REFERENTIAL CHOICE IN THE WRITTEN NARRATIVES OF INDONESIAN ADULTS*

Katharina Endriati Sukamto
Universitas Katolik Indonesia Atma Jaya
katharina.sukamto@atmajaya.ac.id

Abstract
This study aims to describe the referential choice of two protagonist animate characters in a silent six-minute film entitled The Pear Story (Chafe, 1980). A total of 80 undergraduate and graduate Indonesian students were asked to watch the film and then retell the story by writing a narrative about the film in Indonesian. Findings indicate that when the protagonist animate referents are mentioned for the first time, a classifier seorang ‘a person’ is always used before the NP. When they are reactivated, they are mostly expressed by zero, pronouns dia or ia ‘he’, clitic –nya ‘his’ or ‘him’, NPs with determiners ini ‘this’, itu ‘that’, tersebut ‘aforementioned’, tadi ‘mentioned before’, NPs with relative clauses and NPs with definite articles si or sang ‘the’ which are often used in fables or tales. This study also demonstrates that the choice of anaphoric expressions of the protagonists is determined by factors such as referential distance and referential interference. This study confirms other cross-linguistic studies about referential choice – that there is a correlation between salience and referring expressions in discourse. Keywords: referential choice, referring expression, protagonist referents, referential distance, referential interference

INTRODUCTION

In narrative discourse, the process of referring to a referent is conventionally called reference. A speaker or writer will typically make a referential choice depending on the level of cognitive assessibility of the intended referent (Fossard et al., 2108). In other words, in referring to a particular referent, the decision to choose a referring expression depends on how accessible the referent is to the hearer (when it is spoken discourse) or to the reader (when it is written discourse). For example, when someone says, *Could you pass the salt*, it is very likely that the hearer can easily associate the linguistic form *the salt* with an object although *the salt* is a first-mention. However, upon hearing *Could you pass it*, the hearer may find it difficult to refer to the correct antecedent.

Referential choice is dynamic (Chafe, 1994), and therefore it should be considered as a “linguistic process of a multi-factorial nature” (Kibrik, 2011, p. 61). That is to say that when referring to particular entity, we have to pay attention to a number of factors, such as animacy and saliency of the targeted referent. Some studies have shown that there is a need to encode more linguistic material for an animate rather than inanimate referent (Dahl & Fraurud, 1996; Fukumura & van Gompel, 2011). As for saliency, the choice of referring expressions is driven by how important the conceptual information is represented in the working memory (Chafe, 1994; Ariel, 1988, 1990; Givón, 1983; Kibrik, 2011). The less accessible the referent is, the more linguistic encoding should be expressed in the referring expression.

Studies about referential choice in discourse have been conducted in a number of languages. They are, among others, referential choice in English and Japanese by adult speakers (Clancy, 1980, 1982), the use of bare reflexive type of anaphora in Chinese (Liu, 2010), referential choice in Russian narrative prose (Kibrik, 2011), the use of referential devices in Spanish newspapers (Casteele, 2013), and the role of accessibility of referential choice in Dutch (Vogels, 2014). The studies reveal that the conceptualization of objects, concepts, or events in the mind of language producers will be expressed in the form of referring expressions. However, before communicating a particular entity, first of all a speaker or a writer has to make a decision, such as which information should be included in the utterance or writing and what kind of referring expression should be used so as to make it accessible to the interlocutor.

To date, not much work has been devoted to referential choice in Indonesian. This study, therefore, seeks to investigate referential choice in the written narratives of Indonesian adults. It attempts to answer the following questions:
1. What anaphoric expressions are used for tracking the two animate protagonists in the narrative?

2. How is linguistic encoding related to the degree of accessibility of the referents?

THEORETICAL BASIS

This present study adopted a number of earlier works by Chafe (1980, 1994), Ariel (1988, 1990), Gundel et al. (1993, 2001), Givón (1983), Huang (2000), and Kirbik et al. (2016). Their theoretical approaches suggested that when someone makes a reference either in spoken or written discourse, there is normally a cognitive characterization of a referent that governs the referential choice. Making reference in discourse is therefore a very complex phenomenon (Huang, 2000), as it requires the speaker (or the writer) to produce “an appropriate type-representation” (Gundel et al., 1993, p. 276). For the purpose of this present study, two relevant issues will be briefly presented in this section, namely (1) cognitive status of referents and referential choice and (2) degrees of saliency and types of referring expressions.

Cognitive status of referents and referential choice

In his seminal work, Chafe (1994) used the term ‘consciousness’ for cognitive states of ideas in the mind. When consciousness is associated with language, it involves the activation of referents in discourse. He proposed that a referent in a piece of discourse may have three different levels of cognitive status: active, semi-active, or inactive. An active referent is one that is in the speaker’s focus of consciousness. A semi-active referent is in the peripheral consciousness of the speaker, while an inactive referent is in the long-term memory at the time of production. In the production of a referent, the speaker assumes that a particular referent can be cognitively active, semi-active, or inactive in the mind of the hearer. In order to make it accessible to the hearer, the activation of a referent will therefore depend on the cognitive status of the referent.

Although Chafe (1994) admitted that consciousness cannot be “publicly observed” (p. 15), he assured that the cognitive characterization of a referent should be based on how active the referent is for the language users. For Chafe, consciousness is “dynamic” (p. 29), in the sense that a referent may move from one cognitive status to another. The following may illustrate two different cognitive statuses of a referent (Chafe, 1994, p. 72).

(a) I talked to Larry last night.

(b) I talked to him last night.
The referent **Larry** in (a) is newly introduced to the addressee, but it is already known to both the speaker and addressee. Unlike **Larry**, which is new information, the referent **him** in (b) has a different cognitive status as it refers to someone who has been previously mentioned. For Chafe (1994), the referent **him** is cognitively active and therefore identifiable. The cognitive states of the referents **Larry** and **him** in the above examples are different from the cognitive state of **a lawyer** in the following (Chafe, 1994, p.71):

(c) I talked to **a lawyer** last night.

In this case, the information expressed by **a lawyer** in (c) is new. Since it is cognitively inactive in the mind of the hearer, it cannot be expressed by a personal pronoun.

Ariel (1990) proposed the term “Encyclopaedic Knowledge” (p. 11) for a referent such as **Larry** in example (a). Unlike Chafe who did not include Encyclopaedic Knowledge as potentially given information, Ariel (1990) was of the opinion that referring expressions such as proper names can be used to refer to people who are already in the mental representation of the addressee. She exemplified it in the sentence **Gandhi is a real man**, where **Gandhi** may be interpreted as Mahatma Gandhi, Indira Gandhi, or Rajiv Gandhi. According to Ariel (1990), natural languages will activate a referent based on the degree of accessibility (p. 10).

Within the cognitive model, Grüning & Kibrik (2005) also believed in activation factors, which in turn will predict the referential choice. Figure 1 below illustrates the production of referential choice.

![Diagram of Discourse Context and Referential Choice](https://via.placeholder.com/150)

**Figure 1.** The cognitive model of reference in discourse production (Grüning & Kibrik, 2005)

As referential choice involves activation process, there are a number of relevant factors which must be addressed. They are, among others, distance to the previously mentioned antecedent, saliency of the referent, properties of the referent (for example, animate vs. inanimate), and potential referential conflict in the discourse (Grüning & Kibrik, 2005). In explaining what is meant by ‘referential conflict’, Grüning & Kibrik (2005, pp. 167-168) provided the following example:
The mechanic started, but immediately returned—**he** began to dig in the box of instruments; they were lying in their places, in full order. **He** pulled out one wrench, dropped it, shook **his** head, whispered something and reached in again. Fedorchuk now clearly saw that the mechanic was a coward and would never go out to the wing. The pilot angrily poked the mechanic at the helmet with his fist <...>

The focus here is **the mechanic**. In case the referring expression **the mechanic** in the last sentence is changed to a pronoun **he**, it is very likely that there will be a referential conflict. For that reason, speakers or writers will ‘filter out’ the possibility of referential conflict by choosing a referring expression that will not cause ambiguity.

For child language, cross-linguistic studies on children’s referential choice have been carried out by a number of linguists. For example, a study which explored the referential choice in Mandarin-speaking children (Huang, 2011) indicated that even children between the ages of 2,2 and 3,1 were quite sensitive to the informativeness features and that their referential choices were made according to the discourse-pragmatic principles (p. 2057). Research on the same topic was also conducted by Guerriero et al. (2006) in English and Japanese, which demonstrated that referential choice in English-speaking and Japanese-speaking children was also dependent upon the pragmatic features of discourse referents.

**Degrees of saliency and types of referring expressions**

Referential choice is a fundamental act in communication (Fossard et al., 2018). In sharing information, therefore, we have to pay attention to the situational as well as the linguistic context in discourse so that the information is accessible for the addressee. This section will briefly review a few studies on how different degrees of referential saliency can be expressed in many different types of referring expressions.

Gundel et al. (1993) proposed six different cognitive statuses which are encoded in different forms. The choice of forms depends on the ‘givenness’ or ‘newness’ of a referent.

![Figure 2. The Givenness Hierarchy (Gundel et al., 1998, 2001)](image)

Gundel et al.’s Givenness Hierarchy depicted the attention states of referents in the speaker’s memory, ranging from the most restrictive (in focus) to the least restrictive (type identifiable).
Gundel et al. (2001, pp. 39-40) explained that the way to read the Givenness Hierarchy is as follows:

The statuses are in a unidirectional entailment relation. If something is in focus (center of attention), it is necessarily activated (in working memory); if it is activated, it is necessarily familiar (in memory); if it is familiar, then the addressee can associate a unique representation; if the addressee can associate a unique representation, then it is referential; and if it is referential, it must be type identifiable. The theory thus correctly predicts that a given cognitive status can be appropriately coded by a form which explicitly signals that status, but also by forms whose meanings are entailed by that status. In the latter case (e.g. use of a definite article for a referent that is in focus) the form is simply underspecified for cognitive status of the intended referent.

For Gundel et al. (2001), the production of language is a matter of coding cognitive statuses of referents with appropriate forms in a given context of use. The linguistic forms “serve as processing signals which assist the addressee in restricting possible interpretations” (p. 39).

In a similar vein, Ariel (1988, 1990) proposed ‘The Accessibility Scale’, which divided referring expressions into low, intermediate, and high accessibility markers, as seen in the following.

![Accessibility Scale](image)

According to Ariel (1990), referring expressions are chosen based on the following factors:

1. Saliency: the antecedent being a salient referent, mainly whether it is a topic or a non-topic
2. Competition: the number of competitors on the role of antecedent
3. Distance: the distance between the antecedent and the anaphor (relevant to subsequent mentions only)
4. Unity: the antecedent being within versus without the same frame/world/point of view/segment or paragraph as the anaphor (Ariel, 1990, pp. 28-29).
In relation to referential salience and types of referring expressions, Chafe (1994, p. 88-89) provided a good example of narrative story in Indonesian. Observe the following.


‘In the past, in the region of West Sumatra, there lived a widow with her son. His name was Malin Kundang. Then Malin Kundang was married with a princess, but Malin Kundang said to his servant, give her some money’.

The above excerpt shows that the form seorang ... ‘a person...’, which is called a classifier in Indonesian, is always used for a first-mention referent (seorang janda ‘a widow’, seorang anaknya ‘her son’, seorang putri ‘a princess’). Malin Kundang, being a protagonist in the story, is repetitively mentioned in the narrative in order to show its importance. Throughout the story, this animate protagonist remains to be the only character that is referred to with a name (Chafe, 1994, p. 88). Secondary importance in the narrative story is Malin Kundang’s mother and his wife. As can be seen in the excerpt, they received a different treatment in their referential expression. After their first-mentions (seorang janda, seorang putri), they are referred to with the clitic –nya.

Some other related works on speakers’ choice of referring expressions in discourse also reveal that the choice of referring expressions is memory-based (Kibrik et al., 2016) and salience-based (Givón, 1983; Grünig & Kibrik, 2005; Orita et al., 2015). Some of the studies used a large corpora (e.g. Kibrik et al., 2016) and they commonly suggest that referential option also depends on other factors such as distance between a referential expression and its antecedent, the syntactic role of the antecedent, as well as the properties of the referent.

**RESEARCH METHOD**

**Research instrument**
The instrument used in this study was a six-minute silent movie ‘The Pear Story’ (http://pearstories.org). This movie was designed by Wallace Chafe – an American linguist – and has been widely used as an instrument to elicit narrative data from speakers of different languages around the world.
Brief synopsis of ‘The Pear Story’ movie
The story started with a farmer harvesting pears, going up and down the tree with a ladder to sort out the pears and put them into the baskets. While picking up the pears, a man leading a goat passed by the tree. The scene continued with a boy on a bike who stopped by the pear tree. The boy intended to steal a basket of pears. He looked up the tree to make sure the farmer was not aware of his presence. After putting the basket on the front part of his bike, he got on his bike and continued his journey. From the other direction there came a girl riding a bike. The boy was looking at the girl when his bike hit a rock. He fell off his bike, and the stolen pears scattered all around. When the boy was picking up the pears, three boys passed by and helped the boy gather the scattered pears. The boy in return picked up three pears for the helpers, who then continued to walk and passed by the pear tree. The final scene reintroduced the farmer (who has been off-screen for a while) who was confused to discover that one of his baskets was missing. The film ended with a scene of the three boys walking in front of him, each was eating a pear.

Research participants
The participants in this study were 80 undergraduate and graduate Indonesian students – 75 females and 5 males – from a private university in Jakarta. Their age ranged from 18 to 44 years old. The students came from different ethnic backgrounds, but all of them claimed that they were fluent speakers of Indonesian.

Data collection procedures
In separate classroom settings, the author played ‘The Pear Story’ film and asked the students to watch the film. After that they were asked to retell the story by writing a narrative about the film in Indonesian. The data were collected between October and November 2012.

Data analysis procedures
This study focuses on the referential choice of two animate protagonists in the film: the pear farmer and the boy on a bike. The procedural steps are as follows:

- Coding the NP clauses that contain two animate protagonists: the pear farmer and the boy on a bike
- Color-coding the two animate protagonist referents
- Listing down the encoding devices of each protagonist by using an Excel spreadsheet
- Listing down the encoding devices for each protagonist when there is an interfering referent
– Listing down the encoding devices for the pear farmer after an episode boundary
– Analyzing the data.

**DISCUSSION**

This section is divided into two sub-sections as an attempt to answer each of the two research questions: (a) Choice of anaphoric expressions and (b) Linguistic encoding and degrees of accessibility.

**Choice of anaphoric expression**

As previously mentioned, this study focuses on the referential choice of two protagonist animate characters in the film: the pear farmer and the boy on a bike. The following pictures illustrate the two main characters in the film.

![The pear farmer](image1)

![The boy on a bike](image2)

The data reveals that for the first mention of the two referents, all participants use either an indefinite N/NP or indefinite N/NP + RC, as can be seen in the following examples:

**Example (1)**

\[\text{Pada suatu pagi seorang bapak memetik buah pir.}\]

at a morning a person father pick fruit pear

‘One morning a man was picking up pears’

**Example (2)**

\[\text{Ada seorang petani pir yang sedang panen buah pir.}\]

exist a person farmer pear REL PROG harvest fruit pear

‘There was a pear farmer who was harvesting pears’

**Example (3)**

\[\text{Seorang anak datang dengan sepedanya.}\]

a person child come with bicycle-3

‘A child came with his bicycle’

**Example (4)**

\[\text{Lewatlah seorang anak laki-laki yang sedang naik sepeda.}\]

pass a person child male REL PROG ride bike

‘A boy who was riding a bike passed by’
When the referent is reactivated, more varieties of referring expressions are used. There are seven forms of NP clause used, namely (1) unexpressed (zero), (2) 3rd person pronoun *dia/ia*, (3) clitic –*nya*, (4) NP + determiner *itu/initi/tersebut/tadi*, (5) NP + RC, (6) article *si / sang* + N/NP, and (7) article *si / sang* + N/NP + RC. Examples are as follows:

**Example (5)**

*Di tengah jalan, dia bertemu dengan Anak perempuan yang bersepeda.*

On his way, he met a girl who is on a bike.

*Karena terpesona dengan anak perempuan itu,*

Because (he) was amazed with the girl,

*sepedanya terantuk batu dan dia jatuh.*

his bicycle hit a stone and he fell down.

Example (5) demonstrates that when the protagonist is continuous or persistent, the form zero, personal pronoun, and clitic –*nya* tend to be used. In this case, the zero form and personal pronouns occupy a subject position.

After an episode boundary, references tend to be specified and lengthy. A definite marker *si* or *sang* is often used.

**Example (6)**

*Sementara itu, sang pemetik buah pir merasa heran.*

‘Meanwhile, the person who picked the pears was surprised’

A past marker *tadi* may also be used, as shown in example (7). This strategy helps the readers to refer back to a referent that has been previously mentioned.

**Example (7)**

*Mereka makan sambil berjalan dan melewati bapak tadi.*

‘They ate while walking and passed by the man mentioned previously’

This study reveals that the basic distributional patterns of anaphoric expressions in the written narrative is as follows:

– Zero, third person pronoun and clitic –*nya* are normally used when the referent is continuous or persistent in an episode.
– When there is an interfering referent, more wording is usually used for the protagonists, usually a lexical NP.
– Minimized lexical expressions with special markers *si* or *sang* are quite effective to make the inactive protagonist referent accessible for the audience. The past marker *tadi* ‘mentioned above’ can also be used as a strategy to reactivate the protagonist that has been active in the readers’ mental state.

**Linguistic encoding and degree of assessibility**

Earlier work on referential choice has discussed that different levels of cognitive status of a referent will basically underlie referential choice (among others, Chafe, 1976; Givón, 1983; Ariel, 1988, 1990; Gundel et al. 1998, Kibrik et al. 2016). The cognitive model developed by those scholars describes that referential choice is governed by how accessible a referent is in the working memory of the speaker/writer. A referent that is highly accessible will require less coding material. On the other hand, a referent that is cognitively less active usually requires more linguistic device.

Based on the data obtained in this study, it is found that a referent which is persistent or continuously mentioned in discourse will need less linguistic material. In contrast, when the referent is reactivated after an episode boundary, it will be highly specified with more linguistic material. In other words, since a protagonist character is the most salient participant in the entire discourse, the continuity of the character throughout the story is central (Givón, 1983). Example (8) illustrates the referential choice in the text.

**Example (8)**

1 Pada suatu hari, **seorang bapak pemetik buah** sedang memanjat pohon pir untuk memetik buah-
2 buah pir yang sudah matang. Sejauh ini **dia** sudah hampir memenuhi dua keranjang penuh. Buah-
3 buah pir yang dipetik **nya** dibersihkan, lalu dipenuhi **nya** keranjang yang kedua. Sementara itu
4 terdengar suara keledai. Ternyata ada **seorang pria dengan keledainya** sedang berjalan menuju
5 ke arah **pemetik buah itu**. Saat **si pemetik buah** sudah memanjat pohon lagi untuk mengisi
6 keranjang ketiga, **si pria dengan keledai** hanya berjalan melewati keranjang-kerja
7 nang tersebut.

‘One day, a **man who is a pear picker** was climbing a pear tree to pick up ripe pears. So far **he** had two baskets almost full of pears. Pears that had been picked by **him** were cleaned, then the second basket was filled by **him** (with pears). Meanwhile, the voice of a donkey was heard. In fact **a man with a donkey** was walking towards the **pear picker**. When the **pear picker** had already climbed the tree again to fill in the third basket, **the man with a donkey** only passed by the pear baskets.'
Referential Choice in the Written Narratives....

Note: The protagonist character is written in bold, whereas the interfering character is written in bold and is italicized.

Example (8) shows that the referent ‘the pear picker’ occurs a couple of times (see lines 1-3). The information flow indicates a referent that is continuous or persistent, and so there is no need for the writer to reactivate the referent with heavier words. However, once there is an interfering referent after the pear picker, the reactivation needs more linguistic material (see line 5).

The writer in example (9) used a different strategy in reactivating a referent. She used a definite marker si to refer to a salient referent that has been previously mentioned (si bapak ‘the man’, si anak ‘the child’). Although the reactivation of the referent is minimized, this choice seems to work well.

Example (9)

1 Pada suatu pagi seorang bapak memetik buah pir. Ada 3 keranjang, 2 penuh dan 1 kosong.
2 Saat si bapak kembali memanjat pohon untuk memetik buah pir, dia melihat seorang priya
3 berjalan dengan kambingnya. Si bapak kemudian memetik buah pir kembali. Saat si bapak
4 berada di atas pohon, seorang anak datang dengan sepedanya. Si anak yang melihat keranjang
5 buah pir merasa tertarik dan dia akhirnya memutuskan untuk mengambil satu keranjang buah
6 priyanya. Namun ada 3 anak yang melihatkeranjang yang berjatuhan. Karena telah membantu membereskan buah pir dan mengembalikan topi
7 si anak yang menemukan 1 keranjang buah pir yang hilang. Tak lama dia melihat 3 anak yang
8 tadi sedang memakan priyanya.

‘One morning there was a man (who) picked pears. There were three baskets, two were full and one empty. When the man returned to climb the tree to pick up the pears, he saw a man walking with his goat. The man then picked up the pears again. When the man was up on the tree, a child came with his bike. The child who saw the pear basket was interested and he finally decided to take a basket of pears and then left. Unfortunately when the child was riding his bike, he fell down with his pear basket, but there were 3 children who helped pick up the fallen pears. Because they have helped picking up the pears and returning the hat of the child, finally the child gave three pears to them. When the man went down the tree, he was surprised to find that one of the baskets was missing. Not long afterwards he saw the three children who were eating pears.’

Apart from the use of definite marker si, the data also shows that some writers used the definite marker sang. Observe example (10).
**Example (10)**

Untungnya ada tiga bocah hampir seumuran dengan sang bocah pencuri yang segera melihat dan menolongnya. Setelah mereka menolong sang bocah pencuri mereka mendapatkan tiga buah pir.

‘Luckily there were three children of about the same age as the child (who is) a thief who saw him and helped him immediately. After they helped the child (who is) a thief they received three pears.

It is interesting to note that the definite markers *si* and *sang*, which are normally used in fables, seem to be used by the respondents to show salience. Although the character is fading away, it can be reactivated by sufficient information by using *si* or *sang* as a strategy.

Findings in this study indicate that narrators will use different strategies to make the intended referents accessible to the readers. This study is consistent with other previous studies (Chafe 1994, Gundel et al. 1993; Ariel 1988) that heavier material will be used to refer to a referent in order to prevent referential conflict. Figure (4) demonstrates the referring expressions which were used by the narrators in this study, ranging from the most accessible to the least accessible.

**Figure 4. Accessibility markers in Indonesian narrative story**

1. Unexpressed (zero)
2. 3rd person pronoun: *dia/ia*
3. Clitic –*nya*
4. NP + determiner
   *itu/ini/tersebut/tadi*
5. NP + Relative Clause
6. Article *si / sang + N/NP*
7. Article *si / sang + N/NP + RC*
8. Classifier *seorang + N/NP*

**CLOSING**

Referents can be expressed in many different ways in written narratives. This study reveals that referential choice of the protagonist characters in narrative stories depends on three important factors:

- Persistence: Zero, third person pronoun, and –*nya* are used when the protagonist is persistent or continuous in the episode.
Different levels of givenness: Less materials will be used when the referent is assumed to be highly accessible to the audience, but more materials will be used when there is a referential competitor or interference.

Salience: A protagonist character is the most salient participant in the entire discourse. Although it is fading away it is the most important character in discourse and thus can be reactivated by sufficient information. In this case writers may use si/sang/tadi as a strategy.

* This article is a revised version of a paper presented at the 23rd Annual Meeting of the Southeast Asia Linguistics Society (SEALS 23) in Bangkok, 29–31 May 2013. It has not been published or submitted elsewhere.

BIBLIOGRAPHY


