The L2 Acquisition of Null and Overt Spanish Subject Pronouns: A Pragmatic Approach

Margaret Lubbers Quesada\textsuperscript{1,2} and Sarah E. Blackwell\textsuperscript{1}

\textsuperscript{1}The University of Georgia and \textsuperscript{2}Universidad Autónoma de Querétaro

1. Introduction

The present study examines the pragmatic rules for Spanish null and overt subject pronoun (SP) use in native Spanish speakers' and second language learners' oral narratives\textsuperscript{1}, focusing exclusively on Spanish first person singular references with tensed verbs. Previous research on the acquisition of Spanish by native English speakers has attested to learners' overuse and inappropriate use of Spanish overt subjects in discourse. For instance, Saunders (1999) observed the excessive use of proper nouns in the oral narratives produced by learners at the Novice and Intermediate-low levels, whereas overt SPs were rarely used until the Intermediate-high level (contradicting a transfer hypothesis); however, speakers at the Intermediate-high level used overt SPs "excessively, using a system that more closely resembles that of English (the learners' L1)" (p. 127). While subject reference may be achieved by a variety of referring expressions (e.g. definite and indefinite NPs, demonstratives, etc.), because our study focuses exclusively on first person singular reference, only first person singular verbs with either null subjects or the personal pronoun yo are examined. The majority of work on the acquisition of Spanish null/overt SPs has been from a generative perspective and has not examined the pragmatic rules learners of Spanish use in narrative discourse.

Based on earlier work on SP expression and discourse anaphora in Spanish, we identify a set of pragmatic rules of SP use in native speaker discourse defined in terms of pragmatic and discourse notions including the following: (1) referent saliency and the cognitive status of in focus (Gundel, Hedberg, and Zacharski, 1993), or the accessibility of the entity in the discourse (cf. Ariel, 1990); (2) focus, or the notion that non-presupposed (new) and contrasted information constitute the stressed (i.e. focused) constituents in utterances (Lambrecht, 1994) and thus must be overt; and (3) pragmatic weight, which assures that Spanish speakers will use overt SPs to make utterances more personally relevant, increasing their stake in what they are saying (Davidson, 1996). The rules we propose attempt to provide a more explanatory account for SP use in Spanish discourse. We then examine learner narratives to ascertain to what extent L2 speakers of Spanish develop these pragmatic rules, which are rarely if ever taught.

2. Use of null and overt SPs by native Spanish speakers

The question of how to account for the use and interpretation of overt versus null SPs in [+pro-drop] languages like Spanish has been addressed within various fields of linguistics, including theoretical syntax, sociolinguistics, and discourse pragmatics. It is a complex problem, given the fact that grammatical, semantic, pragmatic, cognitive, and socio-dialectal/contact factors can influence speakers' choice of null versus overt SPs. Numerous studies on SP expression in different varieties of Spanish have been carried out from a variationist perspective, focusing on the frequencies of

\textsuperscript{1}Data collection was funded by UCMEXUS and CONACYT (Mexico) (Grant No. CN0255). We express our gratitude to research assistant, Matt Kanwit of UGA, to Jien Chen and Wei Zhang of the Statistical Consulting Center of UGA, and to the two anonymous reviewers for their insightful comments. Any remaining errors are ours.

occurrence of SPs ("overts") versus null subjects ("nulls"), as well as probabilities of their occurrence when particular factors or constraints are present/absent (e.g. Cameron, 1994, 1995; Otheguy, Zentella, and Livert, 2007; Ranson, 1991; Silva-Corvalán, 1994, among others). In fact, as Otheguy et al. (2007) point out, the "variable use" of SPs in finite clauses is "one of the most widely studied features of Spanish dialectology," and, "the alternation between overt and null SPs occurs in all geographic and social varieties of the language" (p. 771). Other researchers have addressed SP expression versus omission from a discourse-pragmatic, functional, and/or cognitive perspective, showing how factors including mutual and world knowledge, informativeness and economy of expression, referent accessibility, topicalization, the semantic-pragmatic functions of verb tense and aspect, and negotiation promote the use or non-use of SPs (see Blackwell, 1998, 2003; Davidson, 1996; Hildalgo Downing, 2003; Matos Amaral and Schwenter, 2005; Morales, 1997; Silva-Corvalán, 2001; Stewart, 2003).

The discourse rules that native speakers follow for SP expression are more complex than grammars and textbooks would lead us to believe. In fact, textbook writers and even researchers often suggest that subject expression is a case of variable use or "optional use", though the notion of optionality of SP expression can only really be supported by limiting the analysis to sentences in isolation such as those in (1) (examples from Luján, 1999, p. 1277):

(1) a. Tú/Ø trabajas demasiado.
   b. ¿Crees que yo/Ø tengo alguna duda?

In fact, native speakers perceive a difference between subject use and omission (Luján, 1999), and the term “optional” is valid (if at all) only in a limited number of discourse contexts (Silva-Corvalán, 2001). Silva-Corvalán uses the following example of a context where overt and null subjects could alternate yet be interpreted as coreferential with Pepe (p. 154):

(2) Me vine con Pepe hoy a la oficina; Pepe/él/Ø vive cerca de mi casa.

In spoken discourse, however, there is an overwhelming preference for null subjects in syntactic environments where overt SPs could occur (Silva-Corvalán, 2001), and therefore, subject expression is really the "marked" choice (Blackwell, 1998, 2003). Traditional grammarians use the terms emphasis, contrast, and ambiguity avoidance to explain SP use, but they often do not clearly define what these involve (see, e.g., Bello, 1954; Alarcos Llorach, 1999); and, simply using these terms does not provide a sufficiently explanatory account for SP use in discourse. In the following sections, we discuss the notions of referent saliency, focus, contrastive focus, and pragmatic weight and propose five pragmatic rules for the use of null and overt SPs.

2.1. "In focus" cognitive status or saliency of the referent

Subject expression in Spanish may be partly explained in terms of the notion of psychological focus of attention or the cognitive attention state of in focus as defined by Gundel et al. (1993). An entity is in focus "if the attention of both speech participants can be assumed to be focused on it because of its salience at a given point in the discourse" (Gundel, 1999, p. 294). According to Gundel et al. (1993) the “in focus” status correlates cross-linguistically with the appropriate use of unstressed pronouns and zero (i.e. null pronouns). This means that minimal, unstressed and thus unmarked linguistic forms, including null subjects and unstressed clitic pronouns, will be used when the intended referent is clearly the most salient entity in the discourse context and thus the focus of attention. This status is related to the notion of “antecedent saliency” and to Bosch’s (1983) “aboutness hypothesis”, or the idea that, within a discourse, the most salient entity at any point is always the one that the discourse is “about”, and therefore the one most likely to be referred to anaphorically. Typically, this is

---

2 Silva-Corvalán (2001) hypothesized a lower incidence of SPs with preterit verbs than with present, imperfect, conditional, and subjunctive verbs, because preterits tend to foreground the action, while overt SPs would attract the focus of attention to the subject and away from the event (p. 162). Although Silva-Corvalán points to the need to further examine the possible relationship between tense-aspect and SP expression, analysis of this relationship is beyond the scope of the current study.
achieved in Spanish by the use of the most minimal forms available, null subjects and clitic pronouns (cf. Blackwell, 2003).

The syntactic saliency of a referring expression has been proposed as one of several factors that determine whether an entity is brought into psychological focus. For example, Gundel et al. maintain that “subjects and direct objects of matrix sentences are highly likely to bring a referent into focus, whereas this is not the case for elements in subordinate clauses and prepositional phrases” (1993, p. 279; see also Huang, 1994). Furthermore, the interpretation of references to salient entities in discourse implies that reference interpretation is recoverable on account of, not only the fact that the referent is currently in psychological focus, but also that the context, the mutual knowledge of the interlocutors, the semantics of the utterance, and verb morphology work in tandem to enable addressees to infer the intended interpretation (cf. Blackwell, 2003; Huang, 2000). Our first pragmatic rule states:

(3) Rule #1: "Salient Referent"
Use a null subject whenever the intended referent is in focus, i.e., the attention of both speech participants can be assumed to be focused on it because of its salience at a given point in the discourse, such that the subject reference can be inferred, taking into account factors including the context, the mutual knowledge of the interlocutors, the lexical semantics, and verb morphology.

An example of this pragmatic rule from the native speakers' narratives is the following, where the speaker uses null subjects to refer to both her "primer amor", the school, the boyfriend and herself together, herself alone, and the situation in focus. She introduces "las chavas" while the focus is still on the topic (what the discourse is about--her first boyfriend):


Here, null subjects are the default choice due to the fact that their respective referents are in focus.

2.2. Switch focus of attention

The "switch focus of attention" rule is rooted in the observed tendency in language use, captured by Levinson's (2000) neo-Gricean M(anner)-Principle, which instructs speakers to use a marked form (a less minimal, more complex, less usual expression) to implicate an alternative interpretation, i.e., one other than that which would have been implicated by the use of the unmarked alternative (here, zero) (see also Blackwell, 2003). In other words, overt SPs may be viewed as implicating a meaning (or reference) that a null subject would not (in most cases) encode. Our second pragmatic rule accounts for the fact that native speakers will tend to choose an overt subject expression to switch the focus of attention to another referent in contexts where that referent is (1) not currently in focus and (2) not the same referent as the one to which a null subject in the same context would refer. This rule states:

(5) Rule #2: "Switch focus"
Use an unstressed overt subject to switch the focus of attention from one referent to another, to indicate a change in subject and topic, and to refer to a referent other than the one that would be implied by the use of a null subject.

We refer to this rule as switch focus of attention instead of adopting the term switch reference, which Cameron (1995) defines as "a relationship of same or different reference between two sequentially ordered subjects" (p. 1). Matos Amoral and Schwentener (2005, p. 117) summarize the notion as it is used by Cameron as "essentially mean[ing] that the subject X of a given sentence/utterance S1 is not the same as the subject (Y) of the following sentence/utterance S2." While both notions may account for similar patterns of SP use in discourse, our rule places emphasis on the saliency of the referents in the ongoing discourse and the referentially contrastive power of overt SPs.
Null subjects are used in sequential sentences to refer to different referents, and two or more entities can be in focus simultaneously, yet inferences, as well as verb morphology, allow us to disambiguate reference. But, a switch in the focus of attention to another referent in discourse typically occurs with an overt subject expression. In the following example, the overt yo serves to switch our focus of attention from mi amigo to the speaker, where a null subject would not have achieved this:

(6) De la primera vez que [Ø] me enamoré no tiene mucho. [Ø] Fue el año pasado, de una persona que se llama Jonathan. El, [Ø] lo conoci porque [Ø] llegó un día a saludar a uno de mis mejores amigos y mi amigo entró. Yo estaba en la Preparatoria, y mi amigo entró a una clase y él se quedó haciéndome compañía y [Ø] empezamos a platicar, y así y este.

Initially, the focus is on the speaker herself and the situation at hand, then on both the speaker and Jonathan, licensing reference to them via minimal forms (null subjects and a clitic pronoun). The overt yo is used with morphologically ambiguous estaba, at a point in the discourse when the most salient referring expression is mi amigo. However, switch focus of attention can occur with null subjects as long as other factors (e.g. discourse context, the interlocutor's preexisting knowledge, verb morphology) disambiguate the reference. Example (7) illustrates how the morphology of terminé switches our focus of attention from the events and characters in the narrated film to the speaker himself:


2.3. Focus: New and contrastive

The term focus has been defined in various ways in the literature. According to Lambrecht (1994, p. 206), "focus has to do with conveying new information", and he defines the notion in terms of presupposition and assertion:

(8) [Focus is] "the element of information whereby the presupposition and the assertion DIFFER from each other. The focus is that portion of a proposition which cannot be taken for granted at the time of speech. It is the UNPREDICTABLE or pragmatically NON-RECOVERABLE element in an utterance." (p. 207)

Zubizarreta (1998) also defines focus in terms of the discourse notion of presupposition. It is the non-presupposed part of the sentence, whereas "the presupposed part of the sentence is what the speaker and the hearer assume to be the case (i.e. the shared assumptions) at the point at which the sentence is uttered in a discourse" (p. 1). In the Romance languages, the intonational nucleus is the focused element and the rhythmically most prominent word in the intonational phrase (Zubizarreta, 1998); it answers a relevant wh-question as in (9) (Hidalgo-Downing 2003, pp. 61-62; CAPS indicate the prosodically prominent, focused element, while in c, the entire sentence is the focus):

(9) a. ---¿Qué canta Juan? --- BOLEROS.  
    b. ---¿Qué hace Juan? --- CANTA BOLEROS.  
    c. ---¿Qué pasa en el salón? --- Juan canta boleros.

Zubizarreta maintains that "[i]n Romance NS [nuclear stress] is always assigned to the last (metrically visible) constituent" (1998, p. 78), but this rule is contradicted by the following example from Luján (1999, p. 1280), in which she suggests three acceptable ways of answering the same wh-question:

(10) ---¿Quién va a pagar la factura? a. --*Pago.  b. --Pago yo.  c. --YO pago.  d. --Yo.

In other words, the focal stress on the answer to a wh-question does not have to be utterance final. Hidalgo-Downing (2003) adopts a more general definition of focus, referring to it as the expression of the most essential and most salient information in the utterance. By most essential, the
focus can be new information, but it can also be information already introduced into the discourse and thus both the topic and the focus of the utterance (Hidalgo-Downing, 2003, p. 62). She identifies word order as the factor that distinguishes new from contrastive focus: focused elements that are new information or non-presupposed occur in sentence-final position; contrastively focused elements do not. Therefore, we could potentially answer the question, ¿Quién paga?, with the following:

(11) a. YO pago [y no tú]. (typically called contrastive focus)
    b. Pago YO. (typically called new, non-presupposed focus)

Contrastively focused elements are expressed with prosodic prominence in order to call the hearer's attention to the element for the purpose of contrasting it with other potential members of a contrast set (Gundel, 1999). Contrastive focus can call attention to an entity that is being introduced or reintroduced in the discourse.

For the purposes of this study, we view focused pronouns as prosodically prominent (as opposed to being overt yet neutrally expressed) and contrastive in some way, in that the pronoun is expressed as being in opposition to another element in the discourse (either expressed or implied). The opposition can involve a parallel construction involving two referents, or it may be implicit, between the person referred to by the overt SP and the other relevant entity; in any case, it involves the "contrapositioning" of at least two referents, as Enríquez (1984) aptly notes:

…el hablante contrapone el sujeto a otra u otras personas, bien para ratificar su posición, bien para oponerse a ella o bien para adoptar una posición diferente. En todo caso, lo que exige…es que exista la intención por parte del hablante de establecer una relación determinada entre dos o más actantes. (pp. 113-114)

Examples of such contrapositioning include the following from Enríquez (1984, p. 115, CAPS on focused pronouns added here), where we can easily see how subjects that have contrastive focus can constitute information that may be contrary to our expectations:

(12) a. YO vengo solo. [los demás no sé]
    a'. Vengo solo. [?]
    b. ÉL no se atreve a opinar. [pero su hermano siempre]
    b'. No se atreve a opinar. [?]

We propose the following rule to account for SPs when they constitute the focus of the utterance and thus have prosodic prominence and must be overt:

(13) Rule #3: "Contrastive focus"

Use a stressed overt subject whenever it introduces new/non-presupposed information (typically postverbal), and when it is in (explicit or implicit) opposition to another referent in the discourse.

The following excerpt illustrates this rule as used by a native speaker in our study:

(14) … [Ø] pasamos mucho tiempo juntas además como que [Ø] es una retroalimentación, yo dependo de ella y ella de mí, [Ø] podemos hablar de de tontería y…

2.4. Pragmatic weight

Davidson (1996) uses the term pragmatic weight to account for speakers' use of overt SPs to signal that the utterance is "more personally relevant and more vested with emotion" (p. 555), or to increase the speaker's "stake" in what s/he is saying (p. 551). According to Davidson, overt SPs used for this purpose are topicalized and thus typically utterance initial; they express greater personal involvement and increase the speaker's commitment to his/her utterance. Topicalized overt SPs can also convey that the contribution is a matter of greater importance or facilitates the speaker's taking of the floor at the beginning of a turn. For instance, Davidson, in his analysis of conversational data from
a 1981 study, *El habla de la ciudad de Madrid* (p. 545), observes that speakers often use overt *yo* with verbs of opinion, belief, and claiming (*creer, pensar, decir*) (p. 561), as well as knowledge (*saber*) (p. 559), when they want to add pragmatic weight to their utterances. In the present study, we observed the same function occurring with these verbs, as well as with verbs of emotion such as *sentir* and *querer*. Our native speaker narratives included numerous instances of *yo* to add pragmatic weight, including the following:

(15) a. … O sea verla así, muy muy débil, muy vulnerable, … *Yo* lo sentí muy feo. Ahorita ya gracias a Dios [Ø] ya está recuperándose,…  
   b. La primera vez que [Ø] me enamoré y *yo* creo que [Ø] ha sido la única,…  
   c. … primero [Ø] fue una amistad así muy bonita y como de tres, cuatro meses. De hecho *yo* siento que a veces [Ø] me hacía más caso a mí que a su novio,…

Our pragmatic weight rule for first person singular (i.e. speaker) reference is summarized as follows:

(16) Rule #4: "Pragmatic weight"
Use the overt SP *yo* with speech act verbs of claiming, belief, opinion, emotion, or knowledge to add pragmatic weight to your utterance, to take a firmer stance, to express a greater stake in, or emotional commitment to your assertion or to express that your utterance is highly relevant.

2.5. Epistemic parentheticals

According to Davidson (1996), certain verbs no longer have their truth-functional meaning when used by speakers to make metapragmatic comments about their previous or subsequent utterances. He points out that the verbs most prone to this are those of knowing, seeing, or watching (p. 557). Such expressions have been dubbed "epistemic parentheticals" by Thompson and Mulac (1991). They are often used to modify a speaker's belief in his/her utterance or hedge or mitigate the strength of the assertion. In addition, they can express how the speaker views or "positions him- or herself in relation to their utterance" (Davidson, 1996, p. 557). Spanish epistemic parentheticals such as *no sé* or *digo* reveal speakers' reflections on the content of their narratives. Such expressions constitute "asides" or "evaluative commentary" on the content of the utterance, and as such, they are not foregrounded statements, but rather deemphasized parenthetical expressions. This explains the fact that only null pronouns are used with parenthetical comments. Furthermore, epistemic parentheticals comprised of first person singular verbs and null subjects do not bring a referent back into focus, which explains why they are often followed by overt *yo* to serve this very purpose, as in example (18c) below. Our rule for the use of epistemic parentheticals states:

(17) Rule #5: "Epistemic parentheticals"
Use a null subject in epistemic parentheticals to evaluate the previous or subsequent utterance, or to hedge or mitigate the strength of, or express an evaluative commentary on the utterance.

Examples of the use of epistemic parentheticals from one native speaker's narrative include:

   b. … [Ø] hablo por mis hermanos que también [Ø] *digo*, este, mi mamá es enfermera, una …  
   c. …va a pasar a sexto de primaria. [Ø] *Digo* y *yo* a mi mamá, la admiro por ese ..

3. Research in the second language acquisition of SPs

Research in second language acquisition has made strides in identifying various syntactic properties of the null subject parameter and in determining which of these specific properties can be acquired by adult learners. From a parametric approach several studies have looked at learners' access to the clustering of properties of pro-drop languages, including null subjects, subject-verb inversion and *that*-trace (Liceras, 1989; Liceras and Díaz, 1998, 1999; Al Kasey and Pérez-Leroux, 1998; Lozano, 2002a; Isabelli, 2007). Others, also from a generative approach, have examined learners’
access to Montalbetti’s (1984) Overt Pronoun Constraint (Pérez-Leroux and Glass, 1997, 1999; Lozano, 2002b), a principle not available in the English speakers’ L1; and more recent work has considered the semantic and discursive interfaces with syntax from an Optimality Theory perspective (LaFond, Hayes, and Bhatt, 2001; LaFond, 2003). All have attempted to answer the question of whether or not the principles of Universal Grammar guide the process of acquisition.

One of the earliest studies on null subject acquisition in Spanish L2 is that of Liceras (1989) where she examined the acquisition of null pronouns by English and French speakers learning Spanish in the classroom. Liceras suggested that the acquisition of null subjects was a prerequisite for acquiring other related properties of the [pro-drop] parameter, such as verb-subject inversion and that-trace violations; however, her data could not provide evidence that suggested a clustering of properties for acquisition. Learners were successful in resetting the parameter for the null subject and subject-verb inversion but not for that-trace. Isabelli (2007), who examined the acquisition of these same properties among study abroad learners in Spain, confirmed Liceras’s findings. In a later study, Liceras and Díaz (1998) investigated the possibility that learners’ L1s affect how null subjects are licensed within their L2s. They analyzed the production of learners of three different typological languages (French, English, and Japanese), but did not find conclusive results. The authors conclude that L2 learners depend on the pragmatic functions of null and overt pronouns to convey and retrieve meaning. In Liceras and Díaz (1999), the authors reject their former claim that pronoun acquisition is due to the resetting of the [pro-drop] parameter when they examined data of L2 Spanish learners from different L1s and found again that all learners, regardless of L1, produced null subjects at the beginning and advanced stages.

Al-Kasey and Pérez-Leroux (1998) also looked at the acquisition of a cluster of two distinct domains related to the [pro-drop] parameter, specifically null expletives (e.g. [Ø] llueve) and null thematic subjects (e.g. [Ø] salió). The results showed that learners used both null expletives and null thematic SPs at about the same time, although null expletives were more consistently used. The authors conclude that this is evidence that these two properties are clustered and that learning one implies learning the other. In a more recent study, Lozano (2000a) attempts to establish the role of UG in learners’ acquisition of the obligatorily null expletives (ExpS) and the optional (according to him) referential pronominals (ProS). The results show that learners are sensitive to the differences between null and overt ExpS from the earliest stages of acquisition. Lozano concludes that learners have different mental representations for ExpS and ProS.

Pérez-Leroux and Glass (1997, 1999) explored the effects of the OPC, which states that overt pronouns are avoided when the pronoun is bound to a formal variable, such as the trace from wh-movement or quantifier raising (e.g. Nadie dice que él (Ø) ganará el premio) (Montalbetti, 1984). They found that learners produced more nulls for bound than for referential contexts and more overt pronouns for focus than for topic contexts. In both studies, the authors claim that access to the constraints of OPC are available to learners, and they master semantic contrasts that guide null pronoun use in Spanish even though they are not explicitly taught such a distinction (a finding similar to those we report for the intermediate and advanced learners in this study). In a similar study, Lozano (2002b) considers learners’ knowledge of both the OPC and the CFC (Contrastive Focus Constraint) in Spanish L2 among English and Greek speakers. Lozano’s results showed that English and Greek learners of Spanish performed like native speakers for the OPC cases, but differed on the CFC cases. The author suggests that this is because in English, pronoun coreference is less restrictive. In sum, Lozano views the results as evidence that both UG and the L1 are sources of knowledge in adult L2 learning.

LaFond, Hayes, and Bhatt (2001) and LaFond (2003) examined the constraint interactions of the [pro-drop] parameter from an OT perspective in second language acquisition. LaFond et al. claim that learners overgeneralize topic to mean subject, and therefore tend to drop all subjects, thereby suggesting that variant outputs of interlanguage grammars are the result of differences in the rankings of syntactic and discourse constraints. LaFond (2003) concludes that the learning process is one of recursive constraint demotion which is characterized in terms of the syntax/discourse.

Although early studies have hinted at the need to consider semantic and pragmatic factors in determining overt/null SP use and acquisition, and more recent research has incorporated at least the notions of contrastive focus and reference into their analyses, few studies have drawn upon theoretical work in discourse analysis and pragmatics. As Koike, Pearson, and Witten (2003) note, “[s]tudies utilizing discourse analysis have played an increasing role in the understanding of L2 learners’ gains in SLA proficiency, corresponding to a shift in focus to learners’ performance in communicative interactions” (p. 160). By using concepts from discourse pragmatics (e.g., referent saliency, contrastive
focus, pragmatic weight) and determining some of the rules that reflect native Spanish speakers’
language knowledge and use of SPs in spoken narratives, we attempt to determine whether L2 learners
acquire and use these same rules. In carrying out this study, we respond to a gap in our understanding
of Spanish L2 knowledge and use of null and overt SPs.

4. The present study

The research questions guiding this study are the following: 1) What are some of the pragmatic
rules that native Spanish speakers follow in their use of null and overt first person, singular SPs? 2) To
what extent do native English speakers studying Spanish as a foreign language follow these rules?
And, 3) if Spanish L2 learners follow these rules, which ones are learned early and which are learned
later? Based on our literature review, we have proposed five pragmatic rules for native speaker use of
null and overt SPs (Section 2): salient referent, switch focus, contrastive focus, pragmatic weight, and
epistemic parenthetical. We then examine the oral narratives of both native Spanish speakers and
two groups of L2 learners and compare the native speakers’ and learners’ use of overt and null first
person (1s) SPs in light of these rules.

4.1. Participants

The data were taken from the database of the UAQ-UCD Project for the Study of Spanish as a
The database includes 240 oral narratives produced by learners from five levels and native speakers
from Mexico. For the present study, we included 120 of the narratives of 30 learners from three levels
and 20 narratives from five native Spanish speakers for control. There were ten informants in each of
the three learner groups who were enrolled in first and second year basic Spanish courses and upper
level Spanish linguistics courses at the University of California Davis. The learners from Group 1 were
beginning students completing their third ten-week term at the university. Eight of the ten had had two
or three years of high school Spanish, but their initial placement exam did not place them in a second
year course. Three had had some French. Group 2 was made up of intermediate students enrolled in
the second term of the second year (fifth term). All but one had had from two to four years of Spanish
in high school, and the one who had had no high school Spanish had taken French for six years. The
students in Group 3 were enrolled in upper division linguistics courses, and all had had from two to
four years of Spanish in high school. Two had also studied French from four to six years. None of the
learners had participated in study abroad, and none were heritage speakers of Spanish. The control
group consisted of five Mexican undergraduate Modern Language students.

4.2. Instruments

In the original project, each participant was requested to complete four oral narrative tasks: a
retell of a five-minute segment of the Charlie Chaplin silent film, A Woman, a personal description of
another person, a personal narrative, and a description of the participant’s plans for the future.
Participants completed two forms: a bio-data form with information on language study background,
languages spoken in the home, study abroad experience, major, and other personal information; and a
consent form explaining the purpose of the study, the procedure, confidentiality, and compensation.
All participants were informed that the study was to compare U.S. and Mexican university students’
cultural perspectives on various topics. These forms were in English for the learners and in Spanish for
the control group. In the first session, participants narrated the film (task 1) and described the most
important person in their lives (task 2). The following day, in the second session, they described the
first time they fell in love or felt special about someone (task 3) and described their professional plans
for the future (task 4). The English speakers were given the instructions for the task in English and
were provided with some basic vocabulary in Spanish related to the topic and asked to narrate in
Spanish. The Spanish speakers were provided with the instructions in Spanish but were not given the
vocabulary. Participants were provided with five minutes before each task to plan and take notes (they
left the notes with the interviewer before leaving the session). Each session lasted approximately 45

3 See Appendix for a sample of the instructions for one of the tasks.
minutes. The narratives were recorded on a digital Sony recorder in a sound-attenuated room with only the interviewer (a research assistant) and the participant present. They were subsequently transcribed using conventional orthography. All data were collected and transcribed between November of 2002 and December of 2004. Although there were fewer 1s SPs in the narratives of the film, we decided to include them because there were multiple examples of epistemic parenthetical use.

4.3. Analysis

For this study, all null pronominal referential contexts were tagged in the transcriptions, distinguishing both person and number. Subsequently, using the MonoConc Pro 2.0 software, concordances were run on all overt SPs and the null tags. We therefore generated two lists, one with the sentences with overt SPs and one with null SPs. For our analysis, we focused only on the 1s SPs. Using the collocations generated by the concordance, we identified each referential context according to the five pragmatic rules proposed. If a 1s pronoun (overt or null) referred to a referent which was in focus because of its prominence in the discourse, it was marked as salient referent. For example, in Cuando yo tenía quince años, jugaba al fútbol, it is clear that the subject of jugaba is the same as tenía on account of the salience of the first person (the speaker) at this point in the discourse. We identified a pronoun as switch focus when the referent differed from the one in focus in the preceding discourse, as in, pero él sabe que yo no estaba feliz. For contrastive focus, we considered SPs that occurred in parallel constructions involving one 1s reference, for example: Yo vivo en Davis y él vive en Bakersfield. We marked pronouns with the pragmatic weight function with verbs of claiming, belief, opinion, knowledge, and emotion (e.g. decir, creer, pensar, saber, sentir), as in La primera vez que me enamoré y yo creo que ha sido la única, where the referent is fully recoverable from the previous discourse due to its salience, but the speaker is adding weight to his/her utterance by using overt yo. Finally, we marked pronouns for the epistemic parenthetical use when they combined with verbs of knowing (saber), speaking (decir), or seeing (ver), and were used as evaluative comments on the content of the speakers’ utterances, as in, No sé, pues era muy cómodo, and, hablo por mis hermanos que también digo, este, mi mamá es enfermera.

Each author analyzed 70 of the 140 narratives studied. The first sets of data were done jointly to assure analyst reliability. Subsequent sets were analyzed independently; however, throughout the process, frequent consultation occurred to compare analyses. The final results of the analyses were again corroborated jointly. Once all the narrative sets had been analyzed, frequencies were tabulated and percentages were calculated for each group’s distribution of overt and null 1s SPs for the different pragmatic constraints. Subsequently, chi-square ($X^2$) tests were carried out on the frequencies for each group and each constraint.

5. Results and discussion

The results show that all learner groups and the native speakers prefer null to overt 1s SPs in their narratives (see Table 1). From a count of 7,643 words in the narratives of Group 1, there were 123 1s SPs identified and of these, 67% were null pronouns and only 33% were overt. For Group 2, of 9,315 words, there is even a higher ratio of nulls to overts: 80% of the 199 1s SPs were nulls, and for Group 3, from a count of 16,118 words, 71% of the 377 instances were null. Among the native speakers, 65% of the 289 1s subject references were nulls (from a word count of 10,281). The chi-square test ($X^2=5.4$) reveals that the differences among the groups are not significant.

<table>
<thead>
<tr>
<th>Group</th>
<th>Percentage</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overt</td>
<td>Null</td>
</tr>
<tr>
<td>1–beginners</td>
<td>33</td>
<td>67</td>
</tr>
<tr>
<td>2–intermediate</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>3–advanced</td>
<td>29</td>
<td>71</td>
</tr>
<tr>
<td>4–native</td>
<td>35</td>
<td>65</td>
</tr>
</tbody>
</table>

$df=3$, $X^2=5.4$, $p=0.1449$, Cramer’s $V=0.0759$
As in previous studies (Liceras, 1989; Al-Kasey and Pérez-Leroux, 1998; Isabelli, 2007), we have evidence that even beginning learners are capable of resetting the null pronoun parameter.

5.1. Results for Group 1 (First year)

In general, the first year students prefer nulls to express a salient referent, 70 to 30%, and overts to express contrastive focus, but to a lesser degree, 61 to 39%. They have the most difficulty with switch focus, as learners sometimes switch the focus of attention to a different person via a verb with a null subject, but it is not immediately apparent about whom they are speaking until later in the discourse (depende en él, en la guerra pero [Ø]-1s quisiera viajar en Grecia y Turkey; conoci Melissa cuando [Ø]-1s era niña), although we do see a very slight preference for overts for switch focus, but only for 56% of the instances, and for both morphologically ambiguous verbs and for verbs whose morphology makes clear the reference (Ella tenía pelo largo y yo era en colegio; Mi papa vive en Bakersfield. Yo estudio en Davis y vivo en Bakersfield; pero él sabe que yo no se enamoraba). The data confirm Liceras and Díaz’ (1998) findings that these learners do not simply “transfer” L1 knowledge in the use of overt pronouns for contrastive focus, as can be seen in the lack of overt contrastive 1s SPs in some contexts (?[Ø] Estoy cristian y María está cristian también; me encanta mi papa, pero mi mamá es una mujer y ?[Ø] soy una mujer también). It is evident that these learners have difficulties with the constraints for all uses, except for the epistemic parentheticals (realized mostly as unanalyzed phrases such as no sé; no recuerdo; lo siento), but have begun to distinguish between the use of nulls for salient referents (Yo no sé qué [Ø] voy a trabajar en el futuro; [Ø] Soy de Idaho, pero no [Ø] quiero ir a Idaho para escuela) and overts for contrastive focus (él no llama me y yo también; Mi padre escribía muy bien y yo escribo también). They still do not distinguish pronoun use for the rule of switch focus where native speakers tend to use a higher frequency of overts. The chi-square ($X^2=13.93$, with a $p$-value of .003) reveals that there is a significant difference in the way these learners use pronouns according to pragmatic constraint. These results are presented in Table 2.

Table 2. Pronoun distribution – Group 1 (First year)

<table>
<thead>
<tr>
<th>Referential contexts</th>
<th>Overt</th>
<th>Null</th>
<th>Percentage</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salient referent</td>
<td>30</td>
<td>70</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Switch focus</td>
<td>56</td>
<td>44</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>61</td>
<td>39</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Pragmatic weight</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Epistemic parenthetical</td>
<td>6</td>
<td>94</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total number</td>
<td></td>
<td></td>
<td>123</td>
<td></td>
</tr>
</tbody>
</table>

$df=3, X^2=13.93, p=0.003$, Cramer’s $V=0.3365$

Although it appears that in some ways, the learners in this group are beginning to distinguish the different pragmatic uses of null and overt 1s SPs, morphological and agreement errors are common. Some authors (Liceras and Diaz, 1998) have suggested that by using the overt pronoun, learners are able to produce the correct verbal forms more often. We did not find evidence that there were fewer agreement errors with overt pronouns. However, there were more morphological errors for switch focus contexts than for salient referent and contrastive focus. It may be that lack of control over overt and null pronoun use for this context also contributes to morphological confusion; conversely, the errors for switch focus may be due to the learners’ inability to manipulate verb morphology when changing subject referents. The first year learners’ errors most often included the wrong person (i.e. [Ø] me siento muy bien con él y [Ø]-1s *va (voy) a muchas citas con él; En diez o quince años [Ø]-1s *necesita vive (necesito vivir) cerca de mis padres), but also the wrong verb morphology (No, no [Ø]-1s *viaje (viajo) mucho ahora) or non target-language forms (y yo *salé (sali) mucho con él; yo *conocé (conoci) a él en escuela secundaria). There were no instances of overt yo used for pragmatic weight, which suggests that these learners have not incorporated this use into their pronoun system.
5.2. Results for Group 2 (Second year)

The results of the analysis for Group 2, the second year learners, can be observed in Table 3. There are notable differences between this group and Group 1.

<table>
<thead>
<tr>
<th>Referential contexts</th>
<th>Overt</th>
<th>Null</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salient referent</td>
<td>16</td>
<td>84</td>
<td>131</td>
</tr>
<tr>
<td>Switch focus</td>
<td>68</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>100</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Pragmatic weight</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Epistemic parenthetical</td>
<td>2</td>
<td>98</td>
<td>34</td>
</tr>
<tr>
<td>Total number</td>
<td>199</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The greatest improvement in this group is in the categorical use of overts, 100% of the time, for the contrastive focus contexts (él jugaba deportes y yo también: Él sentía más para mí que yo sentía para él). There is also a slight increase in the use of overts for switch focus (ella tiene 35 años cuando yo naci; ella me da consejos cuando yo las quiero), from 56% usage in Group 1 to 68% in Group 2; and for the use of nulls for salient referent contexts ([Ø] Me siento muy cómoda con ella; mi mamá me enseñó a coser cuando [Ø]-1s era niña), from 70% to 84% for Groups 1 and 2 respectively. Like the beginners, the second year students continue to prefer nulls for epistemic parentheticals, 98% of the time (again in mostly unanalyzed phrases), and do not consider the use of overt pronouns for adding pragmatic weight to their utterances. Again, the results of the chi-square ($X^2=78.29$, with a $p$-value <0.0001) confirm that there is a significant difference in pronoun use among the pragmatic constraints (the results of Fisher’s exact test confirm the chi-square).

For this group there were far fewer morphological errors, although occasionally there were errors of person (pero cuando [Ø]-1s *empezó (empacé) en la escuela del colegio; no [Ø]-1s *quieres (quiero) tres muchachos); of mood or tense ([Ø] sé que todo va a cambiar cuando [Ø] *tengo (tenga) hijos; Si no [Ø] *trabajara (trabajo) como escritor; no [Ø] *trabajará (trabajaré)); and non-native forms (yo *teneré (tendré) bodo (una boda) grande en el verano; yo probablemente *teneré (tendré) niños). Again, like Group 1 these errors occur with both null and overt 1s SPs.

5.3. Results for Group 3 (Third/fourth year)

The results for the third and fourth year learners are presented in Table 4. The most important change in the use of 1s SPs from the previous two groups and this one is for the switch focus context (pero yo no puedo y ellas no entienden; ellas quieren que yo voy (vaya) a San Francisco con ella; empezó a dejar trabajar más como cuando yo nací); this group prefers overt 1s SPs in 96% of the cases, compared to 68% and 56% for Groups 2 and 1, respectively. Indeed, it is in this context where we have seen the greatest change among the three learner groups, from barely distinguishing null and overt 1s SPs among the first year learners, to preferring overt twice as much among the second year participants, to an almost categorical use among the 3rd/4th year learners.

<table>
<thead>
<tr>
<th>Referential contexts</th>
<th>Overt</th>
<th>Null</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salient referent</td>
<td>20</td>
<td>80</td>
<td>245</td>
</tr>
<tr>
<td>Switch focus</td>
<td>96</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>96</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Pragmatic weight</td>
<td>100</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Epistemic parenthetical</td>
<td>0</td>
<td>100</td>
<td>72</td>
</tr>
<tr>
<td>Total number</td>
<td>377</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results for the third and fourth year learners are presented in Table 4. The most important change in the use of 1s SPs from the previous two groups and this one is for the switch focus context (pero yo no puedo y ellas no entienden; ellas quieren que yo voy (vaya) a San Francisco con ella; empezó a dejar trabajar más como cuando yo nací); this group prefers overt 1s SPs in 96% of the cases, compared to 68% and 56% for Groups 2 and 1, respectively. Indeed, it is in this context where we have seen the greatest change among the three learner groups, from barely distinguishing null and overt 1s SPs among the first year learners, to preferring overt twice as much among the second year participants, to an almost categorical use among the 3rd/4th year learners.
The patterns of use we had observed with the first two groups for salient referent and contrastive focus contexts continue with this group, with 80% preference for nulls for salient referent contexts, and 96% preference for overts for contrastive focus (ella está estudiando negocios, yo estoy estudiando ciencias; ella va a tener una gran experiencia y yo no voy a tener una amiga). As for the previous two groups, the results of the chi-square ($X^2=174.69, p<.0001$) show that there is a significant difference according to constraint.

There were very few morphological errors for the learners in this group, only two for person agreement (y [Ø]-1s *va (voy) a verlo en el día de gracias; [Ø]-1s *tiene (tengo) un casa, la casa solamente para mi), none for tense or mood, and only two non-target language forms (no [Ø] puedo decir que that I regret it, que lo *arrepiento (me arrepiento), pero yo *viú él (lo vi) en mi clase de ingeniera). For these errors, there was one overt 1s SP and one English 1s SP. Again, we can not make any claims regarding learners depending on overt SPs to help select verbal morphology.

For the first time, we note that these learners are using overt 1s SPs in some way to add pragmatic weight to their statements (en el aspecto que yo quiero de un hombre...que yo quiero de mi vida; Pues, yo diré, mi madre, ella es mi amiga; no sé, pero yo puedo decir que él fue mi primer amor), although there were only eight instances. The last example contains an epistemic parenthetical, no sé, with an overt yo immediately following to express pragmatic weight, a pattern observed frequently in the native speakers’ narratives.

5.4. Results for Group 4 (Native speakers)

The results for the native speaker group are as we expected: the use of 1s SPs, although not categorical, show strong tendencies and are certainly not in free variation. The native speakers in this study prefer null SPs for salient referent contexts (93% of the time) and for epistemic parentheticals (100%). Overt SPs are preferred for switch focus, contrastive focus, and for adding pragmatic weight to statements throughout the narratives. The chi-square test ($X^2=227.74, p<.0001$) again confirms a statistically significant difference for pronoun use according to pragmatic constraint. We have already provided representative examples for these contexts in Section 2.

<table>
<thead>
<tr>
<th>Referential contexts</th>
<th>Overt</th>
<th>Null</th>
<th>Percentage</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salient referent</td>
<td>7</td>
<td>93</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>Switch focus</td>
<td>90</td>
<td>10</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Contrastive focus</td>
<td>100</td>
<td>0</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Pragmatic weight</td>
<td>100</td>
<td>0</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Epistemic parenthetical</td>
<td>0</td>
<td>100</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Total number</td>
<td></td>
<td></td>
<td>289</td>
<td></td>
</tr>
</tbody>
</table>

$df=4, X^2=227.74, p<.0001$, Cramer’s $V=0.8877$

6. Conclusions

The results of our analysis provide evidence that the use of overt and null SPs in Spanish is neither optional nor random either for native speakers or L2 learners. The L2 learners in this study, even in the beginning stages, produce null SPs for all contexts but prefer overt SPs for contrastive focus contexts, which confirms the results found in earlier studies. However, unlike previous research (Pérez-Leroux and Glass, 1997, 1999; Lozano, 2002b), we have distinguished between contrastive focus and switch focus and the data show that by the end of the second year of study, these learners are still not producing overt SPs to express switch focus to the same extent that the more advanced (third and fourth year) students and native speakers do. The beginning and intermediate learners have not learned the pragmatic rule of overt SP use for topicalizing or increasing the speaker’s “stake” in what s/he is saying, which we, following Davidson (1996), termed pragmatic weight. Interestingly, some of the speakers in the advanced group, without having been taught this use, start to use overt pronouns for
adding pragmatic weight to their utterances. Finally, the results reveal that even advanced learners in the third and fourth years are still using overt SPs to express a referent that is already salient in the discourse. Perhaps this is the result of the often-repeated yet inaccurate explanation that SP use in Spanish is optional. We could not find evidence that verbal morphology for these learners was sufficiently salient to signal the use of a null SP in these contexts. It is evident that they depend on the discourse constraints of a previously expressed and fully recoverable referent to determine (at least to some extent) the use of null and overt 1s SPs. Indeed, further research would need to examine the extent to which referents are recoverable in the discourse and how this impacts the use of overt and null SPs among L2 learners.

7. Appendix: Sample of task instructions

TASK 3:

Describe the first time you fell in love or the first time you felt special about someone. Describe in detail the person and the events related to your experience. Below are some vocabulary items that you may want to use. You will have five minutes to look at these words and to think about what you want to say. Tell the researcher when you are ready. You may NOT take notes for this task. The researcher may ask you a few questions about the person or the experience.

- enamorarse (de) ‘to fall in love’
- sentirse ‘to feel’
- salir con alguien ‘to go out with someone’
- abrazar ‘to hug or embrace’
- novio/novia ‘boyfriend/girlfriend’
- cita ‘date’
- besar ‘to kiss’

References


