds.), *Peripheries*, 165–189. Springer.

Aboh, Enoch Oladé. 2004b. Topic and focus within D. *Linguistics in the Netherlands* 21(1). 1–12.

Aboh, Enoch Oladé. 2010. Information structuring begins with the numeration. *Iberia* 2(1). 12–42.

Aboh, Enoch Oladé & James Essegbey. 2010. General properties of the clause. In Enoch Oladé Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 39–64. Springer. `https://doi.org/10.1007/978-90-481-3189-1_3`.

Abrusán, Márta. 2016. Presupposition cancellation: explaining the 'soft–hard' trigger distinction. *Natural Language Semantics* 24. 165–202. `https://doi.org/10.1007/s11050-016-9122-7`.

Adger, David & Gillian Ramchand. 2005. Merge and Move: Wh-dependencies revisited. *Linguistic Inquiry* 36(2). 161–193. `https://doi.org/10.1162/0024389053710729`.

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68(1). 43–80.

Aissen, Judith. 1996. Pied piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14. 447–491.

Alexopoulou, Theodora & Frank Keller. 2007. Locality, cyclicity, and resumption: At the interface between the grammar and the human Sentence processor. *Language* 83(1). 110–160. `https://doi.org/10.1353/lan.2007.0001`.

Alieva, Natalia. 1991. Morphemes in contemporary spoken Cham: Qualitative and quantitative alternations. *Cahiers de Linguistique Asie Orientale* 20. 219–229.

Alieva, Natalia. 1994. The progress of monosyllabization in Cham as testified by field materials. In Cecilia Odé & Wim Stokhof (eds.), *Proceedings of the Seventh International Conference on Austronesian Linguistics [ICAL]*, 541–549. Amsterdam: Rodopi.

Ameka, Felix K. 2007. Information packaging constructions in Kwa: Micro-variation and typology. In Enoch Oladé Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 141–176. Springer. `https://doi.org/10.1007/978-90-481-3189-1_7`.

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

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

Ariel, Mira. 1990. *Accessing noun phrase antecedents*. Routledge.

Arregi, Karlos. 2003. Clitic left dislocation is contrastive topicalization. In Ellen Kaiser & Sudha Arunachalan (eds.), *Proceedings of the 26th Annual Penn Linguistics Colloquium*, vol. 9, 31–44. University of Pennsylvania Working Papers in Linguistics.

Asher, Nicholas. 1993. *Reference to abstract objects in discourse*. Kluwer Academic Publishers.

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`.

Aymonier, Etienne & Antoine Cabaton. 1906. *Dictionnaire cham-français [Cham-French Dictionary]* (Publications de l'Ȇcole française d'Etrême-Orient 7 [Publications of the French School of Asian Studies 7]). Paris: Leroux.

Baclawski Jr, Kenneth. Forthcoming(a). Eastern Cham optional *wh*-movement is discourse connected-movement. *Natural Language & Linguistic Theory*.

Baclawski Jr, Kenneth. Forthcoming(b). Optional *wh*-movement and topicalization in Eastern Cham. In *AFLA 26: The Proceedings of the 26th Meeting of the Austronesian Formal Linguistics Association*.

Baclawski Jr, Kenneth. 2016. *Multiple fronting restrictions in Eastern Cham: An [ID]-feature account*. Paper presented at the LSA Annual Meeting, Washington DC.

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. Convergence and divergence in Eastern Cham langauge contact. In *New Ways of Analyzing Variation – Asia Pacific (NWAV-AP5)*. Paper presentation, ARC Centre of Excellence for the Dynamics of Language, Brisbane, Australia.

Baclawski Jr, Kenneth. 2018c. 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. 2018d. Subpart of Focus Fronting is discourse subordination. In Wm.G. Bennett, Lindsay Hracs & Dennis Ryan Storoshenko (eds.), *Proceedings of the Thirty-Fifth West Coast Conference on Formal Linguistics*, 87–95. `http://www.lingref.com/cpp/wccfl/35/paper3378.pdf`.

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`.

Baclawski Jr, Kenneth & Zihan Cindy Yang. 2019. Experimental evidence for the status of the Right-Frontier Constraint. In *Discourse Expectations: Theoretical, Experimental, and Computational Perspectives (DETEC 2019)*. Paper presentation, Leibniz Centre for General Linguistics (ZAS).

Baier, Nico. 2018. *Anti-Agreement*. University of California, Berkeley dissertation.

Baker, Mark C. 2008. On the Nature of the Antiagreement Effect: Evidence from Wh-insitu in Ibibio. *Linguistic Inquiry* 39. 615–632.

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`.

Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP (2)*, 16–51. Oxford University Press.

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102(2–3). 87–113. `https://doi.org/10.1016/S0024-3841(96)00046-0`.

Blood, David. 1967. Phonological units in Cham. *Anthropological Linguistics* 9(8). 15–32.

Blood, David. 1977. A three-dimensional analysis of Cham sentences. In David D. Thomas, Ernest W. Lee & Nguyễn Đăng Liêm (eds.), *Papers in South East Asian Linguistics 4*, 53–76. Canberra: Australian National University.

Blood, Doris. 1961. Women's speech characteristics in Cham. *Asian Culture* 3(3–4). 139–143.

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`.

Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33(3). 351–383.

Bossi, Madeline & Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: A Journal of General Linguistics* 4(1). 65. `https://doi.org/10.5334/gjgl.246`.

Brandi, Luciana & Patrizia Cordin. 1989. Two Italian dialects and the Null-Subject Psarameter. In Osvaldo A. Jaeggli & Ken Safir (eds.), *The Null Subject Parameter*, 111–142. Dordrecht: Kluwer Academic Publishers. `https://doi.org/10.1007/978-94-009-2540-3_4`.

Broadwell, George Aaron. 2001. Optimal order and pied-piping in San Dionicio Zapotec. In Peter Sells (ed.), *Formal and empirical issues in optimality theoretic syntax*, 197–224. Stanford, CA: CSLI Publications.

Bruening, Benjamin & Thuan Tran. 2015. The nature of the passive, with an analysis of Vietnamese. *Lingua* 165(Part A). 133–172. `https://doi.org/10.1016/j.lingua.2015.07.008`.

Bruening, Benjamin, Xuyen Dinh & Lan Kim. 2018. Selection, idioms, and the structure of nominal phrases with and without classifiers. *Glossa: A Journal of General Linguistics* 3(1). `https://doi.org/10.5334/gjgl.288`.

Brunelle, Marc. 2005. *Eastern Cham register: Phonetic, phonological and sociolinguistic approaches*. Cornell University dissertation.

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 & Phú, Văn Hẳn. 2018. Colloquial Eastern Cham. In Alice Vittrant & Justin Watkins (eds.), *The Mainland Southeast Asia linguistic area*, 522–557. Mouton de Gruyter.

Bùi, Khánh Thế. 1996. *Ngữ Pháp Tiếng Chăm [Grammar of the Cham Language]*. Hà Nội: Nhà Xuất Bản Giáo Dục.

Bunt, Harry, Rashmi Prasad & Aravind Joshi. 2012. First steps towards an ISO standard for annotating discourse relations. In Harry Bunt, Manuel Alcantara-Plá & Peter Wittenburg (eds.), *Proceedings of the Joint ISA-7, SRSL-3, and I2MRT Workshop on Semantic Annotation and the Integration and Interoperability of Multimodal Resources and Tools*, 60–69.

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

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

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

Carlson, Greg N. 1977. *Reference to kinds in English*. University of Massachusetts dissertation.

Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38(1). 31–66. `https://doi.org/10.1007/s10988-014-9162-8`.

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. 1977. On *wh*-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal Syntax*, 71–132. New York: Academic Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth L. Hale & S. Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. MIT Press. `https://doi.org/10.7551/mitpress/9780262527347.003.0003`.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press.

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.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language,* 1–52. Cambridge, MA: MIT Press.

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

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332. `https://doi.org/10.1162/0024389054396917`.

Cleary-Kemp, Jessica. 2015. *Serial verb constructions revisited: A case study from Koro.* University of California, Berkeley dissertation.

Cole, Peter, Gabriella Hermon & Li-May Sung. 1993. Feature percolation. *Journal of East Asian Linguistics* 2(1). 91–118. `https://doi.org/10.1007/BF01440584`.

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

Constant, Noah. 2011. Re-diagnosing appositivity: Evidence for prenominal appositives from Mandarin. In Carissa Abrego-Collier, Arum Kang, Martina Martinović & Chieu Nguyen (eds.), *Proceedings of Chicago Linguistics Society (CLS) 47*, 47–61. University of Chicago, Chicago Linguistic Society.

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

Coon, Jessica. 2009. Interrogative possessors and the problem with pied-piping in Chol. *Linguistic Inquiry* 40(1). 165–175. `https://doi.org/10.1162/ling.2009.40.1.165`.

Cristea, Dan, Nancy Ide & Laurent Romary. 1998. Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of ACL/COLING,* 281–285. Montreal, Canada.

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

Del Gobbo, Francesca. 2003. *Appositives at the interface.* University of California, Irvine 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.

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

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.

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`.

Fanselow, Gisbert & Damir Ćavar. 2002. Distributed deletion. In Artemis Alexiadou (ed.), *Theoretical approaches to universals*, 65–107. John Benjamins.

Fernald, Theodore B. 2000. *Predicates and temporal arguments*. Oxford University Press.

von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical Perspectives* 22(1). 137–170. `https://doi.org/10.1111/j.1520-8583.2008.00144.x`.

Frascarelli, Mara & Annarita Puglielli. 2007. Focus in the Force-Fin system: Information structure in Cushitic languages. In Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages. The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 161–184. Mouton de Gruyter.

Greenberg, Joseph. 1975. Dynamic aspects of word order in the numeral classifier. In Charles N. Li (ed.), *Word order and word order change*, 27–46. Austin: University of Texas Press.

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`.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. John Benjamins. `https://doi.org/10.1075/la.66`.

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.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. MIT Press.

Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53. `https://doi.org/10.1016/B978-0-12-545850-4.50014-5`.

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.

Holler, Anke & Lisa Irmen. 2007. Empirically assessing effects of the Right Frontier Constraint. In A. Branco (ed.), *DAARC 2007, LNAI 4410*, 15–27. Springer-Verlag Berlin Heidelberg.

Hoonchamlong, Yuphaphan. 1991. *Some issues in Thai anaphora: A government and binding approach*. University of Wisconson-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`.

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

Hunter, Julie, Nicholas Asher, Eric Kow, Jérémy Perret & Stergos Afantenos. 2017. Defining the right frontier in multi-party dialogue. In *19th Workshop on the semantics and pragmatics of dialogue (SemDial 2015 - goDIAL)*, 95–103.

Jackendoff, Ray. 1972. *Semantics in Generative Grammar*. MIT Press.

Jackendoff, Ray. 1977. *X′ Syntax: A study of phrase structure grammar* (Linguistic Inquiry Monograph Series 2). MIT Press.

Jenks, Peter. 2013. Quantifier float, focus, and scope in Thai. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 39, 90–107. `https://doi.org/10.3765/bls.v39i1.3872`.

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

Jensen, Joshua. 2013. *The Structure of Jarai clauses and noun phrases*. The University of Texas at Arlington dissertation.

Jensen, Joshua. 2014. *Jarai clauses and noun phrases: Syntactic structures in an Austronesian language*. Mouton de Gruyter.

Kamp, Hans, Josef van Genabith & Uwe Reyle. 2011. Discourse Representation Theory. In D. Gabbay & F. Guenther (eds.), *Handbook of Philosophical Logic*, vol. 15, 125–394. Springer.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer.

Kayne, Richard S. 1994. *The antisymmetry of syntax (Linguistic Inquiry Monographs 25)*. MIT Press.

Kennedy, Christopher. 2013. A scalar semantics for scalar readings of number words. In Ivano Caponigro & Carlo Cecchetto (eds.), *From grammar to meaning: The spontaneous logicality of language*, 172–200. Cambridge University Press. `https://doi.org/10.1017/CBO9781139519328.010`.

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`.

Koopman, Hilda & Dominique Sportiche. 1986. A note on long extraction in Vata and the ECP. *Natural Language & Linguistic Theory* 4(3). 357–374. `https://doi.org/10.1007/BF00133374`.

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`.

Kotek, Hadas. 2019. *Composing questions*. MIT Press.

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch zeitgenössischer Forschung*, 639–650. Mouton de Gruyter.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The Generic Book*, 125–175. University of Chicago Press.

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.

Kuno, Susumu. 1987. *Functional syntax: Anaphora, discourse, and empathy*. University of Chicago Press.

van Kuppevelt, Jan. 1995. Structure, topicality and questioning. *Journal of Linguistics* 31(1). 109–147.

Kuroda, S.Y. 1965. *Generative Grammatical studies in the Japanese language*. Massachusetts Institute of Technology dissertation.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge University Press.

Lascarides, Alex & Matthew Stone. 2009. Discourse coherence and gesture interpretation. *Gesture* 9(2). 147–180. https://doi.org/10.1075/gest.9.2.01las.

Lasnik, Howard & Tim Stowell. 1991. Weakest crossover. *Linguistic inquiry* 22(4). 687–720.

Lekakou, Marika & Kriszta Szendrői. 2012. Polydefinites in Greek: Ellipsis, close apposition and expletive determiners. *Journal of Linguistics* 48(1). 107–149. https://doi.org/10.1017/S0022226711000326.

Löbel, Elisabeth. 2000. Classifiers versus genders and noun classes: A case study in Vietnamese. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen & Mirja Saari (eds.), *Gender in grammar and cognition*, 259–320. Mouton de Gruyter.

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

Manetta, Emily. 2006. *Peripheries in Kashmiri and Hindi-Urdu*. University of California, Santa Cruz dissertation.

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.

Martí i Girbau, Núria. 2011. *The syntax of partitives*. Universitat Autònoma de Barcelona dissertation.

McCloskey, James. 1990. Resumptive pronouns, A-bar binding and levels of representation in Irish. In Randall Hendrick & Stephen R. Anderson (eds.), *Syntax and Semantics 23: The Syntax of Modern Celtic Languages*, 199–256. San Diego: Academic Press.

McCloskey, James. 2001. The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37(1). 67–100. https://doi.org/10.1017/S0022226701008775.

McCloskey, James. 2011. Resumptive pronouns, A-binding, and levels of representation in Irish. In Alain Rouveret (ed.), *Resumptive Pronouns at the Interfaces*, 65–119. John Benjamins. https://doi.org/10.1075/lfab.5.02mcc.

Mikkelsen, Line. 2015. VP anaphora and verb-second order in Danish. *Journal of Linguistics* 51(3). 595–643. https://doi.org/10.1017/S0022226715000055.

Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese*. New York: Academic Press.

Miyagawa, Shigeru. 2009. *Why Agree? Why move? Unifying agreement-based and discourse-configurational languages*. MIT Press. `https://doi.org/10.7551/mitpress/8116.001.0001`.

Miyagawa, Shigeru. 2017. Topicalization. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 152. 1–29.

Molnár, Valéria. 2006. On different kinds of contrast. In Valéria Molnár & Susanne Winkler (eds.), *The architecture of focus*, 197–234. Mouton de Gruyter. `https://doi.org/10.1515/9783110922011.197`.

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.

Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41(1). 35–82. `https://doi.org/10.1162/ling.2010.41.1.35`.

Munaro, Nicola. 2003. On some differences between exclamative and interrogative wh-phrases in Bellunese: Further evidence for a split-CP hypothesis. In Christina Tortora (ed.), *The syntax of Italian dialects*, 137–152. Oxford University Press.

Nguyễn, Hùng Tưởng. 2004. *The structure of the Vietnamese noun phrase*. Boston University dissertation.

Nguyễn, Hùng Tưởng. 2013. The Vietnamese noun phrase. In Volker Gast, Daniel Hole & Elisabeth Löbel (eds.), *Linguistics of Vietnamese: An international survey*, 57–86. Mouton de Gruyter. `https://doi.org/10.1515/9783110289411`.

Nikolaeva, Irina. 2014. *A grammar of Tundra Nenets*. Walter de Gruyter. `https://doi.org/10.1515/9783110320640`.

Nowak, Ethan Paul. 2016. *Two dogmas about demonstratives*. University of California, Berkeley dissertation.

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.

Onea, Edgar. 2019. Underneath rhetorical relations: The case of Result. In Malte Zimmermann, Klaus von Heusinger & Edgar Onea (eds.), *Questions in Discourse*, 194–250. `https://doi.org/10.1163/9789004378322_008`.

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`.

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.

Peterson, David. 2010. Khumi elaborate expressions. *Himalayan Linguistics* 9(1). `https://doi.org/10.5070/H99123402`.

Polanyi, Livia. 1985. A theory of discourse structure and discourse coherence. In P.D. Kroeber, W.H. Eilfort & K.L. Peterson (eds.), *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistics Society.*

Polanyi, Livia. 1988. A formal model of the structure of discourse. *Journal of Pragmatics* 12. 601–638.

Polinsky, Maria. 2001. *Information structure and syntax: Topic, discourse-linking, and agreement.* Third Workshop on Spoken and Written Texts. Austin: Texas Linguistic Forum.

Postal, Paul M. 1994. Contrasting extraction types. *Journal of Linguistics* 30(1). 159–186. https://doi.org/10.1017/S0022226700016212.

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.

Potts, Christopher. 2005. *The logic of conventional implicatures.* Oxford: Oxford University Press.

Prince, Ellen. 1992. The ZPG letter: subjects, definiteness, and information-status. In Sandra Thompson & William Mann (eds.), *Discourse description: diverse analyses of a fundraising text*, 295–325. John Benjamins. https://doi.org/10.1075/pbns.16.12pri.

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.

Riester, Arndt, Lisa Brunetti & Kordule De Kuthy. 2018. Annotation guidelines for Questions under Discussion and information structure. In Evangelia Adamou, Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser-described Languages: Studies in prosody and syntax, Studies in Language Companion Series*, 403–443. John Benjamins. https://doi.org/10.1075/slcs.199.14rie.

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.

Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond* (The Cartography of Syntactic Structures, Vol. 3), 223–51. Oxford University Press.

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. 1985. *Association with focus.* University of Massachusetts, Amherst dissertation.

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. 2000. Weak crossover as a scope phenomenon. *Linguistic inquiry* 31(3). 513–539. `https://doi.org/10.1162/002438900554424`.

Ruys, E.G. 2004. A note on weakest crossover. *Linguistic Inquiry* 35(1). 124–140. `https://doi.org/10.1162/002438904322793365`.

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`.

Schlenker, Philippe. 2010. Supplements within a unidimensional semantics II: Epistemic status and projection. In Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds.), *Proceedings of the Annual Meeting of the North East Linguistic Society 40*, vol. 2, 167–182. Amherst, MA: GLSA.

Schlenker, Philippe. 2012. Maximize Presupposition and Gricean reasoning. *Natural Language Semantics* 20. 391–429. `https://doi.org/10.1007/s11050-012-9085-2`.

Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts, Amherst dissertation.

Sells, Peter. 1984. *Syntax and semantics of resumptive pronouns* dissertation.

Shlonsky, Ur. 1991. Quantifiers as functional heads: A study of quantifier float in Hebrew. *Lingua* 84(2–3). 159–180.

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`.

Slioussar, Natalia. 2007. *Grammar and information structure: a study with reference to Russian*. Utrecht: LOT.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corrolaries for constituent structure. *Linguistic Inquiry* 19. 425–449.

Starke, Michal. 2001. *Move dissolves into Merge: A theory of locality*. University of Geneva dissertation.

Stickney, Helen. 2009. *The emergence of DP in the partitive structure*. University of Massachusetts, Amherst dissertation.

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`.

Taboada, Maite & Debopam Das. 2013. Annotation upon annotation: Adding signalling information to a corpus of discourse relations. *Dialogue and Discourse* 4(2). 249–281. `https://doi.org/10.5087/dad.2013.211`.

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.

Thurgood, Graham. 2005. A preliminary sketch of Phan Rang Cham. In K. Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 489–512. Routledge.

Thurgood, Graham & Fengxiang Li. 2007. The grammaticalization paths of Proto-Chamic *hu "receive". In *Proceedings of the Chicago Linguistics Society 39*, 205–214.

Tonhauser, Judith. 2012. Contrastive topics in Paraguayan Guaraní discourse. In Anca Chereches (ed.), *Proceedings of Semantics and Linguistic Theory (SALT) 22*, 268–285.

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.

van Urk, Coppe. 2018. Pronoun copying in Dinka Bor and the Copy Theory of Movement. *Natural Language & Linguistic Theory* 36(3). 937–990. https://doi.org/10.1007/s11049-017-9384-x.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A Dinka Bor case study*. MIT dissertation.

van Urk, Coppe & Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46(1). 113–155. https://doi.org/10.1162/LING_a_00177.

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.

Yang, Dong-Whee. 2009. Edge feature, D-effects, and movement. *Studies in Generative Grammar* 19(1). 83–122.

Zeldes, Amir. 2018. *Multilayer corpus studies*. Routledge. https://doi.org/10.4324/9781315112473.

Zimmermann, Malte. 2007. Overt existential closure in Bura (Central Chadic). In T. Friedman & M. Gibson (eds.), *Proceedings of SALT 17*, 333–350. Cornell, CLC Publications.

Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. MIT Press.

Zyman, Erik. 2018. *On the driving force for syntactic movement*. University of California, Santa Cruz dissertation.