

# Cross-linguistic Variations of *pro* Licensing Conditions

Jinghua Zhang

College of Foreign Languages, Yanbian University, Yanji, China

**Abstract**—Null subject languages are generally called *pro*-drop languages. According to the generative syntactic analysis, languages can be divided into *pro*-drop and non-*pro*-drop languages. Subject Agreement (AgrS) is argued to be the licensing condition for the null subject languages. Nevertheless, *pro*-drop licensing conditions may vary in different languages. This paper analyzes various null subject languages with different *pro*-drop licensing conditions from cross-linguistic perspective.

**Index Terms**—*pro*-drop, null subject, partial *pro*-drop, discourse-based *pro*-drop

## I. PRELIMINARY

Null subject phenomenon is universal in many languages. For example, Italian, Spanish, Portuguese and Brazilian Portuguese etc. allow null subjects. According to the research, Subject Agreement (AgrS) in INFL licenses null subjects, while English and French disallow null subjects for the lack of AgrS. Chinese, Korean and Japanese etc. also allow null subjects even though there isn't AgrS. Therefore, Agreement is not the only standard in showing distinctive features of null subjects. In this paper, the author will discuss the conditions of null subjects, focusing on the Chinese null subjects.

Generally speaking, null subject phenomenon is divided into two subcategories according to the occurrence of Agr. One is null subjects with Agr in finite clauses called *pro*-drop (Italian and Spanish etc.), the other one is null subject without or with meager Agr called non-*pro*-drop. The former is subdivided into full *pro*-drop and partial *pro*-drop languages according to the degrees of richness of the AgrS. In this case, null subjects are recoverable according to the form of the verb which is based on the Agreement. The latter case includes the null subjects in the imperatives, truncated subjects in informal expression and PRO, which often occur in non-finite clauses etc. The most typical example is English. Languages of these types of null subjects are not to be considered as *pro*-drop languages.

Actually, there is one more type of *pro*-drop languages without Agr in which Chinese is the most typical representative. Chinese allows *pro*-drop though there is no Agr. It poses problem for the criteria of licensing *pro* which is based on the Agreement. It is reported that there also exist similar *pro*-drop phenomena in Korean and Japanese.

### A. Null Subject Languages

Null subject language refers to the *pro*-drop (pronominal drop) in the subject position in finite clauses, such as Italian and Spanish. Therefore, it is also called as a *pro*-drop language. "*pro*-drop languages refer to those which allow a pronominal subject to be left unexpressed, they 'drop' the subject pronoun." (Haegeman, 1994)

Null pronominal subject (*pro*) is allowed in finite clauses because there is AgrS in INFL which licences *pro*. In other words, the AgrS is rich enough so that the null subject can be recovered by the verbal inflection (Agr). This is the condition which licences *pro* confirmed by many researchers. Examples (1)-(3) show that the form of null subject is recoverable because of the richness of AgrS in INFL. Therefore, we can infer the number, gender and person etc. according to the verbal inflection. Example (1) and (2) show the null subjects in finite clauses, while in (3) the null subject refers to the matrix subject *Gianni* which is in the discourse (It is also influenced by the Agr).

Italian:

(1) Lui ha telefonato<sup>1</sup>  
he has telephoned.  
he has phoned.

(2) Ha telefonato.  
Has telephoned.  
\*(He) has telephoned.

(3) (a) Gianni ha detto che lui ha telefonato  
Gianni has said that he has telephoned  
(b) Gianni ha detto che ha telefonato  
Gianni has said that has telephoned

In addition to the full *pro*-drop languages mentioned above, there are certain numbers of languages that are partially *pro*-drop, i.e., they partially have verbal inflection which will lead to *pro*-drop. According to Platzack (1987), languages

<sup>1</sup> Examples (1)-(3): Introducing Transformational Grammar from rules to principles (Ouhalla, 1994, P275).

like German and the Scandinavian allow only expletive *pro* (non-referential, subjects to be non-overt). It implies that "INFL in these languages is richer than that of English, but poorer than in Italian". (Haegeman, 1994)

According to Borer (1989), not all types of inflection are strong enough to allow *pro*-drop. He relates *pro*-drop option in Modern Hebrew to the richness of inflection. "In the present tense, only gender and number are overtly realized, person is not. The third person is the unmarked form in the other tenses too."

Furthermore, Slavic languages and American Sign Languages also exhibit frequent occurrence of *pro*-drop feature. For example, Croatian which is one of the Slavic languages shows *pro*-drop features in (4). Turkish, which is one of the Altaic languages is also a partial *pro*-drop language. For instance,

Croatian

(4) Vidim ga. Dolazi.

See-1st PERS-SING he-ACC. Come-3rd PERS-SING.

(I) see him. (He) is coming.

Turkish

(5) Geldiğini gördüm.

Coming-POSSESSIVE saw.

(I) saw (you/him/her/it) come.

In (5), the verb *gör-mek* is declined in the first person simple past tense form. The object pronoun can be deduced from the context though it is not clear enough.

Esperanto, which is a constructed language sometimes also exhibits *pro*-drop phenomena in certain context where the pronoun has been used before. The dropping of the subject-pronoun equivalent *li* in the subsequent sentence is syntactically correct in (6).

Esperanto

(6) Ĉu vidas lin? e Venas nun.

Question-particle see him? (li) Comes now.

Do (you) see him? (He) is coming now.

Besides, Brazilian Portuguese, Finnish, and Marathi are also partial *pro*-drop languages in certain discourse or context according to the research of Holmberg (2005). According to the research statistics, among the 847 sample tokens from rural variety of Brazilian Portuguese, there are about 46% null subjects and 54% of non-null subjects concerning with undetermined reference, taboo reference, referent in the border of definition, generic collective as a referent and defined reference etc. (Holmberg, 2005). Therefore, *pro*-drop may also be influenced by the discourse or the context. For instance,

Brazilian Portuguese

(7) (a) Pedro disse que *e* ganhou na loto.

Pedro said that won on the-lottery

Pedro said that (he) won on the lottery.

(b) Aqui n ão pode nadar.

here not can swim

(One) can't swim here.

Finnish

(8) (a) Pertti sanoi ett äe voitti lotossa.

Pertti said that (he) won on-lottery

(b) T ääll äei voi uida.

here not can swim

(One) can't swim here.

Marathi

(9) (a) Ram mhanala ki *e* lotteri jinkla.

Ram said that (he) lottery won

(b) Hya khurchiwar aaramani bushushakto.

this chair-on (one) comfort-with sit

(One) can sit comfortably in this chair.

From the analysis on the full *pro*-drop and partial *pro*-drop, we can conclude some common features of *pro*-drop languages. 1) There is Agreement which makes subject recovered. 2) Dropped subjects are pronominals. 3) They occur in finite clauses. It is reported that many Romance languages (except French) etc. also partially show *pro*-drop features. They are Occitan, Catalan, Portuguese, and Romanian etc. To sum up, we can categorize various types of *pro*-drop languages chiefly depending on the richness of the Agreement as the following table. Group A refers to full *pro*-drop language and Group B refers to partial *pro*-drop languages. These two groups of *pro*-drops are possible because of the AgrS in INFL though there are differences in the sufficiency of richness. Group C refers to different types of *pro*-drop phenomena from what we have discussed and this is to be discussed in part III by focusing on the Chinese null subject phenomenon.

Group	<i>pro-drop languages</i>	<i>Agr</i>
A	Italian, Spanish etc.	++Agr
B	German, Scandinavian, Modern Hebrew, Turkish, Esperanto, Occitan, Catalan, Portuguese, Romanian (except French), Croatia, Brazilian Portuguese, Finnish, Marathi etc.	+Agr
C	Chinese, Korean, Japanese etc.	-Agr

B. *Non-null Subject Languages*

Some non-null subject languages such as English and French have meager verbal inflection that it is hard to differentiate person, number and gender by the Agr in INFL. It is disallowed to drop the *pro* in the subject position in English. However, sometimes we can encounter the null subjects in English as well as in other languages. It is also one of the null subject phenomena, but they are essentially different from what we mentioned above. As a matter of fact, the null subject is truncated informally under certain circumstances for various reasons. It is very common in informal or spoken languages such as in English. (PRO is also one of the null subject, but it occurs in non-finite clauses.)

English

- (10) \*Has spoken.  
\*John has said that *e* has spoken.
- (11) A: Where did you go last night?  
B: (I) Went to see the movie.
- (12) A: (You) open the window.  
B: Yes, Sir.

II. ANALYSIS ON PRO-LICENSING CONDITIONS

Why is it possible to drop the pronominal subject in the sentences? We have mentioned that Italian type languages can drop pronominal subjects in finite clauses. The reason is that there is overt AgrS in INFL. It means that the dropped pronominal subjects can be recovered according to the verbal inflection, i.e., "*pro* is licensed by an overt Agr category co-indexed with it". (Ouhalla, 1994) In other words, *pro-drop* must have a co-indexed Agr category which governs it and makes *pro* recovered.

According to Chomsky (1981, 1982), one of the main features of *pro-drop* languages is recoverability. This condition also emphasizes the function of Agr. Following this rule, *pro-drop* in Italian is recoverable because of the overt AgrS in INFL. Because there is no overt lexical antecedent in Italian type *pro-drop* languages, syntactic feature of *pro* is [-anaphoric, +pronominal]. In contrast, in English type languages, *pro-drop* is disallowed because the pronominal subject is irrecoverable according to the agreement.

- |  |  |                     |
|--|--|---------------------|
| Spanish:   | French:  | English             |
| (13) (a) <i>pro</i> Vimos a Juan.<br>(we) see Juan.          | (14) (a) * <i>pro</i> Voyons Jean.<br>(we) see Jean.           | (15) (a) *(I) see.  |
| (b) <i>pro</i> Baila bien.<br>(he/she ) dances well.         | (b) * <i>pro</i> Danse bien.<br>(he/she) dances well.          | (b) *(we) see.      |
| (c) <i>pro</i> Estamos cansadisimos.<br>(we) are very tired. | (c) * <i>pro</i> Sommes tres fatigues.<br>(we) are very tired. | (c) *(you) see.     |
|  |  | (d) *(you) see.     |
|  |  | (e) *(he/she) sees. |
|  |  | (f) *(they) see.    |

Recent years, more attention is focused on the conditions of null subjects. It is proved that not all the *pro-drop* languages have overt Agr in INFL. Chinese is one of the most typical examples. According to Huang (1984), Chinese is a *pro-drop* language even if there is no overt Agr. It implies that there might be some other factors licensing null subjects. Detailed analysis on the conditions of null subjects in Chinese will be elaborated in the following section.

III. DISCOURSE-BASED PRO-DROP LANGUAGE—CHINESE

We have seen that *pro-drop* languages largely depend on overt Agr in INFL. According to James Huang (1984), "*pro* is possible either in languages with rich agreement (AgrS) or no agreement at all" (e.g. Chinese, Japanese and Korean etc.). Unlike typical *pro-drop* languages such as Italian and Spanish which are based on agreement marking, Chinese *pro* has no overt Agr which can make the dropped subject recovered.

Chinese has no verbal inflection. Huang (1984, 1989) proposes a generalized control theory and argues that "identification hypothesis is essentially correct, but that it must be more broadly interpreted than is assumed in the agreement-based theory".

Chinese has no overt agreement to recover the form of the subject. However, it is a *pro-drop* language. What licenses *pro-drop* in Chinese? Let's look at some of the examples in detail.

Chinese:

- (16) Speaker A: Zhangsan kanjian Lisi le ma?  
Zhangsan see Lisi ASP Q?
- Speaker B: (ta) kanjian (ta) le.  
(he) saw (him).

- (17) Zhangsan shuo [*e* hen xihuan Lisi].  
 Zhangsan say (he) very like Lisi  
 Zhangsan said that (he) liked Lisi.

In (16), the null subject refers to *Zhangsan* which is mentioned in the discourse. In speaker B, the subject position can be null or non-null. If it is non-null, the subject could be *Zhangsan* or the pronominal *he*. That implies it is the place where case can be assigned by government (case filter). In (17), the null subject in the embedded clause refers to either the matrix subject *Zhangsan* or other lexical subject. (*say* is not an obligatory control verb) It also implies that the position of embedded subject is governed and assigned case. Therefore, we can prove that null subject of Chinese is a governed *pro*.

Generally speaking, finiteness of Chinese sentence is determined by the occurrence of element of AUX such as aspect marker such as *le* or modal etc. (Huang 1989). However, it is not an obligatory factor deciding finiteness. In many cases, the element of AUX doesn't occur as in (17). In addition, Huang argues that Chinese *pro* has similarities with PRO because it is anaphoric as in (17). In the light of similarities between PRO and *pro* in Chinese, Huang put forward Generalized Control Rule<sup>2</sup>. In (17), we can find that the null subject is controlled in its control domain<sup>3</sup> and has [+anaphoric, +pronominal] (PRO) syntactic feature. Therefore, we can conclude that Chinese *pro* has the feature of PRO. More precisely, Chinese *pro* is controlled in its control domain. Huang argues that "PRO and *pro* are instances of the same category, and that the properties of PRO/*pro* fall under a proper theory of generalized control." (1989). In (18), subject of the sentential complement is controlled by the matrix subject within its control domain and the embedded subject can be null.

Chinese:

- (18) Zhangsan shefa [PRO bangmang wo].  
 Zhangsan try help I  
 Zhangsan tried to help me.

Chinese *pro* is controlled within its control domain. As a matter of fact, control domain is closely related to the discourse as in (19).

- (19) Zhangsan qi ma qi de [*pro* hen lei].  
 Zhangsan ride horse ride till very tired  
 Zhangsan rode a horse until he got very tired.

However, not all the *pro*-drop in the control domain is controlled properly in Chinese. It is closely related to the matrix control verbs. According to Huang (1989), verbs in Chinese *try*, *manage*, *fail*, *condescend*, and *decide* etc. require obligatory control, while *say*, *ask*, and *wonder* etc. don't require obligatory control especially when it is followed by a sentential complement or contained in a sentential subject. Therefore, *pro* has optional control as in (20).

- (20) Zhangsan shuo [*pro*/PRO mingtian bu bi lai].  
 Zhangsan say tomorrow not need come.  
 Zhangsan said that (he/she/we one...) need not come tomorrow.

To sum up what has been stated above, Chinese *pro* is different from agreement-based *pro* in Italian. It is not only governed, but also controlled in its control domain like PRO. Therefore, Chinese *pro* is similar to PRO.

We also can find from the examples that Chinese *pro*-drop occurs in control domain which is in the scope of discourse. In many cases, matrix subject serves as the antecedent of *pro* and *pro* is anaphoric. Chinese *pro* is also governed for the reasons of case theory. We can infer from the perspectives of Huang that control domain is a crucial concept in defining Chinese *pro*. Generalized Control Rule (GCR) provides the rationale for analyzing Chinese *pro*. Huang's analysis is essentially based on the relationships between the syntactic categories. Admittedly, Huang's theory provides a good approach in explaining the special case of Chinese *pro*.

In addition to the syntactic relationship between categories, we also can consider it from pragmatic aspects. Let's go back to the very beginning and start from the analysis on the agreement-based *pro*. We have seen that Italian type *pro*-drop languages have sufficiently rich agreement in INFL which licences *pro*. This type of *pro* is available whether there is discourse or not. In Chinese, pronominal subjects without discourse (viz. no control domain) can't be dropped because of the arbitrary references. Unlike typical *pro*-drop languages such as Italian, Chinese lacks Agreement. To have discourse for *pro*-drop is basically consistent with Huang's GCR theory. Compared with Italian *pro*, Chinese *pro* requires discourse. Therefore, it is plausible to define this type of languages as discourse-based *pro*-drop languages.

Chinese

- (21) (a) \*(wo) kanjian le.  
 (b) \*(women) kanjian le.  
 (c) \*(ni) kanjian le.  
 (d) \*(nimen) kanjian le.  
 (e) \*(ta) kanjian le.

Italian:

- (22) (a) (io) mangio 'I eat'  
 (b) (tu) mangi 'you eat'  
 (c) (lui/lei) mangia 'he/she eats'  
 (d) (noi) mangiamo 'we eat'  
 (e) (voi) mangiate 'you eat'

<sup>2</sup> Generalized Control Rule (GCR) An empty pronominal is controlled in its control domain (if it has one).

<sup>3</sup> According to Manzini (1983) and Nishigauchi (1984),  $\alpha$  is the control domain for  $\beta$  iff it is the minimal category that satisfies both (a) and (b):

a.  $\alpha$  is the lowest S or NP that contains  $\beta$  or the minimal maximal category containing  $\beta$ .

b.  $\alpha$  contains a SUBJECT accessible to  $\beta$ .

(f) \*(tamen) kanjian le. (f) (essi) mangiano 'they eat'

When there is discourse, Italian *pro* also can be recovered (e.g. (3)). This is because they are syntactically in the control domain and pragmatically confined in the scope of discourse which determines what the dropped *pro* is. As a whole, without the implication of discourse, it is very hard to tell the phi-features of the pronominal subjects such as, person, number or gender etc. Therefore, Chinese *pro* is allowed in the discourse/control domain. For example, in (23), *pro* is governed because it is assigned case via case filter and anaphoric because it has an antecedent *Mary* in the matrix clause. The matrix clause serves as the discourse and forms control domain with the embedded clause.

(23) Mali jue ding [e likai] (le).  
Mali decide (she) likai (ASP)  
Mary decided to leave.

We can extend this analysis to the other *pro*-drop languages. Italian and other partial *pro*-drop languages also show similar pragmatic features as in Chinese. Common feature of null subjects in full and partial *pro*-drop languages is that *pro*-drop is not only allowed in the case of Agreement, but also in the case of discourse.

Generally speaking, Chinese allows *pro*-drop quite naturally and frequently. Chinese *pro* is inferable depending on the discourse or context. There is one more evidence which supports the perspective that Chinese *pro*-drop requires discourse or context.

According to the "hot-cool" division of media,<sup>4</sup> languages also can be divided into "hot" languages and "cool" languages. "hot" languages don't need any conscious participation of the readers because all the syntactic categories are to be elaborated clearly such as English, while "cool" languages need more readers' participation to extract the value. Chinese type languages belong to "cool" languages because the omitted syntactic categories are inferable through the discourse or the context (relevant knowledge of the world or the common topic shared by the speaker and the listener). It is reported that Korean and Japanese also belong to "cool" type languages. In contrast, in English type "hot" languages syntactic categories including pronominals in the subject position can't be omitted even there is discourse or the context. The following example is about Chinese *pro*-drop and all the answers provided by B are possible because the null subjects are all inferable in the discourse (Actually, null objects are also inferable in the discourse).

(24) A: Zhangsan kanjian Lisi le ma?<sup>5</sup>  
Zhangsan see Lisi LE Q  
B: a. ta kanjian le.  
he saw e.  
He saw [him].  
b. e kanjian ta le.  
saw him  
[He] saw him.  
c. e kanjian e le.  
saw  
[He] saw [him].

We have discussed that null subject is not only allowed by the agreement, but also by the discourse. We also find some other cases in Chinese, viz. pragmatically context-based null subjects. By depending on the extra knowledge of the speaker and the listener (particularized conversational implicatures), we allow null subjects in Chinese. For example,

(25) A: xiansheng, [nin] laidian shenme? (Sir, what (do you) want to eat?)  
B: [wo] yao yibe kele. ((I) want a cup of Coke.)

(In the restaurant, A is a waiter and B is a customer.)

(26) A: [ni] shenme shihou laide? (When did (you) come?)  
B: [wo] ganglai. ((I) arrived just now.)  
A: [ni] deng le hen chang shijian ma? (Did (you) wait for me for long time?)  
B: [wo] meiyou. (No. (I) didn't.)

(A is late after making an appointment with B to meet at the school gate.)

To summarize what have been illustrated above, Chinese allows *pro*-drop phenomenon. However, the licensing condition is rather different from the subject agreement-based *pro*-drop. Chinese *pro* is much subject to the inference by an antecedent in its control domain/discourse or context. Therefore, Chinese *pro*-drop is allowed when there is discourse or context. Whether Chinese is a full *pro*-drop or partial *pro*-drop needs further verification. Syntactic and pragmatic distribution of *pro* in various languages can be generalized as follows.

Italian type: [+finite clause], [+discourse/control domain], [-anaphoric, +pronominal], [++Agr]

German type: [+finite clause], [+discourse/control domain/context], [-anaphoric, +pronominal], [+Agr]

Chinese type: [+finite clause], [+discourse/control domain/context], [+anaphoric, +pronominal], [-Agr]

Syntactic and pragmatic features of null subjects in different languages can be observed from the above distributional features. Critically speaking, the agreement-based condition of licensing *pro* doesn't have the distinctive feature to

<sup>4</sup> *Understanding Media: The Extensions of Man* (McLuhan, 1964)

<sup>5</sup> On the distribution and reference of empty pronouns, *Linguistic Inquiry* 15: 531-574. (Huang, 1984)

explain null subject phenomena, because it can't explain diversified situations of licensing *pro*. From the fact that Chinese also allows *pro* without agreement marking, we can conclude that *pro*-drop is also licensed by discourse or context.

#### IV. CONCLUSION

According to the research above, we conclude that there are various conditions in licensing null subjects. Null subject languages can be divided into three types on the basis of different *pro*-licensing conditions. Italian type *pro*-drop languages are licensed by strong subject agreement in INFL. Partial *pro*-drop languages like Turkish, German and Esperanto etc. have meager agreement and they show dependence on the discourse or the context to some extent. Chinese type *pro*-drop languages without subject agreement (either strong or meager) completely depend on the discourse or the context. Cross-linguistic variations in *pro*-licensing conditions display idiosyncratic features of individual language.

#### REFERENCES

- [1] Borer, Hagit. (1989). Anaphoric AGR, *Studies in Natural Language & Linguistic Theory* 15: 69-109
- [2] Chomsky, N. (1955). The Logical Structure of Linguistic Theory, mimeographed. Harvard, Plenum.
- [3] Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht: Foris.
- [4] Chomsky, N. (2004). Three Factors in Language Design. *Linguistic Inquiry* 36, 1-22.
- [5] Chomsky, N. (2005). On Phases, ms. MIT, Cambridge.
- [6] Haegeman, Liliane. (1994). Introduction to Government & Binding Theory. Oxford: Blackwell.
- [7] Holmberg, A. (2005). Is There a Little Pro? Evidence from Finnish, *Linguistic Inquiry* 36: 533-564.
- [8] Huang, C.-T. James. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531-574.
- [9] Huang, C.-T. James. (1989). Pro-Drop in Chinese: A Generalized Control Theory. In *The Null Subject Parameter*, ed. By O. Jaeggli and K. Safir, 185-214.
- [10] McLuhan, M. (1964). Understanding Media: The Extensions of Man. New York: McGraw-Hill.
- [11] Ouhalla, Jamal. (1994). Introducing Transformational Grammar. London: Arnold.
- [12] Platzack, Christer. (1987). The Scandinavian languages and the null-subject parameter, *Natural Language and Linguistic Theory* 5: 377-401.
- [13] Radford, Andrew. (2004). English Syntax: An Introduction. Cambridge: Cambridge University Press.

**Jinghua Zhang** was born in Jilin Prov., China in 1974. She received her PH.D. degree in English linguistics from Chungnam National University, Korea in 2008.

She is currently an associate professor in the College of Foreign Languages, Yanbian University, Yanji, China. Her research interests include linguistics and language education.