1. Introduction

This work in progress studies leísmo, the marking of direct objects with the clitic le, as it occurs in two contact varieties of Andean Spanish. The data consists of recorded interviews with speakers of Andean varieties of Spanish, both bilingual and monolingual. The data was compiled in two different moments (1994 and 1997) and in two different regions: Lima, the capital city and Chota, Cajamarca, a monolingual region with a history of extended contact with Quechua about 100 years ago. Andean Spanish is the product of a language contact situation between Quechua and Spanish that began with the arrival of the Spanish language (and its speakers) to America in the 16th century. The Andean Spanish speakers in Lima are originally from the southern Andes who moved into the capital city for economic reasons or to flee the terrorist war during the 80s.

What prompted our interest in leísmo was the similarity of certain occurrences in both bilingual Spanish in Lima and monolingual Spanish from the Northern Andes. It was also striking that data for both varieties accounted for almost the same percentage of occurrences of the clitic LE with an accusative role (15%). Data correspond to 36 speakers of Andean bilingual Spanish and 29 of Andean monolingual Spanish.

Our initial research questions examined whether these occurrences of accusative LE are triggered by the same semantic-pragmatic constraints, and how different or similar are they from the constraints that explain accusative LE in non-contact varieties of Spanish. Also, an important question is whether these constraints indicate a path towards simplification which, at the same time, may point towards a process of grammaticalization of the verbal clitic, as a marker of agreement.

Thus, our working hypotheses are:

(a) The use of acc-LE will be constrained by the same pragmatic-semantic constraints found in other varieties with respect to animacy; and by similar discursive constraints with respect to clitic usage.
(b) The sociolinguistic distribution of leísmo in the monolingual variety of Cajamarca will replicate the distribution of leísmo according to oral proficiency that bilingual speakers show.
(c) The analysis of socio-historical factors will show that the existence of leísmo in Andean Spanish is motivated by a language contact situation.

2. Analysis of linguistic data

As we said before, the bilingual data examined in this paper was produced by native speakers of Quechua whose oral proficiency in Spanish has been aligned on a continuum from less to more proficient in the second language. On the other hand, the monolingual data from the Northern Andes was obtained from speakers whose first language is Spanish living in a region with a history of language contact with Quechua 100 years ago.

---

Given the coincidences in the data previously observed we would like to ask whether the variety in Cajamarca represents the result of the Spanish acquisition process. How do we connect the dots that link the proficiency continuum with the linguistic situation in the Northern Andes?

We expect that the linguistic analysis will help us clarify some of these issues. Finally, we want to examine these data as to what it is telling us in reference to processes of second language acquisition and acquisition of semantic-pragmatic features. If there is simplification is this an indication of grammaticalization of the clitic as a marker of agreement?

The following are samples of the use of acc-LE in the data.

**Bilingual data.**

**Low Proficiency Group**

(1) Sp: El chiquillo quiere agarrarLE ; está siguiéndoLE  
    Q: ¿Al perro?  
    Sp: Sí, al perro, acá está ya.

Sp: The child wants to catch him; (he) is following him.  
Q: The dog?  
Sp: Yes, the dog, here he is already.

**Mid Proficiency Group**

(2) Se hace remojar la ropa, LE hago remojar dos horas, después 0² lavo  
The clothes are soaked. I soak them for two hours, and then I wash them

**Very Low Proficiency Group**

(3) A jóvenes LE abusaban, LE engañaban. En las noches a veces 0 sacaban los militares, o si no otros, ¿no?, otros grupos venían y LO sacaban así.

They abused young people: They deceived them. In the evenings, sometimes, the soldiers grabbed them. Sometimes others did it. Right? Other groups came and grabbed them, like this.

**Monolingual Data**

(4) Q: Aah..¿y los hombres que cosa …?  
    Sp: …Los que estudian también, no solamente estudian, LES hacen trabajar, a [sic] su papás sobre todo los sábados o días que no tienen clases, o sino por las mañanitas LES levantan tempranito  
    Q: Aah.. and the men, what?  
    Sp: …Those who study also, not only study, they make them work, their parents, mainly on Saturdays or days in which they do not have classes, or during the mornings they wake them up early.

(5) Q: ¿Cómo es el asunto de...digamos las jóvenes, estudian?  
    Sp: sus papás esto no LES, no [les] educan, no porque no: no quieran sino, [sic] algunos LES falta dinero  
    Q: how is it with the young women, do they study?  
    Sp: their parents, err, do not educate them, not because they do not want to do it, but because some do not have enough money

² [0] means the absence of a clitic.
Sí, o sea que yo cuando, tenía algo de diez, doce años, ya me imaginaba lo que está(ba) [sic] haciendo mi papá, él siempre, mm, las noches, LE veía que rezaba…

Yes, when I was around ten, twelve years old, I imagined what my father was doing, he always, at night, I saw him that he prayed…

Whereas the monolingual versions tend to distinguish plural from singular, the bilingual data show avoidance of the plural form. The occurrences of plural acc-LE were random in the bilingual data, thus they were included in the analysis together with the singular forms.

Another interesting feature to examine is the correlation with animacy. We observed in both varieties a high tendency to mark the accusative function with the clitic LO (either by itself or doubled by an NP). However, when acc-LE is used, bilingual speakers tend to use it in reference to the level of animacy of the object: the more animate the object referent the higher the likelihood of a choice of LE.

This follows patterns identified in previous studies for other varieties of Spanish (Valdez 2002; Paredes 1996; Franco, 1993; Landa, 1993). Furthermore, it is remarkable that the monolingual data show that, with the exception of [-animate], LE is used exactly half the percentage of what bilinguals do to mark the animacy relationship³. Table 01 shows these results:

<table>
<thead>
<tr>
<th></th>
<th>Bilingual</th>
<th>Monolingual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N/Total/%</td>
<td>N/Total/%</td>
</tr>
<tr>
<td><strong>-animate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO</td>
<td>644/685/94</td>
<td>788/851/93</td>
</tr>
<tr>
<td>LE</td>
<td>41/685/6</td>
<td>63/851/7</td>
</tr>
<tr>
<td><strong>+animate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO</td>
<td>199/266/75</td>
<td>281/322/87</td>
</tr>
<tr>
<td>LE</td>
<td>67/266/25</td>
<td>41/322/13</td>
</tr>
<tr>
<td><strong>+human</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO</td>
<td>157/231/68</td>
<td>1305/1561/84</td>
</tr>
<tr>
<td>LE</td>
<td>74/231/32</td>
<td>256/1561/16</td>
</tr>
</tbody>
</table>

³ The correlation between animacy and accusative le has not developed extensively in monolingual Andean Spanish.
Although we don’t have percentage information on the relationship between information status and choice of acc-Le in monolingual Spanish from Chota, from the data examined we observed a similar pattern in terms of choice of acc-LE and the information status of situationally evoked referents. Observe example (8):

(8) Sp: entonces, (…) formaba parte del comité pro-construcción y aparte donde yo dormía eh, hacían las celebraciones, entonces’ el catequista me, me invitaba pes, mm, (…) cuando llegaban a hacer su celebración, yo echaba llave a mi puerta por donde ingresaba, y me venía a otra mi casa ¿no? *ajá
A: entonces eh, me iba dando cuenta que estoy quedando un poco mal, que con encerrarLES ahí a veces LES dejaba encerrados ahí y me quitaba, y, y entonces eh, seguía más o menos m, invitándome, el catequista. Bueno, yo le’ecía, ya cualquier día LES voy a acompañar.

Sp: then, I was part of the construction committee and besides where I slept, err, they did the celebrations, so, the catequist would invite me, when they arrived to make their celebration, I locked my door where he came in and I would come to my other house.

*yes
A: so, I was realizing that I was giving a bad impression, that locking them up sometimes I left them locked in and I left, and so, err, he continued more or less inviting me, the catequist. Well, I used to say to him, any day I will join you.

Notice that in (8) the referent to the acc-LES is evoked in the discourse. A celebration entails more than one person in a sort of gathering. The speaker makes a comment of having been invited by the catechist to the gatherings. He also mentions the fact that he would just open the door, let them in, and, then go to his room. He would leave them alone in the room. But, he also tells the catechist that any day he would join THEM. Also, notice the use of agreement number both for the acc-LES and the dative-LE (le decía).

What explanation can we draw to this point? Could we outline a pattern for both varieties that would connect them? We consider that acc-LE is triggered by the same semantic-pragmatic features: animacy and information status of the referent in both varieties. Moreover, we suspect that it is the combination of both features that is playing a significant role in the selection of acc-LE by our speakers. Let’s look at (9), a sample from a monolingual speaker:

(9) eh, por la propia experiencia, por la propia experiencia, o sea que, yo, hm., tenía, un caballo ¿no?, un caballo más o menos, de regular estampa, y LE quería y LE quiero al caballo, entonces eh, en época de lluvia yo LE hacía dormir, debajo de, del techado pues ¿no? de la casa

Err, from my own experience, my own experience, it is, that I, hmm, I had a horse, a regular size horse, and I loved him, I love the horse, so in the rain season I made him sleep underneath the roof of the house

In (9) the object referent is also accessible to the speaker; in addition, it is animate. The question is, however, why has accusative le not developed more extensively in monolingual Andean Spanish?

In an attempt to answer this question, we need to look at possible origins of acc-LE in these speakers. Bilingual speakers undergo a process of second language acquisition in which they use different learner strategies to process data in order to communicate. In the context of the object pronoun system, these Quechua speakers encounter the Spanish direct object (DO) system as well as the indirect object (IO) system.

4 The use of acc-LE clitic doubling could be explained by the change of verb tenses (imperfect and present).
Let’s examine the IO system, bearing in mind the assumption that dat-LE marks agreement. Givón (1984) has proposed that dative case-role marking is related to topical entities, in that dative clitic doubling will be required mostly when the NP referent is not easily accessible in the data. At the same time, due to the topicality feature of the NP referent, this will have high degrees of discourse saliency. Belloro (2007) explores this question and examines a total of 180 IO structures in the habla culta of Buenos Aires. Belloro finds that 85.6% of the dative verbs are encoded with a clitic, whereas only 12.8% are clitic doubled. Consequently, in the discourse, it is most likely to have constructions such as (10) than those represented in (11)

\[(10) \text{ dat-LE } \text{ di el libro} \]
\[(11) \text{ dat-LE } \text{ di el libro a María};\]

\[(10) \text{ I gave (him/her) the book} \]
\[(11) \text{ I gave the book to María} \]

This is the regular usage unless there is the need to access the referent of the clitic (the referent must be inaccessible). This is a key point for us to suggest a possible path in the reanalysis of the object system by bilingual speakers.

Before, however, it is necessary for us to examine another feature that both bilingual Andean Spanish and monolingual Andean Spanish from Chota share: (post verbal) direct object clitic doubling. Its manifestation varies in frequency, but in general we observe a neutralization of gender and number features into the acc-LO form. This structure is more extended than LEÍSMO, and has been studied extensively from different perspectives. An example of this structure is:

\[(12) \text{ acc-LO sembramos la arveja ‘We plant peas’} \]

An analysis of this structure in bilingual Andean Spanish shows that when the object referent is distant from the verb in the discourse, there will be a tendency to duplicate it with the acc-LO. (Paredes, 1996). Likewise, we can predict a similar trend in the monolingual data from Chota as the occurrence of DO clitic doubling replicates what happens in the bilingual data examined for this paper. Although there are not statistics for these occurrences in the variety of Chota, let’s observe (13) from Valdez’s (2002) data:

\[(13) \text{ Y todos creían pues los habitantes de acá que era la Virgen del Rosario, así lo pasaron varios años, con la Virgen del Rosario, se quedó luego en nada…} \]
\[\ldots \text{ Y trataron de sacarLO pues a la virgencita así todavía desfigurada con su capita vieja, todo tal y como había salido de la iglesia. LO llevaron a la iglesia, de la iglesia ya LO andaban todas las comunidades, de comunidad en comunidad LO andaban a la virgencita} \]

And they all believed, the inhabitants from here, that she was the Virgin of the Rosary, and in that way several years passed, with the (name) Virgin of the Rosary, then nothing… And they tried to take her out the Virgin, still disfigured with her old cape, everything as she had come out of the church. They took her to the church, from the church they took her to all the communities, from one community to the other they took the virgin.

In a way, clitic doubling seems to resemble the pattern found for dative doubling: an NP is doubled when the object referent is thought of being inaccessible or when the speaker wants to make it more accessible to the hearer. The second language learner, strategizing the process of acquisition of a complex object system, identifies LE(s) as the one form with the fewest complications: it does not mark gender, number does not seem to be that crucial and in terms of realization, the semantic constraints for the use of dat-LE and dat-LE doubling are easily processed.

---

5 Topicality indicates a unit whose features qualify it to be the topic of a sentence, i.e. people tend to talk about it.
6 The use of andaba is transitive in this sentence.
Similarly, the learner has the accusative forms that are more complex in that they need to mark not only gender, but number as well. LO is chosen as the accusative marker, disregarding gender or number. Moreover, direct object clitic doubling emerges favored by the same semantic-pragmatic features that explain Indirect Object doubling; i.e. saliency and topicality. Meanwhile, acc-LE is introduced in accusative structures as a competing acc-marker triggered by the same features that trigger dat-LE.

The reanalysis of the object system would follow the pattern:
(non-doubling) LE dative > dative- doubling
a) reanalysis
(non-doubling) LO acc > accusative-doubling
b) extension
(non-doubling) LE acc > accusative-LE doubling

3. Sociolinguistic trends

An examination of oral proficiency and schooling factors in our bilingual and monolingual data showed an interesting pattern between these two varieties. There seems to be a mirror image that connects the groups by proficiency levels and by schooling. Observe table 3 below.

<table>
<thead>
<tr>
<th>Bilingual Data Proficiency Groups</th>
<th>N/Total/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Prof. Group</td>
<td>LO 521/619/ 84</td>
</tr>
<tr>
<td></td>
<td>LE 98/ 619/ 16</td>
</tr>
<tr>
<td>Moderate Prof. Group</td>
<td>LO 270/312/ 87</td>
</tr>
<tr>
<td></td>
<td>LE 42/312/ 13</td>
</tr>
<tr>
<td>Low Prof. Group</td>
<td>LO 147/179/ 82</td>
</tr>
<tr>
<td></td>
<td>LE 61/226/ 18</td>
</tr>
<tr>
<td>Very Low Prof. Group</td>
<td>LO 63/74/ 85</td>
</tr>
<tr>
<td></td>
<td>LE 11/74/ 15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monolingual Data Schooling</th>
<th>N/Total/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-secondary</td>
<td>922/1098/ 84</td>
</tr>
<tr>
<td></td>
<td>176/1098/ 16</td>
</tr>
<tr>
<td>Secondary</td>
<td>180/226/ 80</td>
</tr>
<tr>
<td></td>
<td>46/226/ 20</td>
</tr>
<tr>
<td>Primary</td>
<td>539/621/ 87</td>
</tr>
<tr>
<td></td>
<td>82/621/ 13</td>
</tr>
<tr>
<td>No formal education</td>
<td>733/789/ 93</td>
</tr>
<tr>
<td></td>
<td>56/789/ 7</td>
</tr>
</tbody>
</table>

Notice that each level of oral proficiency in bilingual speakers seems to resemble each level of schooling in the monolingual Andean Spanish speakers, or vice versa. In other words, there is a connection between monolingual Andean Spanish and the bilingual variety that we would like to investigate in a future study.

4. The language contact situation

In order to understand the different constraints that influence variation and change in a variety that originated from language contact, we need to look at the different factors that contribute to the situation of bilingualism and shift. Thomason and Kaufman (1991) believe that it is the sociolinguistic history of the speakers that will determine the direction a change will take. In order to determine whether an external explanation is plausible for a certain variety it is necessary to consider factors related to intensity of contact (Thomason 1996). Intensity of contact involves factors of time and level of bilingualism. Factors that contribute to intensity of contact are a high level of bilingualism, socioeconomic and/or political pressure on the one speaker group to shift to the other language, length of contact, and relative sizes of the speakers’ populations.

The language contact situation between Quechua and Spanish originated over 500 years ago with the Spanish conquest. Andean Spanish is the product of this contact (Cerrón Palomino 1992, Rivarola
Andean Spanish is spoken as a native variety by monolingual speakers and by bilingual speakers who also speak Quechua. It is also spoken as a second language by some of those bilinguals (Escobar 1990). In this paper, we analyze data from bilingual and monolingual speakers. The monolingual sample underwent language shift about 100 years ago (Valdez 2002). Bilingualism in the Andean area has taken place over several centuries, an extended period of time that has allowed for extensive bilingualism. In the situation of conquest, the majority of the native population did not have access to education and consequently to more standard varieties of the language. Some of the learners’ errors fossilized and became part of a native variety transmitted by parents to children (Cerrón-Palomino 1989; Muysken 1984; Rivarola 1990). An example of this is the monolingual sample from Chota.

Since the 16th century, there have been pressures on the native population to shift into Spanish. Spanish has been the language of education, law and social advancement. Due to the situation of conquest, there was a large native population and a relatively smaller native Spanish speaking population, that of the conquerors. For this reason, access to the speakers’ native variety of the target language (i.e. Spanish) was limited. Speakers of what Muysken (1984) refers to as “interlanguage outputs” were in contact with each other, thereby transmitting this variety to each other and to their children.

The bilingual sample is comprised of speakers who speak Quechua and Spanish and who live in Lima. There is language contact at the individual level. The stages of language acquisition are manifested in the proficiency groups. These speakers, due to their occupations, are in contact with monolingual varieties of Spanish. They are in service jobs such as housekeeping, construction workers, gardeners, artisans. There is language contact between Andean Spanish and monolingual Spanish as well.

On the other hand, in the monolingual sample speakers are no longer in contact with Quechua in the area. However, their speech presents features of shift-induced interference (Thomason and Kaufman 1991; Valdez 2002). Intensity of contact in the monolingual sample in Chota is understood as a past situation in which two languages coexisted. Due to the fact that the region underwent language shift at least 100 years ago, shift-induced interference features entered the Spanish of this population. Geographical isolation has greatly contributed to the maintenance of these features in the Spanish of these speakers. At the time of the data collection, to arrive in Chota one had to fly into Cajamarca city and then take a bus on an unpaved road for eight hours. Once in the region, to reach the villages and speakers in that area one had to walk or ride a horse up the mountain, for as long as an hour and a half. These geographical conditions have caused this variety to be isolated from monolingual standard varieties of the language.

We use the concept of shift-induced interference (Thomason and Kaufman 1991) to explain some of the constraints that favor the occurrence of leísmo in our sample. Shift-induced interference is interference that comes about through imperfect learning of the language. We understand imperfect learning in our speaker sample as an ‘incomplete learning’ in which speakers have not mastered the constraints that relate to the DO and IO in the TL, in this case, Spanish.

Shift-induced interference can occur whether shift occurred or not. “The crucial sociolinguistic factor is not whether or not shift takes place, but whether or not there is imperfect learning by a group of people” (Thomason 1996:7). In our sample, both groups show imperfect learning of the language. We see that imperfect learning of the language took place whether there is shift (the sample from Chota) or not (the sample of bilinguals in Lima).

We have established here that the social conditions for the occurrence of leísmo can be traced to a language contact situation. However, the linguistic constraints that favor the use of leísmo in our sample conform to the use of LE in non-contact varieties of Spanish as well.

---

7 Or Aymara.
8 See definition below. For a mention of other features see (Valdez 2002)
9 TL refers to the target language
10 We have inherited Thomason’s terminology, however we are aware the implications it might have in the current trends of language acquisition. We will address that issue in another paper.
5. Leísmo in a language contact situation

In this section, we would like to address the question of external vs. internal causation as a plausible explanation for a given change (Thomason and Kaufman 1991; Silva-Corvalán 1994). We will show that the study of leísmo in Andean Spanish gives us information on processes of language contact as well as on how internal and external constraints interact in a variety of Spanish which originated from a language contact situation. The factors that contribute to linguistic change in bilingual communities have been extensively discussed (Thomason and Kaufman 1991; Silva-Corvalán; Mougeon and Beniak 1991 among others). We propose multiple causation as an explanation for the use of leísmo in this variety of Spanish.

As we have seen from the analysis discussed above, speakers are more likely to choose accusative LE when the NP referent is [+ human] [+ animate]. This conforms to what has been found for non-contact varieties of Spanish (Franco 1993; Klein 1981). In addition, speakers will more likely choose accusative LE for [+ human] [+ animate] referents when the referent has been situationally evoked or evoked in the discourse. This correlation between animacy, information status and the choice of LE has also been found for non-contact varieties of Spanish (Franco 1993). In this respect the change observed here constitutes a possibility internal to the Spanish system.

Among the changes that have an internal motivation are those motivated by the “relative complexity of a given paradigm” (Silva-Corvalán 1994:92). DO marking in Spanish constitutes an instance of this complexity. While leísmo is a feature present in different areas of the Spanish speaking world (Quilis et al. 1985; Klein 1981) it is not the norm for Peruvian Spanish. Klee and Caravedo (2005) find that leísmo is used in Lima in 4% by natives of Lima. We see that there is a higher percentage of leísmo in our sample: 15% for monolingual and bilingual speakers. The reinterpretation of the IO restrictions by bilingual speakers gives as a result a feature that while it is a possibility in the Spanish language it is not a frequent feature in Peruvian Spanish.11

At the same time, we find that the origin of accusative LE takes place due to a reinterpretation of constraints in a process of second language acquisition by these speakers. In this process, the second language learner, when coming into contact with DO and IO systems, identifies LE as the form with the fewest complications. The semantic constraints for the use of dative LE and LE doubling are easily processed. Moreover, using LE the speaker can avoid gender marking and number. In this process, the speakers introduce accusative LE as triggered by the same features that trigger dative LE. This reinterpretation of IO constraints applied to the DO constitutes an instance of external causation for the change under study due to the bilingual nature of the environment in which it originates.

In the acquisition process, speakers internalized the two forms, i.e. LE or LO, however the pragmatic restrictions that govern the occurrence of these forms in Spanish were reinterpreted. We see that speakers use both of the forms, as direct and indirect object. In other words, we see here a process of variation in which LE and LO coexist.

6. Conclusion

The analysis presented here and the discussion on oral proficiency and schooling points out some of the connections between bilingual Spanish in Lima and monolingual in Chota. In addition to having a similar percentage of occurrences, the acc-LE is favored in both varieties by the same discursive (accessibility) features. This, in addition to the fact that animacy is not a major trigger of acc-LE leads us to believe that the monolingual variety is more advanced in the grammaticalization process of the clitic as an object agreement marker. This idea is supported also by the presence of number agreement in that variety, which we do not observe in the bilingual Spanish.

11 We consider leísmo as a feature of Andean Spanish, and we consider Andean Spanish as a variety of Spanish, contrary to ideas that the use of le that is brought about by contact with a substratum indigenous language is not a part of the Spanish syntactic system. For this see De Mello (2002). We base this statement in the fact that what our speakers reinterpret are the constraints found to function in the realization of accusative and dative pronouns in Spanish.
Moreover, intensity of contact has been established by explaining the social and historical aspects that are involved in a language contact situation such as that of the Peruvian Andes. We have shown that the occurrence of leísmo is explained by an interaction between linguistic constraints shared by non-contact varieties of Spanish and by the socio-historical conditions that have led to the appearance of this feature in a language contact context. Therefore, we have shown that the interaction of internal and external causes for change can explain the use of leísmo in Andean Spanish.

References

Escobar, Anna Maria. 1990. Los Bilingües y el castellano en el Perú. Lima: IEP.
Rivarola, José Luis. 1990. La Formación Lingüística de Hispanoamérica. Lima: Pontificia Universidad Católica del Perú.