

MĀN ON THE REFERENTIALITY CONTINUUM IN THAI¹

Jinawat Kaenmuang
 Chulalongkorn University
 jinawat.k@chula.ac.th

Piroon Piyamahapong
 Thammasat University
 ppiroon9@tu.ac.th

Pittayawat Pittayaporn²
 Chulalongkorn University
 pittayawat.p@chula.ac.th

Abstract

Pronouns are traditionally defined as a pro-form of an explicit antecedent. However, the pronoun *mān* in Thai sometimes occurs without any co-referring explicit nominal expression, leading previous studies to consider them as non-referential. This study argues that, despite the absence of an explicit antecedent, such instances of *mān* have implicit referents that are inferable from context. One thousand instances of *mān* functioning as subject or object from the Thai National Corpus were analyzed in a usage-based approach. They were categorized according to their referentiality using three criteria: explicitness of a nominal antecedent, concreteness of an antecedent, and inference of a referent. The analysis reveals that the referentiality of the pronoun *mān* is not dichotomous but instead lies on a continuum in which one end expresses semantic referentiality with an explicit antecedent with a higher level of referent concreteness, and the other expresses pragmatic referentiality with an implicit antecedent with a lower level of referent concreteness. Additionally, different types of referential expressions—verb phrases, clauses, and discourse—and ambiguous cases among them strongly support the notion of gradience of referentiality.

Keywords: semantic referentiality, pragmatic referentiality, the pronoun *mān*
ISO 639-3 codes: tha

1 Introduction

In a traditional point of view, pronouns are referential in nature. The phenomenon in which expressions refer to the antecedents indicating entities in the real world is labelled *reference*, while *referentiality*, or referential meaning, is defined as the linking of expressions to referents in the universe of discourse. As studies on referentiality are abundant among European languages, it is generally assumed that morphosyntactic devices play an important role in exhibiting specific types of meaning, including a referential one. Spanish, for example, makes use of different syntactic constructions to encode the referentiality of the pronoun *se*. The reflexive *se* (SE_{REFL}) in (1) refers to the actor of the event *vio* ‘saw’, while the impersonal *se* (SE_{IMPERS}) and the passive *se* (SE_{PASS}) in (2) and (3) are non-referential interpretations of an unpronounced subject. These examples illustrate that different syntactic constructions play a vital role in differentiating referentiality from non-referentiality.

- (1) *Juan se vio (a sí mismo) en el espejo.*
 Juan SE_{REFL} saw DOM self same in the mirror
 ‘Juan saw himself in the mirror.’

¹ This paper forms an integral part of the first author's dissertation entitled “*Change in referentiality of the pronoun /man1/ in Thai during the Rattanakosin Period*”.

² Corresponding author. Head of Center of Excellence in Southeast Asian Linguistics, Faculty of Arts, Chulalongkorn University.

- (2) *Se conoce a María como buena madre.*
 SE_{IMPERS} KNOWS DOM Maria as good mother
 ‘Mary is known as a good mother.’
- (3) *Se destruyeron las casas.*
 SE_{PASS} destroyed the houses
 ‘The houses were destroyed.’ (McDonald 2017)

While referentiality seems to be clearly dichotomous in European languages such as Spanish, it is vaguer and harder to differentiate in isolating languages such as Thai. According to Pipalova (1988:196), the distinction between referential and non-referential uses is too complex a matter to be reduced to a dichotomy model. Thus, it seems more appropriate to speak about the degree of referentiality as a result of a combination of syntactic, morphological, lexical, and referential type factors.

In Thai, the following examples from the Thai National Corpus, a general corpus of standard Thai comprising approximately 33 million words which is designed to be balanced in its domain and medium proportions (Aroonmanakun 2007), clearly show that it is difficult to separate the referential use of the pronominal units from the non-referential uses. In (4), the pronoun *mān* explicitly refers to the antecedent *hīmá?* ‘snow’, and as a result, the referent is easily retrieved. However, *mān* in (5) lacks an explicit antecedent and is therefore ambiguous. It can be interpreted as referring either back to the antecedent *k^hâ:ŋbōn* ‘upstairs’ or to the atmospheric environment upon inference. Contexts therefore play a significant role in understanding the referential meaning of this word.

- (4) *hīmá? tòk nāk mān pòkk^hlum t^hā:ŋ k^hāw bā:n*
 snow fall heavy PRO cover way enter house
 ‘It snowed heavily, covering the entryway of the house.’
- (5) *cà? sàj fāj k^hâ:ŋbōn p^hró? mān mû:t*
 TAM put fire upstairs because PRO dark
 ‘(I) will have the upstairs lit because it is dark.’

Indrambarya (2011; 2012) and Angkapanichkit (2011), on the other hand, based on their formal criteria of verb subcategorization and antecedent occurrence, respectively, contend that referential *mān* and non-referential *mān* are dichotomous. However, these studies have not yet accounted for contextual effects leading to ambiguity in referential interpretations nor different shades of referentiality of the pronoun *mān*. The referentiality of *mān* in (5) is somehow shaded from the prototypical referential meaning as it may be interpreted as the place mentioned or the environment. Such different interpretations of *mān* arise mainly due to different types of contexts in which it occurs.

In contrast to traditional view on referentiality, the usage-based approach (Bybee 2023; Perek 2023) illuminates the nuanced interpretation of referentiality gradience, offering a new perspective that will captivate semanticists’ interest. In this approach, the grammatical system is not viewed as a fixed set of rules accessible to language users but rather as a system emerging from language use, underscoring the collaborative nature of language use, where each interaction contributes to the evolution of linguistic structures. Sentence meaning is believed to derive from both linguistic elements, such as each constituent’s conventional meaning and specific semantic properties, and non-linguistic factors, including pragmatic meaning, speakers’ communicative intentions, and context.

In this alternative approach, context-encompassing not only the immediate situation of the utterance (e.g., speaker, addressee, location, and time) but also shared information or common ground between interlocutors—plays a pivotal role in determining references. Speakers contribute to effective communication by anticipating listeners drawing necessary inferences from the preceding discourse. Contexts may involve familiar entities, cultural knowledge, and information jointly accessible to the interlocutors in conversation (Roberts 2019; Diessel 2019). Cognitive linguists consider this shared knowledge as part of consensual encyclopedic knowledge (Lee 2001; Stalnaker 2002).

This study embarks on a less-traveled journey, departing from previous studies into the use of the pronoun *mān*. It aims to provide a comprehensive account of the phenomenon associated with *mān*, revisiting more

prototypical cases and exploring the so-called non-referential uses. It thus investigates referentiality gradience from language-in-use data collected from the Thai National Corpus (Aroonmanakun 2007). Adopting the stance that referentiality is gradient (Pipalova 1988; Chen 2009), we hypothesize that the meaning of the pronoun *mān* in Thai does not fall neatly into two distinct categories, referential and non-referential. Instead, it comprises many ambiguous usages affected by the context where semantic and pragmatic referentiality play a role for interpretation. The former pertains to the inherent properties of linguistic expressions to denote entities while the latter emphasizes how those expressions are employed in specific communicative settings to establish reference. These different shades of meaning for *mān* are, therefore, evidence for referentiality as a continuum rather than as a dichotomous distinction. Viewing referentiality as a gradient scale can provide a clear and detailed picture of such pronominal referential properties.

2 Literature review

Personal reference terms in Thai consist of personal pronouns, personal names, kin terms, and occupational titles (Iwasaki & Ingkaphirom 2005). According to Cooke (1968), personal pronouns are composed of 27 first-person pronouns, 22 second-person pronouns, and 8 third-person pronouns. Personal pronouns in Thai are in three-way person contrast (first, second, and third), and two-way contrast in number (singular and plural). Singular and plural personal pronouns may also differ in terms of gender, formality, and intimacy (Kullavanijaya 2000:80).

The term *mān* is normally regarded as a third-person singular pronoun by Thai grammarians and linguists who either treat it as a member of a distinct word class (Upakitsinlapasan 2001; Iwasaki & Ingkaphirom 2005:2; Bandhmedha 2022:22), or designate it as a subtype of the noun category due to shared syntactic distributions (Panupong 1991:65; Warotamasikkhadit 1996:43; Prasithrathsint 2013). Like nouns, it may precede main verbs, auxiliaries, or determiners, and it follows verbs or prepositions, as in examples (6), (7), and (8).

- (6) *mê:/mān* *mā:*
 mother/PRO come
 ‘The mother/It came.’
- (7) *mê:* *tī:* *sūnák/mān*
 mother beat dog/PRO
 ‘The mother beat the dog/it.’
- (8) *mê:* *p^hû:t* *kàp* *lû:kc^hā:j/mān*
 mother speak to son/PRO
 ‘The mother talked to her son/it.’

mān as a personal pronoun has traditionally been described as substituting for people, animals, and things (Upakitsinlapasan 2001), or stated people and entities (Bandhmedha 2022:22). This function illustrates that the pronoun *mān* is prototypically referential. The use of *mān* without any referents in the *mān pēn* ‘it is’ construction, as shown in (9) and (10), was first observed by Chutisilp (1984) who presumed that *mān* in this construction is influenced by English and later Thai speakers calqued the construction. Research focusing instead on the referentiality of *mān* as a pronoun was undertaken by Indrambarya (2011; 2012) and Angkapanichkit (2011). They separated the referential functions of *mān* into two distinct categories: referential *mān* and non-referential *mān*.

- (9) *mān* *pēn* *k^hwā:mk^hít* *t^hī:* *dī:*
 PRO COP thought REL good
 ‘It is a good idea.’

- (10) *mān* *pēn* *rûaŋ* *sānùk*
 PRO COP story fun
 ‘It is a fun story.’

Angkapanichkit (2011) studied the use of the pronoun *mān* in narrative discourses and found that it can be used as a referential or non-referential pronoun. The main difference between the two lies in the fact that the referential use has nominal antecedents, while the non-referential use has none. Regarding non-referential use, *mān* is labeled a dummy subject once it acts as a subject with no referred linguistic expressions. Antecedents can be located either inter-sententially or extra-sententially. In the latter case, *mān* adjacently follows its antecedents, and is termed a ‘shadow pronoun’.

Indrambarya (2011; 2012), further found that referential *mān* can co-occur with any personal verbs, while non-referential *mān* only occurs with two types of verbs: true-impersonal verbs (verbs indicating existence or verbs indicating degree of possibility, accessibility, ease, time, or temperature, as shown in (11)) and pseudo-impersonal verb (verbs indicating psychological state, as shown in (12)). The term *mān* as a non-referential subject may co-occur with pseudo-impersonal verbs. This use changes personal verbs into impersonal ones and indicates alienation from an action or undesirable activity.

- (11) *mān* *k̄:t* *pʰèndīnwāj* *tʰi:* *jī:pùn*
 PRO happen earthquake REL Japan
 ‘There was an earthquake in Japan.’

- (12) *mān* *māj* *ŋá:j* *tʰi:* *cà?* *lū:m* *kʰrāj* *sàk* *kʰōn*
 PRO NEG easy REL TAM forget who QNT CLF
 ‘It is not easy to forget someone.’

In semantics, one of the central aspects of language use is the process of selecting or associating words with specific meanings. This is called referring or denoting (Saeed 2022). Such referring is interpreted as the manifestation of denotation of an entity in a context. It is based on the interaction between a referring expression and its referent within a model of discourse. Among expressions referring to specific referents, pronouns are firmly shown to directly link to referents through their antecedents. The notion of referentiality has received considerable attention among logicians, philosophers, psychologists, and linguists (Chen 2009; Chen, Huang, & Wu 2010; Chen 2015; Fretheim & Gundel 1996). While scholars from these different fields have their own interests here, linguists put emphasis on linguistic encoding and the interpretation of speakers’ intended referent by the addressee or interlocutor (Fretheim & Gundel 1996).

Reference is traditionally defined as the relationship held between an expression and what that expression relates to or stands for in the real world (Lyons 1977; Carlson 2006; Abbott 2010). Following this definition, the referring expression correctly identifies the referent if the reference is successful (Lyons 1977). Referentiality—or referential meaning³—in contrast, is described as a semantic property of nominals involving the speaker’s intent to refer to or mean a nominal expression that exists within a particular universe of discourse (Givón 1978). Therefore, it can be concluded that reference refers to the phenomenon which involves the pairing between a referring expression and its referent, whereas referentiality deals with the inherent property of referring within any expression. It should be noted that a referent is not only restricted within the perceptible real world, according to Lyons’ definition. It also expands to cover both concrete and abstract entities in the universe of discourse, in Givón’s definition.

Referentiality has long been seen as a dichotomy—referential versus non-referential—in the relevant literature (Finegan 2014; Huddleston & Pullum 2015; Levinson 1983; Lyons 1977; Yule 1996). For example, Li and Thompson (1976) famously claim that topic-prominent languages do not require a dummy subject, whereas subject-prominent languages do so to satisfy syntactic constraints. Most conventional methods for detecting non-referential pronouns depend on syntactic heuristics to recognize expletive usages of the pronoun *it*. For example, the pronoun *it* in ***It** is important to maintain a stable price* is a classic example of expletive *it*, where the use of *it* is not connected to any specific reference. As a result, one can infer that it is non-referential

³ Chandler & Munday (2011) illustrate that referentiality and referential meaning refer to the same thing and that both can be defined as a reference to the external world.

(Uryupina et al. 2016). In contrast, the example *Lucy was on holiday. One morning, she decided to go to the mountains* demonstrates that the speaker continues to refer to the subject *Lucy* in the following phrases, explicitly employing the personal pronoun *she* to represent pronominal subjects (Quesada & Lozano 2020).

Some studies, however, reveal that referentiality should be treated as gradient in nature. Pipalova (1988:196) maintains that distinction between referential and non-referential uses is too complex a matter to be reduced to a dichotomy model. Thus, it seems more appropriate to speak about the degree of referentiality as a result of a combination of syntactic, morphological, lexical, and referential type factors. Gundel, Hedberg, and Zacharski (2005) found cases in which the third-person pronouns used in spoken American English lack explicit noun phrase antecedents but are still referential because the referents are inferable from a particular semantic frame or from mutual knowledge between interlocutors (such as event, activity, or situation).

Chen (2009; 2015)⁴ supports the perspective that referentiality is not binary by classifying level of referentiality into three distinct categories: semantic, pragmatic, and discourse thematic. Semantic referentiality is defined in terms of the property of a linguistic expression that links itself with an entity in a particular discourse. While semantic referentiality does not depend on context (i.e., referents are construed by lexical encoding of expressions), pragmatic referentiality is contextually dependent. Three components of pragmatic referentiality are posited: (i) the existence of an entity is presupposed in the universe of discourse established between the speaker and the addressee, (ii) the entity is a specific reference to the speaker, and (iii) speakers intend to use the expression to refer to the specific entity.

Regarding discourse thematic referentiality, Chen emphasizes that it is not simply an alternative form of referentiality but rather a comprehensive approach that encompasses both semantic and pragmatic aspects of reference within discourse. It involves strategically choosing referential expressions that carry thematic importance and contribute to the coherence and progression of the discourse. This approach considers various factors, such as the frequency of using and referring to a single noun phrase, alongside other linguistic and contextual strategies employed in discourse development. Consequently, discourse thematic referentiality is not treated as a distinct category for data analysis in this study given its integration with broader referential frameworks.

The theoretical framework for this study builds on the evolving understanding of referentiality in linguistic research. Traditionally, referentiality has been dichotomized into two distinct categories: referential and non-referential. This binary approach, while foundational, has been increasingly challenged by evidence of intermediate cases that do not conform strictly to either pole. Drawing from usage-based linguistics, which emphasizes the gradience and context-sensitivity of language, this study proposes a continuum model of referentiality. The continuum accounts for varying degrees of referentiality, recognizing the fluidity between highly referential expressions (e.g., specific definite noun phrases) and non-referential uses (e.g., expletive constructions). This framework is further informed by discourse-pragmatic theories that highlight how context and interaction influence referential interpretation.

3 Data and analysis

In this study, we analyzed 1,000 instances of orthographical *mān* from the Thai National Corpus (TNC). Each instance was carefully investigated by the three authors to avoid personal biases. Those in which *mān* occurred in the following syntactic categories were manually removed: (1) nouns, nouns which are parts of nominal compounds and proper names such as *mān^hé:t* ‘cassava’ and *mānmū:* ‘lard’ *ʔāndā:mān* ‘Andaman’, (2) verbs such as *nā:mān* ‘The skin face is oily’ and *kīn ʔā:hā:n mān* ‘eating oily food’, (3) adjectives such as *sāj rō:ŋt^há:w mān wā:w* ‘putting on polished shoes’, and (4) idiomatic expressions such as *c^hāŋmān* ‘Whatever!’ and *mānmū:* ‘to enjoy doing something’. In the end, 954 out of the 1,000 contexts remained for consideration.

The term *mān* in these 954 contexts functions as a subject, object of a verb, object of a preposition, emphasis marker of a subject, and emphasis marker of a verbal object, as shown in (13) to (17). The data reveal that immediate occurrences of *mān* after noun phrases need to be included as they highlight the meanings of referential expressions. For example, *mān* in (16) below is deliberately used to highlight the referent preceding

⁴ Chen (2015) identified a fourth type of referentiality, logic-philosophical, adding to the previous three distinct categories: semantic, pragmatic, and discourse thematic referentiality. This study, however, adopts only the semantic and pragmatic referentiality for its analysis, as the logic-philosophical referentiality is considered to be unbounded from the linguistic perspective, and discourse thematic referentiality is centered on the property of a lexical unit within discourse development.

the word *mān*, *mǎ*: ‘dog’. The use of *mān* in such cases still depicts the referential property regarded as the most typical feature of pronominals.

- (13) *c̄x̄:* *krār̄ô:k* *t̄ua* *n̄uŋ* *mān* *kāmlāŋ* *tʰé?* *tʰua* *j̄u:*
 find squirrel CLF one PRO TAM nibble nut TAM
 ‘(I) found a squirrel. It was nibbling on the nuts.’
- (14) *mē:w* *kʰāmo:j* *plā:* *jā:ŋ* *jā:* *l̄x̄:j* *t̄i:* *mān*
 cat steal fish grill grandmother LNK hit PRO
 ‘A cat stole a grilled fish, so the grandmother hit it.’
- (15) *khǎw* *p̄it* *khà:w* *l̄x̄:j* *māj* *rú:* *ʔārāj* *k̄iawkàp* *mān* *tʰáwràj*
 3SG close news LNK NEG know INDEF about PRO Q
 ‘As they kept it undisclosed, we knew nothing much about it.’
- (16) *mǎ:* *mān* *d̄x̄:n* *tā:m* *t̄ô:j* *t̄ô:j*
 dog PRO walk follow closely
 ‘The dog kept following (me).’
- (17) *ʔum* *l̄u:k* *mān* *n̄òj* *si?* *r̄ô:ŋhâ:j* *dāŋ* *lân* *là?*
 carry child PRO PTC PTC cry loud much PTC
 ‘Carry the baby. It is crying out loud.’

Unlike previous research in which the referential meaning of the pronoun *mān* is primarily based on the existence of nominal antecedents, we adopt three criteria to analyze referentiality: explicitness, concreteness, and inference.

As the analysis of the referential meaning requires supporting linguistic evidence to avoid personal biases, explicitness or existence of linguistic forms is thereby taken as the principal criterion. Referential expressions are divided into two categories, nominals and non-nominals. Nominal antecedents cover nouns, noun phrases, and nominalizations, while non-nominals include verb phrases, clauses, and discourses or groups of relevant clauses.

While linguistic expressions, whether nominal or non-nominal, refer to referents in either the real world or the universe of discourse, they greatly vary in terms of concreteness—the referent property of direct perception through sensory experiences (Bolognesi, Burgers, & Caselli 2020). Tangible real-world referents, such as people, animals, and things, are therefore considered more concrete than those depicting actions, events, or attitudes.

Linguistic expressions are sometimes referential but sometimes not. Those with referential meanings are usually expressed through explicit antecedents. When an explicit antecedent is present in a sentence, it provides a clear reference point for the listener or reader to make an inference. In this way, the antecedent acts as the anchor or starting point for understanding the meaning and intended reference of the subsequent pronoun or referring expression. Without a clearly identifiable antecedent, language users make use of inference to understand the referent the word *mān* conveys.

Implicit antecedent and inference, however, involve inferring or presupposing missing information to establish referential connections. Listeners or readers rely on contextual cues, background knowledge, and the structure of the sentence to make inferences and fill in the gaps, allowing for a complete understanding of the intended meaning. For example, *mān* in (18) to (20) refers to weather, time and surrounding environment, respectively, although there is a lack of obvious referent.

- (18) *mān* *r̄ô:n* *mā:k* *p̄i:* *n̄i:* *f̄õn* *f̄á:* *māj* *t̄òk*
 PRO hot much year DEM rain sky NEG fall
 ‘It’s very hot this year. There’s very little rainfall.’

- (19) *jà:* *c^hák^há:* *mān* *sā:j* *mā:k* *lé:w* *ná?*
 NEG sluggish PRO late much TAM PTC
 ‘Don’t delay. It is very late now.’
- (20) *dā:n* *k^háw* *sā:j* *pāj* *rúa:jrúa:j* *mān* *mú:t* *mā:k*
 walk enter street go leisurely PRO dark mâ:k
 ‘(I) leisurely walked along the street. It was really dark.’

Examples (21) and (22) illustrate a correspondence between explicitness, concreteness, and inference. If the word *mān* refers to an explicit antecedent that is a specific entity with high concreteness, inference is not required. In contrast, if the word *mān* refers to an implicit antecedent without obvious linguistic expressions that is not an entity in the real world, inference, therefore, is required to interpret the referent of *mān*. It is this lack of explicit antecedent that led previous studies to classify instances of *mān* into dichotomous categories: referential and non-referential.

- (21) *sǒncū:nīp̄x̄:* *sājpré:t* *lé?* *p^hā:j* *pēn* *má:j* *júk* *kò:nprāwàttīsà:t*
 juniper cypress and pine COP tree era prehistoric
mān *sù:pp^hān* *dūaj* *mālét*
 PRO reproduce by seed
 ‘Juniper, cypress, and pine trees are prehistoric plants. They reproduce by producing seeds.’

In (21), *mān* refers to the explicit antecedents *sǒncū:nīp̄x̄:* ‘juniper’, *sājpré:t* ‘cypress’, and *p^hā:j* ‘pine trees’, which indicate tangible real-world objects. Interpreting the referents of *mān* thereby requires no inference.

On the contrary, *mān* in the conversation in (22) refers to an implicit antecedent that is not a specific entity. Therefore, inference is required to identify the referent of *mān*.

- (22) A: *ǎwñājđī:* *nā:hùaj* *wà?* *mān* *k̄:t* *ná:mpà:* *lǎjlà:k*
 DM worrisome PTC PRO exist flash flood overflow
kàp *dīnk^hlo:n* *t^hālòm* *t^hé:w* *sājjók* *l̄x̄:j*
 and mud collapse around Sai Yok exactly
 B: *k^hōn* *ǎù:n* *wá:* *ñāj* *pāj* *kān* *jù:* *máj*
 person other say Q go RECP stay Q
 A: ‘Well, it’s quite worrisome. There exist flash floods and mudslides right around Sai Yok.’
 B: ‘What do other people say? Will they still be going?’

However, *mān* in some contexts, exemplified in (23), does not exhibit the same properties shown by those in the two previous examples.

- (23) *p^hú:t* *bè:p* *ní:* *mǎ:jk^hwā:m* *wá:* *jāññāj* *mān* *fāñ* *thāmêñ thāmêñ*
 speak type DEM mean COMP Q PRO listen weird
 ‘What do you mean by saying this? It sounds strange.’

The term *mān* refers to the phrasal verb *p^hú:t bè:p ní:* ‘saying this’. In this case, the antecedent has an explicit linguistic form, but rather than denoting a tangible entity, it refers to an event of speaking.

(24)	[<i>rāw</i>	<i>kô:</i>	<i>pāj</i>	<i>nāj</i>	<i>mā:</i>	<i>nāj</i>	<i>dúaj</i>	<i>kān</i>	<i>bəj</i>	<i>kʰráp</i>
	1PL	LNK	go	Q	come	Q	together	RECP	often	PTC
	<i>rāw</i>	<i>pāj</i>	<i>kīn</i>	<i>kʰá:w</i>	<i>kān</i>	<i>rāw</i>	<i>cʰūan</i>	<i>kān</i>	<i>pāj</i>	<i>cʰəppīŋ</i>
	1PL	go	eat	rice	RECP	1PL	invite	RECP	go	shopping
	<i>rāw</i>	<i>lén</i>	<i>kī:lā:</i>	<i>dúaj</i>	<i>kān</i>	<i>rāw</i>	<i>kʰūj</i>	<i>prəp</i>	<i>tʰúk</i>	
	1PL	play	sport	together	RECP	1PL	talk	relieve	suffering	
	<i>sùk</i>		<i>kān</i>	<i>tālə:t]</i>	<i>mān</i>	<i>pēn</i>	<i>mūan</i>	<i>kʰwā:msāmpʰān</i>		
	happiness		RECP	always	PRO	COP	similar	relationship		
	<i>rāwà:ŋ</i>		<i>fē:n</i>							
	between		boyfriend/girlfriend							

‘We often go out together. We go out to eat together. We go shopping together. We play sports together. We talk to each other all the time. It’s like we are in a relationship.’

The pronoun *mān* in (24) refers to discourse combining various sentences (1) *rāw kô: pāj nāj mā: nāj dúaj kān bəj kʰráp* ‘We often go out together.’; (2) *rāw pāj kīn kʰá:w kān* ‘We go out to eat together.’; (3) *rāw cʰūan kān pāj cʰəppīŋ* ‘We go shopping together.’; (4) *rāw lén kī:lā: dúaj kān* ‘We play sports together.’; and (5) *rāw kʰūj prəp tʰúk sùk kān tālə:t]* ‘We talk to each other all the time.’ Although the mentioned discourse is expressed through an explicit linguistic expression, the referents to which the pronoun *mān* refers are not concrete entities: they are a series of events.

The word *mān* in every context, consequently, does not belong neatly to either referential *mān* or non-referential *mān* as previous studies on *mān* illustrated. There may be other categories between the two posited groups.

4 *Mān* on the referentiality continuum

The referentiality of the pronoun *mān* in Thai can be categorized into two primary types: semantic referentiality and pragmatic referentiality. Semantic referentiality relies on explicit antecedents such as noun phrases, verb phrases, clauses, and discourse. On the other hand, contextual elements, namely, context-based referents, speaking subject referents, and common ground referents, shape pragmatic referentiality. A total of 954 instances of *mān* were analyzed, with ten ambiguous cases excluded due to their potential interpretation as both semantic and pragmatic.

Table 1 illustrates the raw frequency of different antecedent types to which the pronoun *mān* refers.

Table 1: Frequency of antecedent types to which the pronoun *mān* refers

semantic referentiality (865)				pragmatic referentiality (79)		
noun phrase	verb phrase	clause	discourse	context-based referent	speaking subject referent	common ground referent
721	63	49	32	35	17	27

This breakdown provides insights into the varied contexts in which *mān* is used. Examining both linguistic structures and contextual cues gives a deeper understanding of how referentiality operates in Thai discourse.

4.1 *mān* referring to explicit antecedents with semantic referentiality

Explicitness of expressions functioning as antecedents of the pronoun *mān* is found to correlate positively with the concreteness of referents to which the expressions refer. For example, the word *mān* in (25) and (26) refers to the nominals *mā: tʰi: bā:n* ‘my dog’ and *kā:nwīnícʰāj rô:k nī:* ‘diagnosing this disease’, respectively.

(25)	NP[<i>mā:</i>	<i>tʰi:</i>	<i>bā:n]</i>	<i>hàw</i>	<i>kəj</i>	<i>mā:k</i>	<i>té:</i>	<i>mān</i>	<i>máj</i>	<i>kàt</i>
	dog	at	home	bark	good.at	a lot	but	PRO	NEG	bite
	‘My dog barks a lot, but it does not bite.’									

- (26) NP[kā:nwīníc^hǎj rō:k nī:] t^hām jā:k tē: mān pēnpājdā:j
 diagnose disease this make difficult but PRO possible
 ‘Diagnosing this disease is a difficult task, but it is possible.’

In addition to nominal antecedents, *mān* can refer to verb phrases (VP), clauses (CL), and discourse (DISC), as shown in sentences (27) to (29). The analysis reveals that the pronoun *mān* can function as a substitute for other syntactic categories, such as verb phrases, single clauses, and multiple clauses, opposing the traditional understanding of Thai pronouns as limited to nominal categories.

- (27) VP[rō:ŋhā:j bəj bəj] māj dī: mān t^hām hāj dū: ʔə:nʔē:
 cry often often NEG good PRO make give lookweak
 ‘Crying often is not good, as it makes you look weak.’

- (28) CL[phū:nām mā: māj p^hrómprīāŋ kān] mān bóŋc^hí: nājǎáʔ bā:ŋjā:ŋ
 leader come NEG together RECP PRO indicate sense INDEF
 ‘That all leaders from various nations did not attend the meeting signaled some abnormalities.’

- (29) DISC[kə:t p^hèndīnwǎj rūnrē:ŋ tē: k^hlú:nlōm sāŋòp ná:mt^hālē: lót
 happen earthquake severe but storm calm sea.water decrease
 lōŋ jā:ŋrúatrēw p^hítpòkkāti sət tà:ŋtà:ŋ mī: p^hrúttikām plē:k
 down rapidly unusual animal various have behaviour strange
 pāj cà:k pòkkāti] mān sādē:ŋ wá: kāmīāŋcàʔ kə:t sū:nā:mí?
 go from normal PRO show COMP TAM happen Tsunami
 ‘There was a severe earthquake, but the sea is calm. The level of sea water decreases unusually rapidly. Animal behaviors are strange. It indicates that a tsunami is going to take place.’

The word *mān* in (27) refers to the verb phrase *rō:ŋhā:j bəj bəj* ‘crying often’, while it refers to the clausal antecedent *phū:nām mā: māj p^hrómprīāŋ kān* ‘all leaders from various nations did not attend the meeting’ in (28) and the discursial antecedent *kə:t p^hèndīnwǎj rūnrē:ŋ tē: k^hlú:nlōm sāŋòp ná:mt^hālē: lót lōŋ jā:ŋ rúatrēw p^hítpòkkāti sət tà:ŋ tà:ŋ mī: p^hrúttikām plē:k pāj cà:k pòkkāti* ‘There is a severe earthquake, but the sea is calm. The level of sea water decreases unusually rapidly. Animal behaviors are strange’ in (29).

Explicit antecedents can be categorized into two major groups, nominal and non-nominal, both of which differ according to the referents to which each category refers, where the former refers to concrete entities or abstract ideas as expressed by the noun phrase or the nominalization, while the latter refers to event(s) as communicated by verb phrases, clauses, or discourse. In spite of different semantic categories of referents, *mān* referring to all of the explicit antecedents exhibits semantic referentiality.

4.2 *mān* referring to implicit antecedents with pragmatic referentiality

From the data, we find that implicit antecedents crucially need inference to understand their implied or inferable antecedents as there are no direct linguistic expressions. The pronoun *mān* referring to implicit antecedents functions as a subject of the predicate.

In general, communication between interlocutors sometimes omits linguistic expressions to represent references, whether they are lexical, phrasal, or clausal. Instead, referential meanings are intuitively grasped through linguistic contexts, metalinguistic contexts, and encyclopedic knowledge. Comprehending the referents of a pronoun that lacks linguistic antecedents requires interpretation, inference, consideration of the speaker’s viewpoint and attitude towards a mentioned event, and an awareness of the background knowledge that signifies the relationship between interlocutors. This process results in mutual knowledge which serves as a mechanism for facilitating effective communication. Therefore, understanding such referential meaning cannot solely rely on linguistic knowledge (Sainsbury 2005).

The term *mān* as a subject referring to implicit antecedents can be divided into three subtypes according to its inferential referents, be they context-based referents, speaking subject referents, or common ground referents.

Context-based referents refer to those inferred from the linguistic structure, which communicates ideas through general reasoning. Consequently, the expression of meaning indicates what is considered a typical occurrence, encompassing in this determination general truths, logical reasoning, and societal norms.

- (30) A: *kān sáʔ kʰé:k wānkʰ:t lé:w càʔ cʰô:kdī:*
 eat PTC cake birthday LNK TAM lucky
 B: *mān kʰūn mā: tʰūŋ kʰɔ: là pʰi: jā:k nɔ:n*
 PRO up come to neck PTC ADR want.to sleep
jā:ŋdīaw lɔ:j tɔ:nni:
 only LNK now
 A: ‘Eat this birthday cake and you will be lucky.’
 B: ‘I am full up. I just only want to lie down now.’

- (31) A: *jā:k sú: ʔārāj ʔik máj*
 want.to but INDEF additional Q
 B: *máj kʰrǎp mān mət krāpǎw láʔ*
 NEG PTC PRO used.up wallet PTC
 A: ‘Do you want to buy anything else?’
 B: ‘No, I used up all my money.’

The term *mān* in (30) refers to food eaten, although there is no linguistic expression referring to something eaten, while in (31), it can be interpreted as the speaker’s money.

The referents inferred from the semantic features of predicative verbs demonstrate that each verb possesses a specific semantic property that signifies a particular referent. When combined with other contextual information within a sentence, language users can accurately deduce the referents associated with the pronoun *mān*, even in the absence of an explicit antecedent. Here are a few examples illustrating referential inferences derived from the semantic characteristics of verbs.

- (32) *tɔ:n klā:ŋkʰū: mān mu:t mā:k máj klā: ʔə:k pāj nǎj lɔ:j*
 when night PRO dark a.lot NEG dare go.out go INDEF LNK
 ‘It’s very dark at night. (I) don’t dare go outside.’

The term *mān* in this example refers to the surrounding environment because the stative verb *mu:t* ‘dark’ subcategorizes a nominal with a specific semantic property, circumstance.

In addition to context-based referents, there is another category known as speaking-subject referents. When examining events from a subjective standpoint, they exhibit evaluative and judgmental meanings expressed by speakers. This phenomenon diminishes the transparency of the grammatical subject’s role and intensifies the expression of the speaker’s attitudes, ideas, beliefs, and presence. As a result, subjective meaning experiences a corresponding augmentation (Benveniste 1971), as in (13).

- (33) *mān pēnpājmājdā:j tʰi: càʔ ju:t cʰi:wít mǎ: tūa ní: pāj*
 PRO be.impossible REL TAM prolong life dog CLF DEM go
ʔi:k 40-50 pī:
 additional 40-50 year
 ‘It is impossible to prolong this dog’s life by more 40-50 years.’

The term *mān* in (33) is the grammatical subject of the evaluative verb *pēnpājmājdā:j* ‘to be impossible’ referring to the speaker’s judgment towards a certain action *ju:t cʰi:wít mǎ: tūa ní: pāj ʔi:k 40-50 pī:* ‘to prolong this dog’s life for more 40-50 years.’ Notably, the pronoun *mān* occurs in a specific construction, referred to as an ‘evaluative construction’, with a syntactic pattern shown below. This instance illustrates a shift of focus from *mān* as a grammatical subject to that of a speaking subject.

Figure 1: Structural schema of *mān* in an evaluative construction

<i>mān</i>	+	Evaluative verbs	+	<i>tʰi:</i>	Clause
------------	---	---------------------	---	-------------	--------

Lastly, *mān* can also convey a common ground referent as shown in (34).

- (34) A: *ʔāwŋājādī: nā:hùaj wàʔ mān k̄:t ná:mpà: lǎjlà:k*
 DM worrisome PTC PRO exist flash flood overflow
kàp d̄ĩnkʰlo:n tʰālòm tʰě:w sājǰók l̄̄:j
 and mud collapse around Sai Yok exactly
 B: *kʰōn ʔù:n wá: ŋāj pāj k̄n jù: máj*
 person other say Q go RECP stay Q
 A: ‘Well, it’s quite worrisome. There exist flash floods and mudslides right around Sai Yok.’
 B: ‘What do other people say? Will they still be going?’

The term *mān* in the example above, in contrast to the pleonastic subject classified by previous studies, indexes a common ground referent with pragmatic referentiality, as this pronoun embodies the meaning of shared knowledge between the speaker and the listener, specifically in the context of a prearranged excursion to Sai Yok Waterfall with friends. Therefore, the utilization of the pronoun *mān* in A’s utterance serves to evoke this shared knowledge with B, thus exhibiting intersubjective meaning.

In essence, events can manifest a heightened level of subjective meaning which denotes the speaker’s inclination to establish a shared understanding with the listener. Once the interlocutors are in comparable circumstances, they have mutual knowledge. The speaker uses a linguistic expression containing an intersubjective meaning to communicate with the addressee so that the addressee can comprehend the referent to which the word *mān* refers.

The pronoun *mān* indexing a common ground referent functions as the grammatical subject in an existential construction with verbs such as *k̄:t* ‘to occur’, *mī:* ‘to have’, and *prā:kòt* ‘to appear’. Upon analysis, it becomes apparent that the word *mān* acts as a grammatical subject without explicit or implicit antecedents. However, employing *mān* as the subject of these verbs imbues the discourse with subjective meaning, as it integrates the speaker’s thoughts, beliefs, and attitudes into the spoken expression. In other words, the presence of *mān* before these verbs indicates the shared knowledge between interlocutors, thus imparting a relative effect on the speaker’s utterance. If *mān* before the verbs *k̄:t* ‘to occur’, *mī:* ‘to have’, and *prā:kòt* ‘to appear’ is omitted, the sentence would be purely informative.

The examples (30) to (34) illustrate that the presence or absence of an explicit antecedent is not sufficient in comprehending the meaning of *mān*. Moreover, they highlight the importance of context in interpreting the referents of the pronoun as either context-based referents, speaking subject referents, or common ground referents. The coordination between interlocutors, known as a joint activity, significantly impacts the interpretation of this pronoun in addition to factors like speaker, addressee, location, and time. Bezuidenhout (2019) contends that reference (or referring, in her terminology) is a joint activity, emphasizing that it is not words themselves but people who refer by using words in appropriate contexts. This cooperative process of reference, which is not just something the speaker(s) does but also something that results from the combined efforts of two or more interlocutors, produces impossible results for a single person to achieve.

To illustrate this process, speakers initially formulate a clear idea of a specific referent they intend to communicate. Subsequently, they seek an appropriate linguistic device to express that concept, often resorting to pronouns. Ideally, speakers aim to provide sufficient information for addressees to comprehend the intended referent(s). Moreover, speakers can rely on the assumption that addressees can retrieve the intended referent, given the implicit relationship between the linguistic antecedent and the referent. This assumption is grounded in the shared common ground among interlocutors, which encompasses parts of encyclopedic knowledge and is crucial for effective communication.

Upon producing and transmitting the utterance, addressees perceive and decode the message, gaining access to the same conceptual understanding that the speakers intended. This mutual understanding, achieved through effective communication, is not just about speech participants focusing on the same object or scene but also about aligning their knowledge and beliefs. As Bezuidenhout (2019), Roberts (2019), and Diessel

(2019) have noted, communication presupposes that interlocutors share a common ground, and this shared understanding is the ultimate goal of effective communication.

4.3 Referentiality continuum

Excluded from the categorization in Table 1 are ten cases that cannot be unambiguously classified as either semantic or pragmatic referentiality. They crucially serve as compelling evidence affirming the gradient nature of the referentiality continuum. Among these ten instances, notable examples such as (35) and (36) stand out, where discerning whether they represent semantic or pragmatic referentiality is not straightforward.

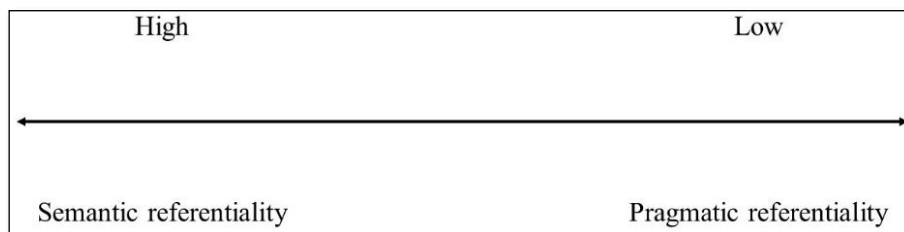
- (35) A: *pī: nùŋ dā:j kʰā:w kì kwīan*
 ear one get rice QNT CL
 B: *tʰā: dī: nəj kô: dā:j pè:t kwīan tʰā: lēw kô: hòk*
 if good PTC LNK get eight CL if bad LNK six
kwīan tʰúŋ cà? dī: jāŋŋāj mān kô: cōn jù: wānjāŋkʰām
 CL to TAM good QNT PRO LNK poor TAM as.always
 A: ‘How many carts of paddy rice do you get each year?’
 B: ‘Eight if it is good or six if it is bad. No matter how much one/I/we get(s), one/I/we still stay(s) poor.’

- (36) *mǔ: lō: lá 120 sú: 20 mān cà? tàt jāŋŋāj*
 pork kilogram each 120 buy 20 PRO TAM cut how
 ‘Pork is 120 baht per kilogram. How should it be cut/how should I cut it if you buy just twenty baht?’

In these two examples, the pronoun *mān* demonstrates a significant level of ambiguity, as its reference is not immediately clear without additional context. In (35), *mān* could refer specifically to the speaker, who identifies as a farmer and is personally experiencing the described circumstances. Alternatively, it might denote farmers in general, representing a broader group of individuals facing similar situations. Similarly, in (36), *mān* can be understood in two distinct ways: it could refer to the pork that is being cut, emphasizing the object of the action, or it might refer to the butcher responsible for performing the cutting. The dual interpretation of both examples highlights the flexibility and context-dependent nature of the pronoun.

This analysis elucidates the continuum scale of referentiality in Thai and underscores that each case exhibits varying degrees of referentiality. Importantly, it supports Li and Thompson’s (1976) claim that pragmatic-based, topic-prominent languages do not allow for non-referential entities, particularly non-referential or so-called dummy or pleonastic subjects. The continuum of referential meaning is composed of two ends: semantic referentiality on one end and pragmatic referentiality on the other. Semantic referentiality refers to the use of *mān* having explicit antecedents in discourse, while pragmatic referentiality refers to the use of *mān* lacking explicit antecedents and, thereby, requires inference to interpret the meaning of *mān*.

Figure 2: Continuum of referentiality degree of the pronoun *mān*



Contextual ambiguities found in examples (37) and (38) can prove the nature of the continuum posited. In example (37), the pronoun *mān* may be interpreted as either referring to the whole discourse or referring to the speaker himself. This example perfectly displays the unclear boundary between semantic referentiality and pragmatic referentiality.

Example (38), in contrast, does not display an ambiguity between semantic and pragmatic referentiality but shows a gradience between subtypes within the category of pragmatic referentiality, as pronoun *mān* may

be interpreted as a speaking subject referent (the speaker itself) or a common ground referent (the entity mentioned earlier).

- (37) *mûa mē: klàp cāk tʰāmṇā:n rāw rú: sùapʰá:*
 when mother come.back from work 1PL disassemble clothes
mē: pʰúa hǎ: sè:tsātā:ŋ lé:w pāj c̄x̄: súa piak piak
 mother to find coins TAM go find shirt wet
kô: tʰǎ:m mē: wâ: tʰāmmāj ʔà:pná:m lé:w mâj cʰáj
 LNK ask mother COMP Q bathe TAM NEG use
pʰá:chétṭua mē: b̀: k wâ: nân kʰū: ŋua kʰŏŋ mē
 towel mother tell COMP DEM COP sweat of mother
fāŋ lé:w rák mē: càpcāj l̄:j m̄n trūŋ cāj m̄:
 listen TAM love mother whole-heartedly LNK PRO touched come
cōntʰúŋ tʰúk wān ní:
 until every day DEM

‘When our mother came back from working, we tore through her clothes to find some money. Then, we found her wet shirt. We asked her why she did not dry herself with a towel, but she replied that it was her sweat. Upon hearing that, we have been feeling whole-heartedly touched until today.’

- (38) A: *t̄:ŋ cʰáj ŋx̄n pʰx̄:m ʔi:k tʰáwràj b̀: k ná?*
 TAM use money add additional Q tell PTC
 B: *mâj t̄ŋ h̀: aŋ kʰráp m̄n tʰām sèt tʰānwē:lā: nē:n̄:n*
 NEG tam worry PTC PRO make finish in.time certainly
 A: ‘Let me know if you need more money.’
 B: ‘No worries. I will get it done/it will be done in time.’

Based on the semantic and pragmatic referentiality and ambiguous cases illustrating a combination of the two types, the pronoun *mān* conveys a referentiality that lies along the continuum below, thereby supporting the claim that referential property is not dichotomous but occurs by degree.

Figure 3: Referentiality continuum of the pronoun *mān* in Thai

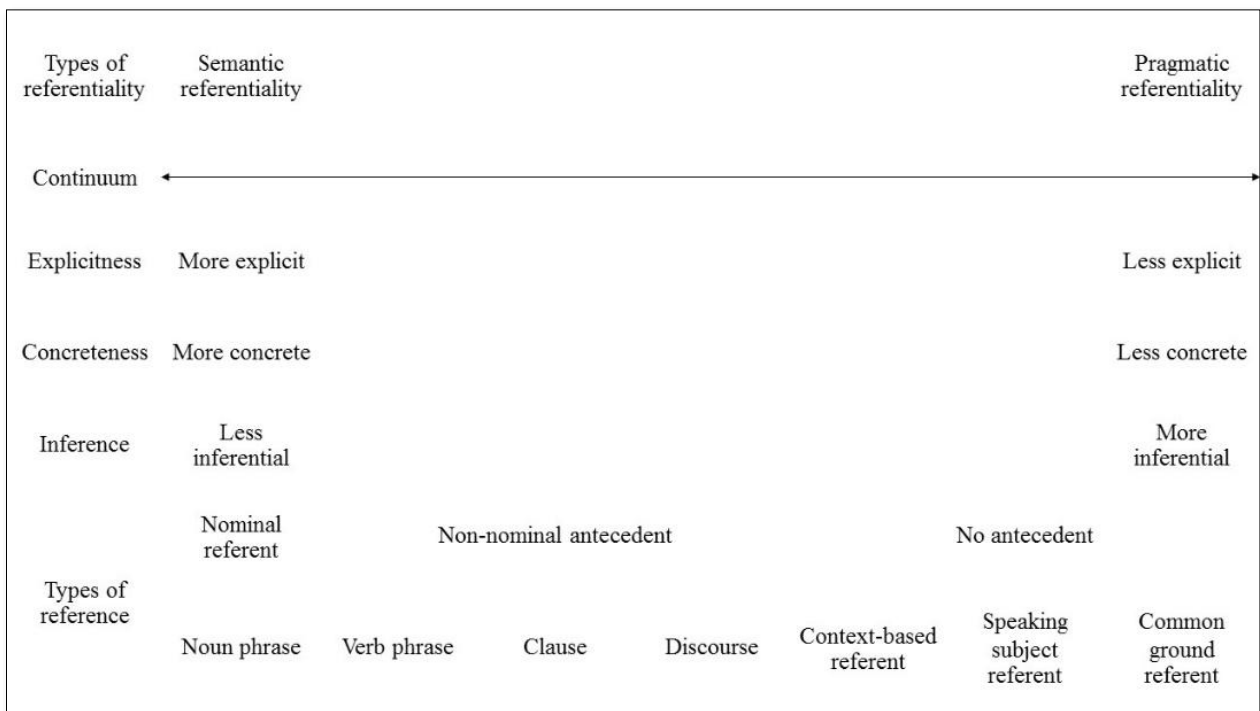


Figure 3 reveals the gradient nature of the referentiality of *mān*. It demonstrates a gradual shift from semantic referentiality, characterized by the presence of explicit antecedents that denote concrete referents, to pragmatic referentiality, which entails implicit antecedents and indicates less tangible or abstract referents. This pragmatic referentiality encompasses the expression of the speaker’s attitude, viewpoint, and shared background with listeners.

5 Discussion

As shown in the previous section, degree of referentiality varies with regard to the type of context in which the pronoun *mān* occurs. These contexts include two main groups: linguistic and non-linguistic. While the former deals with the occurrence of explicit linguistic elements within the discourse, the latter is a matter of the physical surroundings and mutual knowledge among interlocutors beyond a given utterance.

The referential pronoun *mān*, when occurring with an explicit linguistic antecedent, exhibits semantic referentiality. This type of referentiality always co-occurs with linguistic contexts that refer to specific parts of an utterance adjacent to a focused unit, such as a sound or word (Crystal 2008:108-109). The pronoun *mān* without explicit antecedents has also been proven to be referential in some cases when a given context is sufficient for its referent retrieval. Such a case therefore exhibits pragmatic referentiality, which is defined by its property of context-induced inference as shown in the example (30) reduplicated in (39).

- (39) A: *kīn sá? kʰé:k wānkʰ:t lé:w cà? cʰòkdī:*
 eat PTC cake birthday LNK TAM lucky
 B: *mān kʰūn mā: tʰūŋ kʰɔ: là pʰi: jà:k nɔ:n*
 PRO up come to neck PTC ADR want.to sleep
jà:ŋdāw lɔ:j tɔ:nní:
 only LNK now
 A: ‘Eat this birthday cake and you will be lucky.’
 B: ‘I am full. I just want to lie down now.’

It is obvious that the pronoun *mān* in this example has no explicit antecedent. However, the situational context relating to an eating event invites us to interpret *mān* as any food being eaten. This strongly confirms that the use of the pronoun *mān* without explicit antecedent is not normally identical to its non-referential use. It is thus unacceptable to treat all antecedent-lacking cases as non-referential uses of *mān* as claimed in previous studies. Referentiality degree of *mān* should, rather, be divided into two subcategories lying on a continuum between semantic and pragmatic referentiality.

The assumed pragmatic referentiality of the pronoun *mān* is theorized to be grounded in its usage patterns. Specifically, when employed as a subject, *mān* refers to an entire discourse characterized by an indistinct syntactic boundary and multiple events. When *mān* is used as a subject, its referential function goes beyond a single syntactic unit to include the whole discourse. This interpretation of pragmatic referentiality shows how determining the pronoun *mān*’s functions in its usage is complicated. The nuanced nature of this reference becomes apparent in its ability to encapsulate diverse events within a discourse, contributing to the complexity of the pragmatic functions attributed to the pronoun *mān*. This analysis highlights the dynamic nature of referentiality in linguistic usage. It stresses the importance of looking at both syntactic and contextual factors when studying the pragmatic aspects of pronoun usage, especially when the referential scope goes beyond normal syntactic limits. This is perfectly aligned with what Gundel, Hedberg, and Zacharski (2005) found in their study on third-person pronoun use in a casual English conversation. Specifically, their findings exhibited that it is a somewhat difficult task to judge whether a personal pronoun is referential or not, as some inferable pronouns were found to be referentially vague.

The findings contribute to cross-linguistic studies on expletive pronouns, particularly in the context of pro-drop languages, by providing a nuanced understanding of their roles in discourse. In pro-drop languages such as Thai and Vietnamese, the third-person, non-gendered, and inanimate pronouns *mān* and *nó* often serve expletive-like functions, enhancing discourse coherence without being constrained by the syntactic necessity of subject filling. For instance, *mān* in Thai parallels *nó* in Vietnamese, which Greco, Phan, and Haegeman (2017) describe as functioning effectively without explicit antecedents, yet still contributing meaningfully to

textual clarity. This highlights a broader typological pattern where expletive-like pronouns facilitate cohesion in languages that, unlike English, do not mandate overt syntactic subjects.

These findings have theoretical implications for the typology of expletive pronouns. By situating *mān* within a continuum framework, this study highlights how languages such as Thai achieve discourse coherence through flexible pronominal systems that operate independently of strict syntactic requirements. This framework challenges the traditional binary distinction between expletive and referential pronouns, suggesting a gradience in their discourse roles. Furthermore, a comparative analysis of *mān* in Thai and *nó* in Vietnamese, particularly in contrast to true syntactic expletives such as *it* in English, reveals cross-linguistic variation in how coherence is maintained.

6 Summary

The pronoun *mān* exhibits referentiality, which can be divided into two categories: semantic and pragmatic. The former refers to cases in which *mān* substitutes a nominal expression of directly involved concrete entities in the real world, while the latter covers cases in which the pronoun in question refers to an implicit antecedent, thereby requiring inference. These two distinct categories of referentiality are not an either/or alternative but lie on a continuum with semantic referentiality on one end and pragmatic referentiality on the other. This study brings fresh insight to *mān*, which refers not only to nominal expressions but also non-nominal categories verb phrases, clauses, and discourses, thereby supporting the gradient nature of referentiality.

By framing *mān* as fulfilling a pragmatic rather than strictly syntactic role, this study paves the way for future research into the functions of expletive-like pronouns in topic-prominent and pro-drop languages. This perspective opens new avenues for examining how such pronouns contribute to discourse coherence and textual cohesion across typologically diverse languages. Future research could test the proposed continuum model in other topic-prominent or pro-drop languages, exploring parallels and divergences in the functions of expletive-like pronouns. For instance, extending the analysis to other Southeast Asian languages, such as Khmer or Burmese, could provide a broader understanding of regional typological patterns.

Contributions of Co-authors

Jinawat Kaenmuang conducted the literature review, developed the theoretical framework, designed the methodology, interpreted the data, and drafted all sections of the paper. Piroon Piyamahapong contributed to data interpretation and drafting all sections. Pittayawat Pittayaporn supervised the project, reviewed the data analysis and interpretation, provided critical feedback during the drafting process, and assisted in revising the manuscript for final submission.

Acknowledgments

Acknowledgments are extended to Chulalongkorn University Graduate School for their generous support through the Chulalongkorn University Graduate Scholarship to Commemorate the 72nd Anniversary of His Majesty King Bhumibol Adulyadej and the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) awarded to the first author.

Additionally, we extend our sincere appreciation to Mark Alves, Mathias Jenny, and the anonymous reviewers associated with the Journal of the Southeast Asian Linguistics Society. Their critical evaluation, insightful feedback, and constructive comments have significantly contributed to the refinement and academic rigor of this paper.

References

- Abbott, Barbara. 2010. *Reference*. Oxford: Oxford University Press.
- Angkapanichkit, Jantima. (จันทิมา อังคพานิชกิจ). 2011. waiyakon phasa thai kap ayu khong phuphut. (ไวยากรณ์ภาษาไทยกับอายุของผู้พูด) [Thai grammar and age of speakers]. Bangkok: Thailand Research Fund.
- Aroonmanakun, Wirote. 2007. Creating the Thai National Corpus. *Manusaya: The Journal of Humanities*. 13:4-17.
- Bandhmedha, Navavan. (นาวรรณ พันธุเมธา). 2022. waiyakon thai. (ไวยากรณ์ไทย) [Thai grammar]. 9th edn. Bangkok: Faculty of Arts, Chulalongkorn University.

- Benveniste, Émile. 1971. Subjectivity in language. In Mary Elizabeth Meek (ed.), *Problems in general linguistics*, 223-230. Florida: University of Miami Press.
- Bolognesi, Marianna., Burgers, Christian., & Caselli, Tommaso. 2020. On abstraction: Decoupling conceptual concreteness and categorical specificity. *Cognitive processing* 21.3:365-381. <https://doi.org/10.1007/s10339-020-00965-9>
- Bybee, Joan. 2023. What is usage-based linguistics? In Manuel Diaz-Campos & Sonia Balasch (eds.), *The handbook of usage-based linguistics*, 7-29. New Jersey: Wiley-Blackwell.
- Carlson, Gregory. 2006. Reference. In Laurence Robert Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 74-96. Oxford: Blackwell.
- Cooke, Joseph Robinson. 1968. *Pronominal reference in Thai, Burmese, and Vietnamese*. Berkeley: University of California Press.
- Chandler, Daniel & Munday, Rod. 2011. *A dictionary of media and communication*. Oxford: Oxford University Press.
- Chen, Ping. 2009. Aspects of referentiality. *Journal of Pragmatics* 41.8:1657-1674. <https://doi.org/10.1016/j.pragma.2009.03.002>
- Chen, Ping. 2015. Referentiality and definiteness in Chinese. In William S-Y Wang & Chaofen Sun (eds.), *The Oxford handbook of Chinese linguistics*, 404-413. Oxford: Oxford University Press.
- Chen, Jing., Huang, Huaxin, & Wu, Yicheng. 2010. Aspects of pragmatic referentiality. *Journal of Pragmatics*, 42.3:870-874. <https://doi.org/10.1016/j.pragma.2009.08.001>
- Chutisilp, Pornpimol. (1984). *A sociolinguistic study of an additional language: English in Thailand*. Illinois: United States. (Doctoral dissertation).
- Crystal, David. 2008. *A dictionary of linguistics and phonetics*. 6th edn. Oxford: Blackwell.
- Diessel, Holger. 2019. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.
- Finegan, Edward. 2014. *Language: Its structure and use*. 7th edn. Connecticut: Cengage Learning.
- Fretheim, Thorstein., & Gundel, Jeanette K. (eds.). 1996. *Reference and referent accessibility*. Amsterdam: John Benjamins.
- Givón, Talmy. 1978. Definiteness and referentiality. In Joseph Harold Greenberg (ed.), *Universal of human language: Vol. 4: Syntax*, 291-330. California: Stanford University Press.
- Greco, Ciro., Phan, Trang., & Haegeman, Liliane. 2018. On *nó*, an optional expletive in Vietnamese. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in Generative Grammar*, 31-52. Oxford: Oxford University Press.
- Gundel, Jeanette., Hedberg, Nancy., & Zacharski, Ron. 2005. Pronouns without NP antecedents: How do we know when a pronoun is referential. In António Branco, Tony McEnery, & Ruslan Mitkov (eds.), *Anaphora processing: Linguistic, cognitive and computational modelling*, 351-364. Amsterdam: John Benjamins.
- Huddleston, Rodney., & Pullum, Geoffrey Keith. 2015. *The Cambridge grammar of the English language*. 8th edn. Cambridge University Press.
- Indrambarya, Kitima. 2011. *On the non-referential subject /man/ 'it' in Thai* [Paper presentation]. The 21st Annual Meeting of the Southeast Asian Linguistics Society, Bangkok, Thailand, 11-13 May 2011.
- Indrambarya, Kitima. 2012. *On impersonal subjects in Thai* [Paper presentation]. The 22nd Annual Meeting of the Southeast Asian Linguistics Society, Agay, France, 30 May-2 June 2012.
- Iwasaki, Shoichi, & Ingkaphirom, Preeya. 2005. *A reference grammar of Thai*. Cambridge: Cambridge University Press.
- Kullavanijaya, Pranee. 2000. Power and intimacy: A contradiction in a Thai personal pronoun. In Videia P. De Guzman & Byron Wilbur Bender (eds.), *Grammatical analysis: Morphology, syntax and semantics: Studies in honor of Stanley Starosta*, 80-86. Honolulu: University of Hawaii Press.

- Lee, Benny P. H. 2001. Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics* 33.1:21-44. [https://doi.org/10.1016/S0378-2166\(99\)00128-9](https://doi.org/10.1016/S0378-2166(99)00128-9)
- Levinson, Stephen. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Li, Charles N., & Thompson, Sandra A. 1976. Subject and topic: A new typology of language. In Charles N. Li (ed.), *Subject and topic*, 457-489. New York: Academic Press.
- Li, Xiaoshi., Chen, Xiaoqing., & Chen, Wen Hsin. 2012. Variation of subject pronominal expression in Mandarin Chinese. *Sociolinguistic Studies*, 6.1:91-119. <https://doi.org/10.1558/sols.v6i1.91>
- Lyons, John. 1977. *Semantics, Volume I*. Cambridge: Cambridge University Press.
- MacDonald, Jonathan E. 2017. An implicit projected argument in Spanish impersonal- and passive se-constructions. *Syntax*, 20.4:353-383. <https://doi.org/10.1111/synt.12146>
- Panupong, Vichin. (วิจิตร ภาณุพงศ์). 1991. khongsang khong phasa thai: rabop waiyakon. (โครงสร้างของภาษาไทย: ระบบไวยากรณ์) [The structure of Thai: Grammatical system]. 11th edn. Bangkok: Ramkhamhaeng University Press.
- Perek, Florent. 2023. Construction grammar and usage-based theory. In Manuel Diaz-Campos & Sonia Balasch (eds.), *The handbook of usage-based linguistics*, 215-231. Oxford: Blackwell.
- Pipalova, Renata. 1988. Some remarks on delimiting reference in English. *Philologica Pragensia*, 70.4:192-198.
- Prasithratsint, Amara. (อมรา ประสิทธิ์รัฐสินธุ์). 2013. chanit khong kham nai phasa thai: kan wikhro thang wakkayasamphan. (ชนิดของคำในภาษาไทย: การวิเคราะห์ทางวากยสัมพันธ์) [Parts of speech in Thai: A syntactic Analysis]. Bangkok: A. S. P. Publishers.
- Quesada, Teresa., & Lozano, Cristóbal. 2020. Which factors determine the choice of referential expressions in L2 English discourse?: New evidence from the COREFL corpus. *Studies in Second Language Acquisition* 42.5:959-986. <https://doi.org/10.1017/S0272263120000224>
- Saeed, John. I. 2022. *Semantics*. 5th edn. Oxford: Blackwell.
- Sainsbury, Richard Mark. 2005. *Reference without referents*. Oxford: Clarendon Press.
- Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25:701-721. <https://doi.org/10.1023/A:1020867916902>
- Yule, George. 1996. *Pragmatics*. Oxford: Oxford University Press.
- Warotamasikkhadit, Udom (อุดม วโรตม์สิขกคิตต์). 1996. waiyakon thai nai phasasat. (ไวยากรณ์ไทยในภาษาศาสตร์) [Thai grammar in linguistics]. Bangkok: Ramkhamhaeng University Press.
- Upakittisinlapasan, Phraya. (พระยาอุปกิตศิลปสาร). 2001. lak phasa thai. (หลักภาษาไทย) [Principles of Thai grammar]. 10th edn. Bangkok: Thai Watthana Phanit.
- Uryupina, Olga., Kabadjoy, Mijail., Poesio, Massimo. 2016. Detecting non-reference and non-anaphoricity. In Massimo Poesio, Roland Stuckardt, & Yannick Versley (eds.) *Anaphora resolution: Theory and applications of natural language processing*. Heidelberg: Springer. https://doi.org/10.1007/978-3-662-47909-4_13

Appendix: List of abbreviations

ADR	addressive	CL	clause
CLF	classifier	COMP	complementizer
COP	copula	DEM	demonstrative
DISC	discourse	DM	discourse marker
DOM	direct object marker	IMPERS	impersonal
INDEF	indefinite	LNK	linker
NEG	negative	NP	noun phrase
PASS	passive	PL	plural
PRO	pronoun	PTC	particle
Q	question	QNT	quantifier
RECP	reciprocal	REFL	reflexive

REL	relativizer	SG	singular
TAM	tense-aspect-mood	VP	verb phrase
1,2,3	grammatical person		

Reviewed: Received 23 July 2024, revised text accepted 22 December 2024, published 1 February 2025

Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Sigrid Lew, Dr Mathias Jenny, Dr Kenneth Van Bik