The Distribution of Null Subjects in Non-native Grammars: Syntactic Markedness and Interface Vulnerability

Juana M. Liceras, Anahí Alba de la Fuente, and Cristina Martínez Sanz
University of Ottawa

1. "Illicit" null subjects

1.1. "Illicit" null subjects: Heritage speakers

This study originated as a reaction to the proposal that the "illicit" null subjects identified by Montrul (2004)—the ***pro*** with an asterisk which appears in (1) – (3) below—evidenced how Heritage Spanish speakers are affected by the vulnerability of the syntax/pragmatics interface. Montrul and Rodríguez-Louro (2006) maintain that these data provide evidence against the unidirectionality of pragmatic deficits (Sorace 2004) according to which overt subjects but not null subjects would be overproduced. In other words, Montrul and Rodríguez-Louro (2006) maintain that null subjects are also overproduced.

Since null subjects are not supposed to bear a [+topic shift] feature, in example (1), the change of topic (el lobo, *the wolf*) should trigger the presence of an overt subject, because otherwise we do not know who filled the wolf's stomach with stones. However, the overt pronoun ELLOS (*they*) would not make the sentence less ambiguous because the agreement marker, the "N" of llenaro**N** is as explicit as the overt pronoun. This implies that this type of null subject is not "illicit" because it is null but because it creates an ambiguity which could neither be resolved by using an overt pronoun. We would like to argue that this type of ambiguity could also be present in native Spanish and therefore, the interface vulnerability attributed to these null subjects would not account for the specific distribution of null subjects in non-native grammars.

(1) Entró el abuelito con el perro y *pro* cortó el estómago del lobo y *pro* sacó a la abuelita y a el Caperucita Roja. Mientras estaba dormido el lobo ***pro** le llenaro**N** el estómago con piedras, y la abuelita estaba lista pa coserle el estómago. […]
"The grandfather came in with the dog and pro cut the wolf's stomach and pro took out the grandmother and Little Red riding Hood. While was sleeping the wolf, ***pro** filled the stomach with stones, and the grandmother was ready to sew him the stomach." [….]

[HS #205, intermediate (Montrul 2004 (30): 133)]

In the case of example (2), the overt subject pronoun ELLA (*she*) would not prevent this sentence from being ambiguous. Consequently, this example also shows that it is not the null subject as such

that is "illicit" but the need to establish whether it is Red Riding Hood or the grandmother who goes home with the grandfather. Once again, unless the specific DP is used, ambiguity will remain.

(2) … Despertó el lobo cosido, lleno de piedras, y el abuelito, la abuelita y la caperucita roja estaban riendo de él. Al fin todo salió bien y ***pro** se fue a casa con el abuelito.
"Woke up the wolf sewn, full of stones, and the grandfather, the grandmother and the Little Red Riding Hood were laughing at him. In the end everything went well and ***pro** went home with the grandfather."   [HS #205, intermediate (Montrul 2004 (30): 133)]

As for the null subjects identified as "illicit" in (3), there is a change of referent which makes the paragraph pragmatically incoherent. Here, once more, the presence of the overt masculine pronoun ÉL (*he*) would not make the sentence more coherent than the third person singular agreement marker "Ó" of the Spanish verb comiÓ. However, the null subjects are not ambiguous because the actual meaning of the verbs and the explicit feminine direct object (DO) clitic LA (*her*), as well as the presence of the DP direct objects la abuelita and la Caperucita, resolve the ambiguity.

 (3) Entonces la Caperucita roja encontró, *pro* fue a ver quién estaba en la cama, entonces *pro* encontró que era el lobo. ?*pro* estaba corriendo del lobo y entonces *pro* salió fuera, o *pro se la comiÓ, *pro se comiÓ a la abuelita y a la Caperucita.
"So, the Little Red Riding Hood found, pro went to see who was in the bed, so pro found that was the wolf. ? Pro was running away from the wolf and the pro went out, or *pro ate her up, *pro ate up the grandmother and the Red Riding Hood".

[HS #209, intermediate (Montrul 2004 (32): 133)]

Therefore, we would like to hypothesize that this type of discourse incoherence may also be present in native Spanish.

1.2. "Illicit" null subjects: Non-native grammars

In spite of their assertion that null subjects are overproduced due to the vulnerability of the syntax/pragmatic interface, in Montrul and Rodríguez-Louro's (2006) study, the intermediate group of learners of Spanish produces an insignificant percentage of "illicit" null subjects, while it is the group who produces the largest percentage of redundant overt subjects (Table 1) and also, crucially, the group who produces the largest percentage of agreement errors (Table 2). In other words, the production of agreement errors does not seem to trigger an increase in the production of "illicit" null subjects.

Table 1: Percentage of pragmatically redundant overt subjects and illicit null subjects: Native and non-native speakers[1]

	Overt Subjects		Null Subjects	
	Correct	*Redundant*	*Correct*	*Illicit*
Intermediate (n= 16)	77	22.9	99	1
Advanced (n= 16)	92.4	7.6	91.6	8.4
Near-native (n = 16	99.7	0.3	94.5	5.5
Control (n = 20)	100	0	99	1

[1]Adapted from Montrul and Rodríguez-Louro's (2006), table 3.

Table 2. Percentage accuracy on person and number verbal agreement: Native and non-native speakers[1]

	% Agreement Errors	
Intermediate (n= 16)	12.55	(61/468)
Advanced (n= 16)	2.28	(15/659)
Near-native (n = 16	0.13	(1/794)
Control (n = 20)	0.49	(5/1019)

[1]Adapted from Montrul and Rodríguez-Louro's (2006), table 2.

Using an experimental grammaticality judgment task (see 4) to test the choice of topic versus non-topic connected subjects by Spanish learners at different levels of competence, Lafond *et al*. (2001) found that non-native speakers chose (percentages were always rather low) null subjects as in B(1) when B(2), with overt subject Beth, is the pragmatically licit option here because a non-topic connected subject has to be overt. Note that the referent is a subject but the intervening subject (*you*) makes the null subject in B(1) "illicit".

(4) Sample of Grammaticality Judgment Task used to test choice of topic versus non-topic connected subject (Lafond *et al*. 2001: 126)

A:	Hola John	"Hi, John"
B:	Hola Ana	"Hi, Ann"
A:	¿Te gustaría almorzar conmigo?	"Would you like to eat lunch with me?
B:	Sí, me gustaría. ¿Puede venir Beth también?	"Yes, I would. Can Beth come too?
A:	Seguro, ¿a qué hora quieres ir?	"Sure, when do you want to go?
B:	1) Bueno, está en clase ahora. ¿Está bien a las 12:30?	"Is in class now. Is 12:30 ok?
	2) Bueno, Beth está en clase ahora. ¿Está bien a las 12:30?	**"Beth is in class now. Is 12:30 ok?**

However, when there is no intervening subject, null pronouns obey what Carminati has defined as the Position of Antecedent Hypothesis (PAH). This hypothesis states that "The null pronoun prefers an antecedent which is in the Spec IP position, while the overt pronoun prefers an antecedent which is not in the Spec IP position" (Carminati 2002:33). In other words, the null pronoun prefers an antecedent which is a subject. Violations of this hypothesis will result in pragmatically deviant null subjects.

Several studies, among them Sorace and Filiaci's (2006) and Belleti *et al*.'s (2007), have tested Carminati's PAH using a picture verification task. They found that near-native speakers significantly differed from native speakers when choosing antecedents for the overt subjects, both with Forward Anaphora, as in (5), and Backward Anaphora, as in (6), but did not differ at all when choosing antecedents for the null subjects.[1]

(5) Mentre lei$_{ki}$/*pro*$_i$ si mette il cappotto, la mamma$_i$ dà un bacio alla figlia$_k$
  'While she/*pro* is wearing her coat, the mother kisses her daughter'
(6) La mamma $_i$ dà un bacio alla figlia$_k$ mentre lei$_{ki}$/*pro*$_i$ si mette il cappotto
  'The mother kisses her daughter while she/*pro* is wearing her coat'

[Examples taken from Sorace and Filiaci 2006: 352]

Belleti *et al.* specifically state that while there were significant differences between the near-native and the native speakers in the interpretation of pronominal subjects, null pronominal subjects were produced at a comparable rate in spontaneous production by both the near native and the control groups (Belleti *et al*. 2007: 674).

1.3. "Illicit" null subjects: Native grammars

It seems that Carminati's PAH is not always respected by native speakers. In fact, we have found so-called "illicit" null subjects like the **pro* in example (7) produced by a native speaker of Dominican Spanish.

(7) Entonces cuando el gigante lo vio a él, David le dijo a él: tú vienes a mí con espada y jabalina, yo vengo contra ti en el nombre de Jehovah y cogió una honda. Puso una piedrecita así. Pero el gigante tenía todas esas cosas puestas y ahí mismo él agarró la honda, le tiró la piedra, y ahí mismo lo mató. Le dio ahí y cuando **pro* cayó, cogió la misma espada de él y le mochó la cabeza. "Then, when the giant saw him, David told him: you come to me with sword and javelin, I come against you in the name of Jehovah and (he) took a sling. (He) put a little stone like that. But the giant was wearing all those things and right there he grabbed the sling, (he) threw the stone against

---

[1] Accuracy (the hypothesis is not defined categorically, but in terms of "preference") was never 100%, neither in the case of the non-native speakers nor in the case of the native speakers.

him, and right there (he) killed him. (He) hit him there and when ***pro*** fell, (he) took his own sword of him and cut off his head."

<div align="right">[Martínez Sanz (forthcoming): Eliser, 8UW]</div>

In this example, the antecedent of ***pro*** is not the subject of the main clause and the null subject bears the [+topic shift] feature. While it is obvious that the one who fell is Goliath, the giant, no overt subject pronoun has been used there. This seems to imply that, under special circumstances, Spanish null subject pronouns can bear the [+topic shift] feature (contra Sorace 2000). The special circumstances may be that the narrator assumes that the interlocutor is familiar with the story, as in the case of Little Red Riding Hood, used by Montrul (2004) and Montrul and Rodríguez-Louro (2006). This is a rather interesting example because in spite of the fact that Dominican Spanish is a variety of Spanish which depicts a clear overuse of overt pronouns, here the speaker chooses to assign the feature [+topic shift] to a null subject pronoun.

## 1.4. Null subjects in Romance-derived Creoles

According to Lipski (1999) some null subjects are possible in Philipinie Chabacano Spanish, Mauritian Creole, Papiamento and Palenquero. However, null subjects in Romance-derived Creoles are an exception because they are always rare and seem to have different sources. Furthermore, the majority of these Creole null subjects are produced in main clauses, whereas there are very few instances of null subjects in embedded clauses which are coreferential with the subject in the matrix clause (double null subject configurations).

The fact that Creoles do not have Spanish-like bound pronominals (verbal agreement markers) which would identify null subjects, as well as the presence of null subjects mainly in main clauses leads Lipski (1999) to argue that these subjects are null constants (Lasnik and Stowell 1991). In other words, according to Lipski, rather than being instances of the null category *pro* which represents null subjects in languages such as Italian and Spanish, the Creole null subjects are instances of the null category which has been said to occur in English child language (Rizzi 1994) and in the English and French of diaries (Haegeman 1990).[2]

2. Null subjects: Minimalist accounts

In the GB framework, null subjects have to be licensed and identified (Rizzi 1986). In languages such as Italian and Spanish (pro-drop languages), they are licensed via the [+strong] feature of INFL and identified via the agreement markers (the subject bound morphemes). In fact, there are analyses of French which consider it a pro-drop language with subject clitic pronouns which act as identifiers, as shown in (8).

(8) Estudian español / They study Spanish / Ils étudient l'espagnol
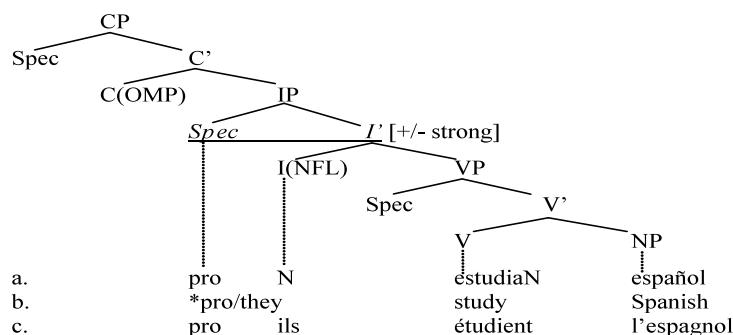
<div align="right">[Liceras *et al*.1998: 264]</div>

---

[2] This category is different from either pro, PRO, t, or NP-t in that it is [-pronominal], [-anaphoric], [-variable]. We have already shown (Liceras *et al*. 1998) that Spanish non-native null subjects are not null constants.

Since Creoles in general, and Romance-based Creoles in particular, do not have [+strong] agreement, they are not supposed to have Spanish-like or Chinese-like null subjects.[3]

Taking a Minimalist approach, Alexiadou & Anagnostopoulou (1998) go a step further in their analysis of I(nflection) in languages such as Italian or Spanish and propose that the set of phi-features of I is interpretable. For these authors, AGR is a referential pronoun and, consequently, there is no need for *pro*. Following Rohrbacher (1992) and Speas (1994) these authors propose that strong morphemes, the Spanish agreement markers such as the "N" in (8a), "…have individual lexical entries in the numeration while weak morphemes do not" (Speas 1994: 185). In fact, Alexiadou & Anagnostopoulou (1998) re-formulate the null subject parameter in relation to EPP checking, which implies that how the EPP feature is checked will determine whether a language is a [+null subject] or a [-null subject] language. The adoption of Speas' (1994) account of the distribution of null NPs in terms of economy principles leads them to propose that a given head is projected only if it has semantic and phonological content. Thus, in Spanish-like languages EPP is checked via merge, as shown in (9).[4]

(9) Pedimos   el alto el fuego
    ask for$_{1st.\ person\ plural}$ the cease to the fire
   "We are asking for a cease fire"
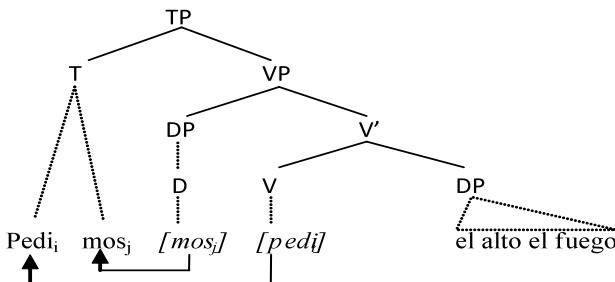
[From Liceras, Fernández-Fuertes & Alba de la Fuente, forthcoming, example (11)]

We would like to adopt this view of Spanish agreement markers as well as Roberts' (2001) proposal according to which this merge operation is an unmarked operation of core grammar.

EPP checking in English consists of moving the overt subject to the Spec-T position as shown in (10).

(10) We asked for a cease fire
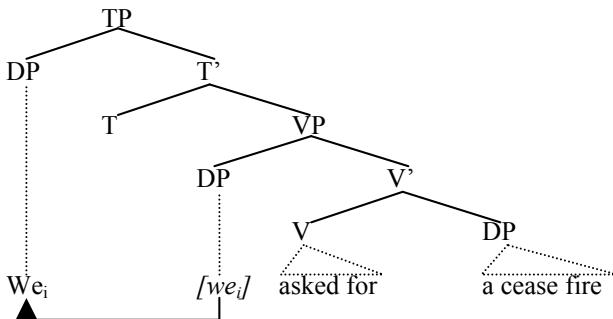
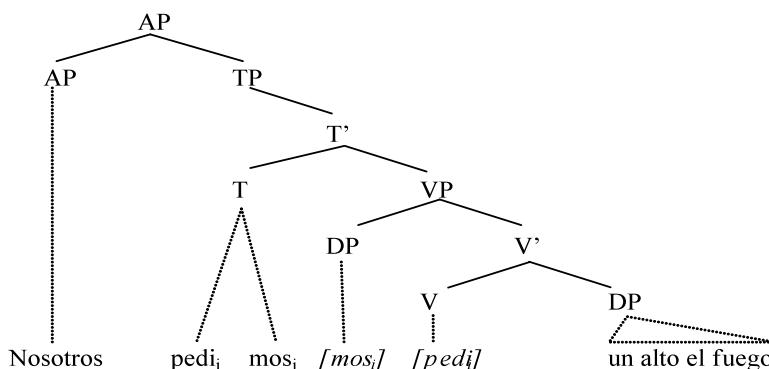[From Liceras, Fernández-Fuertes & Alba de la Fuente, forthcoming, example (12)]

---

[3] Creoles neither seem to have the discourse-bound operator which has been said to be the licensing mechanism for null subjects in Chinese-like languages (Huang 1984).

[4] There are other Minimalist accounts but basically they all consider AGR markers as clitic pronouns. For instance, Holmberg (2005) argues that the GB theory of *pro* cannot be maintained because phi-features are uninterpretable in the verb and cannot specify (identify) *pro*. Languages such as Spanish have an inherent null deficient pronoun which must enter an Agree relation with I containing D. The difference is that Holmberg (2005) argues that there is a pronoun even if it is deficient or not pronounced.

This is considered a marked operation by Roberts (2001) because it creates structure (the TP layer which hosts the subject DP).

As for the operation which adjoins Spanish overt pronouns to TP—to choose one possible analysis, specifically Kato's (1999)—we assume that it is a marked operation too because it creates an extra layer of structure (as shown in 11)—.

(11) Nosotros pedimos un alto al fuego
      "We are asking for a cease fire"

```
                        AP
              ┌──────────┴──────────┐
             AP                     TP
              │                      │
          Nosotros                   T'
                          ┌──────────┴──────────┐
                          T                      VP
                        ┌─┴─┐            ┌────────┴────────┐
                      pedi_i  mos_j     DP                V'
                                         │        ┌────────┴────────┐
                                      [mos_i]      V               DP
                                                 [pedi]        un alto el fuego
```

[From Liceras, Fernández-Fuertes & Alba de la Fuente, forthcoming, example (14)]

In fact, taking this as a marked operation seems to be also empirically justified because all the data that we have presented above indicates that these pronouns are problematic for non-native speakers, including near-native speakers.

Based on these syntactic analyses we would like to propose that:
• Learners of Spanish have to master two different sets of subject pronouns: (i) strong subject pronouns, which, by default, have the [+topic shift] feature, and (ii) weak subject pronouns (the agreement markers) which do not seem to have such a feature. As we have indicated above, previous research has shown that both sets of pronouns are present in non-native Spanish grammars from the early stages of development.
• Merge is the operation that leads to the incorporation of the weak subject pronouns in the structure. These pronouns are interpretable and part of the numeration. This merge operation is unmarked and, as such, it is not expected to be problematic for language learners. There seems to be a correlation between the pragmatically deviant use of overt pronouns and the production of agreement errors (Montrul and Rodríguez-Louro 2006). This correlation did not show in the case of the use of weak pronouns (Tables 1 and 2).
• The strong pronouns are adjoined and occupy a focus position. This is a marked operation and, as such, is expected to be problematic for language learners.
• Both native and non-native speakers produce and accept some pragmatically deviant null subjects (Sorace and Filiaci 2006; Belletti *et al*. 2007). However, non-native speakers' intuitions are not as clear-cut as native speakers' intuitions when judging pragmatically deviant null subjects (Lafond *et al*. 2001).

3. Our study

3.1. Research questions

Based on the proposal according to which Spanish has two sets of subject pronouns and on previous findings, we formulated three general research questions:

(i) Will native and non-native speakers' narratives show different patterns in terms of the distribution of overt and non-overt subjects?

(ii) Will native and non-native speakers' narratives show different patterns in terms of the relationship between the use of overt (strong) and bound (weak) pronouns?

(iii) Will null subjects with switch reference produced by native and non-native speakers in written narratives be different in terms of both quantity and pragmatic coherence?

3.2. Hypotheses

1. Null subjects will not be problematic for non-native speakers because the EPP checking via Merge is an unmarked operation of core grammar. However native and non-native speakers of Spanish may differ in terms of how they resolve ambiguity. Consequently:

• Non-native speakers of Spanish will produce fewer instances of weak subjects with switch reference than native speakers because they may have more problems with Spanish weak pronouns (the agreement morphemes).

• Intermediate and advanced learners of Spanish will produce fewer instances of ambiguous null subjects than native speakers because they will use overt pronouns instead.

• Given the findings of previous studies, the number of pragmatically deviant ("illicit") null subjects (weak pronouns that are instances of unresolved ambiguity) will be small both in the case of native and non-native speakers.

2. These non-native grammars will not show the pattern of distribution of null subjects which characterizes Romance-based Creoles because weak pronominals (agreement markers) are part of these grammars from the early stages. Consequently:

• Null subjects with switch reference will occur in both main and subordinate clauses.

• Instances of null subjects in embedded clauses which are coreferential with null subjects in the matrix clauses will not be rare.

3.3. Methodology: Participants, data elicitation and data analysis

We have analyzed data from two different types of narratives. The non-native (NN) participants were asked to write freely about the character that would result from choosing from a list of randomized characteristics regarding age, marital status, profession, hobbies, place of residence, etc. Each narrative had an average of 500 words. The Native participants (N) were asked to give their impressions about a short film they were shown.

We analyzed narratives from 15 intermediate and 15 advanced non-native speakers who studied Spanish at two North American universities. They were classified as intermediate and advanced on the basis of the placement tests used by the two universities and of the amount of classroom contact hours they had had.

We also analyzed narratives from 15 native speakers of peninsular Spanish. They were university students from various faculties at a Spanish university in Spain.[5]

We have chosen to analyze written narratives which do not tell previously known stories and do not have an immediate interlocutor which is assumed to be familiar with the story in order to avoid the potential reliance on the familiarity of the interlocutor with the story, as in the case of Red Riding Hood. This type of narrative was also chosen because the data are not highly dependent on performance issues due to, for instance, pressure to retrieve vocabulary.

To carry out our analysis we classified all inflected sentences with strong or weak subjects according to the type of sentence (main or subordinate), type of reference (same reference or switch reference), subject type (weak versus overt subjects), tense, person, means of identification of the weak subjects and type of ambiguity.

Overt subjects were further subdivided into personal pronouns versus other categories such as DP, CP, etc.

In terms of identification, we determined whether the weak pronoun was enough to establish the reference (AGR identification) or whether it had to be identified by some other means because it was potentially ambiguous (AMB). When a weak pronoun was ambiguous we indicated whether this was due to the actual weak pronoun (AMB-FLEX) or whether it was a case of discourse ambiguity (AMB-DISC), as in the examples (1) to (3) above where there are several potential referents (Red Riding Hood, the grandmother, the grandfather, the wolf). AMB-FLEX and AMB-DISC were resolved by several means: discourse cues such as verb meaning or initial referent, previously or subsequently mentioned DPs, proper nouns, overt subject pronouns, object clitic pronouns, reflexives or possessives.

3.4. Results

As Figure 1 shows, the Intermediate group's production of overt versus weak subjects was larger than that of the other two groups.
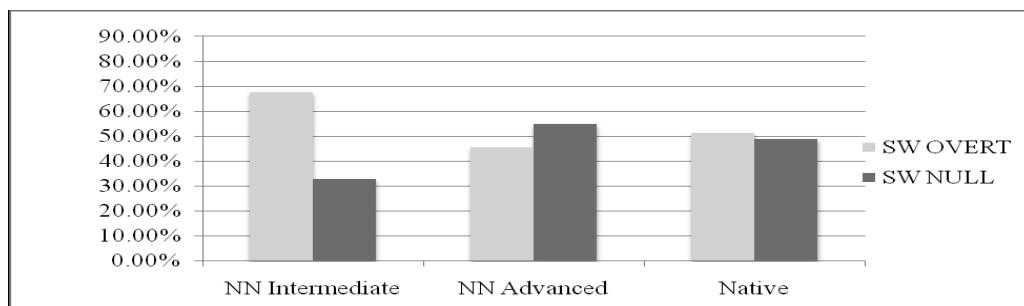

Figure 1. Production of overt and weak (null) subjects in switch reference contexts

The results of a Two-way ANOVA on the use of weak and overt subjects showed a significant effect of Group[1] $(F(2,42)=8.855, p= 0.001)$. A post-hoc analysis, using the Bonferroni Correction, shows that the NN Intermediate group differs significantly from both the NN Advanced and the Native groups, both in the use of weak and overt subjects (p< .004 in all cases). The Native and NN Advanced groups do not differ significantly in the use of weak (p= 1.00) and overt subjects (p= 1.00).

As for the identification of weak pronominal subjects, Figure 2 shows that weak pronominal subjects (bound morphemes = AGR) with switch reference which were non-ambiguous were used at a higher level by the NN intermediate group followed by the NN Advanced group. This group used the least number of weak pronominal subjects with multiple-referential value (AMB-FLEX) and the least number of third person weak morphemes which could create discourse ambiguity.[6]
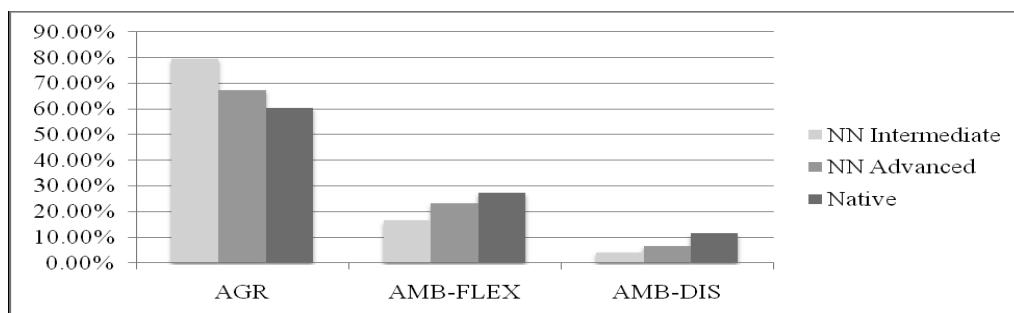

Figure 2. Identification of weak (null) subjects

The results of a Three-way ANOVA showed a significant effect of Group for AGR (p= 0.025), but not for AMB-FLEX (p=.087) and AMB-DISC (p=.086). A post-hoc analysis, using the Bonferroni Correction, shows that the Intermediate group differs significantly from the Native group with respect to the use of AGR (p=.027). None of the other comparisons yielded any significant differences (p>.08 in all cases).

The number of weak subjects which could not be clearly identified was extremely small (Table 3).

Table 3 . Percentage of "illicit" weak (null) subjects

GROUP	Total "illicit"	Total weak (null)	%
NN Intermediate	6	187	3.20%
NN Advanced	5	252	1.98%
Native	7	208	3.36%

---

[6] It may be that variety in the use of tenses increases with proficiency; it may also be that non-native speakers at the Intermediate level use fewer imperfect forms and also fewer means of identification (clitics, reflexives, etc.) and that Advanced speakers use them more and avoid strong pronouns when AGR is transparent in terms of establishing reference.

Examples of "illicit" weak subjects produced by the NN Intermediate group are listed in (12) to (14).

(12) Si (el) ballet no *estuve [fuera] parte de mi vida no *sabe [sabría] como *estoy[estar]: (el) ballet es mi y **pro** (yo) soy (el) ballet [502-W081-BUC]

If ballet were not part of my life I would not know how to be: ballet is me and I am ballet.

(13) Queremos dar a nuestros hijos la vida que mis padres me dieron. Ahora **pro** sabe (usted???) un poco acerca de Vicente. [511-W081-HAW]

We want to give our children the life that my parents gave me. Now you (singular-formal???) know a little bit about Vicente.

(14) Pero, si **pro** (uno/se) quiere ser una psiquiatra buena, la primera cosa que **pro** (uno/se) tiene que aprender es como hablar con muchas personas diferentes y tipos. [512-W081-HOO]

But, if (you) want to be a good psychiatrist, the first thing that (you) have to learn is how to speak with many different people and types (of people).

In (12), Spanish requires the overt pronoun YO even if the verb form SOY can only refer to first person because there is contrast. In (13), the interlocutor was never addressed directly, which makes it impossible to establish a referent for that null subject. In (14), the referent has not been mentioned before and that third or second person formal form of the verb seems to have a generic, "impersonal" value, which cannot be conveyed by the third person agreement marker of Spanish (it happens to be Ø).

Some instances of "illicit" weak subjects produced by the NN Advanced Group are listed in (15) to (17).

(15) Hace poco cumplí 33 años y empecé a tener dudas que había un lugar para mi o que **pro (yo?)** podía cambiar el rumbo rápido de mi vida. [602-W071-BLA]

Not too long ago I turned 33 and began to have doubts that there would be a place for me or that (I?/the rapid course of my life?) could change.

(16) Espero que **pro (ustedes?/ellos?)** comprendan y que **pro (ustedes?/ellos?)** me lo perdonen. [612-W071-GUA]

I hope that (you-plural-formal?/they?) understand and that (you-plural-fromal?/they?) forgive me for that.

In (15), it is not clear whether the narrator (YO) could change the course of his life or whether the course of his life itself could change. In (16), it was not possible to establish whether the narrator was addressing his parents as ustedes (as in previous sentences) or whether he/she was talking about the parents in the 3rd. person plural.

Some instances of "illicit" weak subjects produced by Native group appear in (17) – (19).

(17) En mi caso particular, la adolescencia se caracterizó por dar una vida excesiva a los papeles. Ante la indiferencia, y los complejos que creaba **pro (yo?/la adolescencia?)** construía mundos. [C06-JUF]

In my particular case, adolescence characterized itself by giving an excessive life to papers. Confronted with the indifference, and the complexes that it created (I? / adolescence?) built worlds.

(18) No era un conjunto de batallitas de diario o de conquistas de lo que **pro (yo?/the diary?/the diary's author?)** presumía. El plagio o la mentira no eran sus líneas. [C06-JUF]

It was not about the set of little diary stories or achievements that (I?/it?) was bragging about. Plagiarism or lies were not (its/the diary's author?) lines.

(19) Me miró sorprendido y me sonrió. En el folio **pro** (**él/yo**) llevaba las respuestas del examen y una nota. Cuando termines rompe el papel. [C10-SOL]

He looked at me surprised and smiled. On the sheet (he/I) was carrying the answers to the exam and a note. When you finish destroy the paper.

In (17), the switch of referent would require YO because CONSTRUÍA is ambiguous between first and second person. In (18), it is not possible to establish the referent for the subject of PRESUMÍA. In

(19), only the interpretation of the third person marker as ÉL (HE), not as YO would be pragmatically appropriate since HE is the subject of the previous verbs (MIRÓ and SONRIÓ).

The production of weak subjects in main clauses is larger in the case of the NN Intermediate group followed by the NN Advanced group, while the opposite trend applies to the production of weak subjects in subordinate clauses. Even though the NN groups produce more weak subjects in main clauses, they still produce a substantial number of weak pronominal subjects in subordinate clauses. This is different from the pattern which characterizes Romance-based creoles where weak subjects occur mainly in matrix clauses.
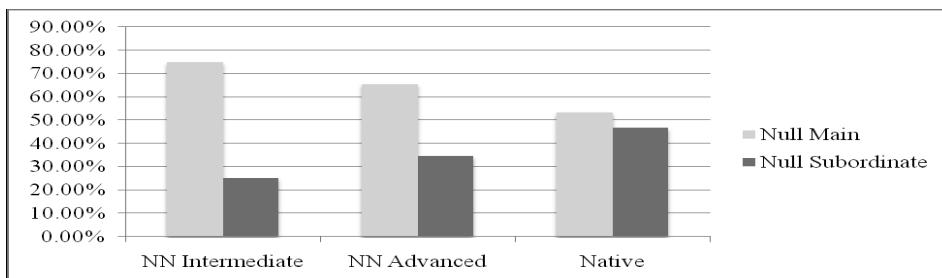

Figure 3. Distribution of weak (null) subjects in main and subordinate clauses

The total number of coreferential weak subjects in embedded and subordinate clauses produced by the three groups of subjects appears in Table 4.

Table 4. Production of coreferential weak (null) subjects: embedded and subordinate clauses

GROUP	TOTAL
NN Intermediate	36
NN Advanced	60
Native	39

Even if the Intermediate group is different from the Advanced and the Native groups in terms of overall production of weak subjects, means of identification of null subjects or use of null subjects in subordinate clauses, clear evidence that this group has mastered the two pronominal systems of Spanish—and specifically the system of weak pronouns—is provided by their use of coreferential null subjects in embedded and subordinate clauses (similar to that of native speakers). And even more telling is the fact that tense or agreement errors do not interfere with the establishment of coreference, as shown by examples (20) and (21).

(20) Supe la primera clase que *amé [amaba] bailar   (502-W081-BUC)
    (I) found out the first class that (I) loved-$_{preterite}$ [loved-$_{imperfect}$] to dance
(21) La *encuentra [encontré] en la calle el primer día que he ido al trabajo   (633-W071-DEN)
    (I) her find [found] in the street the first day that I have gone to work

As for the use of strong subject pronouns, Figure 4 shows that the use of strong subject pronouns by the NN-Intermediate group is different from that of the other two groups.
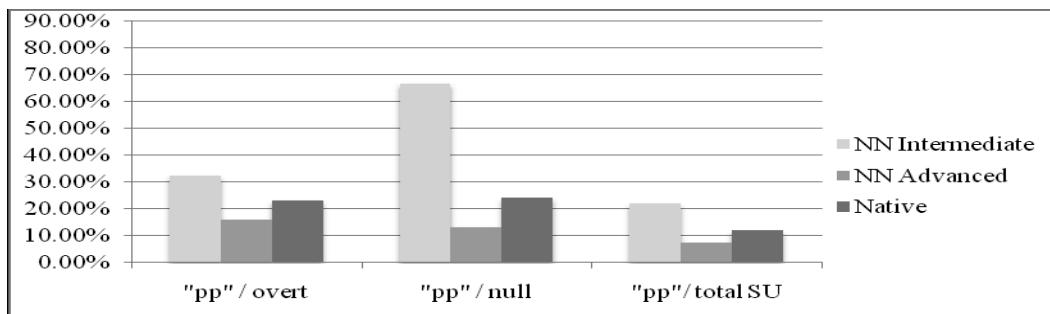

Figure 4. Production of strong (overt) versus weak (null) personal pronouns (pp)

The results of a One-way ANOVA indicate that there are significant differences among the groups. A post-hoc analysis using the Bonferroni Correction shows no significant differences between the Native and Advanced groups. There are significant differences, however, between the Intermediate group and both the Advanced ($p< .001$) and the Native groups ($p= .009$). This confirms the previous results (both obtained through narratives and via a picture verification task) which showed that overt pronouns are problematic for non-native speakers.

4. Conclusions

The first research question that we formulated was whether native and non-native speakers' written narratives would show different patterns in terms of the relationship between the use of overt (strong) and bound (weak) pronouns. Our data show that Native and NN speakers are different with respect to the overall production of overt subject pronouns versus weak subject pronouns (agreement markers) but only in the case of the NN Intermediate group.

As for whether native and non-native speakers' written narratives would be different in terms of both overall quantity and use of ambiguous weak pronominal with switch reference, our data show that the NN Intermediate group is significantly different from the NN Advanced group and the native group in terms of the overall quantity of weak subjects they produce.

Based on the analysis of narratives that we have carried out, we can conclude that Native and NN speakers of Spanish **do not differ** with respect to the production of "illicit" weak subjects—null subjects which lead to unresolved ambiguity—and this is so in spite of the problems that NN speakers have with the use of weak pronouns (agreement markers) and tenses. The production of "illicit" weak subjects is extremely small in all cases and the fact that NN Intermediate speakers use more overt pronouns does not seem to affect the number of "illicit" weak subjects. Our data support the conclusions of previous studies (Sorace and Filiaci 2006; Belleti *et al.* 2007, among others) where it has been shown that the number of pragmatically deviant ("illicit") weak subjects—instances of unresolved ambiguity—was not an issue for non-native speakers. We would like to argue that weak subjects are not problematic for NN speakers because EPP checking via Merge is an unmarked operation of core grammar.

We can also conclude that Native and NN speakers of Spanish **differ** in terms of how they resolve ambiguity. First of all, NN speakers produce fewer instances of null subjects with switch reference than Native speakers. However, this does not seem to be related to the fact that the NN speakers have problems with Spanish weak pronouns (agreement morphemes), which is precisely what the data analyzed by Montrul and Rodríguez-Louro's (2006) already showed.

Our data have also shown that, at least in written narratives, both Native and NN Spanish weak pronouns can bear a [+topic shift] feature.

We have also shown that even though weak subjects with switch reference occur in both main and subordinate clauses and that instances of weak subjects in embedded clauses which are coreferential with weak subjects in the matrix clauses are not rare. Both findings confirm our hypothesis that NN Spanish grammars do not exhibit the patterns of distribution of weak pronouns which characterize Romance-based Creoles, which implies that NN Spanish weak pronouns are not null constants.

 As for carrying out further research on these topics, we would like to suggest that the referential status of weak pronominals (agreement markers) has to be taken into consideration when investigating so-called null subjects. Furthermore, weak pronominals—as suggested by Lozano (2008)—should be analyzed in relation to the specific person (first, second, third) and number (singular or plural) they encode. Lozano (2008) has shown that vulnerability—mainly with respect to overproduction of overt pronouns—affects pronouns which carry [3rd person] and [+animate] features but not those which carry [1st person] and [-animate] features. Finally, we would like to propose that a distinction should be made between "illicit" weak (null) subjects which create ambiguity that cannot be resolved via an overt pronoun but requires an overt DP subject, and "illicit" null subjects which bear contrastive focus or violate Carminati's Position of Antecedent Hypothesis.

References

Alexiadou, A. & Anagnostopoulou, E. (1998). Parametrizing AGR: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory*, 16, 491-539.

Belletti, A., Bennati, E. & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language and Linguist Theory*, 25, 657–689.

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. Ph.D. dissertation, University of Massachusetts, Amherst.

Haegeman, L. (1990). Non-overt subjects in diary contexts. In J. Mascaró & M. Nespor (eds.). *Grammar Progress, GLOW essays for Henk van Riemsdijk* (pp. 167-179). Dordrecht: Foris.

Holmberg, A. (2005). Is there a little pro? Evidence from Finnish. *Linguistic Inquiry,* 36 (4), 533-564.

Hulk, A. & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition,* 3(3), 227-244.

Kato, M.A. (1999). Strong and weak pronominals and the null subject parameter. *Probus,* 11 (1), 1-38.

Lafond, L., Hayes, R. & Bahatt, R. (2001). Constraint demotion and null subjects in Spanish L2 acquisition. In J. Camps & C. Wiltshire (eds.). *Romance Syntax, Semantics and L2 Acquisition* (pp. 121-135). Amsterdam: John Benjamins.

Liceras, J. M., Fernández Fuertes, R. & Alba de la Fuente, A. (forthcoming). Overt subjects and copula omission in the Spanish and the English grammar of English-Spanish bilinguals: on the locus and directionality of interlinguistic influence. *First Language*.

Liceras, J. M., Laguardia, B., Fernández, Z., Fernández Fuertes, R. & Díaz, L. (1998). Licensing and identification of null categories in Spanish. In J. Lema & E. Treviño (eds.). *Theoretical Analyses on Romance Languages* (pp. 263-282).. Amsterdam: John Benjamins.

Lasnik, H. & Stowell, T. (1991). Weakest crossover. *Linguistic Inquiry,* 22, 687-720.

Lipski, J. (1999). Null subjects in (Romance-derived) creoles: routes of evolution. Paper presented at the Annual Meeting of the Society for Pidgin and Creole Linguistics, Los Angeles, January 8, 1999.

Lozano, C. (2009). Selective deficits at the syntax-discourse interface: Evidence from the CEDEL2. In N. Snape, Y.I. Leung & M. Sharwood-Smith (eds). *Representational Deficits in SLA* corpus (pp. 127-166). Amsterdam: John Benjamins,

Maranz, A. (1995). The minimalist program. In G. Webelhuth (ed.). *Government and binding theory and the Minimalist Program* (pp. 351-382). Oxford: Blackwell.

Martínez, C. (forthcoming). *Syntactic variation and the null subject parameter: the case of Dominican Spanish*. Ph.D. dissertation, University of Ottawa.

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: a case of morpho-syntactic convergence. *Bilingualism: Language and Cognition,* 7 (2), 125-142.

Montrul, S. & Rodríguez Louro, C. (2006). Beyond the syntax of the null subject parameter. A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In L. Escobar & V. Torrens (eds.). *The Acquisition of Syntax in Romance Languages* (pp. 401-418). Amsterdam: John Benjamins.

Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, 17, 501-557.

Rizzi, L. (1994). Early null subjects and root null subjects. In T. Hoekstra & B. D. Schwartz (eds*.). Language acquisition studies in generative grammar* (pp. 151-176). Amsterdam: John Benjamins.

Roberts, I. (2001) Language change and learnability. In S. Bertolo (ed.). *Language acquisition and learnability* (pp. 81-125), Cambridge, Mass.: Cambridge University Press.

Rohrbacher, B. (1992). English AUX-NEG, Mainland Scandinavian NEG-AUX and the theory of V-to-I raising. *Proceedings of the 22nd Western Conference on Linguistics* (WECOL 92).

Sorace, A. (2000). Differential effects of attrition in the L1 syntax of L2 near-native speakers. In *Proceedings of the 24th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data interpretations and methods. *Bilingualism: Language and Cognition,* 7 (2), 143-145.

Sorace, A. & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3) 339-368.

Speas, M. (1994). Null arguments in a theory of economy of projection. *University of Massachusetts Occasional Papers in Linguistics,* 17, 179-208.

Proceedings of the 2009 Mind/Context Divide Workshop

edited by Michael Iverson, Ivan Ivanov, Tiffany Judy, Jason Rothman, Roumyana Slabakova, and Marta Tryzna

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Liceras, Juana M., Anahí Alba de la Fuente, and Cristina Martínez Sanz. 2010. The Distribution of Null Subjects in Non-native Grammars: Syntactic Markedness and Interface Vulnerability. In *Proceedings of the 2009 Mind/Context Divide Workshop*, ed. Michael Iverson et al., 84-95. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2348.