

John Benjamins Publishing Company



This is a contribution from *Interdisciplinary Perspectives on Vietnamese Linguistics*.

Edited by Nigel Duffield, Trang Phan and Tue Trinh.

© 2019. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Effects of grammatical roles and parallelism on referential form production in Vietnamese spoken and written narratives

Binh Ngo, Elsi Kaiser and Andrew Simpson

University of Southern California

This study investigates the use of null and overt pronouns and noun phrases in Vietnamese spoken and written narratives, with a focus on referents' grammatical roles and grammatical parallelism. Looking at Vietnamese allows us to address questions left open in prior work regarding the effects of pronominal form (null, overt) on reference resolution in different grammatical positions. Furthermore, looking at the use of Vietnamese pronouns adds typological breadth to the literature on reference resolution: Vietnamese overt pronouns differ typologically from pronouns in English-type languages as they also function as kin terms. Results from data we collected involving spoken and written narratives show that referential form choice is influenced not only by the grammatical role of the antecedent but also the grammatical role of the pronominal element. When the subject of the current clause refers to the subject of the preceding clause (subject parallelism), we find a high rate of (null and overt) pronouns. Lack of parallelism triggers mostly NPs. When the object of the current clause refers to the object of the preceding clause (object parallelism), more pronouns were produced than in non-parallel cases. Crucially, we find no clear differences in the distribution of null vs. overt pronouns, suggesting that grammatical roles and parallelism have the same effects on both pronoun types. Our results also show no effects of written vs. spoken modality, which indicates that modality does not play a role in the interaction between grammatical factors and referential form choice.

Keywords: reference resolution, pronouns, Vietnamese, subjecthood, parallelism, narrative

1. Introduction

It is widely agreed that entities in a discourse vary in their salience/prominence: At a particular point in time, some entities are more salient or prominent in the discourse participants' mental models than other entities. Prior work suggests that the salience level of entities influences speakers' referential form choice as well as comprehenders' interpretation of referential forms (Ariel 1990; Givón 1983; Gundel et al. 1993). It is frequently suggested that more reduced referential forms tend to be used for highly salient referents while fuller referential forms tend to be used for less salient referents. Thus, if a language has both null and overt pronominal forms, null pronouns are often used to refer to highly salient referents while overt pronouns are used to refer to less salient referents, as shown in (1).

- (1) *Most salient referents* ----- *Less salient referents*
 Null pronouns Overt pronouns NPs

The claim that there exists a relationship between the salience of the referent and the type of referring expressions leads to the question of what influences how salient referents are. Prior work indicates that referents' salience¹ can be influenced by a number of factors, including the grammatical role of the antecedent (for example, subject vs. object) (Chafe 1976; Crawley & Stevenson 1990) and whether the pronoun and its antecedent occupy parallel grammatical roles (i.e. both elements are in subject position or in object position) (Smyth 1994; Chambers & Smyth 1998). The work we report in this paper builds on this insight that referential form use depends not only on the grammatical role of the antecedent but also on the grammatical role of the anaphoric form. As we discuss below, theories of referential form cannot focus solely on a notion of salience derived on the prior realisation of the antecedent but also have to take into account the argument structure of the anaphor-containing sentence.

Many of the fundamental studies on grammatical parallelism effect have largely focused on English and English overt pronouns (e.g., Smyth 1994; Stevenson et al. 1995; Chambers & Smyth 1998). Consequently, even though it is widely known that

1. Terms such as 'referent' and 'referring expression' are standardly used in psycholinguistic work on pronoun interpretation and production. In this paper, we use the term 'referent' to mean not only the entity in the world that a certain linguistic element picks out, but – more relevantly for our present purposes – the linguistic realisation of that entity. Thus, we will often say, for the sake of brevity, that a particular referring expression refers to a preceding subject or object – even though this is not strictly speaking correct, since the referring expression refers not to the grammatical role of subject or object but to the entity which occurs/is linguistically realised in subject/object position.

null and overt pronouns across languages have different properties (e.g., Spanish: Alonso-Ovalle et al. 2002; Italian: Carminati 2002; Japanese: Clancy 1980; Chinese: Li & Thompson 1979), to the best of our knowledge, little is known about the extent to which grammatical parallelism can affect the comprehension and production of null and overt pronouns.

In this paper, we report a narrative study on Vietnamese, a language that allows null and overt pronouns in both subject and object position. We examine how and whether Vietnamese speakers' choice of referential forms, particularly null and overt pronouns, is influenced by (i) *the grammatical role of the antecedent* and (ii) *the grammatical role of the referring expression* – in particular, whether they have the same grammatical role (grammatical parallelism) or not. Thus, the first aim of our work is to shed light on referential form choice in a context where the alternation between null and overt pronominal forms has not previously been systematically considered. Our second aim is to investigate potential differences between spoken and written language in the use of referential forms, i.e. possible effects of language 'modality' – the physical means used to express language, with speech, writing or gestural communication. Prior work suggests that spoken and written language differ with regards to kinds of referential forms that are regularly produced in these different modalities (see Chafe & Tannen 1987 for an overview). Generally speaking, it has been suggested that pronouns are more common in spoken than in written language, and that full NPs are used more frequently in written language (e.g., Biber et al. 1999; Christensen 2000). However, these studies mostly discuss overall counts and many of them contain data from different genres with various levels of formality. Thus, it is difficult to know whether the differences are due to modality *per se* or to other properties that have been correlated with modality in these prior studies.

Thus, in the current study, we carefully consider the effect of modality (written vs. spoken) on the choice of referential form, while keeping the genre and level of formality constant by using explicit instructions. This allows us to test for potential differences between written and spoken language more directly.

The structure of this paper is as follows: In the remainder of Section 1, we discuss previous findings on the effects of grammatical role, grammatical parallelism and modality on referential form interpretation and production. We also discuss the nature of the Vietnamese pronominal system and compare it to other *pro*-drop languages and other pronominal systems. In Section 2, we describe the spoken and written narratives tasks that we used to elicit data as well as how the data was analyzed. Section 3 presents the results from the written and spoken tasks and provides a comparison between the two types of data. Section 4 discusses the implications of our findings, compares them to findings from other languages, and outlines directions for future work.

1.1 Grammatical roles and grammatical parallelism

One well-known factor that influences referents' salience is grammatical role (i.e. being realised in subject or object position) (Chafe 1976; Brennan, Friedman & Pollard 1987; Crawley & Stevenson 1990; see also Gordon, Grosz, & Gilliom 1993; Gordon & Chan 1995; Perfetti & Goldman 1974). To identify salient referents, prior work has often used pronoun interpretation or subsequent mention likelihood as a diagnostic. In one of the earliest works on this topic, Chafe (1976) presented a number of observations and argued that subjects indeed have a special prominent cognitive status – for example, that knowledge about subjects is more readily accessible than knowledge about other parts of sentences. The special status of subjects has been confirmed in many subsequent studies.

Recent work by Fukumura & van Gompel (2010) used sentence-continuation tasks to investigate whether and to what extent the production of pronouns in English is influenced by semantic biases (induced by verbs and connectives such as 'because') and the grammatical roles of potential antecedents (subject *vs.* object). Fukumura and van Gompel found that participants produced more pronouns (relative to names) when referring to the preceding subject than to the preceding object, regardless of the semantic biases of verbs and connectives. These results add to the body of literature showing that grammatical subjects are privileged as antecedents of subsequent pronouns.

The effect of grammatical roles is also reflected in parallelism effects (Smyth 1994; Stevenson et al. 1995; Chambers & Smyth 1998). Chambers & Smyth (1998) found that pronouns, at least in English, tend to prefer antecedents in matching grammatical positions: Pronouns in subject position tend to be interpreted as referring back to preceding subjects, and pronouns in object position tend to be interpreted as referring back to preceding objects. However, to the best of our knowledge, work on grammatical parallelism effects has focused on English (overt) pronouns and has not systematically looked at the null *vs.* overt pronoun distinction.

Although the null *vs.* overt distinction has not been investigated systematically in parallelism configurations, a large body of prior work has investigated the referential properties of null and overt pronouns in subject position. Before continuing on to review this prior work, it is important to note that broadly speaking, languages with both null and overt pronouns come in two types: *pro*-drop languages which have rich subject-verb agreement (i.e. 'agreement *pro*-drop languages'), and 'discourse *pro*-drop' languages (Barbosa 2011; Neeleman & Szendrői 2007), which typically lack verb agreement and permit *pro* in subject and object positions subject to discourse recoverability. Prior work on pronoun interpretation in agreement *pro*-drop languages such as Italian and Spanish has led researchers to conclude that the antecedent's grammatical role is crucial for the use and interpretation

of null and overt pronouns in subject position: while null pronouns tend to refer back to preceding subjects, overt pronouns tend to refer to preceding objects (e.g., Alonso-Ovalle et al. 2002; Carminati 2002, but see Fedele 2016 for some Italian data that points to a more nuanced picture).

1.2 Discourse *pro*-drop languages

In this paper, we focus on referential forms in a discourse *pro*-drop language – Vietnamese – for several reasons.² First, discourse *pro*-drop languages typically have null and overt pronouns occurring in both subject and object position. This distributional property allows us to expand the investigation beyond overt pronouns and subject pronouns. In addition, the availability of null and overt pronouns in both subject and object position means that we can investigate the full range of parallel and non-parallel configurations (as explained below in Section 1.1) with both null and overt pronouns. This would not be possible if we were to investigate *pro*-drop languages which have strict/heavy constraints on the use of null pronouns in object position. Thus, discourse *pro*-drop languages are an ideal tool to explore the interaction between pronominal form, the grammatical role of the antecedent (subject or object), and crucially, also the grammatical role of the referring expression (subject or object).

Previous work suggests that the null vs. overt pronoun distinction in discourse *pro*-drop languages appears to be less clear than in *pro*-drop languages. Several studies looking at Chinese pronouns in narratives suggest that the choice between null and overt pronouns appears to be in free variation and reflects speakers' personal interpretations of the discourse context (Li & Thompson 1979) as well as speakers' personal preferences (Christensen 2000). However, while the results in Li & Thompson (1979) suggest that null pronouns seem to be the common, default form, other work (Chen 1986; Christensen 2000) found that both null and overt pronouns are used frequently in narratives. These studies indicate that speakers' choice and discourse structure have the main influence on the use of null and overt pronouns in Chinese.

In contrast, other work shows that null and overt pronouns in Chinese are strongly influenced by syntactic structure. In terms of comprehension, Yang et al. (1999, 2003) conducted a number of self-paced reading studies and found that

2. Vietnamese is naturally classified as a discourse *pro*-drop language as it is not a language in which verbal agreement licenses the occurrence of *pro*, unlike Spanish and Italian. Vietnamese patterns like other discourse *pro*-drop languages such as Chinese and Japanese, where *pro* is essentially available whenever its content can be recovered from the ongoing discourse context (Barbosa 2011).

grammatical role (subject vs. object) has a strong effect on how rapidly pronouns are read in Chinese. For example, in a reading-time study reported in Yang et al. (1999), participants slowed down when repeated names rather than null or overt pronouns were used to refer to prior-mentioned referents – but only when the repeated names occurred in subject position (see also Gordon et al. (1993) on the repeated name penalty in English). In fact, Yang et al.'s (1999) follow-up study found that slow-downs only occurred when a repeated name in subject position was used to refer back to a preceding subject (subject parallelism). In addition, Yang (2003) found that participants read subject pronouns faster when they referred back to the preceding subject than to the preceding object, even in contexts that favored object interpretations. As a whole, these findings show that Chinese null and overt pronouns in subject position have an interpretation preference toward antecedents in subject position.

In related work, Simpson et al. (2016) examine the comprehension of Chinese overt pronouns in subject position through a series of sentence completion experiments. These studies mostly focus on transfer-of-possession verbs (e.g., *send*, *give*, *kick*). Simpson et al. (2016) found that participants tend to interpret overt subject pronouns in the continuations as referring back to the preceding subject. Although this tendency can be modulated by other discourse factors such as the nature of the event (e.g., perfective vs. imperfective) and the type of coherence relation (e.g., Explanation vs. Occasion), evidence for a subject preference remains strong. Put together, the results in Yang (1999, 2003) and Simpson et al. (2016) emphasise the importance of the antecedent's grammatical role. However, these studies did not explore the production aspect of null and overt pronouns. Furthermore, they have only focused on subject pronouns and have not yet examined object pronouns. Thus, further work is needed to obtain a more complete picture.

Related work has been conducted in Japanese, another discourse *pro*-drop language. Hinds (1975, 1983) and Clancy (1980, 1982) investigated Japanese null and overt subject pronouns by means of questionnaires, conversations and narratives and found that the use of overt pronouns in Japanese is very restricted, compared to null pronouns. One potential explanation for this restriction lies in the fact that Japanese overt pronouns are historically derived from nouns and exhibit semantic and syntactic behaviors similar to nouns (Kuroda 1965), which is different from pronouns in other languages such as Chinese. For example, *kare* in Japanese can function as a pronoun meaning 'he' and a noun meaning 'boyfriend'. *Kare* can also take modifiers and determiners similar to nouns do (e.g., *ureshii kare* 'happy guy') (Hinds 1975). It is important to note that although null pronouns are not found in traditional-styled narratives in Japanese – according to Clancy (1980) – they can occur in daily conversations; see Hinds (1975, 1983); Amano & Kondo (2000).

Null and overt pronouns in subject position in Japanese have also been examined by means of experimental work. Ueno and Kehler (2016) conducted a series of sentence completion studies on the interpretation of Japanese null and overt pronouns. Their experiments employed transfer-of-possession verbs as well as implicit causality verbs (e.g., *surprise*, *praise*). Similar to Kehler & Rohde's (2013) work in English, they had both pronoun-prompt (comprehension) and no-prompt conditions (production): Participants either had to interpret an overt subject pronoun before providing their continuations or they could freely use whatever referential form they preferred. Similar to Simpson et al.'s (2016) study on Chinese, perfective and imperfective aspect were also manipulated. Furthermore, since Japanese has topic marking, Ueno and Kehler also manipulated topichood using topic *vs.* nominative marking on the preceding subject. The results of Ueno and Kehler (2016) show that Japanese overt pronouns in subject position, similar to English overt pronouns, are sensitive to a number of pragmatic factors (e.g., (im)perfective marking, implicit causality bias). In contrast, null subject pronouns have much less sensitivity to pragmatic manipulations, none for the (im)perfective manipulation and only limited sensitivity to the implicit causality manipulation. Nevertheless, both Japanese null and overt pronouns in subject position exhibit a subject bias similar to what has been found for Chinese subject pronouns.

In sum, cross-linguistically, it is unclear whether null and overt pronouns in subject position behave differently and how the grammatical role of the antecedent can affect the use of null and overt pronouns in discourse *pro*-drop languages. Furthermore, null and overt pronouns in object position have not been systematically investigated in prior work.

1.3 Vietnamese

We choose Vietnamese – a discourse *pro*-drop language – as the language of investigation for two reasons. First, Vietnamese allows both null and overt pronouns in both subject and object positions as shown in Example (2). In (2b), null pronouns are used to refer back to both the preceding subject and object while in (2b'), an equivalent of (2b), overt pronouns are used. (Null pronouns are denoted with parentheses in the translation.)

- (2) a. *Vân nhìn thấy Nam trên đường về nhà.*
 Vân saw Nam on way back home
 'Vân saw Nam on her way home.'
- b. *Gọi mấy lần nhưng anh không nghe.*
 call several time but he not hear
 '(She) called (him) several times but he didn't hear (her).'

- b'. *Cô gọi anh mấy lần nhưng anh không nghe cô.*
 she call he several time but he not hear she
 'She called him several times but he didn't hear her.'

Second, unlike many other languages discussed in the pronoun resolution literature, Vietnamese overt pronouns are most commonly derived from kinship terms.³ In Example (3a), the element *ông* is used as a kinship term and is interpreted with its literal meaning 'grandfather.' In (3b), *ông* is used as part of a compound noun and no longer has the literal kin term interpretation 'grandfather' but contributes the meaning of 'old male' to the compound. In (3c), *ông* is used as an overt pronoun, where it again does not mean 'grandfather', but is used in a pronominal way for anaphoric reference to some antecedent in the discourse which has the properties of being male and old.

- (3) a. *Ông của Lan vừa đến.*
 grandfather of Lan just arrive
 'Lan's grandfather just arrived.'
- b. *Ông nông dân đang hái trái cây.*
 old.male.farmer PROG pick fruit
 'The farmer is/was picking fruit.'
- c. *Ông hái từng trái một.*
 old.male.he pick each fruit at once
 'He picked the fruit one by one.'

Thus, many elements which are used in a typically pronominal way in Vietnamese also appear in other linguistic contexts, incorporated into larger compound words frequently depicting professions and as pure kinship terms with relational meanings. The use of such elements as pronouns is established by means of two criteria. First, as pronouns, such elements do not project their literal kin term meaning ('grandfather', 'uncle', 'aunt' etc.), but communicate more general information about gender and age. Second, in their pronominal use, elements such as *ông*, *cô*, *anh* and *bà* occur either bare (i.e. not part of a larger compound word) or with a demonstrative modifier, e.g., *ông ấy* (lit. 'that old male person'). In the pear story narratives investigated in the current study, elements such as *ông*, *cô*, *anh* and *bà* were never

3. In addition to the extensive list of kinship pronouns, Vietnamese also has a small set of pronouns which do not come from kinship terms (e.g., *nó* 'he/she/it', *họ* 'they'). The pronoun *nó* occurred very infrequently (only one participant used *nó*) in the narratives: participants had a strong preference for using kinship pronouns. However, occurrences of *nó* were nevertheless counted as an overt pronoun whenever this element did occur. The pronoun *họ* 'they' was also very infrequent. We did not include the few occurrences of *họ* in our analyses, because this pronoun is often ambiguous in terms of which group it refers to and therefore difficult to code.

used as kin terms encoding a relational meaning to others in the storyline, but occurred either as parts of larger compounds, when a discourse participant was introduced (and sometimes referred back to at a much later point), or as pronouns, when reference was made to some other NP in the discourse.

This kin term pronoun system distinguishes Vietnamese from other discourse *pro*-drop languages such as Chinese and Japanese which have previously been studied. Chinese overt pronouns are similar to English-type pronouns in that they only denote number (also gender in third person pronouns in written Chinese) (Li & Thompson 1981). Meanwhile, as shown in Section 1.2, Japanese overt pronouns have noun-like behaviors (Kuroda 1965; Hinds 1975, 1983). More importantly, the difference in Chinese vs. Japanese overt pronoun systems is also correlated with different patterns of use: Previous work on Chinese narratives shows that both null and overt pronouns are frequently used (Christensen 2000). However, in Japanese narratives, null pronouns are the most frequent form while overt pronouns occur only rarely (Clancy 1980, 1982). Null pronouns in Japanese are considered as the equivalent of English pronouns. In contrast, Japanese overt pronouns have very restrictive use with specific connotations (see Hinds 1975 for a full discussion) and their occurrences are often considered to be due to influence of Western languages such as English. Thus, among discourse *pro*-drop languages, null and overt pronouns vary greatly in their properties and usage. A closer look at the typologically different kinship pronoun system in Vietnamese can contribute valuable information regarding pronoun behavior cross-linguistically.

In this paper, we present our work on narratives as an initial investigation of null and overt pronouns in Vietnamese. We also aim to draw a direct comparison between pronouns in Vietnamese and in other discourse *pro*-drop languages. Since previous studies on pronouns in Chinese and Japanese which also discuss spoken and written modality have used narratives (Christensen 2000; Clancy 1982), we also use a narrative task to keep our study maximally parallel to prior work.

Most importantly, we are interested in the effects that grammatical roles of the antecedent and of the referring expression itself have on referential form choice in *both subject and object positions*. A sentence completion task is typically used to investigate referents' subsequent mentions in subject position but not in object position. Therefore, a narrative task which allows us to examine referents' occurrences in both subject and object positions is better suited for our purposes.

Furthermore, in order to test for potential effects of spoken vs. written modality, we keep the number of referents and the genre constant in both written and spoken modalities. Prior work on narratives and modality only reports overall counts of referential forms without details about the grammatical positions of their occurrences (Clancy 1980; Christensen 2000). Additionally, many of these studies also look at written and spoken data in different genres (news vs. conversational) (Biber

et al. 1999). Thus, the differences found may be due to the discourse type and not modality. Taking these factors into consideration, our study maintains maximal parallelism between our spoken and written narratives in genre and formality. We also include grammatical roles and grammatical parallelism in our analysis. Our goal is to shed light on the mechanisms licensing referential forms and to examine whether modality (i.e. the use of spoken vs. written language) indeed has a direct influence on these mechanisms.

2. Data collection

We used a narrative task based on the Pear film, similar to the narratives used in work on Chinese (Christensen 2000) and Japanese (Clancy 1980). The experiment consisted of two parts, spoken and written. Prior work on Chinese and Japanese either only discussed the overall counts of referential forms (Christensen 2000) or how referential forms are used with regards to discourse structure; e.g., the number of intervening clauses, number of intervening referents (Clancy 1980). In contrast, our study focuses on the mechanisms licensing referential form choice (null *vs.* overt). Thus, we incorporate factors such as (i) *grammatical roles of the antecedent and of the pronominal element* and (ii) *grammatical parallelism* into our analysis and examine their influence on referential form use in both spoken and written modalities.

2.1 Method

Twenty native speakers of Vietnamese (living in Vietnam) participated in the experiment. First, each participant was shown the Pear film (Chafe 1980) about a boy stealing pears. There are sound effects in the film but no spoken words. After watching this film, participants were first instructed to recount the story as if they were speaking to a friend who had not seen it. The narratives were recorded. This made up the spoken task of the experiment. After verbally narrating the story, participants were instructed to recount the story as if they were writing to a friend who had not seen the film. This made up the written task of the experiment.

2.2 Data analysis

To prepare the data for further analysis, we transcribed the spoken narratives orthographically. We also included features of spoken language such as hesitations, pauses, false starts, repetitions and self-corrections in the transcription. When repetitions and self-corrections occurred, we only considered the final occurrence in the

analysis, under the assumption that this is the version with which participants were most satisfied. In the next step, we divided the spoken narratives into utterances. Following Hurewitz (1998), Passonneau (1998) and others, we define an utterance as a finite clause (that is, containing a finite verb) but do not consider relative clauses as separate utterances for purposes of discourse segmentation. Relative clauses are grouped with the main clause whose components they modify, following Hurewitz (1998) and others. As a consequence, given that our analysis focuses on subjects and objects of the main clause, referents that are only mentioned inside relative clauses are not included in the analysis (see also Bel et al. 2010, Walker et al. 1998).⁴ We adopted these same criteria to divide the written narratives into utterances. Thus, similar to an utterance in the spoken narratives, each utterance in the written narratives consists of a finite verb and may include a relative clause. It is important to note that in this analysis, we only report cases in which referents occur in *adjacent utterances*. We did not encounter ambiguous pronouns in this dataset (with the exception of *họ* ‘they’, which was not counted due to its ambiguity, see footnote 3).

We coded all mentions of *singular third-person human referents in adjacent utterances* for (i) *grammatical role* and (ii) *referential form*. Regarding (i) *grammatical role*, we coded referents’ *grammatical roles in both the preceding and the current utterances* (e.g., subject, object, possessive, etc.). In other words, we coded the grammatical roles of the antecedent and of the anaphoric element. For the purposes of the current work, we only discuss Subject and Object roles in our analysis. Four grammatical configurations were established based on referents’ preceding and current grammatical roles as shown in Table 1. (See footnote 1 regarding our use of the term ‘referent’ in this paper.)

Table 1. Four configurations based on grammatical roles in preceding and current clause

Preceding clause (antecedent)	Current clause (anaphoric element)	Grammatical configuration
Subject	Subject	Subject-Subject (Subject parallelism)
Subject	Object	Subject-Object
Object	Subject	Object-Subject
Object	Object	Object-Object (Object parallelism)

4. Although arguments might be made that referents that are only mentioned inside relative clauses should be included in studies of anaphor-antecedent relations, following the norms adopted by previous investigations allows us to create a profile of Vietnamese which can be compared directly with studies of other languages. As for complement clauses, these were included in the current study of Vietnamese (though they were very rare – only one relevant token in each of the spoken and written narratives).

Regarding (ii) *referential form*, since our goal was to observe how grammatical roles can influence the current choice of referential form (*i.e.* null pronoun, overt pronoun, and NP), we only coded referents' *referential form in the current clause*. Examples (4)–(7) illustrate how data is coded with regards to the four grammatical configurations. The referents of interest are marked in **bold**. Null pronouns are indicated in the English translations by pronouns in parentheses.⁵

- (4) a. *khi **cậu bé** này đi ngang qua một con đường*
 when CL boy this go past a CL road
 'when this boy went past a road'
- b. *thì (Ø) gặp một cô bé cũng đi một chiếc xe đạp.*
 then (Ø) see a CL girl also ride a CL bike
 'then (he) saw a girl who also rode a bike.'
- Configuration: *Subject-Subject*
 Referential form: *null pronoun*⁵
- (5) a. *cậu thấy ba **cậu bé** đang đứng trước mặt mình*
 he see three CL boy PROG stand front face self
 'he saw three boys standing in front of him.'
- b. *một **cậu bé** đỡ **cậu** dậy.*
 a CL boy pull he up
 'a boy pulled him up.'
- Configuration: *Subject-Object*
 Referential form: *overt pronoun*
- (6) a. *thì (Ø) đã đỡ **cái cậu bé** này dậy*
 then (Ø) PAST pull CL CL boy this up
 'then (they) pulled this boy up.'
- b. ***cậu bé** này lúc này đau chân*
 CL boy this time this hurt leg
 'at this time, this boy hurt his leg.'
- Configuration: *Object-Subject*
 Referential form: *NP*
- (7) a. *thì nó gặp một **bé gái** đi ngược chiều*
 then he see a CL girl go opposite direction
 'then he saw a girl going on the opposite direction'

5. Note that the element *cậu* which appears in Examples (4)–(7) occurs either as part of a larger compound *cậu bé* meaning 'boy', or as a pronoun meaning 'he' (young male). Its original lexical meaning is the kin term relation 'uncle' (mother's brother). Other speakers used the pairs *chú bé* 'boy' and *chú* 'he' (young male) for the same discourse referent. The original lexical meaning of *chú* is also 'uncle' (father's younger brother).

- b. *và do (Ø) mãi nhìn bé gái*
 and because (Ø) busy look CL girl
 ‘and because (he) was busy looking at the girl’
 → Configuration: *Object-Object*
 Referential form: *NP*

When counting null pronouns, we excluded those that occur in coordinate constructions with “and”, “but” and so on. We did this to avoid inadvertently inflating the number of null pronouns. Even in languages like English, standardly analyzed as not allowing *pro*-drop, coordination structures like “Lisa went home and made a sandwich” and “Lisa went to the library but could not find her friend” allow what superficially looks like a missing pronoun/*pro*. As a result, excluding these types of structures in our analyses ensures that all null pronouns reported in our results are ‘proper’ null pronouns and not analyzable in terms of coordination.

3. Results

In this section, we first present some general information about the narratives. We then discuss how referential forms are used with regards to the four grammatical configurations in Table 1. In the present paper, we focus on the details of the written task. We also provide a summary of the spoken task which is presented in more detail in Ngo & Kaiser (2018). Finally, we will draw a comparison between written and spoken results.

Let us first look at the length of the narratives. We removed hesitations, pauses, repetitions and self-corrections from the spoken narratives prior to performing the word count to keep them parallel with the written narratives. Table 2 shows that on average, the spoken narratives are longer than the written narratives considering both the average number of words and the average number of utterances. We also calculated the average number of words per utterance for each participant and averaged them across all participants. The result shows that spoken utterances are longer than written ones. This is in line with prior work on written vs. spoken differences, specifically that spoken language tends to be more elaborate while written language is more concise (e.g., Drieman 1962; Horowitz & Newman 1964; Tannen 1980).

Table 2. Average length of the narratives by word count, utterance count, and average number of words per utterance among participants

	Avg. word	Avg. utterance	Avg. words per utterance
Written	317.2	35.45	9.15
Spoken	381.45	39.85	9.91

Table 3 shows the overall use of null pronouns, overt pronouns and NPs in the narratives based on the four configurations discussed in Table 1 above. As seen in Table 3, among the three types of referential forms, null pronouns and NPs occur slightly more frequently than overt pronouns. Additionally, we found no difference between written and spoken narratives with regards to referential form use. These patterns might seem to suggest that, at least on this broad level, referential form choice occurs randomly/at chance since there is no clear preference for any of the forms. However, as we show later in this paper, this is not the case. When grammatical roles and grammatical parallelism are taken into account, clear patterns of preference start to emerge. Thus, it is importance to not only look at the overall frequency of referential form use but also to consider the environment in which the forms occur.

Table 3. Overall percentages of null pronouns, overt pronouns and NPs used in written and spoken narratives

	Null pronouns	Overt pronouns	NPs	Total
Written	35%	30%	35%	100%
Spoken	34%	30%	36%	100%

3.1 Referential forms in written and spoken narratives

Let us first look at the *written* results. We first examined how frequently participants used each type of grammatical configuration in Table 1 in their narratives. When participants produced an NP or a (null or overt) pronoun in subject position or object position, we noted what position the antecedent was in. In Table 4 and Figure 1, as in Table 1, the first part of each label refers to the grammatical role of the antecedent and the second part refers to the grammatical role of the pronoun or NP (e.g., Subject_{antecedent}-Subject_{anaphoric_element}). We found that re-mentioning of the same referent is mostly likely to occur in the Subject-Subject configuration (Subject parallelism). As seen in Table 4, the Subject-Subject configuration occurs at a rate of 76%, far more frequently than any of the other three configurations.

Table 4. Percentage of each configuration

	Subject-Subject	Subject-Object	Object-Subject	Object-Object	Total
%	76	7	9	8	100

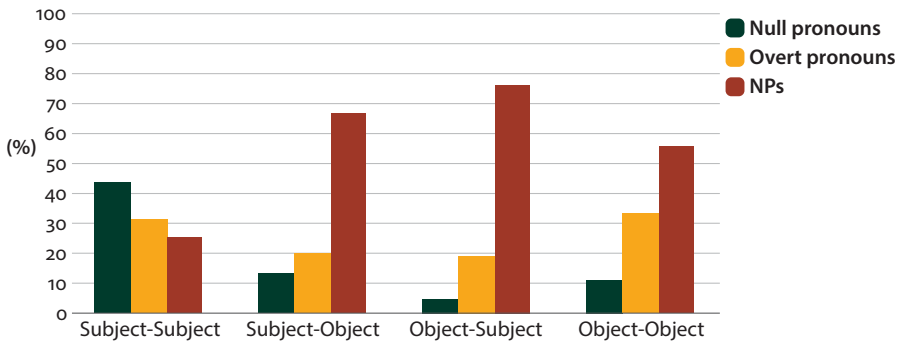


Figure 1. Percentages of referential forms in four grammatical configurations in written task. (The first part of each label refers to the grammatical role of the antecedent and the second part refers to the grammatical role of the pronoun or NP (e.g., Subject_{antecedent}-Subject_{anaphoric_element}))

We also investigated referential form choice (*i.e.* null pronouns, overt pronouns, and NPs) in the current clause in each grammatical configuration. Figure 1 shows the percentages of each referential form in the four configurations.

As shown in Figure 1, the Subject-Subject configuration (Subject parallelism) mostly occurs with pronouns (null pronouns + overt pronouns = 74.71%), whereas the other three configurations consist of mostly NPs (> 55% NPs in each configuration). To examine the pattern of pronoun vs. NP across the four configurations, we conducted a series of chi-square tests.⁶ The results suggest that the distribution of pronouns vs. NPs in the Subject-Subject configuration differs significantly from the other three – as expected from the patterns in Figure 1. Specifically, participants produced significantly more pronouns (null + overt pronouns) relative to NPs in the Subject-Subject configuration than in the Subject-Object configuration ($p < .001$), the Object-Subject configuration ($p < .001$), and the Object-Object configuration ($p < .01$). We also compared the use of null vs. overt pronouns in the

6. We used *chi*-squared test for the statistical analyses, although we realise that aspects of our data are not ideal for this statistical test. Our elicited-narration technique yielded a corpus of multiple narratives and thus involves multiple observations from each participant. However, our open-ended task differs from the standard, more narrowly-controlled within-subjects design often used in psycholinguistics, and although we have multiple observations from each person, the nature of these observations is highly variable across participants. This, as well as the fact that our analysis of pronominal forms involves analyzing responses dependent on the syntactic configuration that a participant chose to produce, lead us to opt for the *chi*-squared analysis over other options, although *chi*-square assumes independence.

Subject-Subject configuration and found no significant difference between the two forms ($p = .06$) – as the patterns visible in Figure 1 lead us to expect.

A closer look at the other three configurations, Subject-Object, Object-Subject and Object-Object (Object parallelism) shows that the proportion of pronouns vs. NPs used in these configurations are not significantly different from each other ($p = .39$). However, the proportion of pronouns in the parallel Object-Object configuration is numerically slightly higher than those in the non-parallel Subject-Object and Object-Subject configurations, 44% compared to 33% and 24% respectively.

Let now turn to the *spoken* results. A detailed discussion of the spoken data can be found in Ngo & Kaiser (2018). Overall, patterns of the spoken narratives resemble those of the written narratives as seen in Figures 2 and 3. Participants mostly use the Subject-Subject configuration to refer to referents in adjacent utterances (78.85%). With regards to referential form choice in the current utterance, results of a series of *chi*-square tests confirm that the distribution of pronouns vs. NP use in the Subject-Subject configuration (Subject parallelism) differs significantly from the other three configurations. Specifically, pronouns (null pronouns + overt pronouns = 73.18%) are the dominant choices in the Subject-Subject configuration. In contrast, the other three configurations Subject-Object, Object-Subject and Object-Object consist of mostly NPs (> 60% NPs in each configuration). With regards to the null vs. overt pronoun choice in the Subject-Subject configuration, participants show no preference for either null or overt pronouns ($p = .13$).

We also examined the pronoun vs. NP choice in the other three configurations, Subject-Object, Object-Subject and Object-Object (Object parallelism). No significant difference was found in the distribution of pronouns and NPs ($p = .08$) among these configurations. Nevertheless, the parallel Object-Object configuration has slightly more pronouns (39.13%) than the other non-parallel Subject-Object and Object-Subject configurations (35.29% and 13.33%, respectively). More interestingly, while the non-parallel Subject-Object and Object-Subject configurations have *no* null pronouns at all, the parallel Object-Object configuration elicits 26.1% null pronouns.

3.2 Comparing written and spoken results

In this section, we examine the effects of modality (*i.e.* written, spoken) on (i) *the use of grammatical configurations* as well as (ii) *the choice of referential forms in each configuration*. For this purpose, we provide a side-by-side comparison of the written and spoken results in Figures 2 and 3.

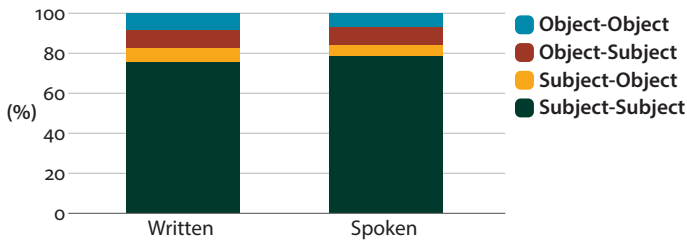


Figure 2. Proportions of the four grammatical configurations in written and spoken narratives

Figure 2 shows the proportions of four types of grammatical role configurations in the written and spoken narratives. We observe the same patterns in both types of narratives. In particular, the Subject-Subject (Subject parallelism) configuration is the most frequent (more than 75% of all occurrences). The other three configurations occur at a similar rate as seen in Figure 2. In short, there is no effect of modality on the occurrence of the four different types of configurations.

In terms of referential form use, we conducted a series of chi-square tests to compare the numbers of null pronouns, overt pronouns and NPs in each grammatical configuration between written and spoken narratives. The results show that Vietnamese participants do not differ in their referential form use in writing and in speaking (p 's = n.s.). In both types of narratives, the Subject-Subject configuration differs significantly from the other three configurations. Figure 3 shows that in the Subject-Subject configuration, pronouns (null + overt pronouns) are the preferred forms. However, in the other three configurations, participants exhibit a preference for NPs over pronouns. This preference for NP use is very clear in the non-parallel Subject-Object and Object-Subject configurations. Interestingly, the parallel Object-Object configuration, although still yielding a high number of NPs, has slightly more pronouns than the non-parallel configurations. Most prominently, in the spoken narratives, null pronouns are found in the parallel Object-Object configuration, but they did *not* occur at all in the non-parallel configurations. In sum, modality does not affect Vietnamese participants' choice of referential form across all four grammatical configurations. Nevertheless, in both modalities, we find that the four configurations elicit different kinds of referential forms (as can be seen in Figure 3, and as we previously discussed in Section 3.1).

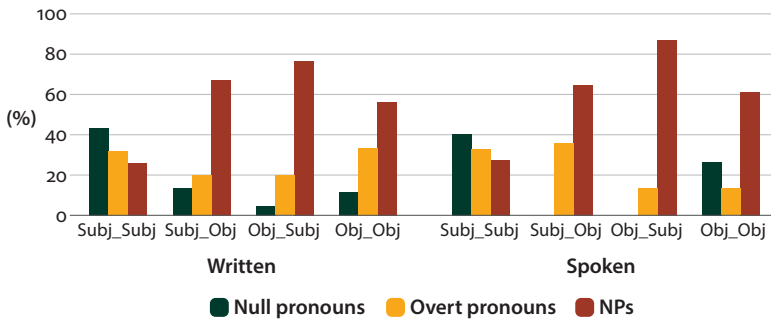


Figure 3. Percentages of referential forms in the four grammatical configurations in both written and spoken narratives⁷

Taken together, our results show no effects of written vs. spoken modality on how Vietnamese participants use either grammatical configurations or referential forms with respect to these configurations. The lack of modality effect on referential form choice in the current study contrasts with previous claims that pronouns and NPs occur at different rates in written and in spoken language (Biber et al. 1999; Christensen 2000).

4. General discussion

In this paper, we have reported a narrative experiment investigating the effects of (i) *grammatical roles*, (ii) *grammatical parallelism* and (iii) *modality* on speakers' referential form choices in Vietnamese. We are particularly interested in how Vietnamese null and overt pronouns are used. This interest stems from the fact that Vietnamese not only allows null and overt pronouns in both subject and object positions but also has a complex kinship pronoun system that differs from other discourse *pro*-drop languages such as Chinese and Japanese. Thus, this paper aims to add to our understanding of pronoun behaviour in typologically different languages.

The narrative experiment has two parts, spoken and written. We instructed participants to recount the Pear film first by speaking, and then by writing. We analyzed the narratives taking into account (i) *referents' grammatical roles in the preceding and current utterances* and (ii) *their referential forms in the current utterance*. This method

7. The absence of columns for null pronouns in the Subj-Obj and Obj-Subj configurations in the spoken narratives is due to the fact, noted at the end of Section 3.1, that speakers did not produce such elements in these two configurations in the spoken narratives.

allowed us to investigate the extent to which grammatical roles and grammatical parallelism affect referential form choice. The results of both spoken and written narratives show that grammatical role and grammatical parallelism play a key role in Vietnamese speakers' choice of referential form. Specifically, Vietnamese speakers use significantly more pronouns (null and over pronouns combined) when the grammatical subject role is maintained across utterances (i.e. Subject parallelism). In contrast, the non-parallel configurations (i.e. Subject-Object, Object-Subject) result in mostly NPs. Interestingly, we also detect hints of a parallelism effect in the Object-Object configuration (i.e. Object parallelism). Although NPs are the most frequent choice, Vietnamese speakers produced more pronouns (null and overt pronouns) in the Object parallelism configuration than in the non-parallel ones. We also observed parallelism effects in the patterning of null pronouns in the spoken narratives: null pronouns only occurred in Subject and Object parallelism configurations. Nevertheless, the Subject and Object parallelism configurations differed in their overall patterns with Subject parallelism favoring pronouns and Object parallelism favoring NPs. These patterns indicate that grammatical role still has a strong impact on referential form choice.

We are also interested in the potential role of modality (i.e. spoken vs. written) on the production of Vietnamese referential forms. Our results show that modality has no significant effect on Vietnamese speakers' referential form choice when the level of formality and subject matter being described are kept parallel in spoken and written descriptions. The patterns of pronoun and NP use are similar in spoken and written narratives. Moreover, Vietnamese speakers also use null and overt pronouns similarly in both modalities. At first glance, this finding seems to contradict prior claims that written language utilises more NPs than spoken language (Biber et al. 1999) and that null pronouns are used increasingly more in written than in spoken narratives (Li & Thompson 1979; Christensen 2000). However, there is a major difference between these studies and our work. Whereas previous studies report the number of tokens without specifying the environment of occurrence (Christensen 2000; Clancy 1982), our study report these numbers with respect to grammatical roles and grammatical parallelism. Crucially, including grammatical factors in the analyses allows us to obtain a clearer view of the underlying mechanism licensing use of different referential forms, particularly null and overt pronouns. Thus, the lack of modality effects in our results suggests that the same underlying mechanism guides production of referential forms in both spoken and written Vietnamese, which we regard as a desirable conclusion.

Another focus of attention in the current study is the choice of null vs. overt pronouns in Vietnamese. In previous, highly influential work on the discourse *pro*-drop language Chinese, Givón (1983) has proposed that there is a strong preference for the use of null pronouns ('zero anaphora') rather than overt pronouns

when the antecedent for such elements is highly salient/prominent in a discourse. Givón (1983) and a broad range of functional studies adopting Givón's approach suggest that 'the more accessible a referent is within a discourse, the less overt coding it will be given, hence that highly accessible antecedents will be referenced with zero anaphora, less accessible antecedents with (overt) pronouns, and very weakly accessible referents with the use of a full NP' (Simpson et al. 2016: 2). Similar observations about the relationship between the form of referring expressions and the salience/prominence of the antecedent are made by Ariel (1990) and Gundel et al. (1993).

A large number of studies have shown that grammatical role has a significant influence on referents' salience and thus, the choice of referring expression (Chafe 1976; Brennan, Friedman & Pollard 1987; Crawley & Stevenson 1990; see also Gordon, Grosz & Gilliom 1993; Gordon & Chan 1995; Perfetti & Goldman 1974). In particular, referents in subject position are more salient than those in object position. These studies along with the salience hierarchy – Givón (1983), Ariel (1990), Gundel et al. (1993) – predict that more reduced referential forms are preferred for highly salient subject antecedents while fuller forms are frequently used for less salient object antecedents. These predictions have been supported in English (e.g., Fukumura & van Gompel 2010) as well as in agreement *pro*-drop languages such as Italian and Spanish (Carminati 2002; Alonso-Ovalle et al. 2002).

The finding of the current study on Vietnamese that speakers employ broadly equal amounts of null and overt pronouns in situations where the grammatical roles of the antecedent and anaphoric element are the same – poses a clear challenge to the salience/accessibility hierarchy. Null pronouns in Vietnamese – being the more reduced referential form – are expected to be chosen much more frequently than overt pronouns to refer to highly salient subject referents, but this was not observed in either the spoken or written narratives. As a clear preference for null pronouns was not found in Subject-Subject coreference relations, the conclusion can be drawn that there is no necessary cross-linguistic application of the salience hierarchy in the choice of referential forms, automatically favoring more reduced forms in instances of reference to recent, highly salient elements within a discourse.

The absence of a straightforward mapping between more reduced forms and more salient elements is in line with Kaiser & Trueswell's (2008) form-specific multiple-constraints approach. Based on data from Finnish overt pronouns and anaphoric demonstratives, Kaiser and Trueswell argue against the assumption that different kinds of referring expressions can be straightforwardly mapped onto a unified salience hierarchy.

The discovery of broadly equal use of null and overt pronouns in the current study of Vietnamese interestingly converges with the results of a recent investigation of Chinese, described in Christensen (2000), which also found that speakers

tend to use null and overt pronouns equally in similar conditions, at least in spoken Chinese.⁸ This suggests that the connections posited between salience and representational form in instances of anaphoric reference should carefully be reexamined in other *pro*-drop languages, to establish which of these follow the Vietnamese/Chinese patterning, and which may perhaps show stronger preferences for null pronouns when these are licensed by the context.

Our results also distinguish pronouns in Vietnamese from those which occur in Japanese in a potentially informative way (Clancy 1980, 1982). It has previously been claimed that the observed, highly restricted use of Japanese overt pronouns may be due to the fact that they are historically derived from nouns and are rich in semantics. The latter property is suggested to constrain their use, resulting in a significantly lower frequency of occurrence than that of null pronouns in the language (Hinds 1975, 1983). Comparing Vietnamese and Japanese, it can be noted that Vietnamese kinship overt pronouns in Vietnamese are also semantically rich, but this does not seem to restrict their use in the same ways as in Japanese.

There are two factors that may influence pronoun use in Vietnamese and null/overt pronoun alternations which have *not* been explored in the current study, warranting further investigation. First, although the grammatical subject of a sentence is also often the topic of a particular stretch of discourse and is highly salient (Givón 1983), being a grammatical subject does not always entail being a topic (Lambrecht 1994). As a result, the second subject in our Subject parallelism configuration is likely to be a topic but does not have to be one. If speakers favour the use of null subjects for reference to the discourse topics and were to use overt pronouns for other instances of anaphoric reference, this might account for some of the variation between null and overt pronouns attested in the study. Indeed, another experiment reported by Ngo and Kaiser (2018) shows that when the topicality of an element in subject position is deliberately increased by means of passivisation, Vietnamese speakers seem to strongly prefer null pronouns to overt pronouns. Consequently, one may question the degree to which grammatical subjects regularly function as topics in Vietnamese. If such a relation does not exist strongly in Vietnamese, this might allow for a more nuanced account of the distribution of null and overt pronouns in patterns of Subject-Subject co-reference.

8. Christensen's investigation of oral and written narratives recounting the pear story in Chinese showed that null and overt pronouns were used at nearly the same rate in the oral narratives. However, unlike Vietnamese, this patterning was not maintained in the written narratives, where null pronouns occurred 55% of the time, while overt pronouns were used at a rate of less than 20%. There is thus a clear effect of modality at play in Chinese, which does not seem to occur in Vietnamese.

Second, another factor which our analysis has not accounted for is the role that coherence relations potentially may play in anaphoric reference. Previous work shows that the production and comprehension of pronouns can be influenced by the type of coherence relation which exists between two clauses (Kehler 2002; Kehler & Rohde 2013). With regards to discourse *pro*-drop languages, Simpson et al. (2016) confirm the effects of coherence relations on the likelihood of mention and referential form use in Chinese. They found that the Explanation⁹ relation results in more continuations referring back to the preceding subjects than the Occasion¹⁰ relation does. This indicates that the subjects in Explanation relations are more likely to be continuing discourse topics. Although we have not computed the details regarding coherence relations in our data, an initial preliminary examination suggests that there was a high amount of Occasion relations in our narrative data. According to Kehler (2008), Occasion is the typical relation used in narratives. In this light, the subjects in our narratives might not be “strong topics”, which could be a reason why null pronouns were not the dominant referential form choice – perhaps null pronouns are only used to refer to very strong discourse topics, and are less commonly used in subject chains which do not involve topics of such strength. We aim to disentangle these factors in future work.

In sum, our results have indicated that grammatical role and grammatical parallelism play an important role in how Vietnamese speakers choose referential forms. We found that not only subjecthood but also grammatical role parallelism increase pronoun use. In contrast, if the referring expressions and its antecedent are not in parallel grammatical roles, and in particular if they are not both subjects – we observe a significant increase in the production of NPs. Unlike prior work, our study found no effects of written vs. spoken modality, indicating that the effects of grammatical roles and parallelism on referential form use are not affected by modality. These results highlight the importance of considering referents’ grammatical roles in adjacent utterances when investigating speakers’ choice of referential form. Regarding the distinction between Vietnamese null and overt pronouns, no differences were found in the current study (but see Ngo & Kaiser 2018 regarding topicality).

We conclude that in Vietnamese, grammatical roles and parallelism have similar effects on both null and overt pronouns. Interestingly, despite the fact that Chinese, Japanese and Vietnamese are all discourse *pro*-drop languages, overt pronoun use varies cross-linguistically. Although Vietnamese overt pronouns are

9. An Explanation relation occurs when a follow-on sentence is used to provide an explanation of the content of a preceding sentence.

10. An Occasion relation occurs with a temporal sequencing of events, the content of one sentence preceding that of a second sentence in time.

semantically rich kinship terms, they are very frequently used similar to Chinese overt pronouns (Christensen 2000). This contrasts with Japanese overt pronouns which are historically derived from nouns, and as claimed in Hinds (1975, 1983), are used restrictively due to their semantics. Finally, to our knowledge, this is the first experimental investigation of a kin-term-based pronoun system.

References

- Alonso-Ovalle, Luis, Fernández-Solera, Susana, Frazier, Lyn, & Clifton, Jr., Charles. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics* 14: 151–170.
- Amano, Shigeaki, & Kondo, Tadahisa. 2000. *Nihongo-no Goi-tokusei* (Lexical properties of Japanese), Vol. 7. Tokyo: Sansei-do.
- Ariel, Mira. 1990. *Accessing Noun-phrase Antecedents*. London: Routledge.
- Barbosa, Pilar. 2011. *Pro-drop and theories of pro in the minimalist program, Part 2: Pronoun deletion analyses of null subjects and partial, discourse and semi pro-drop*. *Language and Linguistics Compass* 5(8): 571–587. <https://doi.org/10.1111/j.1749-818X.2011.00292.x>
- Bel, Aurora, Perera, Joan, & Salas, Namyé. 2010. Anaphoric devices in written and spoken narrative discourse: Data from Catalan. *Written Language and Literacy* 13(2): 236–259. <https://doi.org/10.1075/wll.13.2.03bel>
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Brennan, Susan E., Friedman, Marilyn W., & Pollard, Carl. 1987. A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, Stanford, California*, 155–162. Stroudsburg PA: Association for Computational Linguistics. <https://doi.org/10.3115/981175.981197>
- Carminati, Maria Nella. 2002. *The Processing of Italian Subject Pronouns*. PhD dissertation, University of Massachusetts at Amherst.
- Chafe, Wallace L. (ed.). 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood NJ: Ablex.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and Topic*, Charles Li (ed.), 25–56. New York NY: Academic Press.
- Chafe, Wallace L. & Tannen, Deborah. 1987. The relation between written and spoken language. *Annual Review of Anthropology* 16(1): 383–407. <https://doi.org/10.1146/annurev.an.16.100187.002123>
- Chambers, Craig & Smyth, Ron. 1998. Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language* 39(4): 593–608. <https://doi.org/10.1006/jmla.1998.2575>
- Chen, Ping. 1986. *Referent Introducing and Tracking in Chinese Narratives*. PhD dissertation, University of California, Los Angeles.
- Christensen, Matthew B. 2000. Anaphoric reference in spoken and written Chinese narrative discourse. *Journal of Chinese Linguistics* 28(2): 303–336.

- Clancy, Patricia M. 1980. Referential choice in English and Japanese narrative discourse. In *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, Wallace L. Chafe (ed.), 127–202. Norwood NJ: Ablex.
- Clancy, Patricia M. 1982. Written and spoken style in Japanese narratives. In *Spoken and Written Language: Exploring Orality and Literacy*, Deborah Tannen (ed.), 55–76. Norwood NJ: Ablex.
- Crawley, Rosalind A. & Stevenson, Rosemary J. 1990. Reference in single sentences and in texts. *Journal of Psycholinguistic Research* 19(3): 191–210. <https://doi.org/10.1007/BF01077416>
- Drieman, G. H. J. 1962. Differences between written and spoken languages: an exploratory study. *Acta Psychologica* 20: 78–100. [https://doi.org/10.1016/0001-6918\(62\)90009-4](https://doi.org/10.1016/0001-6918(62)90009-4)
- Fedele, Emily. 2016. Discourse-level Processing and Pronoun Interpretation. PhD dissertation, University of Southern California.
- Fukumura, Kumiko & Van Gompel, Roger P. 2010. Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language* 62(1): 52–66. <https://doi.org/10.1016/j.jml.2009.09.001>
- Givón, Talmy (ed.). 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study* [Typological Studies in Language 3]. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.3>
- Gordon, Peter C. & Chan, Davina. 1995. Pronouns, passives and discourse coherence. *Journal of Memory and Language* 34: 216–223. <https://doi.org/10.1006/jmla.1995.1010>
- Gordon, Peter C., Grosz, Barbara J., & Gilliom, Laura A. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17(3): 311–347. https://doi.org/10.1207/s15516709cog1703_1
- Gundel, Jeanette K., Hedberg, Nancy, & Zacharski, Ron. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274–307. <https://doi.org/10.2307/416535>
- Hinds, John. 1975. Third person pronouns in Japanese. In *Language in Japanese Society*, Fred C. Peng (ed.), 129–157. Tokyo: University of Tokyo Press.
- Hinds, John. 1983. Topic continuity in Japanese. In *Topic Continuity in Discourse: A Quantitative Cross-Language Study* [Typological Studies in Language 3], Talmy Givón (ed.), 47–93. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.3.02hin>
- Horowitz, Milton W. & Newman, John B. 1964. Spoken and written expression: An experimental analysis. *The Journal of Abnormal and Social Psychology* 68(6): 640. <https://doi.org/10.1037/h0048589>
- Hurewitz, Felicia. 1998. A quantitative look at discourse coherence. In *Centering Theory in Discourse*, Marilyn A. Walker, Arvind K. Joshi & Ellen F. Prince (eds), 273–291. Oxford: Clarendon Press.
- Kaiser, Elsi & Trueswell, John C. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23(5): 709–748. <https://doi.org/10.1080/01690960701771220>
- Kehler, Andrew. 2002. *Coherence, Reference, and the Theory of Grammar*. Stanford CA: CSLI.
- Kehler, Andrew. 2008. Rethinking the SMASH approach to pronoun interpretation. In *Reference: Interdisciplinary Perspectives*, Jeanette K. Gundel & Nancy Hedberg (eds), 95–122. Oxford: OUP. <https://doi.org/10.1093/acprof:oso/9780195331639.003.0005>
- Kehler, Andrew & Rohde, Hannah. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39(1): 1–37. <https://doi.org/10.1515/tl-2013-0001>

- Kuroda, Shige-Yuki. 1965. *Generative Grammatical Studies in the Japanese Language*. PhD dissertation, MIT.
- Lambrecht, Knut. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: CUP.
<https://doi.org/10.1017/CBO9780511620607>
- Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley CA: University of California Press.
- Li, Charles N. & Thompson, Sandra A. 1979. Third person pronouns and zero anaphora in Chinese discourse. In *Syntax and Semantics 12: Discourse and Syntax*, Talmy Givón (ed.), 311–335. New York NY: Academic Press.
- Neeleman, Ad & Szendrői, Kriszta. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4): 671–714. <https://doi.org/10.1162/ling.2007.38.4.671>
- Ngo, Binh & Kaiser, Elsi. 2018. Effects of grammatical roles and topicality on Vietnamese referential form production. *Proceedings of the Linguistic Society of America* 3(57): 1–12.
<https://doi.org/10.3765/plsa.v3i1.4354>
- Passonneau, Rebecca J. 1998. Interaction of discourse structure with explicitness of discourse anaphoric noun phrases. In *Centering Theory in Discourse*, Marilyn A. Walker, Arvind K. Joshi & Ellen F. Prince (eds), 327–358. Oxford: OUP.
- Perfetti, Charles A. & Goldman, Susan R. 1974. Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior* 13(1): 70–79.
[https://doi.org/10.1016/S0022-5371\(74\)80032-0](https://doi.org/10.1016/S0022-5371(74)80032-0)
- Simpson, Andrew, Wu, Zoe, & Li, Yan. 2016. Grammatical roles, coherence relations, and the interpretation of pronouns in Chinese. *Lingua Sinica* 2(1): 1–20.
<https://doi.org/10.1186/s40655-016-0011-2>
- Smyth, Ron. 1994. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23(3): 197–229. <https://doi.org/10.1007/BF02139085>
- Stevenson, Rosemary J., Nelson, Alexander W. R., & Stenning, Keith. 1995. The role of parallelism in strategies of pronoun comprehension. *Language and Speech* 38(4): 393–418.
<https://doi.org/10.1177/002383099503800404>
- Tannen, Deborah. 1980. Spoken/written language and the oral/literate continuum. In *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, 207–218. Berkeley CA: BLS.
<https://doi.org/10.3765/bls.v6i0.2133>
- Ueno, Mieko & Kehler, Andrew. 2016. Grammatical and pragmatic factors in the interpretation of Japanese null and overt pronouns. *Linguistics* 54(6): 1165–1221.
<https://doi.org/10.1515/ling-2016-0027>
- Walker, Marilyn A., Joshi, Arvind K., & Prince, Ellen F. 1998. *Centering Theory in Discourse*. Oxford: OUP.
- Yang, Chin Lung, Gordon, Peter C., Hendrick, Randall & Wu, Jei Tun. 1999. Comprehension of referring expressions in Chinese. *Language and Cognitive Processes* 14(5–6): 715–743.
<https://doi.org/10.1080/016909699386248>
- Yang, Chin Lung, Gordon, Peter C., Hendrick, Randall, & Hue, Chih Wei. 2003. Constraining the comprehension of pronominal expressions in Chinese. *Cognition* 86(3): 283–315.
[https://doi.org/10.1016/S0010-0277\(02\)00182-8](https://doi.org/10.1016/S0010-0277(02)00182-8)

