The Particle *Thì* in Vietnamese

A Thesis Submitted to

The Faculty of the College of Arts and Sciences

In Candidacy for the Degree of

Master of Arts in English

By

Mai C. Nguyen

Aug 9th, 2021

Liberty University
College of Arts and Sciences
Master of Arts in English

Student Name: Mai C. Nguyen

_____

Dr. Jaeshil Kim                                                                                          Date
Thesis Chair

_____

Dr. Albert Bickford                                                                                     Date
First Reader

_____

Dr. Paul Müller                                                                                          Date
Second Reader

i

# CONTENTS

# ABBREVIATIONS

The abbreviations in this thesis appear here.

| | |
|---|---|
| 1 | First person |
| 2 | Second person |
| 3 | Third person |
| AP | Adjectival phrase |
| COMP | Complementizer |
| CLF | Classifier |
| DEM | Demonstrative |
| FUT | Future |
| NEG | Negation |
| NP | Noun phrase |
| PASS | Passive voice |
| PL | Pluralization |
| POSS | Possessive |
| PP | Prepositional phrase |
| PROG | Progressive |
| PST | Past tense |
| PTL | Particle |
| RECP | Reciprocal |
| REFL | Reflexive |
| REL | Relativizer |
| S | Sentence |
| SUP | Superlative |
| TOP | Topic |
| VP | Verb phrase |

# ACKNOWLEDGEMENTS

My utmost appreciation goes to Dr. Jaeshil Kim. Dr. Kim has been more than just a professor to me. She is also someone that I turn to for life advice. Every time I felt insecure about my progress, my background, or my ability, she was always the first to get me back on my feet. In every thesis meeting, she showed me various ways to look at the same linguistic phenomenon; she opened my eyes to the fascinating and challenging world of Linguistics. Working under her supervision has sharpened my critical thinking skills, given me a deeper respect and love for my own language, and most importantly, shown me how having a caring mentor could leave such an impact on students. She kindled in me the love for Linguistics and constantly reminded me the importance of self-confidence and humility in doing research. Her patience to refine my rough edges is beyond imagination, and her expertise on the subject-matter is, simply, impeccable.

I would also like to express my gratefulness to Dr. Albert Bickford and Dr. Paul Müller, my other committee members. Dr. Bickford is one of the most thoughtful linguists I have ever met. His encouraging emails, detailed feedback, and kind support continuously inspire me. I would also never forget Dr. Müller's insightful observations, introspective questions, and exceptional wisdom. My English pronunciation would also not be where it is today without Dr. Müller's help.

Besides my professors, friends have also greatly comforted and helped me. I first want to thank Nguyễn Lê Quang Trung (Trung Nguyen). Trung has witnessed and walked with me through many ups and downs. Every time I need a participant for my piloting research, a guinea pig for a curriculum I designed, or a volunteer for any project of mine, Trung is always the first to join. Whenever I am on the verge of giving up, he always reminds me of why I started. I also

# ABSTRACT

This thesis examines the functions of the particle *thì* in Vietnamese. The previous literature mainly describes *thì* as a phrasal particle; the inter-clausal *thì* receives one previous description (Clark, 1992a), but the collected data in this thesis revealed that *thì* is predominantly used sentence-initially. Among 155 collected tokens of *thì*, 48 (31%) occurs phrasally, 38 (30%) occurs inter-clausally, and 69 (44%) occurs sentence-initially. The phrasal *thì* is analyzed as a topic and/or contrastive topic marker (Cao, 1991/2004;[1] Clark, 1992a), or "a discourse template marker" (Tran, Forthcoming). The inter-clausal *thì* is considered "an inchoative conjunction" (Clark, 1992a). A description and an analysis of the sentence-initial *thì* remain unaddressed. Thus, this thesis furthers the literature on *thì* by providing a description of the particle and an analysis of its functions, with focus on the sentence-initial *thì*. The analysis shows that the sentential *thì* is best treated as a Discourse Marker (DM), in the spirit of Fraser (2009). With adaptations from Fraser's framework, the category of DMs provides a unified analysis that can dissect the usages of both the inter-clausal and sentence-initial *thì*. More specifically, the sentential *thì* is a Contrastive Discourse Marker (CDM), Elaborative Discourse Marker (EDM), and Logical Discourse Marker (LDM) that signals various semantic relationships between surrounding propositions. The CDM *thì* indicates a contrastive relationship; the EDM *thì* is subcategorized into Sequential and Non-sequential EDM and conveys additional details; the LDM *thì* connects three types of logical relationships: Abductive, Inductive, and Deductive. The framework of DMs allows for a systematic understanding of the patterned usages of *thì*.

---

[1] The first year refers to the year that the original work was published in. The second year refers to the edition that was cited in this thesis. The page numbers reflect the numbers in the edition cited. The same notation was applied to Thompson (1965/1987).

# CHAPTER 1

## Introduction and Language Overview

### 1.1. Introduction

This thesis unravels the underlying linguistic properties of the particle *thì* in Vietnamese. This particle is under-acknowledged and misunderstood in the Vietnamese language. Specifically, *thì*, and two other particles, *là* and *mà*, forms the notorious triplet whose usage is heavily discouraged in formal language communication. In schools, students are usually criticized for their usage of *thì* in academic papers and class presentations. Other formal communication means, such as the national news broadcasts and newspaper articles, also contain very little usage of *thì*. Paradoxically, despite the linguistic taboo on *thì*'s usage, the particle is prevalently used in everyday speech. In fact, Nguyen (2013), in his book *Nỗi oan thì, là, mà* ('*The slander against thì, là, mà*'), emphasizes the contribution of these particles to the Vietnamese language in spite of their stained reputation. He highlights that *thì* signals unique semantic connotations in communication that speakers would otherwise struggle to convey with other linguistic means. Cao (1991/2004) satirically stresses that if the usage of *thì* should truly be frowned upon, the nationally-renowned poet Nguyễn Du would have undoubtedly become the worst author in Vietnamese literature given his frequent usage of *thì* (Cao 1991/2004:252-253).

Besides the under-acknowledgement of *thì* among language speakers, the status of the particle is also unbalanced in academia. In general, its linguistic properties still remain a conundrum for scholars of Vietnamese. In the literature, *thì* is mostly documented as a phrasal particle marking topics, as in (1), and contrastive topics, as in (2) (Cao, 1991/2004; Clark, 1992; Tran, 2016).

(1) Trường      đó    ***thì***   học sinh      học    giỏi.
    School       DEM   PTL     student       study  good
    'That school, the students study very well.'

(2) Nam ***thì***   thích táo,   nhưng Mai   ***thì***   thích  cam.
    Nam PTL     like   apple  but    Mai   PTL     like   orange.
    'Nam likes apples, but Mai likes oranges.'

The most recent analysis of the phrasal *thì* is proposed in Tran (Forthcoming). In illustrating how

the theoretical framework of Neeleman and van de Koot (2008) works in Vietnamese, Tran

observes that the particle *thì* demonstrates an interesting interface between the syntax and

information structure. Specifically, he claims that the information structural category of contrast

is an autonomous category that triggers left dislocation in Vietnamese, and the left-dislocated

constituents are usually followed by the phrasal *thì*. Following Neeleman and van de Koot's

fusion of the terms, contrastive topics and contrastive foci, Tran resorts to analyzing *thì* as "a

discourse template marker" which "partitions the sentence into the topic-comment or

background-focus construction, and it marks the remaining constituent after the dislocation as

the comment or background" (Tran Forthcoming:7). Besides the phrasal *thì,* Clark (1992a)

documents that *thì* can also occur inter-clausally and provides a description of the inter-clausal

*thì*. She provides the below illustrative examples:

(3) (Clark 1992:93)
    Nếu    chị    thích  ***thì***   chúng ta      sẽ    đi    viếng Seattle.
    If     2      like   PTL     1             FUT   go    visit Seattle
    'If you like, we'll go visit Seattle.'

(4) (Clark 1992:101)
    Chị    đi    cuối   đường Lê Lợi       ***thì***   đi bộ  tới    chợ.
    2      go    end    street Le Loi      PTL     walk   reach  market
    'You get to the end of Le Loi street and then walk to the market.'

As shown in (3), the inter-clausal *thì* is used in conditionals. In (4), the inter-clausal *thì* occurs

between independent clauses to connect sequential actions. Even though the phrasal *thì* has been

2

widely dissected, to my knowledge, Clark (1992a) is the only work to propose an analysis of the inter-clausal *thì*. Clark treats the inter-clausal *thì* as "an inchoative conjunction" (Clark 1992:106), whose function is to conjoin clauses that are semantically related.

Intriguingly, the data findings in this thesis betray that the particle *thì* is more prevalently used on the sentential level as a sentence-initial particle. In fact, out of the 155 collected tokens of *thì* from 3 groups' oral narratives of the Pear Film movie, 107 tokens of *thì* (69%) are the sentential *thì* while the phrasal *thì* only accounts for 48 tokens (31%). On the sentential level, *thì* is more commonly used sentence-initially, accounting for 69 tokens. Even though the sentence-initial *thì* is prevalently used, its existence is only passingly mentioned in past research. The under-acknowledgement of its usage results in a lack of analysis of its functions. The below examples are illustrative of some of its usages in the collected data, which are to recall sequential events, as in (5), and to provide additional commentary, as in (6).

(5)　　NG36(U)[2]:　　Mình　kể　lại　câu chuyện　như　sau.
　　　　　　　　　　　　1　　tell　again　story　　　like　after
　　　　　　　　　　　'I retell the story as follows.'

　　　　NG37(U):　　***Thì***　ở đây　có thể　là　　vào　một　buổi　gần　chiều.
　　　　　　　　　　　PTL　　here　maybe　be　in　　one　CLF　near　afternoon
　　　　　　　　　　　'Here, maybe it is in the afternoon.'

---

[2] This refers to the group code, line number, and the participant who uttered this sentence. See 3.1.3 for a detailed explanation of the data collection notation.

(6) SG38(P): Ba người đó là cùng một ấy ra.
Three person DEM be same one DEM out
'Those three appeared at the same time.'

SG39(P): ***Thì*** bác nghĩ là ba cha con.
PTL 1 think be three father son
'I thought that they were father and sons.'

SG40(Q): ***Thì*** thấy một ông già và hai đứa bé
PTL see one man old and three CLF kid

nên nghĩ là ba cha con
so think be three father son

'I saw an old man and two boys, so I thought they were father
and sons.'

The prevalence of the sentence-initial *thì* deserves an adequate analysis that can tease out its

underlying linguistic properties. While it is plausible to propose that the phrasal *thì* and sentential

*thì* can take on different functions, it is unclear why both the inter-clausal *thì* and especially the

sentence-initial *thì* have not been more carefully analyzed.

Thus, the goal of the thesis is essentially two-fold. First, I will provide a description of

how the particle *thì* is used; this description integrates what has been previously reported of its

phrasal and inter-clausal usages and adds to the current literature by providing a detailed

description of the sentence-initial usages of *thì*. Second, I will propose an analysis that can

capture the properties of the sentential *thì*, especially those of the sentence-initial *thì*; this

account of the particle's usage on the sentential level thus complements previous work that has

focused on its phrasal usages. The analysis treats the sentential *thì* as a Discourse Marker (DM),

in the spirit of Fraser (2009). As a DM, *thì* is not restricted to be classified as any specific

syntactic category, which can accurately capture its syntactic flexibility. Furthermore, Fraser's

framework of DMs also provides a strong foundation to further subcategorize relevant categories

that can represent the various types of semantic relationships that both the inter-clausal *thì* and

4

the sentence-initial *thì* can take on. As varied as the functions of the sentential *thì* can appear on the surface level, the adaptations of Fraser's framework elucidate that, underlyingly, the sentential *thì* showcases properties of a Contrastive Discourse Marker (CDM), Elaborative Discourse Marker (EDM), and Logical Discourse Marker (LDM). Since the scope of semantic coverage of the sentence-initial *thì* is more diverse, these umbrella categories can be subdivided to further define the nuanced differences between the inter-clausal *thì* and sentence-initial *thì*.

The thesis is organized as follows: section 1.2 of this chapter provides a brief overview of the Vietnamese language. Chapter 2 presents a set of definitions of pragmatic notions, topic, focus, and contrast, which will be adopted in this thesis (2.1) and a literature review of past works on the particle (2.2). Chapter 3 consists of a discussion on the data collection process (3.1) and a description of the particle *thì* as a phrasal, inter-clausal, and sentence-initial particle (3.2). Afterwards, chapter 4 details the need to adopt the category of discourse markers to analyze the sentential *thì* (4.2), Fraser's conceptualization of discourse markers (4.3), and the detailed analysis of how Fraser's original and adapted categories of discourse markers can capture the underlying properties of the sentential *thì* (4.4). Lastly, chapter 5 offers concluding remarks (5.1), and comments on aspects that can be addressed in future research on this topic (5.2).

## 1.2 Overview of Vietnamese

Vietnamese is a Mon-Khmer language, which is a branch of the Austro-Asiatic language family. Vietnam is located in Southeast Asia and is home to 54 ethnic groups. However, with the Kinh people accounting for the majority of the country's population, their language, Vietnamese, is chosen as the official language of the country. Similar to the neighboring Southeast Asian and East Asian languages, such as Thai and Chinese, Vietnamese is a tonal language. The language has six lexical tones. Each word carries a tone, and the change of tones alters the meaning of

words, as illustrated in the table below, which is taken from my work on Vietnamese tonal perception (Nguyen 2019:10). The tonal diacritics are kept as in the standard orthography for Vietnamese.

Table 1. Vietnamese Tones (Nguyen 2019:10)

| Number | T1 | T2 | T3 | T4 | T5 | T6 |
|---|---|---|---|---|---|---|
| Vietnamese name | ngang | huyền | sắc | nặng | hỏi | ngã |
| Contour pattern | high (H) | high-low (HL) | low-high (LH) | high-low-glottalization (HLG) | high-low-high (HLH) | High-glottalization-high (HGH) |
| English description | Level tone | Falling tone | Rising tone | Falling tone with ending glottalization | Contour tone with falling and rising pattern | Contour tone with falling and rising pattern with mid-glottalization |
| Example | [ma] 'ghost' | [mà] 'but' | [má] 'cheek' | [mạ] 'rice seeds' | [mả] 'tomb' | [mã] 'horse' |

As seen in the table, the segment [ma] changes its meaning every time a different tone is superimposed on the vowel [a]. Tonal production is susceptible to dialectal variation.

The language has three main dialects: the Northern, Central, and Southern dialects. These dialects exhibit the largest differences in terms of phonetics and phonology, along with some areas of morphology and syntax. For this reason, as will be addressed in chapter 3, the data collection involves speakers from these three dialects in order to have a holistic understanding of how the particle *thì* is used and to see whether dialectal variations affect the usage of the particle.

In terms of its morphological typology, Vietnamese is an analytic language well-known for its monosyllabic words and non-inflectional morphology. Syntactically, it heavily relies on word order to determine grammatical relations, as seen in (7). The example shows that a noun

phrase functions as the subject pre-verbally and object post-verbally. The non-inflectional morphology is observed in the verb 'like'; the verb is not required to agree with the subject in terms of person and number. The example also shows that the verb conjugation remains the same, whether the subject is a third-person singular noun, as in *Nam* of (7a), or a first-person plural noun, as in *họ* 'we' of (7b).

(7)    a. Nam       thích  họ.                       b. Họ  thích   Nam.
           Nam       like    3                           1    like     Nam
          'Nam likes them.'                  'They like Nam.'

The above example also shows that the default word order is S(ubject) V(erb) O(bject). However, past research indicates that Vietnamese also uses topic-comment structure. More specifically, many scholars emphasize that the topic-comment structure is indeed more favored (Cao, 1991/2004; Tran, Forthcoming) and can take precedence over the subject-predicate structure. For this reason, Vietnamese is known as a topic-prominent language, in the spirit of Li and Thompson (1976). For instance, the topic-comment structure is exemplified in (8)-(10).

(8)    Topic             Comment
       [Ở     Việt Nam]    [có    nhiều  xe máy].
       In      Vietnam       have    PL     motorbikes
       'In Vietnam, (there) are many motorbikes.'

(9)    Topic             Comment
       [Thằng Nam]     [nó    rất    thích  táo].
       CLF   Nam         3      very   like    apple
       'Nam, he likes apples very much.'

(10)   Topic             Comment
       [Loại  táo     này]  [tôi    rất    thích  ăn].
       Type  apple  DEM    1      very   like    eat
       'This type of apples, I really like eating (it).'

The topic-prominent nature of Vietnamese allows various types of syntactic constituents, such as noun phrases and prepositional phrases, to be topics of their sentences. It also explains the possible absence of dummy subjects, which is why (8) is not ungrammatical. As in (9), it allows

the presence of the resumptive pronoun that refers back to the topic. Lastly, it also shows that the topic is preferred sentence-initially, which explains the left dislocation of the topic in (10). The left dislocation affects the canonical word order of SVO and changes it to OSV. Whereas this structure is considered syntactically marked, it is pragmatically neutral and is considered a basic sentence structure in Vietnamese, mainly because of the topic-prominent nature of the language. Chapter 2 will show further interaction between the syntax and pragmatics in the Vietnamese language.

# CHAPTER 2

## Definitions and Literature Review

### 2.1 Chapter Overview

As mentioned in Chapter 1, the particle *thì* can occur as a phrasal and sentential particle. The phrasal *thì* has received various analyses; it has a close-knit interaction with information structure. On the sentential level, *thì* can occur inter-clausally and sentence-initially, but the properties of the sentential *thì* have not been widely examined. This chapter first presents a set of definitions for information structural categories of topic, focus, and contrast that will be adopted in this thesis. These notions have been defined quite distinctly depending on each scholar's research agenda. With specific regard to the discussion on the particle *thì,* this chapter also provides a literature review of seminal works related to the particle *thì*. Each previous work on the particle lays a foundation to tease out the underlying functions of the particle and provides critical insights into the remaining questions that still need to be addressed. The literature review shows the need to extend the description of the particle to cover its usages as a sentence-initial particle. The critique of past literature comes from the stance that the phrasal *thì* and sentential *thì* take on different functions, and the critique also confirms the necessity to propose an analysis that can explain the functions of the inter-clausal *thì* and especially the sentence-initial *thì*.

Chapter 2 is organized as follows: section 2.2 provides definitions for the terms, topic, focus, and contrast, respectively. Section 2.3 is a critique of Thompson (1965/1987), Cao (1991/2004), Clark (1992a), and Tran (Forthcoming), with special attention to their description, or lack thereof, of the particle *thì,* and their corresponding analyses of the particle.

## 2.2 Definitions of Topic, Focus, and Contrast

### 2.2.1 Topic

The defining properties of topic have been a controversial matter.[3] In this thesis, topic is defined as denoting given (or old) information and providing a frame for the rest of the sentence or the corresponding discourse. Given information is mutually accessible information whose existence and truth values are accepted and shared by both the speaker(s) and the hearer(s) in a discourse. For a working definition, Chafe's definition is adopted: "The topic limits the applicability of the main predication to a restricted domain. . . [The topic] is the frame within which the sentence holds" (Chafe 1976: 50-51). In Vietnamese, the givenness property of the topic can be established in various ways. Given information can be established through linguistic cues, such as prior mention, or extra-linguistic cues, such as unique references, common knowledge, or contextual cues. The examples below are illustrative, with the topicalized constituents (T) italicized.

(11)    Context: Nam and Ba are housemates. Ba calls Nam on the phone.

Ba:    Có      một    con    chó    ở      đây.
       Have    one    CLF    dog    at     here

*Con    chó*T    đang    đứng    cạnh    cửa.
CLF     dog      PROG    stand   near    door

'There is a dog here. The dog is standing by the door.'

---

[3] The topic is commonly defined as "what the sentence is about," or 'aboutness topic' (Reinhart, 1981). Chafe (1976), Cao (1991/2004), Dooley and Levinsohn (2000), and some others, argued that the topic takes on a scope-setting or scene-setting property. I find that the topic in Vietnamese can both be what the sentence is about, and it also sets a frame for the rest of the sentence (comment) to be applied to.

(12)    Context: Nam and Ba are housemates. Nam just came home from work.

        Nam:    *Con    chó    của    tôi*$_T$    đang    làm  gì?
                      CLF    dog    POSS    1       PROG    do    what
                      'What is my dog doing?'

        Ba:     *Con   chó*$_T$    đang    ngủ    trên    ghế    sô-pha.
                      CLF   dog     PROG    sleep   on    chair    couch
                      'The dog is sleeping on the couch.'

In (11), the NP *con chó* is accessible through a linguistic cue, which is the prior mention of the NP in the existential sentence. In (12), the NP *con chó* is accessible through an extra-linguistic cue, which is the roommates' mutual knowledge about the existence of the dog. Because of the accessibility constraint, the topic must be referential, which is why NP topics are normally specific indefinite (as in (11)) or definite (as in (12)).[4] Generic noun phrases whose references are unique can also take on topichood, as in (13).

(13)    *Mèo*$_T$  ngủ    rất    nhiều  trong  ngày.
        Cat    sleep  very    much   in    day
        'Cats sleep a lot during the day.'

These topics, in anaphoric cases, can be referred back to with pronominalization, as in (14). They also commonly co-occur with deixis, as underlined in (15), and with resumptive pronouns, as in (16). These anaphoric references are bold and italicized.

---

[4] Traditionally, the cross-linguistic topics and Vietnamese topics in particular are normally understood to be definite NPs, with no specific mention regarding its specificity (Chafe, 1976; Cao, 1991/2004). The claim that specific indefinite NPs can also take on topichood is proposed by Tran (Tran Forthcoming:2-3). In agreement with Tran, this thesis also believes that specific NPs are qualified as topics, for their existence is established between the speaker and the hearer and is stored in the Common Discourse Ground (Krifka, 2006) even though their references may not be known to the hearer yet. A nonspecific, indefinite NP, on the other hand, as in the NP *con chó* in the existential sentence, is not qualified as the topic because (1) no new information about the NP is provided (the comment is absent), and (2) the existence of the NP is yet to be conscious in the hearer's mind or agreed by the hearer.

(14)  Ba:    Có     một    con    chó    ở      đây.
             Have   one    CLF    dog    at     here

             *Con    chó*ᴛ   đang    đứng   cạnh   cửa.   **Nó**    sủa    rất    to.
             CLF     dog     PROG    stand  near   door   3        bark   very   loud

             'There is a dog here. The dog is standing by the door. It's barking loudly.'

(15)  *Khu vực*        *Đông Nam Á*ᴛ        trái cây        đa dạng.
      Region           Southeast Asia        fruit           diverse

      ***Nơi***   **đó**   khí hậu        cũng   rất    nóng ẩm.
      Place      DEM     climate        also   very   humid

      'The Southeast Asian region: the fruits are diverse. There, the climate is also humid.'

(16)  *Thằng bé    Nam*ᴛ   tôi    biết   **nó**    từ     rất    lâu.
        CLF    kid   Nam    1      know   3        from   very   long
       'That kid Nam, I have known him for a very long time.'

Evidently, the topic is syntactically marked through word order; it is favored sentence-initially,
which is why the topic in (16), an object NP, is left dislocated to be sentence-initial. Another
syntactic means for identifying the topic is ellipsis (deletion). Vietnamese marks givenness
through ellipsis (Tran 2016:54-55). Hence, to further mark its givenness property, the topic can
be elided in subsequent occurrences, as illustrated by both examples below.

(17)  Nam:   *Con    chó    của    tôi*ᴛ   đang    làm    gì?
              CLF    dog    POSS   1      PROG    do     what
              'What is my dog doing?'

      Ba:    ∅ᴛ     vừa    ăn     xong.   Giờ    ∅ᴛ     đang    ngủ
             TOP    just   eat    finish  now    TOP    PROG    sleep

             trên    ghế    sô-pha.
             on      chair  couch

      '(It) just ate and (it) is now sleeping on the couch.'

(18)  *Khu vực*     *Đông Nam Á*ᴛ      trái cây       đa dạng.
      Region       Southeast Asia       fruit          diverse

      ∅ᴛ    khí hậu      cũng    rất    nóng ẩm
      TOP   climate      also    very   humid

'The Southeast Asian region: the fruits are diverse. (There,) the climate is also humid.'

All of these topic NPs can be followed by the particle *thì* to further mark their topicality

nature. Even though word order is preferred, the presence of *thì* for topic marking purposes is

most critical when the topic-comment boundary is ambiguous (Cao 1991/2004:239), which is

usually when the topic needs more qualifying information to be narrowed down to a specific

and/or referential reference. Below is an illustrative example.

(19)  *Người* học sinh       áo      đỏ     đang    đứng   cạnh   cửa    kia
      CLF    student        shirt   red    PROG    stand   by     door   DEM

      ***thì***   học    rất    giỏi.
      PTL    study   very   well

'The student in the red shirt standing by that door studies very well.'

The NP *người học sinh* 'a/the student' needs the relative clause to be narrowed down so as to

take on the topic role. The length of the NP poses some challenges for the hearer to accurately

identify the end of the topic and the start of the comment, especially if the sentence is uttered in

running speech.

*2.2.2 Focus*

The notion of focus is defined as denoting new information, which is information that is

yet to be placed in the Common Discourse Ground. In line with the current literature on focus,

this thesis also refers to focused information as an answer to a wh-question of the corresponding

wh-question phrase (Chafe, 1976; Cao, 1991/2004; Tran, 2016; Tran, Forthcoming). An

illustrative example is provided below, with the focus of each sentence (F) in brackets and bold.

13

(20)  a.  A:  Ai    rớt    cây    bút?          B: **[Nam]**F   rớt    cây    bút.
             Who   drop   CLF    pen              Nam      drop   CLF    pen
             'Who dropped a pen?'                 'Nam dropped a pen.'

      b.  A:  Nam   làm    gì?              B: Nam      **[rớt   cây    bút]**F
             Nam   do     what                Nam      drop   CLF    pen
             'What did Nam do?'                 'Nam dropped a pen.'

      c.  A:  Nam   rớt    gì?              B: Nam      rớt    **[cây   bút]**F
             Nam   drop   what                Nam      drop   CLF    pen
             'What did Nam drop?'               'Nam dropped a pen.'

      d.  A:  Chuyện        gì    vậy?       B: **[Nam      rớt    cây    bút]**F
             Story         what   that         Nam      drop   CLF    pen
             'What's happening?'                'Nam dropped a pen.'

As seen in (20), any type of constituent can be a focus of its corresponding sentence, such as subject NP (20a), predicate (20b), object NP (20c), and a full sentence (20d). This example also reveals some key distinctions between topic and focus: (1) topic denotes old information whereas focus refers to new information, and (2) topic can be elided because of its givenness property whereas focus cannot, for it would not be felicitous to elide new information. In fact, the answers below where only the focus remains are more common and favored to the questions in (20) (except for (20d) when the full sentence is in focus).

(21)  a.  A:  Ai    rót    cây    bút?          B: **[Nam]**F
             Who   drop   CLF    pen              Nam
             'Who dropped a pen?                  'Nam.'

      b.  A:  Nam    làm    gì?                 B: **[Rót    cây    bút]**F
             Nam    do     what                   drop   CLF    pen
             'What did Nam do?'                   '(Nam) dropped a pen.'

      c.  A:  Nam    rót    gì?                 B: **[Cây    bút]**F
             Nam    drop   what                   CLF    pen
             'What did Nam drop?'                 'A pen.'

Another distinctive nuance is the syntactic marking of topic and focus. Whereas topic is

preferred sentence-initially, which usually results in left dislocation, focus surfaces in-situ.[5]

Because focus surfaces in-situ, a natural question arises as to how focus is marked in Vietnamese.

Cross-linguistically, focused constituents are known to be prosodically marked with intonation or

pitch variation (Chafe, 1976; Büring, 2003; Jannedy, 2007). Given that Vietnamese is a language

with lexical tones (see 1.2), a slight pitch variation can alter the meaning of a word and can even

result in incomprehensible sentences. In other words, it is logical to raise the following question:

If Vietnamese is extremely susceptible to pitch variation and intonational nuances, does it

employ prosodic means to mark focus? If so, what is it? There is, in fact, positive evidence that

Vietnamese does mark focus with prosody, specifically, with loud amplitude, also known as

intensity (Jannedy, 2007).[6] Jannedy recorded two native speakers' answers of five questions that

---

[5] Tran (Forthcoming) disagrees with this position. From his perspective, topic and focus in Vietnamese can both surface in-situ and will only be left dislocated when they are construed contrastively. In other words, Tran firmly believes that it is the information structural category of contrast that triggers left dislocation in Vietnamese; topicalized or foci constituents, when undergone left dislocation, yield a contrastive interpretation, both of which can be followed by the particle *thì* (Tran Forthcoming: 1-2). This thesis disagrees with Tran. A detailed critique of Tran is located in 2.3.4.

[6] Some other works reported that Vietnamese marks focus with intonation and pitch variation (Tran Forthcoming). Tran (Forthcoming) cites Cao (1991/2004) that Vietnamese marks the focus by "placing the stress on the focused element" (Cao 1991/2004 qtd. in Tran Forthcoming:7). Jannedy (2007) says older literature sporadically claims that Vietnamese employs prosodic means to mark the focus (Thompson, 1965/1987; Nguyen, 1990; Dung et. al., 1998) and quotes Thompson (1965/1987) as follows:

> Heavy stress singles out the syllable or syllables of each pause group which carry the heaviest burden of conveying information. Weak stress accompanies syllables, which bear the lowest information-conveying load in the pause group. They often refer to things which have been brought up earlier or which are

15

elicit five different foci: sentence focus, subject NP focus, object NP focus, verb focus, and

predicate focus; she then conducted a raw acoustic measure of the answers. The waveform

variation shows that native speakers put a louder emphasis (louder amplitude) on focused

constituents (Jannedy 2007:216).[7] Other than intensity, the focused constituent can be marked

with other focus sensitive words such as *chỉ* 'only', as illustrated below.

(22)　a.　Tôi　**chỉ**　**[ăn**　**cơm]**F.
　　　　　1　　only　eat　　rice
　　　　　'I only eat rice (not watch TV).'

　　　b.　**Chỉ**　**[tôi]**F　ăn　　cơm
　　　　　Just　1　　　eat　　rice
　　　　　'Only (but no one else) I eat rice.'

　　　c.　Tôi　ăn　**chỉ**　**[cơm]**F.
　　　　　1　　eat　only　rice
　　　　　'I eat only rice (not with vegetables).'

This, however, is not a focus marker. To my knowledge, Vietnamese does not employ any

morphological means to mark focus. The particle *thì* cannot occur with focused constituents, or

else the utterances would be infelicitous. As indicated in the free translations, when the focused

constituents co-occur with *chỉ* 'only', they are more likely to be construed contrastively. In fact,

notions of focus and contrastive focus are intertwined, and sometimes, confusingly inseparable.[8]

---

expectable in the general context. Other syllables are accompanied by medium stress. (Thompson
　　1965/1987:106)
Most of these claims are mainly empirical. Relying on prosodic cues alone, I found that Vietnamese speakers tend to
mark contrastive focused constituents with intensity. However, teasing out this aspect of focus lies beyond the scope
of this thesis. See Jannedy (2007) for more details and related references.
[7] She also conducted further research on this topic in the rest of her article. See "Prosodic Focus in Vietnamese"
(Jannedy, 2007) for more details.
[8] Chafe (1976) acknowledges the fused distinction between the category of focus and contrastive focus, quoting a
position put forth by Bolinger (1961). Bolinger believes that every focused constituent is essentially being
contrasted with constituents that are not it. And the more narrowed the set of alternatives to the focus is, the more
contrastively the focus is construed. He elaborates, "Clearly, in 'Let's have a picnic,' coming as a suggestion out of
the blue, there is no specific contrast with 'dinner party', but there is a contrast between picnicking and anything else
the group might do. As the alternatives are narrowed down, I get closer to what I think of as [a] contrastive accent"
(Bolinger 1961:87). This view is definitely grounded with valid insights, but this construal of the contrastive
interpretation is relatively too broad compared to the scope of it in this thesis.

The subsequent section shows that in order for topic or focus to be construed contrastively, a contextually implied or an explicitly stated set of alternatives to topic and focus must be present.

*2.2.3 Contrast*

As briefly touched on towards the end of the previous section, the information structural category of contrast is a notoriously elusive concept. The confusion mainly stems from how it is defined and used in the literature. Contrast has been treated as either being in complete dependence[9] upon the other two information structural categories: topic and focus (Chafe, 1976; Zimmerman, 2008) or as having its own pragmatic and semantic reflexes (Molnar, 2006; Frey, 2006; Tran, Forthcoming). Tran (Forthcoming) strongly advocates for the latter. This thesis defines the category contrast as not being completely contingent upon the topic and the focus, nor as being entirely independent of these two categories. That is, not all topics or foci are supposed to be construed contrastively, unless a contrastive interpretation is contextually licensed. In line with Tran (Forthcoming), this thesis treats contrast as a pragmatic concept with semantic reflex, but with some nuanced differences from Tran. Semantically, similar to Tran's position, a contrasted constituent, whether of topic or focus, must either contextually imply or explicitly state a set of alternatives to the contrasted entity, and an implication that at least one of the alternatives does not meet the requirement laid out for the contrasted entity.[10] With the

---

[9] In this view, contrastive interpretation is purely marked by pragmatics, also known as contextual cues.

[10] Tran (Forthcoming) made even finer distinctions between the semantic implication of contrast onto the topic and the focus. The below synopsis captures the distinction:

> Contrast is an independent category that semantically functions as a quantifier, and syntactically licenses left dislocation . . . What does it mean by contrastive interpretation? . . . [T]he partition of a sentence into focus-background as a result of focusing yields two sets: one is the set of alternative elements to the focus, and the other is the set represented by the background. The role of contrast is to give information about the relation between the two sets: at least one member of the set of (contextually relevant) alternatives is not contained in the set denoted by the background. . . Unlike contrastive focus, the contrast component in contrastive topics indicates that the speaker for some reason cannot make a claim about alternative topics. As a result of the contrast component effect, upon hearing an utterance that involves a contrastive topic the hearer is still in the dark with respect to the information about a topic alternative to the expressed topic. (Tran Forthcoming:10-13).

presence of contextually determined alternatives, each contrastive utterance carries what

Lambrecht (2004) refers to as "conversational implicatures." Syntactically, contrasted entities in

Vietnamese can be realized in-situ, especially if a set of alternatives is explicitly stated, as in

correction (23), or denial (24) cases. The examples show that the contrastive focus NP *táo*

'apples' can be realized in-situ, just like the other foci; nonetheless, a contrastive interpretation is

identified because of an explicitly mentioned set of alternatives. In this fashion, contrast can be

identified purely based on pragmatic (contextual) cues. Contrastive foci are surrounded by

brackets and are bold and underlined.

        Context: A and B are discussing what fruit Nam likes.

(23)    A:    Nam   thích   nho.
                  Nam    like    grape
                  'Nam likes grapes.'

        B:    Không! Nam   thích   **[táo]**CF!
                  NEG       Nam  like    apple
                  'No (that is not right)! Nam likes apples.'

(24)    A:    Nam   không thích  xoài,  Nam   không  thích  ổi,
                  Nam   NEG   like    mango Nam   NEG    like    guava

                  nhưng Nam   thích   **[táo]**CF.
                  but     Nam   like    apple

                  'Nam does not like mangoes, Nam does not like guavas, but Nam likes apples.'

The in-situ contrastive constituents resemble those of the foci. Interestingly, there is also

evidence that contrasted entities can be syntactically marked through left dislocation, landing

onto a site that looks like the one for topic, at least on the surface level. The key difference lies in

the fact that whereas topic left dislocation allows for resumptive pronouns or anaphoric

references, contrast left dislocation leaves a gap.[11] Moreover, when the contrasted entity is left

dislocated, it should be followed by the particle *thì*, even though the contrastive entity does

denote new information. The examples below are illustrative, with the contrastive topics (CT)

bold and italicized.

(25)    ***Phim***   ***kinh dị***<sub>CT</sub>    ***thì***    *Nam*<sub>T</sub>  thích  Ø<sub>CT</sub>,   nhưng  *anh*    *ấy*<sub>T</sub>   nhát.
Movie  horror         PTL    Nam  like           but     3       DEM   shy
 'Horror movies, Nam likes, but he's a wimp.'

(26)    *Ở*      *miền Bắc*<sub>T</sub>   có     bốn    mùa   nhưng  *ở*     ***miền Nam***<sub>CT</sub>  ***thì***
At      Northern   have   four    season  but     at     Southern      PTL

    Ø<sub>CT</sub>   chỉ    có     hai     mùa.
        only   have   two    season

'In the North, there are four seasons, but in the South, there are only two seasons.'

## 2.3 Literature Review

### 2.3.1 Thompson (1965/1987)

Laurence Thompson gives a brief description of pragmatic particles existing in

Vietnamese, one of which is the particle *thì*. Most of these particles are not syntactically required,

but pragmatically licensed. The particle *thì* is reported to occur with "focal complements." Focal

complements are defined as "substantives [nouns], substantival phrases [noun phrases], and even

predicates [verb phrases] . . . [and they] "establish for their predicates points of reference," such

as the specific person, thing, concept, place, or point in time (Thompson 1965/1987:239-240).

---

[11] If this is the case, a question arises as to whether the contrastive focus can be left dislocated. The answer is
positive, but only in marginal cases. For instance, in order for the NP *táo* 'apples' to be left dislocated and followed
by the particle *thì*, only the limited context below can license the focus left dislocation:
      Nam      không    thích    trái cây  tí nào,   nhưng  **[táo]**<sub>CF</sub> *thì*    Nam    lại     thích   **[t]**<sub>CF</sub>.
      Nam      NEG      like     fruit     at all    but      apple  PTL    Nam    yet     like
      'Nam does not like fruits at all, but (if it were) apples, Nam would like.'
The CF can still remain in-situ in this example, though the left dislocated CF is contextually possible. However,
even in the latter case, as the free translation stated, the constituent prior to the particle *thì* presents a conditional
framework, only through which can the rest of the sentence 'Nam likes X' would apply. In this fashion, the CF
shares similar traits to a topic in terms of its scope-setting. Also, because of this reason, the distinction between
these terms is enigmatic to be completely teased out from one another, if that is even possible.

This documentation of *thì* resembles the phrasal *thì* in my system. An illustrative example from

Thompson for this category is below.

(27)   (Thompson 1965/1987:261)
Rừng **thì** rậm   đường lối   đi     lại    **thì**   khó khăn.
Forest PTL   dense  path         go     back   PTL      difficult
'The forest was dense, the routes of communication difficult.'

In a different section of his book, Thompson also acknowledges the existence of the inter-clausal

*thì*. He shows that *thì* can occur inter-clausally in conditionals, providing the below example. He

also accurately notes that the inter-clausal *thì* and the subordinating conjunction *nếu* 'if' can both

be used to mark conditionals, but only one of them is needed for making the sentence

grammatical.

(28)   (Thompson 1965/1987:279)
Nếu   tiểu thư       thuận tình   **thì**   chúng ta    sẽ     kết
If     lady           agree        PTL      2           FUT    unite

làm   vợ     chồng.
do    wife   husband

'If you [young lady of good family] consent, you and I will unite as man and wife...'

Lastly, Thompson does not explicitly claim that the particle *thì* can occur sentence-initially, but

one of his examples documents the existence of the particle. That example is put below.

(29)   (Thompson 1965/1987:280)
A:   Ba     giờ    tôi    phải   đi     làm.
Three  time   1      must   go     work
'I have to go to work at 3.'

B:   **Thì**   nên     mau    lên!
PTL      should  hurry  up
'(You) should hurry up.'

Even though Thompson briefly acknowledges the existence of the phrasal, inter-clausal, and

sentence-initial *thì*, he does not distinguish them or analyze their functions as part of his

objectives. Given that his work is mainly for reference purposes, Thompson provides a brief and insightful account of how the particle *thì* is used in Vietnamese.

### 2.3.2 Cao (1991/2004)

Cao Xuân Hạo (Hao Cao) is a pioneering domestic scholar who dedicated almost all of his lifetime to unraveling the linguistic principles of the Vietnamese language. He is among the most well-respected scholars nationally. He advocates for the understanding of the Vietnamese language system from a functionalist perspective. In general, Cao believes that every part of speech and component of language must carry a function, whether it is a syntactic, semantic, or pragmatic function, and it must serve communicational purposes (Cao 1991/2004:101). Cao firmly supports the view of treating Vietnamese as a topic-prominent language and examines aspects of the language with the topic-comment structure as the basic sentence structure. He believes that analyzing Vietnamese sentences based on a topic-comment structure is more advantageous, reasoning that the subject-predicate structure is more suitable for understanding Indo-European languages, or languages with inflectional morphology.[12]

Conceptually, Cao treats topic as "the center of attention"; furthermore, topic must be definite and referential. Different types of constituents can take on topichood, including clauses (Cao 1991/2004:156). Moreover, he also divides the umbrella term topic into two subcategories:

---

[12] The following quote is extracted from Cao (1992):

> The appearance of a grammatical subject was due to the morphologization of the features formerly characterized by the logical subject and gradually attached themselves to the argument which most frequently expresses it. This process does not take place in languages without such morphologization, e.g. Chinese or Vietnamese . . . What is commonly called 'subject' in these languages is rather the prime actant of the predicate nucleus . . . The first immediate syntactic member of the sentence always represents the logical subject (the Theme), while the second one always represents the logical predicate (the Rheme). (Cao 1992:137-138).

*khung đề* 'range topic', and *chủ đề*, 'topic'.[13] He provides the examples below to illustrate his

point.[14] The range topic (RT) is in small capitals, and the topic (T) is italicized.

(30)  (Cao 1991/2004:110)
TRONG CÁI  BÌNH  NÀY~RT~ nhiệt độ      lên    đến    39  độ       C.
In      CLF   bottle  DEM   temperature  rise   to     39  degree   Celsius.
'In this bottle, the temperature can reach 39 degrees Celsius.'

(31)  (Cao 1991/2004:110)
*Cái    bình   này*~T~ nhiệt độ      lên    đến    39    độ    C.
CLF    bottle  DEM   temperature  rise   to     39    degree Celsius
'This bottle, the temperature can reach 39 degrees Celsius.'

Cao defines range topic as "the spatial, temporal, and conditional frame, within which the

information in the comment is applicable" and topic as "a specific entity that the comment talks

about" (Cao 1991/2004:160). Whereas the subdivision of the term topic can explain cases where

Vietnamese allows a sentence as (32), this distinction will not be featured prominently in this

work.

(32)  Ở       VIỆT NAM~RT~  nhiều  xe.
in      Vietnam      PL     motorbikes
'In Vietnam, (there) are many motorbikes.'

In terms of focus, Cao's position is in line with what most scholars assume: He believes that (1)

any constituent can be construed as a focus, (2) a focused constituent is usually an answer to a

wh-question, (3) focus is "the most important part in the comment that the speaker wants to

highlight," and (4) it is prosodically marked with strong emphasis (Cao 1991/2004:84-85).[15]

Cao does not provide a set definition for the notion of contrast, but it is implied that (1) contrast

---

[13] His division of these terms aligns with Dooley's distinction of setting vs. topic (Dooley, 2000), McCawley's range topic vs. instance topic (McCawley, 1970), Dik's theme vs. topic (Dik, 1987) and Halliday's conjunctive/modal theme vs. topical theme (Halliday, 1970). Dooley later refined the term setting further and considered it the point of departure. He elaborates, "The term point of departure designates an initial element, often fronted or left-dislocated, which cohesively anchors the subsequent clause(s) to something which is already in the context" (Dooley 2000:35).
[14] All of the examples by Cao are in Vietnamese only, so I provided the interlinear gloss and free translation for all of his examples as documented in this section and throughout this thesis.
[15] He did not specify what prosodic cues are used to mark the focus in Vietnamese, but he did insinuate that strong emphasis on the focused constituents separates it from the rest of the constituents in the sentence.

is contextually construed, (2) contrasted entities are normally followed by the particle *thì*, and (3)

contrast is prosodically realized with high pitch (Cao 1991/2004:120-122).

With respect to the particle *thì*, Cao claims that *thì* is a topic marker which can also take

on the role of a contrastive topic marker as a subsidiary function. Upon close examination, Cao's

understanding of *thì* as a topic marker is quite unique; to be more precise, he considers *thì* a

topic-comment boundary particle. Specifically, he is convinced that if one is to test the topic-

comment boundary of a sentence is, *thì* can be used as a test word: where the topic ends and the

comment starts in a sentence is where the particle *thì* can be used (Cao 1991/2004:239).[16] This

stance states that *thì* is not the only mean of topic marking, but its role as a topic marker is most

pragmatically felicitous when the topic-comment boundary cannot be specified through syntax

alone. In terms of contrastive topics, he notes that the clearer the topic-comment boundary is, the

more likely a topicalized constituent is interpreted contrastively when being followed by *thì*.

Because of this side effect, Cao claims that contrastive topic marking is a subsidiary function that

*thì* can take on. Illustrative examples are provided below.

(33)    (Cao 1991/2004:240)
        Món    này    ***ăn***$_{CT}$   ***thì***    ngon        nhưng ***làm***$_{CT}$ ***thì***   mất công.
        Dish   DEM    eat     PTL     delicious   but     work    PTL     laborious
        'This dish, eating is delicious but making (it) is laborious.'

(34)    (Cao 1991/2004:245)
        ***Tôi***$_{CT}$ ***thì***    sẽ      đi      sau.
        1       PTL     FUT     go      after
        'I will go later (you can go first).'

Besides the examples above, some of Cao's examples where he claims that *thì* is a topic marker

show that *thì* can be used inter-clausally.

---

[16] He also mentioned that *là* 'be', commonly known as the copula in Vietnamese, can also be used as a topic-comment boundary particle, specifically, a comment marker. However, accounting for the particle *là* is beyond the scope of this thesis. See Cao (1991/2004) or Clark (1995) for further discussion on *là*.

(35)    (Cao 1991/2004:241)
       Anh    mời    ***thì***    tôi    đến.
       3       invite  PTL   1     come
       '(If) you invite me, I will come.'

As seen in this example, Cao defines the term topic very broadly. Under his framework, clauses can take on topichood though he does not specify reasons for adopting this view. In other words, Cao does not make a distinction between the phrasal and inter-clausal *thì*. Consequently, from his viewpoint, the token of *thì* in (35) is treated the same as the inter-clausal *thì* in this thesis' framework. With respect to the sentence-initial *thì*, Cao also acknowledges its existence through the following examples, claiming that this usage needs contextual licensing.

(36)    (Cao 1991/2004:192)
       ***Thì***    tôi    có    nói    gì    đâu?
       PTL   1     yes    say    what   where
       'I did not say anything?'

(37)    (Cao 1991/2004:192)
       ***Thì***    anh    giỏi!
       PTL   2     good
       'You must be good!'

Cao considers the usage of *thì* as a sentence-initial particle, however, as marginal and does not offer an analysis of it in his system. Given that Cao's objective is to offer an understanding of Vietnamese from a functionalist perspective and to emphasize the nature of Vietnamese as a topic-comment language, his work gives critical insights into how the Vietnamese language works. His passing mention of the sentence-initial *thì* and his intuitive understanding of the particle lays the foundation for this thesis to explore the usages of the sentence-initial *thì* more in-depth.

### 2.3.3 Clark (1992a)

Marybeth Clark is another scholar who attempts to provide a detailed description of the particle *thì* and is among the first to propose an analysis of the inter-clausal *thì*. In a nutshell,

Clark describes *thì* as "an inchoative conjunction" (Clark 1992a:105-107). This particle can

occur in various positions. She specifically highlights the inter-clausal *thì* and the types of

clauses it can conjoin, which was by-and-large lacking in previous research on the particle. In

terms of its function, Clark resorts to analyzing it as a conjunction with a topicalizing function.

She proposes this approach in order to offer a unified analysis of the phrasal *thì* and inter-clausal

*thì*. To account for the topicalizing function of the inter-clausal *thì*, Clark follows Haiman (1978)

and Dyvik (1982) to extend the scope of topic to treat clauses as topics.

     Prior to researching the particle *thì*, Clark reports extensively on the White Hmong

particle *los*, which, according to Clark, shares similar characteristics to those of *thì*. She later

adapts the analysis of *los* to account for the linguistic properties of *thì*. I detail here some brief

remarks of Clark (1988) on the Hmong particle *los*, which give a good background for her claims

on *thì* in Clark (1992a). Clark (1988) provides an in-depth look at how the particle *los* is used in

Hmong; surprisingly, both the syntactic environments of *los* and its functions bear great

resemblance to those of the particle *thì*.[17] Clark believes that both *los* and *thì* show characteristics

---

[17] For example, *los* also occurs (1) with a phrasal constituent marking it as a contrastive topic, (2) in a copula-like
position that relates a stative verb or an adjective in the following constituent to the preceding constituent, (3) inter-
clausally between a subordinate clause and an independent clause, and (4) inter-clausally between independent
clauses. I changed her gloss of *los* from 'then' to PTL, for the sake of consistency.

(1)  Nws    paub    kuv    ***los***    kuj    paub    thiab
     3       know    I     **PTL**    likewise know    also
     'He knows it, and me, I know it, too.'

(2)  Neb   tau     muag   wb    cov    paj-ntaub    poob    lawn    ***los***    zoo    kwang
     You2  get      sell     I      group   stitchery     gone    already  **PTL**    good    very
     'You sold my pieces of needlework already, well that's very good.'

(3)  yog      nws     yuav   tuaj    ***los***    tsis    txhob   qhib    qhov rooj
     Be that  3      will     come   **PTL**    not    don't    open    door
     'If she's going to come, don't open the door (for her).'

(4)  Nws   ua      paj-ntaub    ***los***    tsis    muaj   chaw   muag   lawn
     3      do      stitchery     **PTL**    not    have   place   sell    already
     'She does needlework and/but there's no place to sell it.'

of "inchoation", which is "an indication of the commencement of actual or perceptual change" (Clark 1988:93) and provides the glosses 'then' and 'and then.' She also finds that both *los* and *thì* can occur inter-clausally, and that the clauses they conjoin share some semantic correlation to another, which Clark refers to as "a co-subordinating relationship." The term co-subordination is coined by Foley and Van Valin (1984:265ff), which refers to "clause linkages that do not involve embedding but are not altogether independent" (Clark 1988:96). In other words, even though co-subordinating clauses are grammatically independent, they are still semantically dependent.

Using these as foundational claims, Clark (1992a) provides a description of *thì* followed by an analysis of its function(s). She describes that *thì* always occurs between X1 and X2 and uses the following notation: X1_X2, where "X1 may be a noun phrase (NP), a marked or unmarked subordinate clause ([S]), or an independent clause (S), and X2 is a predicate (VP) or an independent clause (S)" (Clark 1992a:95). She documents various tokens of the particle *thì* used in modern fiction, poems, folktales, proverbs, etc. A notable strength of her work is the detailed description of the prevalence of the particle *thì* in general, debunking the misconception that the particle is sparingly used.

She starts her description with the common understanding that *thì* is mainly used as a contrastive topic marker on the phrasal level, as illustrated below. *Thì* is glossed as 'then' in Clark (1992a), but is changed to PTL here, for the sake of consistency.

(38)  [Anh      Lâm]$_{X1}$   ***thì***   [ở     nhà]$_{X2}$
       Brother   Lam           PTL      stay   home
      'Elder brother Lam stays home.'

(39)  Hôm nay   tôi   bận    nhưng    [mai]$_{X1}$    ***thì***   [đi   được]$_{X2}$
       Today     1     busy   but      tomorrow      PTL       go   able
       'Today, I'm busy, but tomorrow, I can go.'

Clark also highlights the prevalence of the inter-clausal *thì* that past research has not explored. She sub-categorizes the correlation of the clauses conjoined by the inter-clausal *thì* into three types. The first one is most commonly found in conditional clauses where X1 and X2 are grammatically and semantically dependent (Clark 1992a:98), as in (40).

(40)  [Nếu  chị   thích]$_{X1}$ ***thì*** [chúng ta    sẽ    đi    viếng  Seattle]$_{X2}$.
     If    2    like   PTL   1         FUT   go   visit   Seattle
     'If you like, we can go visit Seattle.'

The second type is when X1 and X2 are grammatically independent but semantically dependent, also referred to as "an interdependent relationship" (Clark 1992a:100), as in (41).

(41)  [Người    ấy     đem   mở    ra    xem]$_{X1}$    ***thì***  [thấy  vàng  vẫn
     Person    DEM   take   open  out   see          PTL    see    gold   still

     còn nguyên  như   trước]$_{X2}$.
     intact       like   before

     'He opened it up to see and saw that the gold was still there as before.'

The last one is when X1 and X2 are both grammatically and semantically independent, also referred to as "a sequential relationship" (Clark 1992a:101), as in (42).

(42)  [Chị   đi    cuối   đường Lê Lợi]$_{X1}$    ***thì***  [đi bộ  tới     chợ]$_{X2}$.
     2     go    end    road   Le Loi        PTL    walk   to     market
     'When you get to the end of Le Loi street then walk to the market.'

Teasing out the semantic and pragmatic nuances that X1 and X2 can carry in the case of the inter-clausal *thì* offers critical insights to further the discussion on this particle.

With regard to her analysis, Clark analyzes *thì* as a conjunction with a topicalizing function. She also extends the scope of topichood to cover clauses in order to provide a unified analysis of both the phrasal and inter-clausal *thì*. This approach raises important questions about the nature of *thì,* but it also prompts a reconsideration of the nuanced differences between the phrasal and inter-clausal *thì*, if at all available. As the analysis chapter will expound with more

details, the functions of the phrasal *thì* and inter-clausal *thì* are actually distinct. With specific

regard to the inter-clausal *thì*, its scope goes beyond topic-marking; it can also be used to signal a

causal relationship, as shown in the example below.

(43)    [Tôi    thích    táo]$_{X1}$ ***thì***    [tôi    mua    táo] $_{X2}$.
        1      like     apple  PTL    1      buy    apple
        'I like apples; (therefore,) I bought apples.'

In this example, the inter-clausal *thì* does not just express that X2 is sequentially following X1.

Rather, X2 is a reasonable consequence given a basis provided in X1. Overall, Clark's detailed

description of the inter-clausal *thì* greatly contributes to the literature on it. As for her analysis,

she was restricted to looking at the functions of the particle *thì* on the phrasal and inter-clausal

levels. Given this objective, she was able to highlight the topic-marking nature of *thì*, but some

functions that are exclusive to either the phrasal *thì* or inter-clausal *thì* still need to be accounted

for, which will be addressed in the description and analysis of the particle in this thesis.

*2.3.4 Tran (Forthcoming)*

Thuan Tran is the most recent scholar to comment on the linguistic properties of the

particle *thì*. His analysis is also the most formal analysis of the particle. However, it is important

to note that Tran's main focus is not to analyze *thì*, but to explore the syntax and information

structure interface in Vietnamese through the lens of the mapping rules framework proposed by

Neeleman and Van de Koot (2008). Tran hinges his discussion of how information structural

categories such as topic, focus, and contrast interact with syntax in Vietnamese on Neeleman's

framework. His main claim is that contrast is an autonomous category in information structure

that results in a syntactic reflex (specifically, left dislocation) and a semantic consequence. In

proving this position, Tran argues that the particle *thì* can be used following a left dislocated

constituent, which can be either a contrastive topic or a contrastive focus. Therefore, even though

*thì* is only used as a tool to support Tran's stance, he also inevitably provides an analysis of the particle and its interaction with pragmatic notions of topic, focus, and contrast. Also because of his theoretical focus, he does not provide a broader description of how the particle *thì* is used in Vietnamese. Indeed, his representation of the particle gives the impression that the particle can only be used as a phrasal particle.

Since Tran's framework is inspired by Neeleman and van de Koot's approach, a brief mentioning of Neeleman and Van de Koot's framework is necessary. Neeleman and van de Koot were the pioneering scholars in proposing the mapping rules approach to show the interaction between the syntax and the information structure, opposing the Cartographic Approach proposed by Rizzi and Cinque. The most applicable aspect of Neeleman and Van de Koot's work to the discussion in Tran is the notion of "a discourse template marker." Examples (44) and (45) are provided by Neeleman to represent a contrastive topic and contrastive focus, respectively. The contrastive topic is underlined and the contrastive focus is italicized in small capitals.

(44)    (Neeleman & van de Koot 2008:143)
        Do you know who John gave the book to?
        I'm not sure, but he gave the record to Susan.

(45)    (Neeleman & van de Koot 2008:142)
        Mary bought *A HAT* in Paris (not a shirt).

In their framework, the contrastive topic 'the record' and the contrastive focus 'a hat' disrupts both the surface constituents of their comment 'he gave X to Susan' and the background 'Mary bought X in Paris.' Therefore, when the contrastive topic and the contrastive focus are left dislocated, they no longer "interrupt a comment and a background." Since the topic and focus can be construed as such even when they are in-situ, "movements of topics and foci do not mark the discourse functions of these elements themselves, but rather their comments and backgrounds" (Neeleman & van de Koot 2008:143-144). From this evidence, they argue that left dislocation is

29

not to mark the topic or focus, for there is evidence that the topic and focus can be marked with other phonological or morphological means. Rather, left dislocation marks the remaining constituent as the comment or background, which is the role of "a discourse template marker."

Adopting their definition of a discourse template marker, Tran demonstrates that *thì* is a discourse template marker. His main claim is that "in Vietnamese topic and focus are two basic information structural categories that surface in situ, whereas contrast is an autonomous information structural category that triggers topic/focus left dislocation" and that left dislocated constituents are marked by the discourse template marker *thì* (Tran Forthcoming:2). In terms of definitions such as topic, focus, and contrast, Tran's view of these notions in Vietnamese is fairly narrow. Tran defines the term topic as "aboutness topics" and adopts Krifka's definition (Krifka, 2006) as a working definition for topic: "The topic constituent identifies the entity or the set of entities under which the information expressed in the comment constituent should be stored in the Common Ground content" (Tran Forthcoming:2). For this reason, Tran believes that topic should be referential and identifiable. He also restricts topics to sentence topics and excludes discourse topics, or anaphoric references of topics (i.e. resumptive pronouns, deixis, etc.). As for focus, Tran claims that Vietnamese also construes a focus as a possible answer to a wh-question. He elaborates, "The contribution of focus is to provide a set of alternative elements of the same type as the focus. Replacing the focus with the alternative elements will generate a set of propositions alternative to the expressed proposition" (Tran Forthcoming:7). He also claims that the focus in Vietnamese is prosodically marked, but he does not provide evidence of prosodic cues. In terms of the category of contrast, he is convinced that topics and foci in Vietnamese surface in-situ and are left dislocated only when the category of contrast is involved. Semantically, Tran maintains that the contrast implies a set of alternatives to the chosen topic or

focus. In other words, his framework and understanding of these categories fuse the distinction

between a contrastive topic and focus. An example of a contrastive topic and contrastive focus[18]

under Tran's framework is provided below in (46) and (47), respectively.

(46)  (Tran Forthcoming:10)[19]
| A: | Các | pop stars | thế nào? | Ai | dẫn | họ | đi | tham quan? |
|----|-----|-----------|----------|----|-----|----|----|------------|
|    | PL  | pop stars | how      | who | lead | 3 | go | explore |

'What about the popstars? Who showed them around?'

| B: | *Các* | *pop stars*CT | *nữ* | *thì* | Bill | dẫn | đi | tham quan. |
|----|-------|---------------|------|-------|------|-----|----|------------|
|    | PL    | pop stars     | female | PTL | Bill | lead | go | explore |

'The female popstars, Bill gave a tour.'

(47)  (Tran Forthcoming:11-12)
Context: Nam's parents know their son must read five books to prepare for the exam and
are discussing which one Nam has read.

| A: | Nam | đã | đọc | Quê Người. |
|----|-----|-----|-----|------------|
|    | Nam | PST | read | Que Nguoi |

'Nam read Que Nguoi.'

| B: | Anh | biết. | **Quê Người**CF *thì* | Nam | đã | đọc | nhưng |
|----|-----|-------|------------------------|-----|-----|-----|-------|
|    | 1   | know  | Que Nguoi      PTL    | Nam | PST | read | but |

| **O Chuột**CF | *thì* | Nam | chưa | đọc. |
|---------------|-------|-----|------|------|
| O Chuot       | PTL   | Nam | yet  | read |

'I know. Que Nguoi, Nam has read, but he has not read O Chuot.'

Fusing the notions of contrastive topic and focus has its own merits, one of which is its

ability to account for the close-knit relationship between the particle *thì* and contrast. Indeed, the

presence of *thì* usually implies a contrastive usage. Within the context of the current literature on

this topic, Tran's framework is the first work to account for this contrastive marking property of

---

[18] He does specify a semantic nuance between CT and CF. Regarding contrastive topics, Tran says that the speaker, in uttering a contrastive topic, "is unwilling or epistemically unable to make an alternative utterance" (Tran Forthcoming:15). In the case of contrastive foci, speakers utter the contrastive focus with implication of denial of other alternatives. In other words, whereas a speaker does not want to utter a statement about the alternative contrastive topics, a speaker insinuates that only the contrastive focused constituent is a felicitous choice given the information in the background.

[19] This example is a Vietnamese version of an equivalent example proposed in Neeleman and Vermeulen (Neeleman & Vermeulen 2012:10).

the particle *thì* in a more formal and systematic way. However, the fused definitions of three

information structural categories of topic, focus, and contrast also create some ambiguous cases.

For instance, a constituent that is considered a topic in Tran's framework can be more reasonably

classified as a focus, as in the example below.

(48)     (Tran Forthcoming:4)
         Context: Let me tell you something about Tan.

         Công ty A      đã      nhận   *Tân*~T~
         Company A   PST    accept  Tan
         'Company A accepted Tan.'

Tran defines topics as 'aboutness topics'; therefore, given the context, Tran argues that Tan is the

topic of the sentence. Though this remark does give further information about Tan, which is his

acceptance into company A, it is infelicitous. If the topic of the sentence is truly Tan, the passive

construction would more likely be used so that the topic NP Tan is sentence-initial, as in the

example below.

(49)     Context: Let me tell you something about Tan.

         *Tân*~T~   được   công ty A       nhận.
         Tan    PASS   company A     accept
         'Tan got hired by company A/Tan got accepted into company A.'

As mentioned above, Tran claims that focus can be realized in-situ. If both topic and focus can

surface in-situ, how can they be distinguished? Indeed, using the definitions Tran provided, the

NP Tan in (48) would be more reasonably construed as a focus, an answer to the question 'Who

did company A hire?'.

   With regard to his treatment of the particle *thì,* Tran rejects the past analyses of the

particle as a topic or contrastive topic marker, using the following example.

(50)    (Tran Forthcoming:6)
        Context: A bargaining context in which the seller asks the buyer how much he/she is happy to pay for the merchandise after the customer's rejections of the price offers given by the seller.

| Cái | áo | này | bao nhiêu | ***thì*** | anh | mua? |
|-----|-----|-----|-----------|-----------|-----|------|
| CLF | shirt | DEM | how much | PTL | 2 | buy |

'This shirt, how much are you willing to pay?'

He believes that the ex-situ wh-question licenses the contrastive interpretation, an answer to which is a left-dislocated contrastive focus. This left-dislocated question word is followed by the particle *thì*. Consequently, he disputes the traditional analysis of *thì* as a topic or contrastive topic marker on this basis. However, a question like (50) is of a peculiar nature. That is, rather than a true wh-question asking for information, it is much more of a rhetorical question expressing frustration with regard to the amount that the buyer is willing to pay. As is detailed in the context provided by Tran, the seller uttered this remark after getting a series of rejections from the buyer regarding the price of the shirt. In such a context, the seller is more likely to utter that question to express his/her frustration without actually inviting any answer.[20] In other words, this question may be an outlier in this analytical scheme of topic and focus. In order to base the theory that *thì* can also mark focus and contrastive focus, more examples would need to be provided. In the grand scheme of his work, Tran's approach captures the role of the phrasal *thì* well and raises interesting questions regarding its functions. His discussion on the particle deepens the understanding of *thì*'s nature, especially the particle's interaction with information structure. With regard to the usages of *thì* on a sentential level, since Tran exclusively looks at the phrasal *thì,* the usages of *thì* as an inter-clausal and sentence-initial particle are not considered in his framework.

---

[20] This question is also more complicated. If the question word bao nhiêu 'how much' is placed before *anh mua*, the action of buying the shirt is understood as not happening yet. However, if the question word bao nhiêu 'how much' is placed after *anh mua*, the action of buying the shirt is understood as already happened ('You bought the shirt.

# CHAPTER 3

## Data Collection Process & Description of *Thì*

This chapter first details the data collection process and then provides a description of the particle *thì* as a phrasal, inter-clausal, and sentence-initial particle. The Pear Film movie produced by Wallace Chafe in 1980 is used as the main source for data collection. The rationale for collecting data, the choice of the Pear Film movie, the method of data collection, the data processing steps, and some overall findings will also be discussed in detail in this chapter. The description of the particle *thì* shows how *thì* is used in Vietnamese, including when it is pragmatically felicitous, the types of constituents it can occur with, and the semantic and pragmatic ramifications that it carries. Illustrative examples are drawn from previous works on *thì* (Cao, 1991/2004; Clark, 1992a; Thompson, 1965/1987)*,* constructed examples, and the tokens of *thì* found in the data collected for this thesis. The data collection demonstrates that the particle *thì* is mostly used as a sentence-initial particle.

Chapter 3 is structured as follows. Section 3.1 details the data collection process, including the rationale for collecting data (3.1.1), the procedure (3.1.2), data processing (3.1.3), and overall findings (3.1.4). Section 3.2 describes the usages of *thì.* Each section starts with the possible ways that the particle *thì* can be used in the following syntactic environments: phrasal, inter-clausal, and sentence-initial. Afterward, a documentation of its frequency as documented in the collected data is presented. Overall, my findings contribute to the already enigmatic nature of the particle, which necessitates a reconsideration of the particle's status in Vietnamese and an analysis to capture its functions.

---

How much did you pay for it?). In other words, the left dislocated of the question word followed by *thì* is more complex than it is currently presented.

### 3.1 Data Collection Process

*3.1.1 Rationale*

This section reports the rationale behind the data collection decision, such as the purpose of the pilot study and the main study, and the usage of the movie The Pear Film by Chafe (1980) is used to gather the corpus of data. In the preliminary research stage, a close examination of past literature and my intuition of *thì* as a native speaker gave a strong impression that the particle *thì* was mainly used as a topic and contrastive topic marker occurring with phrasal constituents. With that conjecture, the original motivation for data collection was to extract more tokens of the phrasal *thì* and explore the nuances of topic and contrastive topic marking properties that *thì* carries. Contrary to my expectation, the pilot data revealed an indisputable reality: the particle *thì* is prevalently used as a sentence-initial particle. This finding fueled my desire to use the same source (the Pear Film movie) to collect more tokens of the particle *thì* in the main data collection round.

The main objective for the source of data collection is one that can streamline the data collection process and can extract the tokens of *thì* in natural conversations. The Pear Film developed by Chafe (1980) is an appropriate fit. Wallace Chafe is renowned for his research on pragmatic notions of topic, focus, and contrast. He is also interested in ways that humans verbalize knowledge that they acquire consciously or subconsciously. Chafe's research interests birthed the production of the Pear Film movie. The Pear Film movie project aims to capture the linguistic, cultural, and cognitive ramifications of language through the lens of narrative production. In Chafe (1980), the scholars contributing to the project share their rationale behind the movie as follows:

We assume that much of human knowledge is stored nonverbally, as is evident, for

example, by our often-having difficulty putting thoughts into words, as well as our

seldom verbalizing the same thoughts in the same way on different occasions. . . We also

wanted to collect data in a number of different languages in order to compare our

findings cross-linguistically. . . We wanted something that would include a set of events,

some in sequence and some simultaneous, some highly codable and some not, some

trivial and some salient, with a set of people and objects that participated in the events in

various ways, themselves varying in codability and salience. We wanted to allow for

ambiguity of interpretation, but at the same time, we wanted the film to be easily

interpretable in some way by people in a variety of cultures. (Chafe 1980: xi-xiii)

The final product of the Pear Film project is a 6-minute movie fully acted out without any

verbal language. Characters in the movie express their emotions, attitudes, and actions solely

through nonverbal cues and body/facial expression. Therefore, the various scenes, characters,

and their interactions provide ample opportunities and involve just enough ambiguity for

participants to interpret the movie freely within a confined and consistent context. The detailed

storyline of the Pear Film movie is provided in Appendix A. In the pilot study, participants have

a positive reaction to the film. They find the ambiguous scenes intriguing and fully immersed

themselves in teasing out the ambiguity, which created plentiful opportunities for naturally

produced utterances. Therefore, the same movie is used to conduct more data in the main data

collection round.

### 3.1.2. Procedures

Interested participants were recruited through online platforms and personal connection.

Once participants agreed to participate in the research, they signed a consent form as part of the

IRB process.[21] Each participant was assigned a pseudonym and divided into groups based on the dialect of Vietnamese that they speak. In total, there were three groups of participants: the Northern Group (code: NG), the Central Group (code: CG), and the Southern Group (code: SG). All calls were conducted and recorded via Microsoft Teams, one group at a time,[22] so that all participants in a group saw the movie at the same time on the same video call. Prior to being shown the movie, each group was given the following instruction in Vietnamese, the translation of which is as follows:

> I am about to show you a short movie. The only thing you need to do is to watch the movie in its entirety. You are welcome to unmute your mic and chat among one another as you watch, but please be courteous not to distract others from watching the movie. Afterwards, I will ask you to share your recollection of the story to the best of your memory. There is no pressure to give an accurate recollection of the movie, so please be as relaxed and honest as possible.

The set of questions used to ask the participants to guide the conversation is located in Appendix B. Each call was recorded, video and audio.

---

[21] The Liberty University IRB decided that this research (IRB-FY20-21-561) is not considered as a human research subject, according to 45 CFR 46.102(l)(1): "The Liberty University Institutional Review Board (IRB) has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study does not classify as human subjects research. This means you may begin your research with the data safeguarding methods mentioned in your IRB application."

[22] Chafe did not provide specific instructions or requirements on how the data should be collected (i.e. group interviews, individual interviews, etc.). In his own research, he showed the movie to a group of participants. Afterwards, each participant met with an interviewer separately who claimed to not have seen the movie in order to tell the interview what they just watched. While this method was effective for Chafe's research agenda, he also noticed that different cultures may react to this data collection method differently and provided no further instructions on how the data should be best collected. Therefore, I relied on my understanding of Vietnamese culture and people in order to come up with the most appropriate way to collect the data so as to extract the most natural utterances possible. Culturally, Vietnamese people tend to feel more at ease to share their honest thoughts and opinions when they are not put on the spot, especially in a virtual setting. They enjoy group conversation, for they do not feel like they must speak up at all times, and the group dynamic also helps with the flow of the conversation. Therefore, to extract more natural tokens of *thì*, I decided that a group call would bring about more positive results.

## 3.1.3 Data Processing

Table 2 details the number of participants, length of conversations, and the word count of each call.

Table 2. Data Collection Information

| Group Codes | Dialects | Number of Participants | Length of Conversations | Word count |
|---|---|---|---|---|
| SG | Southern | 2 | 10 minutes | 1655 words |
| CG | Central | 3 | 11 minutes | 1763 words |
| NG | Northern | 4 | 22 minutes | 3158 words |

Other demographic information, such as age and gender, is provided in Appendix C.

Each conversation was recorded, transcribed, and translated into English with interlinear gloss translation. For confidentiality purposes, each participant was assigned a pseudonym. Due to the length of the conversations and the various participants involved, each line of the conversation is coded for easier reference. Each line is referred to in the following format: Group Code-Line Number(Pseudonym). An illustrative example is below. CG2(N) refers to line 2 of the Central Group data and is spoken by participant N.

| CG2(N): | Trong | cái | video | nớ | là | giống | như | kiểu |
|---------|-------|-----|-------|----|----|-------|-----|------|
|         | In    | CLF | video | DEM | be | similar | like | type |

| | là | ở | nước | | ngoài | á. |
|-|----|---|------|-|-------|----|
| | be | in | country | | foreign | PTL |

'In the video, it looked like the setting was in a foreign country.'

To preserve the naturalness of each utterance, stuttering, fragments, repetition, hesitation, etc. were all transcribed and translated.[23] After the data was transcribed, all utterances containing *thì* were noted.

The following classificational method was used to classify the different tokens of *thì* into its corresponding subcategories. Each token of *thì* was noted in terms of its syntactic environment, which included a documentation of what it precedes and follows. With this, the following classification is made:

i.   If the token is preceded by any phrase and followed by a predicate or a full sentence, it is treated as a token of the phrasal *thì*. An illustrative example is below.

| | NP | | | VP | | |
|-|----|--|--|----|--|--|
| NG194(T): | [Cây | đấy] | ***thì*** | [giống | cây | táo |
| | Tree | DEM | PTL | resemble | tree | apple |

| | hoặc | cây | lê | thôi]. |
|-|------|-----|----|-------|
| | or | tree | pear | only |

'That tree resembles an apple tree or a pear tree only.'

ii.  If a token is preceded by a phrasal constituent with a resumptive pronoun of that constituent in the following constituent of the token, it is treated as a token of the phrasal *thì*, as illustrated below.

---

[23] For this reason, some tokens of *thì* are repeated due to accidental repetition. However, the repetition is marginal, so it does not skew the overall statistical documentation.

<div align="center">NP</div>

SG51(P): Không! [*Quả lê của mình*] ***thì***
NEG    fruit  pear  of    1      PTL

S[24]
[[**nó**]pro đâu có cái thắt ở giữa].
3      NEG have CLF knot at middle

'No, our pears, they do not have the knot in the middle.'

iii. If a token occurs after a dependent clause ([S]) and before an independent clause (S), it is classified as a token of the inter-clausal *thì*.

[S]
NG18(S): [**Lúc** trên đường đi xe đạp
When on road go bicycle

chở cái đống lê]
bring CLF pile pear

S
***thì*** [gặp một cô bé
PTL meet one CLF kid

đang đi hướng ngược chiều]
PROG go direction opposite side

'While he was on the road bringing the pears, (he) met a girl who was going in the opposite direction.'

iv. If a token is followed by a complete sentence with no immediately preceding sentence, it is classified as the sentence-initial *thì*.

---

[24] This refers to the rest of the sentence without the topicalized constituent.

|        | S |
|--------|---|
| CG68(M): | [Anh nghĩ trái đó là trái lê]. |
|        | 1 think fruit DEM be fruit pear |
|        | 'I think that fruit is pears.' |

CG69(CM): Vì sao ạ?
'Why?'

|        | S |
|--------|---|
| CG70(M): | ***Thì*** [ở bên Nga có nhiều trái |
|        | PTL at over Russia have PL fruit |

giống        như   vậy].
similar      like   that
'In Russia, there are fruits similar to that fruit.'

The initial classificational method sorted out the majority of the tokens, except for tokens that occurred in between independent clauses and were produced by the same speaker. Those tokens were sorted out further with prosodic cues. Specifically, tokens that were uttered with a long and obvious pause between the clauses were classified as tokens of the sentence-initial *thì*. Tokens that were uttered with no or a very brief (almost inaudible) pause between the clauses were treated as tokens of the inter-clausal *thì*.[25]

## 3.1.4. Overall Findings

With the above classifications, this section reports some overall findings. In the data collection, 155 tokens of *thì* were documented: 48 tokens (31%) were the phrasal *thì*, 38 tokens (30%) were the inter-clausal *thì*, and 69 tokens (44%) were the sentence-initial *thì*. The following graph visualizes the finding.

---

[25] This subjective classification of pause length was considered adequate for this study. Conceivably, acoustic measurements of the length of any preceding pause could have been made, and a more precise cutoff defined, but in most cases, it was quite obvious whether the pause was long enough to count as a sentence break. The number of uncertain cases was small enough that their classification would not have significantly affected the conclusions.

Graph 1. The Distribution of *Thì* in the Collected Data

Distribution of the Particle *Thì* in the Collected Data

- Phrasal thì: 31%
- Inter-clausal Thì: 30%
- Sentence-initial thì: 44%
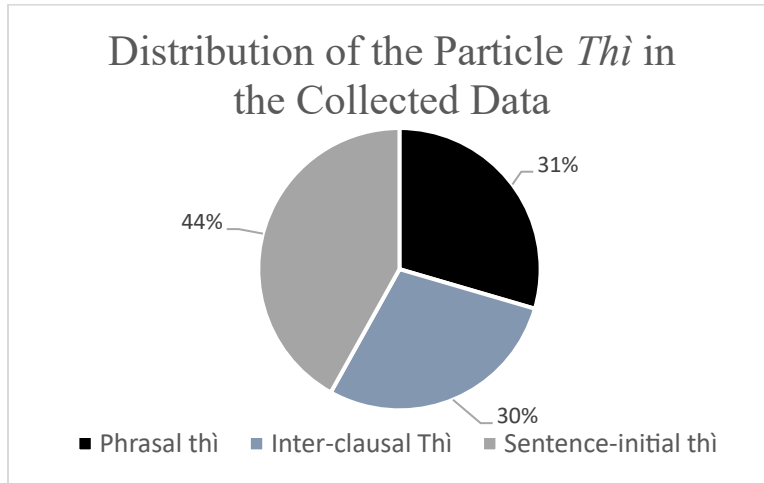
The tendency of *thì*'s usage was also examined in different dialects. The finding is shown in Graph 2. As the graph indicates, the particle *thì* was also used most as a sentence-initial particle across dialects.

Graph 2. Distribution of *Thì* based on dialects[26]

Distribution of *Thì* based on Dialects

| Dialect | Phrasal Thì | Inter-clausal Thì | Sentence-initial Thì |
|---|---|---|---|
| CG | 16 | 10 | 29 |
| SG | 8 | 14 | 16 |
| NG | 24 | 12 | 26 |

---
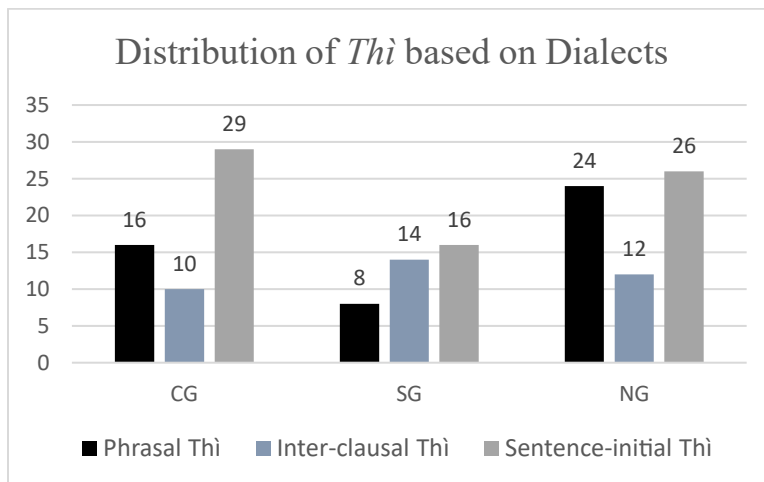
[26] The tokens of *thì* in the Northern dialect is more than those in the other dialects, for the length of the conversation form the Northerner's participants are twice as long compared to those of the other groups.

With regard to gender, a similar finding is observed. Male participants utterred more tokens of the sentence-initial *thì* than their female counterparts, but overall, the sentence-initial *thì* is generally more favored than the phrasal and inter-clausal *thì*. The findings are located in Graph 3.

Graph 3. Tokens of *Thì* based on Gender



Distribution of *Thì* based on Gender
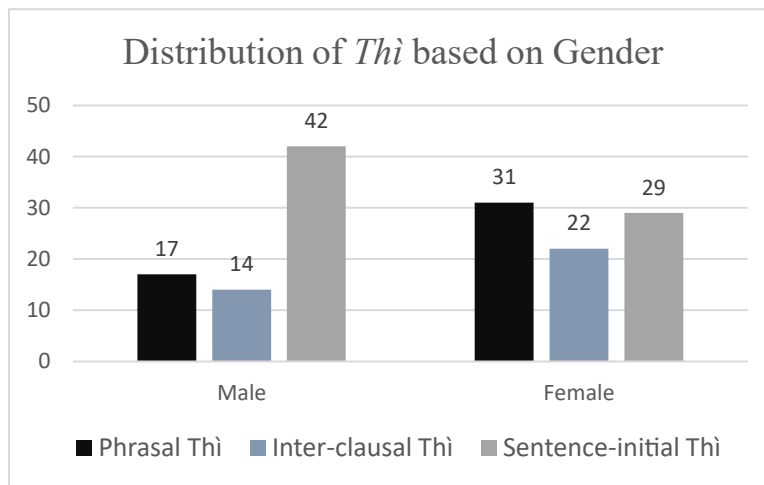
Another variable is age. Three arbitrary age groups were formed, consisting of participants who are between 18-25 years old, 26-50 years old, and above 50 years old. Graph 4 visualizes the findings. Across three groups, the sentence-initial *thì* has the highest frequency.

Graph 4. Tokens of *Thì* based on Age Groups



Distribution of *Thì* based on Age Groups

As shown in all three graphs, the prevalence of the sentence-initial *thì* is consistent across

dialects, gender, and age. Given this finding, it is puzzling that the properties of the sentential *thì*,

especially those of the sentence-initial *thì*, are still under-researched. In the following section, a

description of how the particle *thì* is used as a phrasal, inter-clausal, and sentence-initial particle,

will be discussed in detail.

## 3.2 The Description of *Thì*

### 3.2.1 Phrasal Thì

3.2.1.1 The Phrasal Thì as a Topic Marker

Section 2.2.1 highlights that Vietnamese marks topics through word order. Besides word

order, the particle *thì* can also be used as a topic marker based on our definition; the topic

follows various types of phrasal constituents. Since subject NPs are most likely to take on

topichood in Vietnamese, the particle *thì* is usually found following topicalized subject NPs, as

in the example below.

(51)　　NP　　　　　　S
　　　　[*Nam*]　　***thì***　　[(nó)　thích　táo　lắm]
　　　　Nam　　　PTL　　3　　like　　apple　very
　　　　'Nam, he likes apples very much.'

Besides topicalized subject NPs, the phrasal *thì* can also follow topicalized object NPs, as in (52),

PPs, as in (53), and VPs, as in (54).

(52)  NP                                      S
      [*Măng cụt   kia*]        ***thì***   [mẹ    mua   (đó)]
      Mangosteen  DEM          PTL          mom    buy   DEM
      'Those mangosteens, mom bought them.'

(53)  NP                       S
      [*Trên  bàn*]  ***thì***   [mẹ    để    táo    (trên  đó)].
      On    table  PTL          mom    put   apple  on     DEM
      'On the table, mom put the apples there.'

(54)  VP                       VP
      [*Nấu ăn*]   ***thì***   [không phải   ai     cũng   giỏi   (làm)]
      Cook        PTL          NEG    true   who    also   good   do
      'Cooking is not something anyone can be good at (doing).'

The examples show that topic is usually referred back to with resumptive pronouns or pro-nominalization, creating a topic-comment structure for their sentences with the particle *thì* occurring right at the topic-comment boundary. The presence of *thì* in topic marking cases is most critical when the topic needs to be narrowed down with qualifying information, resulting in a blurry boundary between the topic and the comment. Cao (1991/2004) strongly advocates for this role of *thì*, claiming that separating the topic from the comment is the main role that the particle *thì* takes on. In running speech, such a boundary can be challenging to identify without *thì*. Furthermore, given that the topic-comment structure is a pragmatically neutral sentence structure in Vietnamese (see 1.2), an ambiguous topic-comment boundary can hinder communication. As such, the presence of the particle *thì* can be helpful for topic identification. (55) and (56) are representative examples of those cases.

(55)  NP             S
      [*Cá*]  ***thì***   [cá    chép   nhiều  xương].
      Fish   PTL          fish   cod    PL     bone
      'Fish, cod fish have many bones.'

In (55), the topic *cá* 'fish' and the subject of the comment *cá chép* 'cod fish' are phonologically similar; without the presence of *thì*, the utterance can be phonologically challenging and

pragmatically awkward, and the topic-comment boundary is also ambiguous. Therefore, to avoid

these issues, using *thì* following the topic *cá* 'fish' is pragmatically advantageous to identify the

topic of the sentence instead of just relying on word order.

(56) NP                                        VP
     [*Ngày  nào  mà     anh  rảnh*] ***thì***     [đến  thăm  tôi     nhé!]
     day    what  REL   2      free   PTL     come  visit  1      PTL
     'Whichever day that you are free, please come visit me.'

In (56), the indefinite NP *ngày nào* 'any day' needs to be narrowed down with the relative clause

*mà anh rảnh* 'that you are free' to be qualified as the topic of the sentence. The presence of the

relative clause extends the length of the topic, making it harder for the hearer to distinguish when

the topic ends and when the comment starts. Again, the presence of *thì* helps clarify that

boundary.

3.2.1.2 The Phrasal *Thì* as a Contrastive Topic Marker

      When *thì* is used as a contrastive topic marker, it must be triggered by contextual cues.

The examples below provide the contexts where the usage of the particle *thì* as a contrastive

topic marker is appropriate. The contrastive topics (CT) are bold and italicized.

(57)    Context: A and B are talking about Nam.

    A:    Hôm qua    Nam  ăn     hết    táo    Lan   được  tặng!
         Yesterday   Nam  eat    all     apple  Lan   PASS   receive
         'Yesterday, Nam ate all the apples that Lan was given!'

                                         NP         VP
    B:    Nam  thích  táo    lắm    nhưng [***Lan***] ***thì***  [không thích  táo].
         Nam  like    apple  very   but    Lan   PTL    NEG  like    apple
         'Nam likes apples very much, but Lan does not like apples.'

In this example, both Nam and Lan are mentioned in A's utterance. From there, speaker B

contrasted Lan's dislike of apples to Nam's predilection for apples. The particle *thì* follows the

contrastive topic, which is Lan in this case.

46

(58)   A:   Anh   làm   sai   rồi!          Anh   không biết   nấu ăn sao?
            2     do    wrong already         2     NEG   know   cook  PTL
            'You did it all wrong. Don't you know how to cook?'

            NP          AP            NP          AP
      B:   [*Cô*]  *thì*  [giỏi],   còn   [*tôi*]  *thì*  [dở]. Được chưa?
            2      PTL   good   remain   1     PTL   bad   okay yet
            'You are good at that, but I am extremely bad. (Sarcastic) Are you fine with that?'

Besides following contrastive topics that are subject NPs, the phrasal *thì* also follows

other types of syntactic constituents and marks them as the contrastive topics of their sentences.

In the examples below, the phrasal *thì* follows an object NP in (59), a predicate in (60), and a PP

in (61).

(59)   Context: A and B came home and saw apples on the table.

      A:   Ủa,   táo   này   ở đâu   ra          vậy?
            PTL   apple DEM   where   become      PTL

            Có   phải   táo   Nam   mua   không?
            Yes   true   apple Nam   buy   NEG

            'Where do these apples come from? Are these the apples that Nam bought?'

            NP                    S
      B:   [*Táo đó*]   *thì*   [tôi   mua].
            Apple DEM   PTL    1     buy
            'Those apples, I bought (Nam bought the other apples).'

(60)   Học   tiếng Việt   mất   nhiều thời gian      quá!
      Study Vietnamese   cost   PL    time          very

            VP                VP
      Còn   [*nấu ăn*]   *thì*   [dễ   làm   hơn   nhiều].
      Remain   cook        PTL   easy   do   more   much

'Studying Vietnamese costs so much time! But cooking is so much easier to do!'

47

(61) A:  Chiều      qua     kẹt xe   gần   hai    tiếng!
        Afternoon   past    traffic  close  two    hour
        'Yesterday, (I got stuck) in a traffic jam for almost two hours.'

         PP                                          VP
    B:  [*Vào  năm  giờ   chiều*]        *thì*    [hay     kẹt xe].
         At    five  hour  afternoon      PTL      usually  traffic

        Bảy    giờ    tối    sẽ    ít     xe    hơn.
        Seven  hour   night  FUT   less   car   more

        'At 5 P.M., there is usually a traffic jam. Around 7 P.M., there will be fewer cars.'

All of these examples, however, only contain one contrastive topic. What if the context invites the contrast of more than two topics? The question arises as to whether the phrasal *thì* can follow all of the contrastive topics. Let us consider the example below.

(62)    Context: A came to visit Lan's family and talked to B, their neighbor.

    A:  Cả     nhà     Lan    đâu?   Họ      không  có     nhà    à?
        All    family  Lan    where  3       NEG    yes    home   PTL
        'Where is Lan's family? Are they not home?'

                             NP                   VP
    B:  *Gia đình*   nóᴛ    [*ba*]   (*#thì*)²⁷   [đi     làm],
        Family       3      dad      PTL          go      work

        NP            VP               NP                    VP
        [*mẹ*]   (*#thì*)   [đi  chợ],  [*em trai*]   (*#thì*)  [đi     học],
        mom       PTL        go  market  brother       PTL       go      study

              NP                        VP
        nhưng [*Lan*]  *thì*    ∅_CT   [có     nhà].
        but    Lan     PTL              yes     home

        'The father went to work, the mom went to the market, the brother went to
        study (at school), but Lan is home.'

Besides Vietnamese, some languages exclusively mark contrastive topics with overt markers, an example of which is Korean. In this example, an equivalent Korean example would use the

---

²⁷ This notation means that the usage of the particle *thì* here is pragmatically infelicitous. That is, it is not ungrammatical to have the particle *thì* there, but its presence is pragmatically unncessary and quite redundant.

contrastive topic marker *un/nun* following all four members of the family, without which the utterances are infelicitous.[28] On the other hand, as (62) shows, the particle *thì* can only be used after Lan. This example shows that in Vietnamese, the contrastive topic marker *thì* does not contrast all the alternative contrasted entities to one another. Rather, it subdivides the alternatives into subsets. In this case, Lan forms a subset that is contrasted with the other three family members, who all form another subset of their own.[29]

3.2.1.3 The Tendency of the Phrasal *Thì*'s Usages in the Collected Data

The data findings demonstrate both the topic-marking and contrastive topic-marking natures of the phrasal *thì*. With respect to its topic-marking function, the data show that *thì* is most commonly used as a topic marker following prepositional phrases or noun phrases that refer back to the previously established background context, whether temporal or spatial, as illustrated below.

(63)             PP               S

| NG20(S): | [*Sau* | *đấy*] | ***thì*** | [có | một | đám | trẻ | lại |
|----------|--------|--------|-----------|-----|-----|------|-----|------|
|          | After  | DEM    | PTL       | have | one | bunch | kid | arrive |

| | giúp | cái | bạn | trai | đấy | nhặt | mấy | quả | lê]. |
|-|------|-----|-----|------|-----|------|-----|-----|------|
| | help | CLF | friend | boy | DEM | pick | PL | CLF | pear |

'After that, there was a group of kids helping that boy pick up the pears.'

---

[28]The equivalent example is consulted through personal connection with Korean native speakers.
[29] If there are only two contrasted entities mentioned, *thì* can be used after both, as illustrated in (58).

49

(64)

|  | NP |  |  | **S** |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
| NG40(U): | [*Lúc* | *này*] | ***thì*** | [*có* | *một* | *người* | *đàn ông* |  |
|  | moment | DEM | PTL | have | one | CLF | man |  |

| *dắt* | *con* | *dê* | *hay* | *một* | *con* | *gì* | *đó* |
|---|---|---|---|---|---|---|---|
| lead | CLF | goat | or | one | CLF | what | DEM |

| *đi* | *ngang* | *qua*]. |
|---|---|---|
| go | past | |

'This moment, there was a man walking a goat or some animal passing by.'

Furthermore, the data also show that participants tend to use *thì* in the topic-comment structure where the topic does not hold any grammatical relation in the sentence.

(65)  CG32(O):

| *Mình* | *để ý* | *thứ nhất* | *là* |
|---|---|---|---|
| 2 | notice | first | be |

|  | NP |  |  | **S** |  |  |
|---|---|---|---|---|---|---|
| | [*cái* | *video* | *đó*] | ***thì*** | [*cái* | *cây* | *đó* |
| | CLF | video | DEM | PTL | CLF | tree | DEM |

| *là* | *một* | *cái* | *cây* | *cổ thụ* | *nha*]. |
|---|---|---|---|---|---|
| be | one | CLF | tree | old | PTL |

'I noticed the first thing was that that video: the tree was a century-old tree.'

Lastly, the phrasal *thì* is also documented to mark topics that have resumptive pronouns in the comment, as in (66).

(66)
|  | NP | | | | S | | | |
| NG208(U): | [*Cái* | *bộ phim* | *này*] | *thì* | [tính ra | [*nó*]pro | | không | có |
| | CLF | movie | DEM | PTL | calculate | 3 | | NEG | yes |

| tốn | kinh phí | để | đầu tư | hay | là | làm | gì | hết]. |
| consume | expense | to | invest | or | be | do | what | all |

'This movie, to estimate, it did not cost much money to invest or to be made.'

As for its role as a contrastive topic marker, the data also illustrate that the phrasal *thì* follows noun phrases and marks the noun phrases as contrastive topics.

(67)    Context: Participants were asked what they thought the fruit was.

CG71(N): 'I guess those were apple-pears. They look like pears, guavas, guava-pears. They look like guava-pears.'

CG72(M): 'Guava pears are pears. So, they are pears, right?'

| | | | | NP | | VP | |
| CG73(N): Không | biết | được, | nhưng | [*bên* | *này*] | *thì* | [có]. |
| NEG | know | able | but | over | DEM | PTL | have |

| Hắn | nhìn | giống | | quả | lê, |
| 3 | look | similar | | fruit | pear |

| | NP | | | | VP | | |
| nhưng | [*da* | *bên* | *ngoài*] | | *thì* | [giống | quả | ổi]. |
| but | skin | over | outside | | PTL | similar | fruit | guava |

'I don't know, but over here where I live, I have (that fruit). It looks like a pear but the skin looks like a guava's skin.'

However, the data show that the most common usage of *thì* as a contrastive topic marker is when it follows prepositional phrases. The favored prepositions are *với* 'with', *theo* 'follow', and *còn* 'remain'.[30] Illustrative examples are below.

---

[30] Following Clark (1992a), I consider *theo* 'follow' and *còn* 'remain' as prepositions that are derived from the verbs 'follow' and 'remain'.

51

|       |          | PP |      |      | S |       |     |
|-------|----------|----|------|------|---|-------|-----|
| (68)  | SG96(Q): | **[Theo** | **bác]** | *thì* | [đó | chắc | là |
|       |          | Follow | 1 | PTL | DEM | sure | be |

| đạo diễn | | muốn | truyền tải | một | cái | thông điệp | sống]. |
|----------|---|------|-----------|-----|-----|-----------|--------|
| director | | want | convey | one | CLF | message | life |

| Đó | là | người ta | sống | đơn sơ |
|----|----|----------|------|--------|
| DEM | be | people | live | simple |

| và | người ta | ở | đúng | vị trí | của | mỗi | người. |
|----|----------|---|------|--------|-----|-----|--------|
| and | people | live | exact | position | of | each | person |

'As for me, that probably meant that the director wanted to convey a life lesson. The lesson was that people live a simple life and they stay in their place.'

|       |         | PP |      |      | PP |      |      |      |     |
|-------|---------|----|------|------|----|------|------|------|-----|
| (69)  | NG2(R): | **[Như** | **em]** | *thì*... | **[với** | **góc nhìn** | **của** | **em]** | *thì* |
|       |         | Like | 1 | PTL | with | viewpoint | of | 1 | PTL |

| S |       |     |       |     |      |            |     |         |
|---|-------|-----|-------|-----|------|------------|-----|---------|
| [em | thấy | nó | kiểu... | Em | chưa | hiểu | gì | thật sự!] |
| 1 | see | 3 | like | 1 | yet | understand | what | true |

'Like me...from my viewpoint, I think that it was like...I genuinely do understand anything yet!'

As shown in the examples, the prepositions are mostly followed by the first-person singular pronoun where speakers refer to themselves; doing so helps them express their opinions more clearly, and the use of *thì* implies a contrast with others' opinions.

The table below provides the breakdown of the phrasal *thì* and the corresponding syntactic environments that it tends to occur in.

Table 3. The Tendency of the Phrasal *Thì's* Usage in the Collected Data (total tokens: 48)

| Phrasal *Thì* | Number of Tokens |
|---|---|
| Topic Markers | |
| Topic-Comment Structure | 6 |
| Co-occurrence with Deixis | 8 |
| Presence of Resumptive Pronouns | 4 |
| Contrastive Topic Markers | |
| Occurring with Prepositions | 16 |
| Occurring with other Constituents | 14 |

*3.2.2 Inter-clausal Thì*

3.2.2.1 The Inter-clausal *Thì* in Conditionals

The inter-clausal *thì* is commonly used in conditional sentences, occurring between the conditional clause, also known as the protasis, and the consequent clause, also known as the apodosis. The protasis is optionally marked with subordinating conjunctions, such as *nếu* 'if', *cứ/hễ* 'whenever', *khi* 'when', *giá* 'if only'.[31] The types of conditionals that the inter-clausal *thì* occur in are also diverse. In the first type, *thì* can conjoin conditions where the apodosis is a natural or automatic progression of the conditional premise.

(70)   [S]                                  S
       [Nếu   anh   nhấn   nút   này]   ***thì***   [đèn   sẽ   sáng].
       If     3      push   button   DEM   PTL   light   FUT   light
       'If you push this button, the lights will turn on.'

---

[31] Grammatically, only one of the two needs to be present, but Vietnamese still allows the presence of both the subordinating conjunctions and *thì* in the same utterance.

(71)    [S]                                                    S
        [Nếu    lái    xe    khi    say rượu]    **thì**    [bị    phạt    rất    nặng].
        If      drive  car   when   drunk        PTL       PASS   fine    very   heavy
        'If (you) drive when you are drunk, (you) will be heavily fined.'

As in the above examples, in both cases, *thì* relates the conditional clause with the consequent

clause where the consequent will surely happen as long as the condition is fulfilled. In this type

of condition, *nếu* 'if' is usually replaceable with *hễ* 'whenever' or *khi* 'when'. Alternately, *nếu* 'if'

can be dropped and the main verb in the protasis can be modified by the adverb *cứ* 'just', as

illustrated below.

(72)    [S]                                                    S
        [Khi    anh    nhấn    nút    này]    **thì**    [đèn    sẽ    sáng].
         when   3      push    button DEM    PTL        light   FUT   light
        'Whenever you push this button, the lights will turn on.'

(73)    [S]                                                          S
        [Hễ            lái    xe    khi    say rượu]    **thì**    [bị    phạt    rất    nặng].
        whenever       drive  car   when   drunk        PTL       PASS   penalize very  heavy
        'Whenever (anyone) drives when they are drunk, (that person) will be heavily fined.'

(74)    S                                                      S
        [Cứ     nhấn    nút    này]    **thì**    [đèn    sẽ    sáng].
        Just    push    button DEM    PTL        light   FUT   light
        'Just push this button and the lights will turn on.'

Clark (1992a) also points out that in general, conditionals in Vietnamese only either need the

subordinating conjunctions or *thì*. The above examples can achieve the same conditional effect

without the subordinating conjunctions *nếu* 'if' or *hễ* 'whenever', or the adverb *cứ*. (70) is

equivalent to (75).

(75)    S                                                      S
        [Anh    nhấn    nút    này]    **thì**    [đèn    sẽ    sáng].
        3       push    button DEM    PTL        light   FUT   light
        'You push this button, and the lights will turn on.'

As illustrated in the free translations, the absence of the subordinating conjunctions

communicates the strongest degree of certainty, almost to that of a factual and indisputable

consequence. Furthermore, when the subordinating conjunctions are omitted, the main verb in both the protasis and apodosis clauses can also be modified by adverbs or auxiliary verbs to communicate the high degree of certainty of the conditional bond. These usages become rather idiomatic, as exemplified below.

(76)     [S]                     S
       [(<u>Nếu</u>) *đã*    học]   ***thì***    [không được  lười biếng].
       If     PST   study  PTL    no     able    lazy
       'You already (chose to) study, so you must not be lazy'

(77)     [S]                     S
       [(<u>Nếu</u>) *đã*    làm]   ***thì***    [*phải*  làm    đến    nơi    đến    chốn].
       If     PST   do     PTL    must  do    to     place  to     place
       'You (agreed to) do (this work), so you must do it to the best of your ability.'

(78)                  [S]                     S
       Trẻ con lạ     lắm. [(<u>Nếu</u>) *càng* cấm    nó]   ***thì***    [nó    *càng*  làm].
       Child   strange  very  if      more  forbid  it     PTL    it     more   do

       'Children are very strange. The more you forbid them from doing something, the more likely they will do it.'

The examples (76) and (77) communicate a degree of obligation for the consequent to happen under the condition stated in the protasis. In (78), the speaker utters that the consequent is intensified under the provided condition. In this type of conditional relationship, the most probable outcome is located in the main clause and is almost guaranteed to take place as long as the premise in the condition is met.

      Another type of condition that *thì* can occur in is a hypothetical condition. This condition shows that speakers are unsure if the premise in the conditional clause will take place, but if it does, the consequent in the main clause is a logical and a likely consequence of it.

(79)   Context: A host is talking to a guest.
       [S]                                S
       [Nếu   anh   thích   táo]   **thì**   [ăn   nhiều   táo   nhé]!
       If     3     like    apple   PTL      eat   PL      apple  PTL
       'If you like apples, please eat more!'

(80)   Context: A mom is talking to her child.
       [S]                                S
       [Nếu   trời   mưa]   **thì**   [nhớ       mang   dù        nha   con]!
       If     sky    rain   PTL       remember   bring  umbrella  PTL   child
       'If it rains, remember to bring your umbrella!

(81)   [S]                                          S
       [Nếu   mày   đụng   đến   em        tao]   **thì**   [tao   đánh   mày]!
       If     2     touch  to    brother   1      PTL       1      beat   2
       'If you dare to hurt my brother, I will beat you up!'

A speaker uttering (79) may be unsure of the truth condition of the proposition 'you like apples',

but the speaker offers the invitation anyway in case the condition turns out to be true. Similarly,

(80) and (81) also show such a conditional relationship. As the examples show, the consequent

clause can take on various illocutionary forces, such as invitation, as in (79), advice, as in (80),

and threat, as in (81). The inter-clausal *thì* can occur in this hypothetical condition.

Similar to the previous conditional type, this conditional type also has idiomatic

structures to signal the varied hypothetical conditions, as illustrated below. In this idiomatic

usage, the condition in the protasis and the consequent in the apodosis contain dichotomous

entities. (82) is a representative example of this case.

(82)   Việc   này   (nếu)   **không** sớm   **thì**   muộn   cũng   sẽ     hỏng.
       Work   DEM   if      NEG      early  PTL       late   also   will   break
       'This work, sooner or later, it will not work out.'

This usage is commonly found in literature, as in (83). (83) is a famous poem mocking the

superstition of paying people to hear future predictions about one's life. The person prophesying

the young woman's life in this poem provides obvious answers that will be applicable regardless

of the situation. The conditional structure used in these examples contains dichotomous entities

56

in the conditional clause and the consequent clause, such as *sớm-muộn* 'early-late', as in (82), and *giàu-nghèo* 'rich-poor', *gái-trai* 'male-female', as in (83).

(83)    Số      cô     **chẳng** *giàu* ***thì*** *nghèo*
           Fate   2       NEG   rich   PTL    poor
           'Your fate, if not rich, then poor.'

           Ngày  30     Tết    thịt    treo    trong  nhà
           Day   30     tet     meat   hang   in      house
           'On Lunar New Year's Eve, you'll have some meat in the house (to celebrate).'

           Số      cô     có     mẹ     có     cha
           Fate   2       have   mom   have   dad
           'Your fate, you have parents.'

           Mẹ     cô     đàn bà  cha   cô     đàn ông
           Mother 2     woman  father 2     man
           'Your mother is a woman and your father is a man.'

           Số      cô     có     vợ     có     chồng
           Fate   2       have   wife   have   husband
           'Your fate, you will get married.'

           Sinh   con    đầu lòng     **chẳng** *gái* ***thì*** *trai*
           Birth  child  first-born   no     girl   PTL   boy
           'Your first child, if not a girl, then a boy.'

    Besides being used in hypothetical conditionals, the inter-clausal *thì* can also be used in contrafactual conditionals where the condition in the protasis is almost impossible or impossible to be true. In this usage, there is no clear distinction between a high or low uncertainty level; the apodosis is simply treated as a logical and probable consequence given the condition in the protasis, although recognizing that the condition did not, in fact, occur. Illustrative examples are below.

(84) [S]                                                S
    [<u>Nếu</u>  Nam  học  chăm]     ***thì***  [bây giờ  Nam  đã  có  việc làm].
    if    Nam  study  diligent   PTL   now     Nam  PST  have  job
    'If Nam had studied diligently, Nam would have a job now.'

(85) [S]                                                S
    [<u>Nếu</u>  Nam  đặt  báo thức]   ***thì***  [Nam  đã  không  ngủ  quên].
    If    Nam  set    alarm     PTL   Nam  PST  NEG     sleep    forget
    'If only Nam had set an alarm, Nam would not have overslept.'

The conditions in (84) and (85) exemplify contrafactual conditions that are obviously opposite of

the reality. For instance, in (84), it is presupposed that Nam did not study diligently, and in (85),

Nam did not set any alarm. However, the apodosis clauses in both examples express what the

outcome would have been if the condition had been met. In this fashion, the condition still

logically leads to the consequence even if it is contrafactual.

       Similar to the previous cases of conditionals, besides the usage of *nếu* 'if' in a

contrafactual conditional, *giá* 'if only' is commonly used, as shown below.

(86) [S]                                                S
    [<u>Giá</u>        Nam  học  chăm]  ***thì***  [bây giờ  Nam  đã  có  việc làm].
    If only     Nam  study  diligent  PTL  now    Nam  PST  have  job
    'If only Nam had studied diligently, Nam would have a job now.'

Under the contrafactual conditional, the past tense is normally used in the consequent clause,

which distinguishes it from the hypothetical conditional.

3.2.2.2 The Inter-clausal *Thì* in Sequential and Additional Commentary Usages

Besides the usages of the inter-clausal *thì* in conditional clauses, the inter-clausal *thì* can also connect sequential events and signal relevant additional commentary. Clark (1992a) provides the following example to illustrate the usages of the inter-clausal *thì* in sequential events, referring to them as signaling "an inchoative" relationship.

(87)  (Clark 1992a:101)
  S                                        S
  [Chị   đi     cuối   đường Lê Lợi]     ***thì***   [đi bộ   tới     chợ].
   2     go     end    road  Le Loi      PTL      walk    to      market
  '(When) you get to the end of Le Loi street then walk to the market.'

Besides occurring between independent clauses as Clark documents, the inter-clausal *thì* can also communicate sequentially related events when the clause preceding *thì* is marked with *lúc* 'when' or *khi* 'when'.

(88)  [S]                              S
  [Lúc   mẹ    về]    ***thì***   [tôi   đang   học].
  When  mom   arrive PTL      1     PROG   study
  'When mom came back, I was studying.'

Besides the sequential usage, the inter-clausal *thì* can be used to communicate additional commentary that the speaker wishes to convey in the clause following *thì* in addition to what is uttered in the clause preceding *thì*.

(89)  S                              S
  [Nam   thích  táo]   ***thì***   [chắc   Nam   ăn     táo     trên    bàn].
   Nam   like   apple PTL      sure   Nam   eat    apple   on      table
  'Nam likes apples; he must have eaten the apples on the table.'

In (89), the conjecture that Nam must have eaten the apples is commented based on the speaker's knowledge of Nam's predilection for apples. The presence of *thì* strengthens the relationship between the two clauses, indicating that the additional commentary in the latter clause is made

possible thanks to the information in the former clause. To my knowledge, this thesis is the first to document this usage of *thì*.

### 3.2.2.3 The Tendency of the Inter-clausal *Thì* in the Collected Data

In the collected data, the inter-clausal *thì* is documented to be used in conditionals, to express sequential events, and to provide additional commentary. When used in conditionals, the inter-clausal *thì* is most used in hypothetical conditionals.

(90)    Context: Participants were asked about the relationship between the man harvesting pears and the boy taking the pear basket.

CG62(N):    Chắc chắn    không phải    cha    con    nớ!
            Absolutely    NEG    true    father    son    DEM

            [S]                                S
            [Nếu    là    cha    con]    ***thì***    [phải    đi    cùng    nhau    chứ]
            if    be    father    son    PTL    must    go    together    RECP    PTL

            'Definitely not father and son. If they were father and son, they would have been together.'

(91)    SG31(Q):    Tại vì    thấy    nó    nhìn    lấm lét    á!
                    Because    see    3    look    sneaky    PTL
                    'Because he looked sneaky.'

                    [S]
        SG34(Q):    Chứ    [nếu    nó    là    con    ổng]
                    but    if    3    be    son    3

                    S
                    ***thì***    [nó    phải    kêu    ổng].
                    PTL    3    must    call    3

                    'If he were the man's son, he would have called the man.'

Since participants are unsure of the relationship between the bike boy and the man picking pears, they could only provide hypothetical conditions of their relationship and provide a logical consequence given that hypothetical condition.

Another common usage of the inter-clausal *thì* is the sequential usage, both with the presence of *lúc* 'when' and with independent clauses alone. Statistically, the sequential usage of the inter-clausal *thì* has the highest frequency.

```
                    [S]
(92)   CG23(N):     [Thì³² lúc    đi      xuống],
                    PTL    when   go      down

                          S
                    ***thì*** [mới   nhìn    xuống  ba      cái     rổ       của     mình].
                    PTL    just    look    down   three   CLF     basket  of      self

                    'When he went down, he looked at his three baskets.'


                    S                        S
(93)   SG33(Q):     [Ổng   quay   đi]   ***thì***   [nó    khiêng        đi].
                    3      turn   away  PTL    3      bring          go
                    'The man turned away; the boy took the basket away.'
```

Lastly, the data collection also shows the usage of the inter-clausal *thì* in providing additional commentary, as illustrated below.

```
                    [S]                                    S
(94)   NG98(T):     [Vì        cái     mũ]   ***thì***   ***thì***   [mới    được   tặng   lê].
                    Because    CLF     hat   PTL    PTL    just    able    receive pear
                    'Because of the hat, they were given pears.'
```

The table below summarizes the usage of *thì* in conditional, sequential, and additional commentary cases.

Table 4. The Tendency of the Inter-clausal *Thì's* Usage in the Collected Data (total tokens: 38)

| | Inter-clausal *Thì* |
|---|---|
| Conditionals | 13 |
| Sequential | 21 |
| Additional Commentary | 4 |

---

[32] The sentence-initial *thì* will be discussed in the subsequent section. Therefore, its presence here is not highlighted.

*3.2.3 Sentence-initial Thì*

Since the sentence-initial *thì* is mainly uttered in running speech and documented in the data collection rather than extracted from constructed examples, the majority of illustrative examples in this section comes from the collected data, with only a few constructed examples. For this reason, this section is structured slightly differently from the two previous sections: in the section that details the tendency of usage of the sentence-initial *thì* in the collected discourse, only the number of tokens is provided without illustrative examples. Examples will instead be found in the descriptive section.

3.2.3.1 The Sentence-initial *Thì* Expresses Sequential Events

Similar to the inter-clausal *thì*, the sentence-initial *thì* can conjoin events that happen in succession of one another. This usage of *thì* is most common when speakers recall an incident, retell a story, or verbalize a subconscious thought with the main goal to set up the common discourse ground for the conversation. In the context of the Pear Film oral narratives, speakers most commonly use this towards the beginning of the discourse when they were asked to share with the groups their memory of the movie's storyline.

(95)

| | | S | | | | | |
|---|---|---|---|---|---|---|---|
| SG11(P): | [Nó | bị | đổ-ra]. | | | | |
| | 3 | PASS | drop | | | | |

| | | S | | | | | |
|---|---|---|---|---|---|---|---|
| SG12(P): | *Thì-* | [có | ba | cha | con | ông | này] |
| | PTL | have | three | father | son | man | DEM |

| | | S | | | | | |
|---|---|---|---|---|---|---|---|
| SG13(P): | *Thì-* | [thấy | nó | vậy] | | | |
| | PTL | see | 3 | DEM | | | |

| | | S | | | | | |
|---|---|---|---|---|---|---|---|
| SG14(P): | *Thì-* | [giúp | nó]. | | | | |
| | PTL | help | 3 | | | | |

'He was caused to drop the pear basket. There were three people who were father and sons. They saw him like that. They helped him.'

Example (95) showcases a sequence of events that took place in the movie. Speaker P recalled the scene where the bike boy dropped the pear basket and was helped by the three people, which speaker P assumed to be father and sons. The presence of the sentence-initial *thì* communicates the sequential relationship between these events. The sentence-initial *thì* functions almost like a string that connects the different beads of details in the discourse. In other words, the presence of *thì* not only provides a smoother way of storytelling but it also shows that these strings of events, characters, and settings are meaningfully connected, at least from the speaker's perspective. Since the sentence-initial *thì* is heavily hinged upon pragmatic licensing, it can only be licensed after at least one immediately preceding proposition is uttered.[33]

---

[33] There is one instance that the sentence-initial *thì* might be arguably considered to be using out of the blue. Upon further examination, the particle is used like a placeholder in this case, as in (1).

| (1) NG36(U): | *Thì* | ờ | mình | kể | lại | câu chuyện | như | sau. |
|---|---|---|---|---|---|---|---|---|
| | PTL | uhm | 1 | tell | again | story | like | after |

'Uhm, I retell the story as follows.'

However, this is not really out of the blue, because speakers only started talking when prompted by a question given to the group by the interviewer.

3.2.3.2 The Sentence-initial *Thì* Provides Additional Commentary

In addition to the usage of the sentence-initial *thì* in linking sequential events, it can also be used in front of a clause that contains additional commentary, as illustrated below.

(96) CG44(M):   'There was the boy arriving.'

|            |         |        | S      |           |          |       |        |       |
|------------|---------|--------|--------|-----------|----------|-------|--------|-------|
| CG45(M):   | ***Thì*** | ***thì*** | [anh | bắt đầu | suy nghĩ | nó | có thể | là |
|            | PTL     | PTL    | 1      | start     | think    | 3     | maybe  | be    |

| con trai | của | ổng | hoặc là | một | đứa | bé  | nào  | đó]. |
|----------|-----|-----|---------|-----|-----|-----|------|------|
| son      | of  | 3   | or      | one | CLF | kid | what | DEM  |

'I started to think that he could be his son, or just some boy.'

(97) SG6(P):   'Then there was a boy riding a bicycle; he looked sneaky.'

|          |         | S      |         |     |    |       |     |     |      |      |
|----------|---------|--------|---------|-----|----|-------|-----|-----|------|------|
| SG7(P):  | ***Thì*** | [bác | tưởng | nó | là | con | ổng | ra | lấy | đồ |
|          | PTL     | 1      | think   | he  | be | child | 3   | out | take | stuff |

| phụ  | ổng]. |
|------|-------|
| help | 3     |

'I thought he was the man's son that came to help him.'

Examples (96) and (97) show that participants reference a scene in the movie in CG44(M) and SG6(P), respectively, and add their personal comment or perception of that scene in CG45(M) and SG7(P). Speaker P in (97) relies on the fact that the bike boy approached the man picking pears to provide his opinion that the bike boy could be the man's son.

This usage is also common in everyday speech, as illustrated below, especially when a speaker is commenting on what someone else has just said.

(98)        S

A:   [Gia đình     ấy     có    nhà    to].
      Family  DEM  have  house  big
      'That family has a big house!'

        S

B:   ***Thì***   [cô    ghen tị     à?]
      PTL   2    jealous    PTL
      'Are you jealous of them?'

Furthermore, this type of comment that speakers is usually made when they are in retrospect of some events that already happened or that are factual observations, and the commentary shows the speaker's perception or thought on such events.

(99)        S

A:   [Trời  đang  mưa   lớn    quá!]
      Sky   PROG  rain   heavy  very
      'It is raining so heavily.'

        S

B:   ***Thì***  [tôi  mới  bảo  anh  mang dù       đi   học!]
      PTL  1    just  ask  2    bring umbrella    go   study
      'I asked you to bring an umbrella to school!'

In (99), B's irritated utterance indicates his knowledge of the rain in advance, and he reprimands A for not following his advice. As illustrated in the above examples, the sentence-initial *thì* can be used by multiple speakers.

The data collection also exemplifies this usage when participants start their utterances with *thì* and provide additional commentary on it while referring back to what a different speaker has said, as in (100).

(100)          S

   CG109(M):  [Mình  nghĩ   đây   là    phía        Nam  của  nước Mỹ
               1     think   DEM  be    direction    South  of   America

               miền  Nam   nước Mỹ].
               region South  America

               'I think this is the south of America, southern America.'

   CG110(CM): Tại sao vậy ạ?
               'Why is that?'

               S
   CG111(N):  ***Thì***   [giống        Mexico  hả?]
               PTL    similar     Mexico  PTL
               'It's because it looks like Mexico, is that right?'

Lastly, the commentary usage of the sentence-initial *thì* is very diverse. Indeed, the types of commentary can take on various types of speech acts, as long as the comments have some correlation with the previously uttered proposition. In the series of possible responses as follows, the sentence-initial *thì* can precede commentary expressing various purposes, such as a command (101a), an implicit agreement (101b), and a prediction (101c).

(101)   Context: A and B are talking about the male students in Class C.

               S
      A:    [Con  trai   lớp    C      xấu    quá!]
             CLF   boy   class  C      ugly   very
             'The boys in class C are ugly-looking.'

               S
  a.  B:    ***Thì***  [đừng  nhìn   tụi nó  nữa].
             PTL    stop   look    3       more
             'Don't look at them anymore.'

               S
  b.  B:    ***Thì***  [đâu    ai      nói    tụi nó  đẹp trai].
             PTL    NEG   who   say    3      handsome
             'No one says they are handsome.'

S
    c.  B:    **_Thì_**    [con gái    lớp    C    chắc    thất vọng    lắm].
            PTL    girl    class    C    certainly  disappointed  very
            'Girls in Class C must be disappointed.'

In the collected data, participants also use the sentence-initial *thì* to present a prediction or a

conjecture based on their common knowledge of certain events, similar to (101c). In this usage,

the type of commentary that the sentence-initial *thì* can signal is somewhat conditionally

dependent upon the speakers' common knowledge and judgment.

(102)            S
    SG76(P):    [Ngày  hôm bữa    bác    đi    coi    xe đạp.      Một    xe
                Day    other    1    go    look    bicycle      one    bicycle

                mà    tới    10    mấy    triệu.    Còn    có    xe    70    triệu
                yet    reach  10    some    million  remain have  bicycle 70    million

                xe    100    triệu].
                bicycle 100    million

                'The other day, I went to shop for bicycles. One bicycle could cost 10
                million VND; there were bicycles that were 70 and 100 million VND.'[34]

                S
    SG77(P):    **_Thì_**    [xe    thằng đó    tương đương  với  $3000 là    ít].
             PTL    bicycle CLF   DEM    equivalent    with    $3000 be    little

             'That boy's bicycle must have been equivalent to $3000 USD, and that
             was even considered cheap.'

Speaker P concluded the price of the bicycle that the bike boy used based on the premise that

other bicycles that she saw were also equally expensive. The presence of *thì* in this example can

be understood as 'If what I experience is true, this assumption or educated guess must also be

probable.' In this fashion, this type of commentary using sentence-initial *thì* is similar to the

usage of the inter-clausal *thì* in conditionals. However, whereas the inter-clausal *thì* commonly

signals a conditional relationship, the sentence-initial *thì* can signal more diverse types of

---

[34] 1 USD is equivalent to approximately 22 900 VND. Therefore, 10 million VND, 70 million VND, and 100
million VND are approximately $435 USD, $3048 USD, and $4355 USD, respectively.

relationships, and if it involves a conditional relationship, it is usually implicit, not explicit, as in the above example.

3.2.3.3The Tendency of the Sentence-initial *Thì*'s Usage in the Collected Data

In terms of frequency, the breakdown of the sentence-initial *thì* as expressing sequential events and providing additional commentary is summed up as follows.

Table 5. The Tendency of Usage of the Sentence-Initial *Thì* (total tokens: 69)

|  | Sentence-initial *Thì* |
|---|---|
| Sequential | 39 |
| Additional Commentary | 30 |

## 3.3 Concluding Remarks

This chapter looked at the various usages and syntactic environments of the particle *thì*. The table below summarizes all of the positions and their frequencies of the particle *thì* as documented in the Pear Film oral narratives.

Table 6. The Tendency of Usage of the Particle *Thì* (total tokens: 155)

| Types of the Particle *Thì* | Number of Tokens |
|---|---|
| **Phrasal *Thì*** | |
| Marking Topics | 30 |
| Marking Contrastive Topics | 18 |
| **Sentential *Thì*** | |
| <u>Inter-clausal *Thì*</u> | |
| Signalling Conditionals | 13 |
| Expressing Sequential Events | 21 |
| Providing Additional Commentary | 4 |
| <u>Sentence-initial *Thì*</u> | |
| Expressing Sequential Events | 39 |
| Providing Additional Commentary | 30 |

# CHAPTER 4

## Analysis

### 4.1 Analysis Overview

The particle *thì* is hitherto documented to be used as a phrasal particle and a sentential particle; on the sentential level, *thì* occurs inter-clausally and sentence-initially. The detailed literature review in chapter 2 showed that the phrasal *thì* is most discussed and analyzed in the literature. However, the data findings, as recorded in chapter 3, betrayed that the sentential *thì* is much more prevalent, accounting for 107 out of 155 collected tokens (69%). On the sentential level, *thì* predominantly occurs sentence-initially, accounting for 69 tokens. However, its dominance and functions still need to be accounted for. For this reason, this chapter prioritizes an analysis to capture the underlying linguistic patterns of the sentential *thì*, especially those of the sentence-initial *thì*. Overall, this approach treats *thì* as a polyfunctional particle: *thì* takes on different functions phrasally and sententially. With specific regard to the sentential *thì*, I claim that treating the sentential *thì* as a DISCOURSE MARKER, in the spirit of Fraser (2009), is advantageous. This category not only serves as a flexible super category to capture the syntactic diversity of *thì*, but it also explains the broad spectrum of semantic relationships that the particle can conjoin for its surrounding propositions.

Chapter 4 is organized as follows. Section 4.2 tackles the need to adopt the category of discourse markers; Clark's analysis of the inter-clausal *thì* will be examined more in-depth, and the fundamental distinction between the inter-clausal *thì* and the sentence-initial *thì* will also be teased out. Section 4.3 details the conceptualization of the notion of Discourse Markers (DMs), as proposed in Fraser (2009), and relevant properties of DMs to lay the foundation for my analysis. Section 4.4 details how I apply Fraser's framework to analyzing the sentential *thì*.

Fraser's framework is first applied as is and adapted accordingly to provide a more accurate account of the underlying patterns of *thì* as observed in the collected data. Overall, I found that the sentential *thì* can signal all types of semantic relationships that Fraser documents in his work regarding DMs. However, whereas the category of Elaborative Discourse Markers (EDMs) is too broad and needs further subcategorization, the category of Inferential Discourse Markers (IDMs) must be broadened in order to capture more nuances that the DM *thì* communicates. Lastly, section 4.5 offers some further discussion.

## 4.2 The Need for a Super Category

In 2.3.3, I looked at Clark (1992a) and her analysis of the inter-clausal *thì* as "an inchoative conjunction."[35] I also specified that this term is replaced with sequential in this thesis, for the sake of consistency. As briefly touched on in 2.3.3, Clark's approach is plausible, but it can only account for some instances of sentential *thì*. As evidently shown in the subsequent sections, 4.2.1 and 4.2.2, the inter-clausal *thì* is more than just a sequential conjunction, and the sentence-initial *thì* takes on a broader scope of coverage compared to the inter-clausal *thì*. Consequently, a super category, which is the category of Discourse Markers, is necessary in order to capture the diverse nature of the sentential *thì*.

### 4.2.1 The Inter-clausal thì is mored than an Inchoative Conjunction

It is true that the inter-clausal thì can be treated as a conjunction, and it also signals a sequential relationship; however, the inter-clausal *thì* also commonly occurs in conditional clauses (see 3.2.2.1). In conditional clauses, *thì* is more than just a conjunction that

---

[35] I also highlight that the usage of the term inchoation was inspired by the work of Clark (1988) on the Hmong particle *los*, which, according to Clark, shares great similarity with the particle *thì* in Vietnamese. Clark understands the term inchoation as the act of becoming and translates it into 'then/and then'; when a conjunction takes on "an inchoative function," it backgrounds information in the preceding clause and foregrounds information in the following clause (Clark 1992:106-107). She believes that the inter-clausal *thì* occurs in the syntactic environment of S1_S2, functions as a conjunction, and signals an inchoative semantic relationship between S1 and S2.

communicates a sequential relationship as Clark described. Rather, it can be argued to signal a type of logical relationship, specifically, abductive logic. That is, it signals a scenario where the consequent in the main clause is the most probable result given the premise in the conditional clause. Furthermore, it was also shown that besides the sequential relationship, the inter-clausal *thì* can also signal a cause-and-effect relationship. Evidently, this relationship cannot be housed under the sequential relationship that Clark's position has put forth. To account for these remaining nuances is the goal of this analysis.

### 4.2.2 The Sentence-initial thì is more than the Inter-clausal thì

In chapter 3, I observed many similarities between the inter-clausal *thì* and sentence-initial *thì*. In fact, the subcategories that are proposed in chapter 3 for the inter-clausal *thì* and sentence-initial *thì* are strikingly similar: expressing sequential events and providing additional commentary (see 3.2.2.2). Another similarity is that besides this semantic relationship, the sentence-initial *thì* can also signal a conditional relationship between its utterances, as represented in the below example.

(103)  S                                              S
　　　[Anh  nhấn  nút   này    chưa?]        ***Thì***  [đèn   sẽ    sáng].
　　　 3    push  button DEM   yet           PTL   light  FUT   light
　　　'Have you pushed this button yet? The lights will turn on.'[36]

Lastly, in the function of providing additional commentary, similar to the inter-clausal *thì*, the

sentence-initial *thì* also exemplifies the property of marking a causal relationship.

Despite the similarities, the sentence-initial *thì* takes a larger pragmatic and semantic

scope than does the inter-clausal *thì*. In other words, some semantic relationships between the

surrounding sentences, hereafter S1 (the sentence preceding the sentence-initial *thì*) and S2 (the

sentence following the sentence-initial *thì*), can only be signaled by the sentence-initial *thì*.

Precisely, the sentence-initial *thì* can signal a deductive relationship between S1 and S2 (see

4.4.2), and a contrastive relationship between S1 and S2 (see 4.4.3). Using the inter-clausal *thì* to

signal those relationships would result in illogical utterances. Consequently, the inter-clausal *thì*

and sentence-initial *thì* are distinct.

Given that Clark's current approach is restricted to accounting for the topic-marking

nature of the phrasal and inter-clausal *thì*, a different approach is necessary to analyze the

functions of the sentence-initial *thì*, whose scope of coverage is broader than that of the inter-

clausal *thì*. The proposed category should address the following criteria. The sentential *thì*

exemplifies characteristics of various functional words but is not fixated on any type of syntactic

---

[36] This claim presumes that the question is equivalent to the protasis in the conditional clause 'if you push this button'. In fact, there is evidence that conditional clauses can be rephrased in the form of a question. Haiman (1978) describes the tendency of languages using interrogatives as conditionals. Jespersen (1940) also notices this pattern in Germanic languages, using English as an example. He rationalizes that conditionals can be understood as interrogatives with an implied positive answer. The following conversation illustrates this point of view.
　　(1)  A:　　　Anh　　nhấn　　nút　　này　　chưa?
　　　　　　　　 3　　　push　　button  DEM　　yet
　　　　　　　　'Have you pushed this button yet?'

　　　 B:　　　(nodding yes)

　　　 A:　　　***Thì*** [đèn　　　sẽ　　　sáng]_{S2}.
　　　　　　　 PTL  light　　　FUT　　light
　　　　　　　'Then, the lights will turn on.'

category; therefore, a proposed category should allow for the flexibility to not restrict *thì* to a specific syntactic category. Additionally, the category must be able to capture the various types of semantic relationships that the sentential *thì* can signal. With these in mind, I propose that the category of Discourse Markers as discussed in Fraser (2009) is the super category that can meet these objectives. In the following section, Fraser's conceptualization of this category will be revisted and examined why the category fits to accounting for the properties of the sentential *thì*.

## 4.3 A Proposed Category: Discourse Markers

### 4.3.1 Fraser (2009)

Fraser's (2009) framework of DMs fits the objectives of this analysis because of its coherent structure as a reference framework and its adaptive nature to capture language-specific properties.[37] As for the conceptualization of DMs, Fraser believes that all languages contain a set of markers that are not syntactically licensed, but their presence pragmatically enriches the discourse flow. These markers are categorized under an umbrella term of PRAGMATIC MARKERS, consisting of four types, one of which is Discourse Markers (DMs). Fraser also details the defining properties exclusive to DMs. Firstly, a DM must be a lexical expression.[38] Secondly, all DMs occur between full sentences; he labels these full sentences as S1 and S2. Fraser specifies that DMs occur proposition-initially, so the construction of DMs is as follows: S1—DM+S2. Lastly, DMs only signal a semantic relationship between S1 and S2 without affecting their

---

[37] The notion of discourse markers is not a novel discovery, but rather a widely discussed category. Indeed, Fraser (2009) was not the first work where the term discourse markers was discussed. Prior to his work in 2009, Fraser had been working on this theoretical framework of DMs since the 1980s, refining and shaping the category to be clearer and more comprehensive. Besides Fraser, scholars during his time and before him have also researched this category for years, starting as early as the year 1977 (Labov and Fanshel 1977:156). Fraser lists many different names that discourse markers have been used under, as well as the scope of definitions in works where the term discourse markers is used. See Fraser (2006:932-937) for more details.

[38] Fraser's definition of lexical expressions is quite peculiar. Lexical expressions are defined as pragmatic expressions that are expressed through words, excluding non-verbal cues. The rationale behind this specific property stems from Fraser's disagreement with Schiffrin (1987), who defines DMs as including both non-verbal cues. Following this definition, the particle *thì* is a lexical expression.

internal meanings. The defining properties of a DM are not too far removed from those of pragmatic markers. The below table summarizes the type of basic relationships that Fraser claims DMs can signal for its surrounding propositions.

Table 7. Subtypes of Discourse Markers

| | Contrastive Discourse Markers (CDMs) | Elaborative Discourse Markers (EDMs) | Inferential Discourse Markers (IDMs) |
|---|---|---|---|
| Definitions | A CDM signals a direct or indirect contrast between S1 and S2 | An EDM signals an elaboration in S2 to the information contained in S1 | An IDM signals that S1 provides a basis for inferring S2 |
| Examples | But, alternatively, although, contrary to expectations, conversely, however, etc. | And, above all, after all, moreover, similarly, likewise, etc. | So, as a conclusion, consequently, it follows that, on this/that condition, etc. |
| Example Sentences | A: Harry is hurrying.<br><br>B: **But** when do you think he will get here? | John can't go. **And** Mary can't go, either. | A: I like him.<br><br>B: **So**, you think you'll ask him out, then? |

The subcategories of DMs can communicate a wide range of semantic relationships between S1 and S2. Moreover, DMs may resemble other traditional functional categories (i.e. conjunctions, adverbs, conjunctive adverbs, etc.), but they constitute their own class as they are not fixated on any of those functional categories. Some other relevant properties of DMs are that most DMs can be absent from the S1—DM-X2 construction without heavily hindering the utterance's comprehensibility. Furthermore, as the examples in the table show, DMs can be used by one or multiple speakers, without affecting the felicity of the utterances. Most importantly, DMs are also allowed in construction with the absence of S1 and/or S2, as long as the contextual cues

allow for their pragmatic licensing.[39] Overall, the scope of DMs as proposed by Fraser provides

valuable insights into a framework to account for the properties of the sentential *thì*.

### 4.3.2 The Sentential thì is a Discourse Marker

The sentential *thì* meets all of the qualifying criteria to be a discourse marker under

Fraser's framework: it occurs proposition-initial in the construction S1—*Thì*-S2, and it signals a

semantic relationship between S1 and S2. Notably, since this only deals with the sentential *thì*,

Fraser's notation of S1 and S2 will be adopted to refer to the clause preceding *thì* and the clause

following *thì,* respectively. The only adaptation of the notation is applicable to the sentence-

initial *thì.* It is true that the sentence-initial *thì* can occur preceding S2 without the actual

presence of S1, meaning S1 does not have to be uttered by any speaker. However, this does not

mean that the sentence-initial *thì* does not need a preceding sentence or that it can be uttered out

of the blue. To be more precise, S1, in the case of the sentence-initial *thì*, is better understood as

being implicitly or contextually implied and pragmatically retrievable, thanks to which the

sentence-initial *thì* can be triggered. To capture this subtlety, I adapt Fraser's notation to S1 *Thì*-

S2. The faded S1 represents its syntactically absent but pragmatically conscious status.

## 4.4 The Properties of The Sentential *Thì* as a Discourse Marker

This section applies Fraser's framework and demonstrates how it can elucidate the

underlying linguistic patterns of how the sentential *thì* is used in the collected data of the Pear

Film's narrative. Each subsection is organized as follows. I will start with Fraser's definition of

each category followed by representative examples of the corresponding category as found in the

---

[39] Illustrative examples are provided below.
    (1)  Context: John, on seeing his roommate walk in smiling.
          John: *So* you aced the exam!                  (Fraser 2009:300)
    (2)  A: John will not take his medicine.
          B: *And*? [Implied X2: What do you want me to do about it?]    (Fraser 2009:300)

collected data. From there, I will address aspects of Fraser's classification that should be adapted in order to capture the specific nuances of the sentential *thì* more accurately, especially subtypes of the semantic relationships that only the sentence-initial *thì* could signal, propose the adapted subcategories, and explain how they are portrayed through the data.

## 4.4.1 The Sentential Thì as an Elaborative Discourse Marker (EDM)

Fraser's definition of elaborative discourse markers (EDMs) is as follows: "[A]n EDM signals an elaboration in S2 to the information contained in S1" (Fraser 2009:301). This definition provides an overarching spectrum of the different types of elaboration that DMs can signal. Following the broad definition, I find the following tokens of the sentential *thì* to be representative of this category.

(104) CG27(N):
[Lúc       ông   đang   đứng   phân vân
When       3     PROG   stand  hesitate

nghi ngờ       không          biết   ai     trộm]$_{S1}$,
suspect        NEG            know   who    steal

***thì-***   [có      ba      đứa    trẻ    mà     lúc     nãy      giúp
PTL          have    three   CLF    kid    REL    when    earlier  help

thằng   cu     ăn trộm     lúc     bổ     xe        đó]$_{S2}$.
CLF     boy    steal       when    fall   bicycle   DEM

'When he was befuddled and unsure of who stole (his fruits), there came the three boys that helped the thief earlier.'

In this example, *thì* elaborates that during the moment when the man was confused, the three boys appeared. The context for the utterance took place when the participant was retelling the story in the interview. Therefore, N naturally listed a temporal frame and provided elaborative information that took place during that instant, using the chronological development of the movie as the overall context for his utterance. This usage usually contains the subordinating conjunction *lúc* 'when' or *khi* 'when' as part of S1. The presence of these subordinating conjunctions, along

with the DM *thì*, enhances a stronger temporal binding between the action/event in S1 to that in S2.

(105)  SG16(P):      [Ba    cha    con    đi    một    lúc]$_{S1}$  ***thì***    [thấy    cái    mũ    rớt]$_{S2}$.
                    Three  father  son  go    one    period  PTL      see    CLF    hat    drop
                    'The father and sons were walking for a bit; they saw a hat.'

Despite the absence of *lúc* 'when' and *khi* 'when', the presence of the inter-clausal *thì* still signals the elaborative relationship between the information stated in S2 onto that provided in S1.[40] Specifically, in (105), after presenting that the three people were leaving the bike boy to go in the opposite direction, participant P added that the three people spotted the hat, which was dropped by the bike boy in the immediately preceding scene. Except for the assumed interpretation that the three people are father and sons, participant P only recalled his memory of

---

[40] Section 3.1.3 emphasized that the classification of tokens like this is decided through prosodic cues. Under that premise, this token meets the criteria to be classified as a token of inter-clausel *thì* because there was no obvious pause between *thì*'s surrounding clauses. Understandably, this subjective classification may invite the question of whether the sentence-initial *thì* actually exists in Vietnamese. The answer is firmly positive. It is critical to point out that the subjective classification would, if any, only affect the counting of the tokens. In other words, the dominant prevalence of the sentence-initial *thì* (69 tokens) as opposed to the inter-clausal *thì* (38 tokens), as chapter 3 reported, would be the only factor affected due to the auditory test. However, this alone cannot invalidate the existence of the sentence-initial *thì*. In fact, I find cases where speakers start their utterances with the sentence-initial *thì* as a continuation for their own assertion to refer to a question made by the interviewer, or where other speakers interrupt by starting their own utterance with the sentence-initial *thì,* as shown below. Furthermore, when speakers recall a series of events, it is also common to see them provide consecutive utterances starting with the sentence-initial *thì*. In this case, even though all utterances are said by one speaker, the consecutive presence of the sentence-initial *thì* is undeniable.

SG11(P):    Nó        bị        đổ-ra.
            3 PASS    drop

SG12(P):    ***Thì***    có        ba        cha        con        ông        này.
            PTL        have      three    father    son        man        DEM

SG13(P):    ***Thì***    thấy      nó        vậy.
            PTL        see        3          that

SG14(P):    ***Thì***    giúp      nó.
            PTL        help      3

            'He fell, there were three people who were father and son they saw him like that
            they helped him.'

Future research on this particle can strengthen the statistical findings through the means of scientific tools, such as Praat to guarantee the objectivity of the classificational method.

the movie and gave additional details that adhere to the movie's chronological order. In other words, in participant P's mind, the detail 'they saw the dropped hat' is the successive action to the action in S1, which is 'The father and sons were walking for a bit'; the participant signals this relationship by using the inter-clausal *thì*.

In (106) below, a similar elaborative relationship is observed, and it is shown that the sentence-initial *thì* can be used multiple times with different S2s, as long as the elaborative information is concerned. Specifically, this example contains two S2s: the first token of *thì* elaborates in S2a that out of the three baskets mentioned in S1a, one basket is full, and the second token of *thì* signals that S2b is a successive event to S1b.

(106)  CG9(N):  [Có    ba     cái    rổ]S1a.
                 Have   three  CLF    basket
                 'There were three baskets.'

       CG10(N): [S1a  **Thì-**  [có    một    rổ      đầy]S2a]S1b.
                      PTL     have   CLF    basket  full
                 'There was a full basket.'

       CG11(N): S1b   **Thì-**  [thằng cu    nớ     coi    như    là     nhìn
                      PTL     CLF    boy   DEM    look   like   be     look

                 liếc    liếc    rồi    trộm    đi     luôn]S2b.
                 glance  glance  then   steal   go     all

                 'That boy, like, looked and he glanced glanced and he stole (the one full basket) and left.'

Noticeably, the types of elaborative relationships in these S2s are slightly different. Whereas the sentence-initial *thì* in CG11(N) signals that S2b chronologically follows the event in S1b, the sentence-initial *thì* in CG10(N) simply specifies an additional detail about the fruit baskets mentioned in CG9(N). In other words, the elaborative relationship in CG9(N) and CG10(N) is not sequential, but merely descriptive. A clearer example of this type of elaborative relationship is also observed in (107) below.

(107) CG2(N):   [Trong  cái    video  nớ,    giống  như   kiểu
                In      CLF    video  DEM    similar like  type

                là    ở      nước   ngoài  á]$_{S1}$.
                be    in     country foreign PTL
                'In the video, it looked like the setting was in a foreign country.'

      CG3(N):   S1   ***Thì-***  [có    một    cái    ông    nớ     thu hoạch  trái cây]$_{S2a}$.
                           PTL    have   one    CLF    man    DEM    harvest    fruit
                'There was a man harvesting fruits.'

      CG4(N):   S1   ***Thì-***  [người họ    đi   qua  đi    về     cái  cánh đồng]$_{S2b}$.
                           PTL    person 3     go   past go    back   CLF  field
                'There were people going back and forth towards the field.'

      CG5(N):   S1   ***Thì-***  [có    một    cái ông dắt   dê     đi     qua]$_{S2c}$.
                           PTL    have   one    CLF man walk  goat   go     past
                'There was a man walking a goat passing by.'

As illustrated in (107), all three tokens of *thì* refer to the same S1, which is CG2(N). Similar to

the previous tokens of *thì* explored in this section, these tokens of *thì* also signal an elaborative

relationship between S1 and S2. Specifically, participant N was asked to recall the storyline of

the movie, which laid the foundation for CG2(N). After uttering CG2(N), the participant gave

further details regarding what happened in the video, which are signalled by the three tokens of

*thì* in CG3(N), CG4(N), and CG5(N). However, unlike the elaborative relationship observed in

the second token of *thì* in (106), these tokens of *thì* do not signal a sequential elaboration

between S1 and S2. This claim is plausible because all three tokens of *thì* are triggered by the

same S1, which is CG2(N), not by their immediately preceding utterances. In other words, in a

true sequential elaborative relationship, each token of *thì* should be pragmatically triggered by its

immediately preceding proposition. Even in cases of multiple occurrences, as in (108) below, all

tokens of *thì* refer back to their corresponding immediately preceding proposition.

(108) SG11(P):      [Nó    bị     đổ-ra]s1a.
                    3      PASS   drop


      SG12(P):      S1a    [**Thì-**   [có    ba      cha     con    ông    này]s2a]s1b
                           PTL    have   three   father  son    man    DEM

      SG13(P):      S1b    [**Thì-**   [thấy   nó      vậy]s2b]s1c
                           PTL    see    3       that

      SG14(P):      S1c    **Thì-**    [giúp   nó]s2c.
                           PTL    help   3

'He was caused to drop (the pear basket). There were three people who were father and son. They saw him like that. They helped him.'

This example shows a sequential progression from the events in SG11(P) to SG14(P); specifically, the sequence starts from where the boy dropped the fruit to when the three people appeared and helped him pick up the basket. Since the sentence-initial *thì* helps to elaborate on sequential details in the story, each token of *thì* refers back to its immediately preceding proposition, and multiple S1s and multiple S2s are present. On the other hand, in a case such as (112), all tokens of *thì* are triggered to elaborate on the detail that participant N uttered in CG2(N) 'In the video, it looks like the setting is in a foreign country.' Since all tokens are licensed by CG2(N), imposing a reading where CG4(N) is contextually licensed by CG3(N) would result in an infelicitous utterance.

Given these nuanced differences in the types of elaborative relationship that the DM *thì* can signal, it becomes necessary that Fraser's category of EDM should be further specified. I propose the following adaptations: The category of Elaborative Discourse Markers will consist of two subcategories, Sequential EDM and Non-sequential EDMs. As a sequential EDM, the DM *thì* signals that S2 chronologically follows S1. As a non-sequential EDM, the particle *thì* signals that the speakers wish to give additional details about what is said in S1; therefore, even though S1 and S2 can be contextually

81

bound, S2 does not necessarily immediately follow S1.[41] With the adapted subcategories and

definitions, I find 60 tokens that fall under the category of EDMs. Both the inter-clausal *thì* and

sentence-initial *thì* can take on this function. Among 60 tokens, 21 tokens (35%) are the inter-clausal

*thì*, and 39 tokens (65%) are the sentence-initial *thì*. The table below details the number of tokens that

are Sequential EDMs and Non-sequential EDMs.

Table 8. Tokens of *Thì* as EDMs (total tokens: 60)

|  | Inter-clausal *thì* | Sentence-initial *thì* |
|---|---|---|
| Sequential EDMs | 15 | 31 |
| Non-sequential EDMs | 6 | 8 |

### 4.4.2 The Sentential Thì as a Logical Discourse Marker (LDM)

4.4.2.1 Fraser's Framework of IDMs and Appropriate Adaptations

Fraser's definition of inferential discourse markers (IDMs) is as follows: "[A]n IDM

signals that S1 provides a basis for inferring S2" (Fraser 2009:301). In contrast to EDMs,

Fraser's definition of IDMs provides a narrower lens to capture the relationship between S1 and

---

[41] There is one instance that a token of EDM *thì* could also be classified as signalling a logical relationship. However,
I believe that this token fits better as a Sequential EDM given the proposed definitions.

| CG20(N): | [Rứa | là | mới | là | cầm | mũ | trả | lại |
|---|---|---|---|---|---|---|---|---|
| | So | be | just | be | hold | hat | return | back |

| | cho | thằng | cu | lái | xe | trộm | hoa quả | trái cây]S1. |
|---|---|---|---|---|---|---|---|---|
| | for | CLF | boy | drive | bicycle | steal | fruit | fruit |

'So the other boy brought the hat back to the boy stealing fruits.'

| CG21(N): | S1 | **Thì-** | [thằng | cu | mới | cảm ơn | bằng | cách | đưa |
|---|---|---|---|---|---|---|---|---|---|
| | | PTL | CLF | boy | just | thank | by | method | give |

| | cho | mỗi | đứa | một | quả | trái cây]S2. |
|---|---|---|---|---|---|---|
| | for | each | person | one | CLF | fruit |

'The boy thanked the other boys by giving them each a fruit.'

S2. A representative example of this category is provided below. The IDM *thì* in investigation is

bold and italicized.

(109)  NG104(T):  [Có biết ông ấy đang ở trên đấy]s1
                   Yes know he DEM PROG on up DEM

                   ***thì*** [mới dám lấy]s2.
                   PTL just dare take

                   'He must have known the man was up [on the tree]. Only then did he
                   take [the fruit basket].'

The earlier context of this example shows that participant T tries to argue that the bike boy is a

thief who stole the fruit basket from the man picking pears. One of the reasoning for this

overarching claim is uttered in NG104(T). In (109), the participant conjectures that the bike boy

schemed his plan to steal the fruit basket by observing that the man is high up in the tree. In this

case, the basis for the conclusion that the boy dared to take the fruit basket is inferred based on

the basis of the boy's awareness of the location of the man. None of these details is provided in

the movie's storyline; therefore, both the conclusion and the basis that the participant proposes

relies strictly on his own inference, given the objective details in the movie. This utterance is also

triggered by a question that invites participants to interpret ambiguous details, specifically, the

relationship between the bike boy and the man, rather than recalling the storyline.

Similarly, example (110) also showcases a representation of how the sentence-initial *thì*

can be used as an IDM.

| (110) | SG38(P): | [Ba | người | đó | là | cùng | một | ấy | ra]s1. |
|---|---|---|---|---|---|---|---|---|---|
| | | Three | person | DEM | be | same | one | that | out |

| | SG39(P): | S1 | **_Thì-_** | [bác | nghĩ | là | ba | cha | con]s2 . |
|---|---|---|---|---|---|---|---|---|---|
| | | | PTL | 1 | think | be | three | father | son |

'Those three appeared at the same time; I thought that they were father and sons.'

Participant P concludes that the three people are father and son, based on the basis that they entered the scene together. The inferential relationship is clear because the participant refers to an objective detail in the storyline, which is the co-appearance of the three boys, to serve as a basis for her claim that the three people have familial relationships. With the same basis, other participants do not reach a similar conclusion, which further confirms the nature of the DM _thì_ as an IDM. So far, the most notable difference between EDMs and IDMs lies in the fact that IDMs signal conjectural conclusion and/or assertion in S2, given a certain basis provided in S1.

Generally speaking, the category of IDMs communicates that S1 and S2 are logically related. Specifically, the presence of the particle _thì_ as an IDM signals the speaker's intention to draw a logical connection between the information in S1 and S2, an example of which is that S2 is deduced or S2 is an effect of the reasoning provided in S1. Even though the current definition and scope of Fraser's category of IDMs can elucidate some underlying patterns to this type of semantic relationship that the DM _thì_ can signal, the collected data findings show that the sentential _thì_ can conjoin a broader spectrum of logical relationships, as illustrated below.

(111) CG58(O):  [Tại vì  nếu như   mà     là    cha   con]s1,
                 Because  if        like   PTL    be    father son

                ***thì***   [lúc   sau    lúc    ông   ấy    nhìn   thấy   cái    rổ
                 PTL        when   after  when   3     that  look   see    CLF    basket

                không còn   ấy,   ông   ấy    không bất ngờ   đâu]s2.
                 NEG   remain that  3     DEM   NEG   surprise  PTL

                'Because if they were father and son, after when he saw that the fruit
                basket was gone, he would not have been surprised.'

In this context, participant O claimed that the bike boy and the man are not family members and

was questioned why. CG58(O) is her reasoning for that overarching claim. To rationalize her

claim, she provided the hypothetical condition under which the bike boy was the man's son and

claimed that if that condition were true, the man would not be surprised when he saw that one

basket was lost. With specific focus to the rationalization of this sub-point, speaker O relies on

her abductive reasoning to argue for her stance. That is, the assumption she arrives in S2 is most

probable given the information she has access to, which is presented in S1. The particle *thì* in

this example signals that type of a logical relationship. In this fashion, (111), as well as other

conditional relationships of *thì* in the collected discourse, can be more uniformly classified as

demonstrating abductive reasoning. This logical relationship is different from inferential logic,

which is Fraser's category of IDMs.

Additionally, besides the two logical relationships addressed above, the DM *thì* on the

sentence-initial position also examines another type of logical relationship, as illustrated in (112).

(112) CG68(M):     [Anh  nghĩ  trái  đó   là    trái    lê]s1.
                   1     think fruit DEM  be   fruit   pear
                   'I think that fruit is pears.'

     CG69(CM):     Vì sao ạ?
                   'Why?'

     CG70(M):   S1   *Thì*   [ở     Nga  có    nhiều trái    giống  như    vậy]s2.
                            PTL   at    Russia have PL     fruit   similar like   that
                   'Because, in Russia, there are fruits similar to that fruit.'

The logical relationship between S1 and S2 signals a reverse logic to the inferential logic that

Fraser proposes: Whereas IDMs signal that "S1 provides a basis for inferring S2", example (112)

reveals that the DM *thì* can also signal that S2 provides a basis for inferring the conclusion

and/or assertion in S1. In other words, it can be understood that participant M used deductive

reasoning. That is, he starts with his knowledge of Russian pears to deduce that the fruit in

question is pears. This logical relationship is different from the inferential relationship that Fraser

proposes, especially with the presence of the follow-up question "Why?"

These two representative examples show that the scope of logical relationships that the

DM *thì* can take on is broader than that proposed in the category of IDM. Therefore, I propose

the following adaptations. First, the umbrella category that houses all three types of logical

relationships that have been observed so far will be coined Logical Discourse Markers (LDMs).

Second, as the representative examples above elucidate, the particle *thì* can signal three types of

logical reasoning: inductive, abductive, and deductive. The subcategory of IDM in Fraser's

framework fits that of inductive reasoning; therefore, IDM will be renamed into Inductive LDMs.

Third, the logical relationship observed in the cases when *thì* occurs in conditionals represents an

abductive logical relationship. Hence, the second subcategory is called Abductive LDMs. This

leaves us to the last adaptation of the third subcategory, the logical relationship of which is

illustrative in (112), as Deductive LDMs. With the newly proposed categories, I present the

definitions of LDMs and its subcategories as follows.

Table 9. Proposed Definitions of LDMs

| Logical Discourse Markers | LDMs signal a logical relationship between S1 and S2 in which speakers propose their arguments through rules of inference. |
|---|---|
| Inductive LDMs | Inductive LDMs signal that S2 is inferred based on the premise(s) in S1. |
| Abductive LDMs | Abductive LDMs signal that S2 is the likeliest and most probable conclusion given the condition provided in S1. |
| Deductive LDMs | Deductive LDMs signal that S2 provides the premise(s) to deduce the conclusion in S1. |

With the adapted definitions and proposed categories, the causal relationship that the DM *thì* can

signal can be characterized and represented with more nuances. The adaptations also helps to

account for the distinction between the inter-clausal *thì* and sentence-initial *thì*. More specifically,

whereas the sentence-initial *thì* display all types of LDMs, the inter-clausal *thì* can only be used

as an Inductive LDM and an Abductive LDM. Enforcing such a semantic relationship in an

utterance containing the inter-clausal *thì* yields an infelicitous response, as provided below.

(113)  #[Tôi  mua  táo]$_{S1}$  ***thì***  [tôi  thích  táo]$_{S2}$.
      1  buy  apple  PTL  1  like  apple
      'Intended meaning: I bought apples because I like apples.'

(114)  [Tôi  thích  táo]$_{S1}$  ***thì***  [tôi  mua  táo]$_{S2}$.
      1  like  apple  PTL  1  buy  apple
      'I like apples, so I bought apples.'

Example (113) confirms that the inter-clausal *thì* cannot be used to signal that S2 provides the

premise(s) to deduce S1 (Deductive LDM), whereas it is perfectly acceptable when used as an

Inductive LDM, as in (114). In conclusion, then, the adapted category of Fraser's IDM is now

Logical Discourse Markers (LDMs), with 3 subtypes: Abductive LDMs, Inductive LDMs, and Deductive LDMs. The first two are observed in both the inter-clausal *thì* and sentence-initial *thì*, whereas the third one is an exclusive semantic relationship that only the sentence-initial *thì* can signal for S1 and S2.

4.4.2.2 Ambiguous LDMs Cases

As mentioned above, the boundary between EDMs and LDMs can be unclear. The collected data contain two tokens that are arguably interpretable as either EDMs or LDMs. However, there is evidence to treat the first token of the DM *thì* is an LDM, but the second token should remain as an EDM. A representative example is provided below. To represent the two tokens of *thì* in one sentence more consistently, the brackets and S1_S2 environment of the first token of *thì* are in bold and those of the second token are italicized.

(115)  NG90(T):   **[**Cái     nón     hình như     đúng     của     bạn     đạp     xe đạp**]S1a.**
                    CLF     hat     maybe         true     of     friend     paddle     bicycle
                    'The hat was probably was truthfully of the boy riding the bicycle'

NG91(T):   S1a   **[*Thì₁-*** *[*ba     cậu     nhặt     hộ*]S1b* ***thì₂***     *[*mới     được     lê*]S2b*]S2a.
                    PTL     three     CLF     pick     stead     PTL     just     able     pear
                    'The three boys picked the hat up and they got the pears.'

In this context, the participants were arguing to recall if the dropped hat belongs to the bike boy or the bike girl. The ambiguity can yield the following alternative reading: the first token of the DM *thì* (*thì₁*) is a Deductive LDM and the second token (*thì₂*) is an Inductive LDM because (1) the first token of *thì* signals that S1a is concluded based on S2a, and (2) the second token signals that S1b provides a basis for inferring S2b. It is reasonable to argue that the first token of the DM *thì* is a Deductive LDM but the second token of *thì* cannot be an Inductive LDM. Firstly, it is true that most Deductive LDMs are triggered in the collected discourse through an intervening "Why?" question. However, even without the "Why?" question, the first token of *thì* in NG91(T)

still qualifies as a Deductive LDM because (1) it signals a logical relationship between S1a and S2a, and (2) it signals that the subjective interpretation of participant T in S1 regarding the hat is deduced based on the detail provided in S2a. However, the second token of *thì* does not exemplify a logical relationship because S1 and S2 is sequentially related and the basis for the overarching hypothesis is signaled through the Inductive LDM. As the scope of S1b and S2b indicated in the example, the second token of *thì*, the inter-clausal *thì*, signals a semantic relationship between the detail 'The three boys picked up the hat' [S1b] and 'They got the pears' [S2]. If that token of *thì* were an Inductive LDM, its S1 should be S1a, which is not attested.

With the expanded categories and more nuanced subdivision of subcategories, I found 44 tokens of *thì* used as LDMs. Out of 44 tokens, 17 tokens (39%) are the inter-clausal *thì* and 27 tokens (61%) are the sentence-initial *thì*. The tokens of *thì* with respect to each subcategory of LDMs is provided in table 10 below.

Table 10. Tokens of *Thì* as LDMs

|  | Inter-clausal *thì* | Sentence-initial *thì* |
|---|---|---|
| Abductive LDMs | 13 | 5 |
| Inductive LDMs | 4 | 16 |
| Deductive LDMs | 0 | 6 |

### 4.4.3 The Sentential Thì as a Contrastive Discourse Marker (CDM)

Unlike EDMs and LDMs, the sentential *thì* is only marginally used as a CDM. Fraser's definition of CDMs is as follows: "CDMs signal a direct or indirect contrast between S1 and S2" (Fraser 2009:301). The collected data only have 3 tokens of the sentential *thì*, specifically, the sentence-initial *thì*, as a CDM. A representative example is presented below.

(116)  SG82(P):     [Nói    túm lại              ở đây   là     như    khoa học            giả tưởng
                     Speak   in general           here    be     like   science             fiction

                     vì                 ở       Việt Nam    là     không   có     cho    đi
                     because            in      Vietnam     be     NEG     yes    let    go

                     một     cách       dễ dàng        vậy    đâu]s1.
                     one     method     easy           that   NEG

                     'In summary, it sounded like a science-fiction because in Vietnam; there
                     is no way that they could go just like that.'


       SG83(Q):     S1    *Thì-*    [nó     truyền tải     cái    thông điệp     như    vậy.
                           PTL       it      convey         CLF    message        like   that

                     mình    cứ      giúp đỡ            người ta     bất kể        nguồn gốc
                     2       just    help              people       regardless    origin

                     của     người ta]s2.
                     of      people

                     '(But) it conveyed that type of message. We just need to help other people
                     regardless of their background.'

In this context, participants were asked to share the morale of the movie if they perceive there is

any. Participant P believed that the depiction of kindness in the movie is rather fictional and

referred to her own experience in her native country that thieves can never get away with bad

deeds as easily. To defend the message of the movie, participant Q used the CDM *thì* to signal

that the movie intended to convey such a message, despite the possible push back as P stated.

Similar to the subcategory of Deductive LDMs, the category of CDM is exclusively observed in

the case of the sentence-initial *thì*. Therefore, even though this usage of *thì* is extremely limited,

the category of CDMs allows us to make a finer distinction between the function that is only

exclusive to the sentence-initial *thì*.

## 4.5 Discussion

### 4.5.1 Analysis Concluding Remarks

By adopting and adapting Fraser's framework of Discourse Markers, this thesis is able to analyze the underlying properties of the sentential *thì* in the collected discourse. In 4.2, I outlined a few criteria that the super category should be able to address. Firstly, the category should be able to capture the various semantic relationships exemplified by the sentential *thì*. Accordingly, I showed that the category of DM allows for such flexibility; through my adaptation of Fraser's framework, the three DM categories, Elaborative DMs, Logical DMs, and Contrastive DMs, can capture the types of semantic relationships that the sentential *thì* signals in the collected data. Nevertheless, the framework also illustrates the nuanced differences between the types of semantic relationships that only the sentence-initial *thì* can capture. The table below gives a recap of the scope of semantic relationships that the inter-clausal *thì* and sentence-initial *thì* can signal.

Table 11. The Scope of Semantic Relationships Signaled by Inter-clausal and Sentence-initial *thì*

| | Inter-clausal *thì* | Sentence-initial *thì* |
|---|---|---|
| Elaborative Discourse Markers (EDMs) | | |
| Sequential EDMs | ✓ | ✓ |
| Non-sequential EDMs | ✓ | ✓ |
| Logical Discourse Markers (LDMs) | | |
| Abductive LDMs | ✓ | ✓ |
| Inductive LDMs | ✓ | ✓ |
| Deductive LDMs | ✗ | ✓ |
| Contrastive Discourse Markers (CDMs) | ✗ | ✓ |

The table details that whereas the sentence-initial *thì* can signal all types of semantic relationships under three categories, EDMs, LDMs, and CDMs, the inter-clausal *thì* cannot signal a deductive relationship and a contrastive relationship between S1 and S2. Advantageously, the adapted framework of DMs from Fraser (2009) can accurately describe possible semantic relationships that the sentential *thì* can communicate as well as the specific scope of semantic relationships that inter-clausal *thì* and sentence-initial *thì* can signal. Besides providing a conceptual understanding of how the sentential *thì* is used in the collected data, this framework also presents an understanding of the statistical distribution of the sentential *thì* in this data. Table 12 shows the overall distribution of the types of DMs that *thì* can take on.

Table 12. Distribution of the DM *thì* (Total tokens of DMs *thì*: 107 tokens)

| | |
|---|---|
| Elaborative Discourse Markers (EDMs) | 60 tokens (56%) |
| Logical Discourse Markers (LDMs) | 44 tokens (41%) |
| Contrastive Discourse Markers (CDMs) | 3 (3%) |

Evidently, the particle *thì* is most used as an EDM or LDM, and is least used as a CDM. Indeed, in the data collection, participants mostly started with the EDM *thì* to establish the context of the conversation, which is usually the storyline of the Pear Film movie. Then, when specific questions regarding the different scenes and characters are raised, participants resort to using the LDM *thì* to make logical connections between their claims or assertions to the details in the movie and/or common knowledge. When participants must contrast or advocate for their positions, they use the CDM *thì* to communicate the contrastive relationship.

Lastly, the specific tokens of each category with respect to the inter-clausal *thì* and sentence-initial *thì* is detailed in table 13.

Table 13. The Distribution of inter-clausal *thì* and sentence-initial *thì*

| | Inter-clausal *thì* | Sentence-initial *thì* |
|---|---|---|
| Elaborative Discourse Markers (EDMs) | | |
| Sequential EDMs | 15 | 31 |
| Non-sequential EDMs | 6 | 8 |
| Logical Discourse Markers (LDMs) | | |
| Abductive LDMs | 13 | 16 |
| Inductive LDMs | 4 | 5 |
| Deductive LDMs | 0 | 6 |
| Contrastive Discourse Markers (CDMs) | 0 | 3 |

This table shows how the functions of *thì* as described in chapter 3 are now recast into a more systematic classification. Whereas the rudimentary classification in chapter 3 offered a general understanding of how the particle *thì* can be used, the framework of Discourse Markers presents a more unified understanding of the specific types of semantic relationships that the inter-clausal *thì* and sentence-initial *thì* can take on, as well as the tendency of its usages.

### 4.5.2 The Historical Development of Thì

The current analysis of *thì* as a Discourse Marker provides a systematic look into the underlying usages of *thì*. As established, the syntactic environments that *thì* occurs in (phrasal, inter-clausal, and sentence-initial) are diverse. Its functions, as analyzed, are extremely versatile. This reality invites further consideration regarding the origin of the particle and raises questions such as how *thì* took on these functions, what functions it took on first, and why its usages are so heavily discouraged in formal communication. This section provides a brief look into the historical development of *thì* and documents some preliminary findings to examine this aspect of

the particle. These observations lay a strong foundation to further the understanding of the particle in Vietnamese.

Upon further examination of historical evidence, it is plausible to claim that the particle *thì* was first used as a noun: it is reported to be a loan word from the Chinese word 时 shí, as in the noun 时间 'time' (Hurinville 2009:3). Illustrative examples of *thì* as a noun are provided below: (117) is provided by Hurinville, and (118) is constructed.

(117)  (Hurinville 2009:3)
| ***Thì*** | giờ | là | vàng | bạc. |
|------|------|------|------|------|
| Time | hour | be | gold | silver |

'Time is precious.'

(118)
| Tức | ***thì***, | có | con | chim | bay | đến. |
|------|------|------|------|------|------|------|
| Immediate | time | have | CLF | bird | fly | approach |

'That immediate moment, there is a bird approaching.'

However, the usages of *thì* as a noun have decreased significantly. In fact, *thì giờ* 'time' is mostly substituted by *thời gian* 'time', and *tức thì* 'immediate moment' is normally replaced by *ngay lập tức* 'immediately.' The decrease in the usage of *thì* as a noun[42] gives rise to the possibility that historically, *thì* has been grammaticalized, which explains the status quo of *thì* as a functional word in modern Vietnamese. The progression of the grammaticalization process of *thì* is also important: does the particle *thì* takes on the topic and contrastive topic marking function as a phrasal particle first, or does it develop its function as a discourse marker on the sentential level first? In other words, it is worthwhile to ask whether the progression was linear, or whether both functions of *thì* were concurrently developed. When examined through historical texts, *thì*'s usage matched what was found in the collected data: the particle *thì* has mainly been used on the sentential level. More specifically, the usage of the sentence-initial *thì* was extremely dominant, even in written texts. The first evidence comes from the book *Thầy Lazaro Phiền*

---

[42] This means that this usage of *thì* as a noun is still permitted but considered rather outdated.

'Lazarus Phiền'[43] by author Nguyễn Trọng Quản (Quan Nguyen). Interestingly, within the book, 259 tokens of *thì* are found, 226 of which (87%) are the sentential *thì*, whereas the phrasal *thì* only accounts for 33 tokens (13%). Upon further research, it was made clear that the author intentionally used colloquial language in his book to present the most realistic depiction of the spoken Vietnamese language at the time.

From this finding, I was intrigued to find the distribution and usage of the particle *thì* in solely written texts. The earliest solely written work written in modern Vietnamese that is publicly accessible is the Bible. Therefore, all tokens of *thì* in the book of Mark as present in four translated version of the Bible were counted. The first version of the Bible in Vietnamese was done in 1925,[44] and the most recent version of the Bible was published in 2011. The book of Mark fits my area of research because of its dynamic actions and diverse scenarios. Comparing the four versions provides a clearer and more gradual change in the usages of *thì*. The result is as follows:

---

[43] This book was the first book to be published in Chữ Quốc Ngữ, also known as The National Alphabet, which is what modern Vietnamese is written in. It is a Latin-based alphabet developed by evangelicals in the 13th-14th centuries. Cao (1991/2004) and Clark (1992) claim that the particle *thì* has already been used even back when Vietnamese people still used the writing systems of *chữ Hán*, the Chinese characters, and *chữ Nôm*, the use of Chinese characters only for sounds while Vietnamese syntax and other linguistic aspects were preserved. Due to my limited command in ancient Chinese and my inability to read *chữ Nôm* (the access to works written in these alphabets are also extremely limited), I resort to analyzing works written in *Chữ Quốc Ngữ*.

[44] Interestingly, the first translated Bible was done by Protestants, not Catholics even though Catholic missionaries arrived in Vietnam much longer than Protestants did. The translated Bible into *Chữ Quốc Ngữ* done by the Catholic priests who contributed to the invention of *Chữ Quốc Ngữ* was not widely accessible. Protestantism made its way into Vietnam much later, but within the first 5 years of arriving in Vietnam, Protestants had finished their work of translating the Bible into Vietnamese.

Table 14. Usage of *Thì* in the book of Mark

| Bible Versions | | **1925** | **1976** | **2002** | **2011** |
|---|---|---|---|---|---|
| Lexical *Thì*[45] | | 24 | 1 | 4 | 1 |
| Phrasal *Thì* | | 31 | 29 | 24 | 15 |
| Sentential *Thì* | Inter-clausal *Thì* | 36 | 62 | 16 | 11 |
| | Sentence-initial *Thì* | 60 | 26 | 21 | 2 |
| Total tokens | | 151 | 118 | 65 | 29 |

The overall frequency of *thì* drastically decreased as shown in the translated versions of the Bible. Indeed, in the most recent translation of the Bible, only 29 tokens of the particle *thì* are found in the book of Mark, whereas 151 tokens are found in the first translated version. Similarly, the frequency of the phrasal *thì* and the sentential *thì* is also on the decline, the latter declining from 96 tokens in the first translated version to only 13 tokens in the most recent translation. The empirical evidence from this corpus gives a strong clue to a possible trigger that dismisses the usage of *thì* in Vietnamese and can elucidate a deeper understanding of *thì* as I have analyzed in this thesis.

---

[45] This is the usage of *thì* as a noun. The only two ways that the lexical *thì* was used is *tức thì* 'immediate moment' and *thì giờ* 'time.' The usage of *thì* as a noun has decreased significantly; on top of that, the phrase *tức thì* is also outdated, as presented above. This explains the drastic change from 24 tokens to 4 or just 1 token in the latter translations. Interestingly, the noun *thì giờ* is only used once (Mark 6:31); the rest is *tức thì* 'immediate moment.'

# CHAPTER 5

## Conclusion

### 5.1 Summary

In the introduction, I established the reason why the particle *thì* can be considered an under-acknowledged particle in Vietnamese by showing the unbalanced description and analysis of its functions. Chapter 2 revealed that past analyses focus on accounting for the nature of the phrasal *thì* and overlook its ability to occur as a sentential particle. Section 2.3.3. discussed Clark (1992a) and credited that Clark was the first scholar, to my knowledge, who presented a detailed description of the inter-clausal *thì,* a subcategory of the sentential *thì,* and proposed the analysis that treats the inter-clausal *thì* as "an inchoative conjunction."

Furthering the discussion on the particle, the data in this study shows two things that were not noted previously. First, *thì* is predominantly used as a sentential particle. Out of the 155 collected tokens of *thì* from the oral narratives of the Pear Film movie, 48 tokens (31%) are the phrasal *thì*, 38 tokens (30%) are the inter-clausal *thì*, and 69 tokens (44%) are the sentence-initial *thì*. That is, 107 tokens alone are the sentential *thì*. The data findings necessitated a description of the sentence-initial *thì* to fill in the gap of past literature's description. Consequently, chapter 3 provided a representation of *thì* as a phrasal, inter-clausal, and sentence-initial particle.

As for an analysis, this thesis treats the sentential *thì* as a discourse marker (DM) as proposed in Fraser (2009). *Thì* is shown to meet all the criteria to be a DM. Fraser's original categories of DMs, which are Elaborative DMs (EDMs), Inferential DMs (IDMs), and Contrastive DMs (CDMs) were first used to analyze the tokens of *thì* as found in the collected data. Later, adaptations were proposed to either capture more nuances of the sentential *thì* or to reckon the functions that are exclusive to the sentence-initial *thì*. Overall, the category of CDMs

can remain as is and the category of EDMs must be divided into two subcategories, which are

Sequential EDMs and Non-sequential EDMs. With regard to the category of IDMs, *thì* has the

ability to connect logically related propositions, and showcases three types of logical

relationships: abductive reasoning, inductive reasoning, and deductive reasoning. Therefore, the

category of IDM is broadened to Logical Discourse Markers (LDMs). LDMs consist of three

subcategories: Abductive LDMs, Inductive LDMs, and Deductive LDMs. These subcategories

show that that whereas the sentence-initial *thì* illustrates all three functions of LDMs, the inter-

clausal *thì* can only be used as an Abductive LDM and an Inductive LDM. Evidently, the

category of DMs serves as a beneficial super category that has the ability to show the underlying

patterned usage of the sentential *thì* despite its varied surface level representations.

## 5.2 Future Research Suggestions

The nature of *thì* opens up promising areas for future research. The current approach

prioritizes the account of the semantic and pragmatic implications of the sentential *thì*. However,

this does not negate the possibility of unraveling the particle's functions from other perspectives.

Accordingly, this section details some suggestion of those viable options. Firstly, as briefly

addressed at the end of chapter 4, the possibility of examining the functions of *thì* by tracing its

historical development can reveal insights to understanding its nature. I showed that looking at

how the usage of the particle changes through the different translated versions of the Bible

elucidates a few tendencies: the particle *thì* has undergone, and is possibly still going through,

the process of grammaticalization, for its usage as a noun radically decreased and it is now used

almost exclusively as a function word. Furthermore, I also confirmed that the usage of the

sentential *thì*, especially the sentence-initial *thì*, was far more prevalent in the past, even in

written language. However, a drastic decrease in the overall tokens of *thì*, and especially of the

sentence-initial *thì,* suggests that it is falling out of favor, which could explain why prescriptive grammarians recommend against its use. Future research can look into the development of *thì* from a historical angle and examine the starting point of when *thì* started to be perceived so negatively. A deeper look from the historical perspective may also reveal which of its three main functions developed first, and how it then expanded into the others.

Furthermore, besides the historical linguistics' research direction, the functions of *thì* can also be studied from a syntax-prosody interface. More specifically, whereas prosodic groupings are often used to help determine syntactic constituency, the particle *thì* shows that its syntactic and prosodic constituent structure do not always align. For example, in the case of the phrasal *thì,* it was observed that prosodic grouping is susceptible to the length of the topic: *thì* is joined prosodically with the following constituents after long topics, but with the topics when they are short. As for the writing and punctuation convention, the comma can only come after the topicalized constituents if they are prepositional phrases. And since the particle *thì* is pragmatically licensed rather than syntactically required, it is challenging to accurately pinpoint its true syntactic constituent. Future research could look into this matter more from the angle of the syntax-prosody interface.

The study of *thì* as a discourse marker enriches the current literature on the particle and could prove beneficial for language learning. Current teaching materials of Vietnamese only treat the particle as a topic marker and a conjunction, but as has been shown, it has many other uses. Given that the particle is prevalently used in everyday speech, having a good understanding of its underlying linguistic patterns can positively contribute to learners' command of Vietnamese. Therefore, the analysis of *thì* as a discourse marker can be adapted and converted to be more friendly to the public, especially to language learners.

# Appendices

Appendix A. The Pear Film synopsis (Chafe 1980:xiii-xiv)

The film begins with a man picking pears on a ladder in a tree. He descends the ladder, kneels, and dumps the pears from the pocket of an apron he is wearing into one of three baskets below the tree. He removes a bandana from around his neck and wipes off one of the pears. Then he returns to the ladder and climbs back into the tree.

Toward the end of this sequence, we hear the sound of a goat, and when the picker is back in the tree a man approaches with a goat on a leash. As they pass by the baskets of pears, the goat strains toward them, but is pulled past by the man and the two of them disappear in the distance.

We see another closeup of the picker at his work, and then we see a boy approaching on a bicycle. He coasts in toward the baskets, stops, gets off his bike, looks up at the picker, puts down his bike, walks toward the baskets, again looking at the picker, picks up a pear, puts it back down, looks once more at the picker, and lifts up a basket full of pears. He puts the basket down near his bike, lifts up the bike and straddles it, picks up the basket and places it on the rack in front of his handlebars, and rides off. We again see the man continuing to pick pears.

The boy is now riding down the road, and we see a pear fall from the basket on his bike. Then we see a girl on a bicycle approaching from the other direction. As they pass, the boy turns to look at the girl, his hat flies off, and the front wheel of his bike hits a rock. The bike falls over, the basket falls off, and the pears spill out onto the ground. The boy extricates himself from under the bike and brushes off his leg.

In the meantime, we hear what turns out to be the sound of a paddleball, and then we see three boys standing there, looking at the bike boy on the ground. The three pick up the scattered

101

pears and put them back in the basket. The bike boy sets his bike upright, and two of the other boys lift the basket of pears back onto it. The bike boy begins walking his bike in the direction he was going, while the other three boys begin walking off in the other direction.

As they walk by the bike boy's hat on the road, the boy with the paddle ball sees it, picks it up, turns around, and we hear a loud whistle as he signals to the bike boy. The bike boy stops, takes three pears out of the basket, and holds them out as the other boy approaches with the hat. They exchange the pears and the hat, and the bike boy keeps going while the boy with the paddle ball runs back to his two companions, to each of whom he hands a pear. They continue on, eating their pears.

The scene now changes back to the tree, where we see the picker again descending the ladder. He looks at the two baskets, where earlier there were three, points at them, backs up against the ladder, shakes his head, and tips up his hat. The three boys are now seen approaching, eating their pears. The picker watches them pass by, and they walk off into the distance.

Appendix B. Questions Asked During Interview

The questions were asked in this order. The Vietnamese version will be put first followed by a free translation of the questions into English

1.  Mọi người có thể kể cho em nghe mọi người vừa xem gì được không ạ? Mọi người nhớ nội dung phim vừa xem như thế nào thì cứ kể, không cần phải chính xác từng chi tiết. 'Can you guys tell me what you just watched? You can retell the story however you remember it; there is no right or wrong answer.'

2.  Mọi người nghĩ mối quan hệ của người đàn ông và đứa bé là gì? 'What do you think the relationship between the man picking pears and the boy is?'

3.  Mọi người nghĩ trái cây đó là trái gì? 'What fruit do you think it is?'

4.  Mọi người nghĩ bộ phim này được quay ở đâu? 'Where do you think the movie was filmed at?'

5.  Mọi người nghĩ thông điệp của bộ phim này là gì? 'What do you think is the message of the movie?'

6.  Mọi người có câu hỏi gì nữa không? 'Do you have any other questions?'

Appendix C. Demographic Information on Age and Gender of Participants.

| Group | Speaker | Age | Gender |
|---|---|---|---|
| CG | M | 25 | M |
| | N | 30 | M |
| | O | 26 | F |
| SG | P | 54 | F |
| | Q | 58 | F |
| NG | R | 20 | M |
| | S | 26 | F |
| | T | 31 | M |
| | U | 24 | F |

# References

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics & Philosophy 26*(5). 511–545.

Bolinger, Dwight. 1961. Contrastive accent and contrastive stress. *Language 37*. 83-96.

Cao, Hao. 1992. Some preliminaries to the syntactic analysis of the Vietnamese sentence. *Mon-Khmer Studies 20*. 137-151.

Cao, Hao. 2004. *Tiếng Việt: Sơ thảo ngữ pháp chức năng.* NXB Khoa Học Xã Hội. (Original work published 1991). [*Vietnamese: A functional analysis.* Social Sciences Publishing House].

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles Li (ed.). *Subject and Topic.* 27-55.

Chafe, Wallace. 1980. *The Pear Stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Clark, Marybeth. 1988. An inchoative conjunction in Hmong: Extra-sentential topic marker? *International Symposium on Language and Linguistics.* Thammasat University, Bangkok.

Clark, Marybeth. 1992a. Conjunction as topicalizer in Vietnamese. *Mon-Khmer Studies 20.* 91-109.

Clark, Marybeth. 1992b. Conjunctions as topicalizers: More on Southeast Asian languages. *Mon-Khmer Studies 20*. 87-107.

Clark, Marybeth. 1995. Conjunction as copula in Vietnamese. *Mon-Khmer Studies 26*. 319-331.

Do-Hurinville, Danh. 2009. Study of the topicalizer *thì* in Vietnamese. *Bulletin de la Societe de Linguistique de Paris 104*. 411-443.

Dik, Simon. 1987. *Functional grammar*. Dordrecht: Foris.

Dooley, Robert & Levinsohn, Stephen. 2000. *Analyzing discourse: A manual of basic concepts.*
North Dakota: SIL-UND.

Dung et. al. 1998. Intonation in Vietnamese, in D. Hirst & A. Di Cristo (eds.), *Intonation System: A Survey of Twenty Languages.* Cambridge: Cambridge University Press.

Dyvik, Helge. 1982. *Subject or topic in Vietnamese?* Nr. 13, Serie B, Institutt for Fonetikk og Lingvistikk, Universitet i Bergen.

Foley, William & Van Valin, Robert. 1984. *Functional syntax and universal grammar*. *Cambridge Studies in Linguistics 38*. Cambridge: Cambridge UP.

Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics 31*. 931-952.

Fraser, Bruce. 2006. On the conceptual-procedural distinction. *Style 40*(1-2). 24-32.

Fraser, Bruce. 2009. An account of discourse markers. *International Review of Pragmatics 1*. 293-320.

Frey, Werner. 2006. Contrast and movement to the German prefield. In Valéria Molnár & Susanne Winkler (eds.), *The Architecture of Focus*. 197-233. Berlin/New York: Mouton de Gruyter.

Haiman, John. 1978. Conditionals are topics. *Language 54*(3). 564-589.

Halliday, Michael. 1970. Language structure and language function. In John Lyons (ed.), *New Horizon in Linguistics*. New York: Penguin Books.

Hope, Edward. 1974. *The deep syntax of Lisu sentences*. Ph.D. Dissertation. Australian National University, Pacific Linguistics, Series B, no. 34.

Jannedy, Stefanie. 2007. Prosodic focus in Vietnamese. In S. Ishihara, S. Jannedy & A. Schwarz (eds.), *Interdisciplinary Studies on Information Structure 8*. 209–230. Potsdam: Potsdam University Press.

Jespersen, Otto. 1940. *A modern English grammar on historical principles.* V: Syntax. London: George Allen & Unwin.

Krifka, Manfred. 2006. Association with focus phrases. In Valerie Molnar & Susanne Winkler (eds.), *The Architecture of Focus*. 105–136. Berlin: Mouton de Gruyter.

Labov, William & Fanshel, David. 1977. *Therapeutic discourse*. New York: Academic Press.

Lambrecht, Knud. 2004. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Li, Charles & Thompson, Sandra. 1976. Subject and topic: A new typology of language. In Charles Li (ed.), *Subject and Topic*. 475-490. New York: New York Academic Press.

McCawley, James. 1970. Where do noun phrases come from? In R.A. Jacobs & P.S. Rosenbaum (eds.), *Readings in English Transformational Grammar*. 166-183. Waltham: Ginn.

Molnár, Valéria. 2006. On different kinds of contrast. In Valéria Molnár & Susanne Winkler (eds.), *The Architecture of Focus*. 197-233. Berlin/New York. Mouton de Gruyter.

Neeleman, Ad & Van de Koot, Hans. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative German Linguistics* 11. 137–189.

Neeleman, Ad & Vermeulen, Reiko. 2012. The syntactic expression of information structure. *The Syntax of Topic, Focus, and Contrast: An Interface-based Approach*. 1-38. De Gryuter.

Nguyen, Hoa Dinh. 1990. *Vietnamese*. Amsterdam & Philadelphia: John Benjamins.

Nguyen, Quan. 1887. *Thầy Lazaro Phiền*. HCMC City: Rue Catinat. Retrieved from cvdvn.files.wordpress.com/2017/10/thay-lazaro-phien.pdf

Nguyen, Dan. 2013. *Nỗi oan thì, là, mà*. NXB Trẻ: TP. HCM. [*The slander against thì, là, mà. The Youth Publication Services: HCMC City*].

Nguyen, Mai. 2019. Do you hear what I hear?: On the elusive nature of Vietnamese tones. *Proceedings of the National Conference of Undergraduate Research (NCUR) 2019*. 971-979. Kennesaw State University.

Schachter, Paul & Fe T. Otanes. 1972. *Tagalog reference grammar.* Berkeley, Los Angeles, and London: University of California Press.

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

Schiffrin, Deborah. 1987. *Discourse markers.* Cambridge: Cambridge University Press.

Thompson, Laurence. (1987). *A Vietnamese reference grammar*. University of Hawaii Press. (Original work published 1965).

Tran, Thuan. 2016. Lone contrastive topic construction: A puzzle from Vietnamese. In Mira Grubic & Anne Mucha (eds.). *Proceedings of The Semantics of African, Asian and Austronesian Languages* 2. 52–64. Potsdam: Potsdam University Press.

Tran, Thuan. Forthcoming. Syntax-information structure interface in Vietnamese." In Elin McCready, Hiroki Nomoto, & Yukinori Takubo (eds.). *Discourse Particles in Asian Languages*. Oxford: Routledge.

Van Dijk, Teun. 1979. Pragmatic connectives. *Journal of Pragmatics* 3. 447 - 456.

Zimmermann, Malte. 2008. Contrastive focus and emphasis. *Acta Linguistica Hungarica* 55. 347–360.

Kinh Thánh tiếng Việt bản dịch năm 1925. [Vietnamese Bible translation 1925 version. kinhthanh.httlvn.org/?v=VI1934]

Kinh Thánh tiếng Việt bản phổ thông 1976. [Vietnamese Bible translation 1976 version. Retrieved from kinhthanh.httlvn.org/?v=BPT]

Kinh Thánh tiếng Việt bản dịch mới 2002. [Vietnamese Bible translation 2002 version.

Retrieved from kinhthanh.httlvn.org/?v=NVB]

Kinh Thánh tiếng Việt bản dịch 2011. [Vietnamese Bible translation 2011 version. Retrieved

from kinhthanh.httlvn.org/?v=BD2011]