



Naturalistic L2  
acquisition and the  
formation of one  
immigrant's variety of  
Spanish: a case study

**J. CLANCY CLEMENTS**

**Abstract** Immigrants in their host country sometimes take advantage of the assistance offered in order to learn formally the host country language in a classroom setting. This is not always the case. This contribution examines the variety of Spanish that was created by an immigrant who arrived in Spain to work in the restaurant industry. Because of her work situation, she did not take formal Spanish lessons. Rather, she constructed her own variety of Spanish based on input she received on the job and in her interactions with native Spanish speakers. Three aspects of her grammatical system are the focus: tense-aspect marking, copula use, and the pronoun system. The solutions the immigrant created for communication show she has developed a grammar that is different from that of the native speakers of the host language, and its nature is largely predictable by appealing to frequency and detectability of forms in discourse and the process of L1 transfer.

**Keywords** naturalistic L2 acquisition, immigrant variety, frequency, detectability, L1 transfer

## 1. Introduction

Since the beginning of this millennium, with the wars and/or natural disasters in the Mideast, South Asia, and Africa, European countries have experienced a relatively strong series of immigration waves. Before these recent waves of immigration, there was of course also immigration, but it was on a different scale and the immigrants arriving were largely in search of a better economic future. The case examined in this article focuses on one immigrant from China, Jenny, who arrived in Spain in the mid 1980s in search of a better economic future. When Jenny arrived in Spain, there was already a Chinese-speaking community in Madrid, and she already had contacts within this Chinese-speaking community. She found employment quickly and settled into her new environment fairly smoothly. One key difference between her situation and that of the immigrants to Europe in the present millennium is that the members of the Chinese-speaking Madrid community then, and now, prefer to depend on community-internal connections for services rather than accept aid from government sources. The result of these circumstances on her learning of Spanish was that she ended up learning as much Spanish

as she needed in order to make a living and interact with Spanish speakers on a daily basis. That is, she was interested, first and foremost, in being able to communicate successfully with Spanish speakers of the host country regarding what she wanted and needed, and she constructed her variety of Spanish based on what words she was able to parse and produce in her successful interactions with Spanish-speaking friends and colleagues.

In this article, I look at the circumstances of this Chinese L1 learner of Spanish. I argue that her process of acquiring Spanish had principally to do with the circumstances in which she carried out her day-to-day life in communicating with Spanish speakers in their respective environments. In section 2, I offer some thoughts on types of contact languages to highlight the situations in which the speech of the Chinese-speaking immigrant in question formed. In section 3, I give a brief description of the Chinese-speaking communities in Spain, followed in section 4 by a description of the immigrant's situation. In section 5, I discuss three of the conventionized patterns found in her Spanish and offer concluding remarks in section 6.

## **2. Contact languages and how to define them**

The speech of naturalistic L2 learners is interesting for various reasons, two of which are: it can show us in some cases the processes through which speakers of one language acquire another that is typologically different, and it can give us insights into the initial stages of pidginization. When one thinks of contact languages such as pidgins, what often comes to mind is the image of plantation pidgins involving several languages in contact and people under the yoke of slavery. However, manuals on contact languages (e.g., Bakker 1994, Holm 1988, 1989, Mühlhäusler 1986, Sebba 1997, Winford 2003) remind us that pidgins also form in other language contact situations, such as those involving interethnic and/or immigrant communities, trade, or tourist situations. In the latter two cases, we also encounter instances of naturalistically learned L2 varieties instead of pidgins (e.g. Hinnenkamp 1984), and it is at times difficult to distinguish between pidgins on the one hand and untutored (aka naturalistically formed) L2 varieties on the other (cf. Blackshire-Belay 1991). Whether a particular variety is labeled a pidgin or an L2 variety depends on the definitions employed. Using Hall's (1966)

definition, for example, the linguistic system resulting from a day-long interaction between a tourist and a tourist guide can be a pidgin. However, by Bakker's (1994) definition (and by that proposed by Winford (2003) to a certain extent), a pidgin must be a stable linguistic system, with its own rules that one can speak to varying degrees of proficiency.

Immigrant situations involving workers (with or without their respective families) are defined by, among other things, long-distance displacement of one or more groups of people who need or want to live and work in another place. With respect to trade and tourist situations, the aforementioned manuals mostly focus on trade situations—that is, situations in which traveling traders develop a communication system with peoples living in the areas they visit. Short- or long-distance displacement of one or more groups is also implicit in the definition of this situation type.

Given that language contact situations concerning immigrants also involve semi-permanent to permanent long-distance displacements, the isolating mechanisms that inhibit immigrants from integration play a large role in the development of immigrant varieties of a given language. Taking isolation mechanisms in biology as a basis, Croft (2000: 199) distinguishes three types of isolation mechanisms for language as well: geographical, ecological, and reproductive. Geographical isolation is self-explanatory, referring to the spatial separation of speakers in two populations 'to a distance beyond the normal geographical mobility of the speakers.' Ecological isolation refers to social separation, whereby Croft equates ecological differences in biology with social differences in language. In the same geographical area, certain individuals may not talk to each other because they belong to different social, ethnic, and/or caste groups. In the case of immigrants in a host country, both ecological isolation and reproductive isolation are relevant. The native Chinese-speaking immigrant whose variety of Spanish is the focus of this paper is ethnically distinct from the Spaniards and she also has a different linguistic background from the Spanish-language speakers of the host country. Thus, social and conversational isolation have played a role in her acquisition of Spanish.

Jenny went to Spain to seek a better life economically, to achieve economic stability. The extent to which immigrants learn the host-country language depends on various factors, such as age of the immigrants when they reach

the host country, status of the immigrants in the country, attitudes of the immigrant and host-country communities, the extent of conventionalization of the immigrant variety within the community, and the typological differences between the language of the immigrant and host-country language.

The age of immigrants influences how well they learn a language and it can also influence the extent to which they interact with host country speakers. Adult L2 learners have been shown to learn languages differently than children. For example, Schuler (2017) found that the Tolerance Principle model predicts productive rule acquisition in children, but not in adults. She argues convincingly that cognitive differences involving memory and cognitive control can account for the different ways in which children vs. adults acquire productive rules in language.

The asymmetrical social status between different communities in contact, in this case Jenny and the host-country Spanish speakers, can lessen the access to Spanish. As will be described in section 3, the conditions in which Jenny worked limited her access to Spanish almost exclusively for the first several years of her residence in Spain.

I address certain features of Jenny's Spanish in detail in section 5 and the manner in which her motivation for learning Spanish and the typological distance between Chinese and Spanish shaped the variety of Spanish she created. First, it is important to give some background of the Chinese-speaking community in Spain and Jenny's work situation the first nine years of her residence in Madrid.

### **3. The ecology of the Chinese immigrant community in Spain and one of its members**

In this section, I present an overview of the Chinese community in Spain and of Jenny's situation leading up to the interviews recorded that were used as a basis for the description of her variety of Spanish.<sup>1</sup>

---

**1** The reader is encouraged to consult Li (2016), which contains rich, insightful studies involving identity and language use in many communities in the Chinese diaspora around the world.

### 3.1. The ecology of Chinese immigrant Spanish in Spain

The Chinese ethnic community in Spain is heterogeneous, breaking down along dialectal and extended familial lines. What they share, as Beltrán and García (2001) point out, is being Chinese in a foreign country. Spain has a history with the Chinese that spreads over several centuries. The Spanish encountered Chinese in their colonization of the Philippines in the sixteenth century. In the nineteenth century, Spain had Chinese workers in the mines and plantations of their territories in Peru and Cuba. In the twentieth century, Spain was home to itinerant Chinese peddlers who sold small items on the streets of the country's major cities in the 1920-1930s. In the 1950s, Madrid had resident Chinese circuses, and in the last half of the twentieth century there was a substantial increase of the Chinese presence in Spain's service industry. As the Chinese restaurant industry grew in Spain, services dependent on this industry also emerged, such as import companies, shops, travel agencies, Chinese vegetable farming, transport companies, skilled and unskilled construction labor for restaurant renovation, etc. Given that their work involved predominately restaurants and restaurant-dependent services, the largest concentrations of Chinese are in urban areas, with Madrid and Barcelona having the two largest communities. Ninety-eight percent of this community works in the service industry (Beltrán & García 2001: 285). As of 2000, there were unofficially around 60,000 Chinese in Spain, though the official number is 10,816. The families established in Spain are related to each other through multiple connections of kinship, economics, and place of origin (Beltrán & García 2001: 284–285).

Chinese communities abroad come from a society in which the superiority of their own values is an accepted assumption, though it is said to be more pronounced in The People's Republic than in Taiwan (Beltrán & García 2001: 291–293). This, along with other factors, has had an impact on the degree to which Chinese assimilate in their host countries. The Chinese in Spain emigrated from their homeland generally to seek a better life in terms of their economic or social status. Were they able to attain these goals in their own country, they would not emigrate to Spain or other countries.

As the Chinese language is the pillar of education, and education is highly prized in Chinese culture, the learning and maintenance of the native language in Chinese immigrant communities is important. According to

Beltrán and García (2001), the Chinese set up their communities to isolate and protect their language and culture from that of the host country where they live, and thus tend not to learn Spanish. Literacy data discussed in Beltrán and García (2001: 289–290) reflect this tendency. Of the 264 Chinese immigrants who applied for legalization in 1991, only 27% (72 people) could read and write Spanish. Indeed, Beltrán and García (2001: 291) note that ‘wherever the Chinese immigrate, they build their own communities in isolation from mainstream society, with their own culture, values, and language.’ One result of this is that they do not take advantage of the assistance the government offers immigrants, preferring to be self-sufficient and unattached to the host country. Integration, generally speaking, is neither a need nor a desire. One of several strategies the Chinese use to preserve their culture is to send their children back to China and then have them return as adolescents or young adults. These children end up learning little Spanish because they are not educated in the Spanish education system (Beltrán & García 2001: 292).

### 3.2. The informant<sup>2</sup>

Jenny came to Madrid in 1985 in her late twenties. Born in Nanking, Jenny is a speaker of Mandarin Chinese. She knew she wanted to emigrate and thus learned Chinese massage and acupuncture before leaving her homeland. Upon arriving in Madrid, she knew no Spanish and had studied Russian and English for no more than one year. During her first nine years in the Spanish capital, she worked in a Chinese restaurant, where she had most of her communication in Mandarin Chinese. For the first years of her residence in Madrid, Jenny lived with other Chinese women but then chose to live alone. Her main goal in Madrid was to become financially stable, and she learned Spanish to the extent that it permitted her to do so. As of 2003, she worked with Spaniards as a professional manicurist, masseuse and acupuncturist. She maintained little contact with the Chinese community in Madrid, preferring instead to spend her time with Spaniards, although at that time she had a

---

2 The information on Jenny's variety of Spanish is taken from data collected by the author. Some of the statistical analysis of her speech was first published in Clements (2003a), (2003b), (2005), and (2009).

sister and some Chinese-speaking friends in Madrid. I extrapolate from Jenny's description of her professional trajectory in Spain that the first nine years of relatively little Spanish input she received had as a result that the Chinese linguistic features in her environment were more prevalent than the Spanish features. Consequently, the Spanish variety she developed contained neither Spanish articles nor clitics, nor features such as gender and number agreement. Since 1994, she has increasingly become more integrated into Spanish society, and as a consequence, her vocabulary is substantial. Even though the richness of language she has had access to contains largely Spanish lexical material and structure, her language development is constrained by the system she created during the first nine years of her residence in Madrid.

The data used for the description Jenny's Spanish variety were collected in a set of interviews (in all three hours) at her residence and in a cafe. The author was the interlocutor. The topics covered were: birth place, education and professional life in China, family, reasons for emigrating, likes and dislikes of Spanish society, food, professional life in Spain, and the future in Spain. A loose phonetic transcription of the interview was used for the analysis.

An examination of the data reveals that Jenny's Spanish displays devoicing of voiced stops (*cuatro* 'framed paintings' instead of *cuadros*), the preference of CV structure (*dosə* 'two' instead of *dos*), the non-distinction of /r/ and /l/ (*abrá* 'speak' instead of *hablar*, *fuela* [< Spanish *fuera* 'outside'] used with the meaning 'leave'), lack of noun and verb morphology (*cuatro* 'framed paintings' instead of *cuadros*, *yo quiere* [lit. 1SG want-3SG.PRS] 'I want'), verb-final word order in certain constructions (*conmigo hablando* [lit. with me talking] 'saying to me'; see example 8 below), lack of articles and clitic pronouns (see examples below), and the reanalysis of certain forms or phrases (*fuela* from Spanish *fuera* 'outside' used as 'leave', *pasa* from Spanish *vas a* [lit. 2SG.PRS go to] 'you are going to' used as 'go' and *pasa* [3SG.PRS pass, go by]; see example 9c below). Her speech also reveals that she has developed a fairly stable pronominal system distinct from what is found in standard Spanish. Her variety represents a case of naturalistic L2 acquisition with restricted input due to her work situation the first nine years of her residency in Madrid. Her Spanish constitutes an individual solution to her

communicative needs in Spain, with clear evidence of L1 transfer, as well as certain tendencies in form-function mapping not traceable to her L1.

#### **4. Frequency, detectability, and stages in the formation of a naturalistic L2 variety**

As immigrants interact with host-language speakers, as Jenny has done with native Spanish speakers in this case, the emerging immigrant variety of Spanish can display conventionalized patterns, which Klein and Perdue (1992, 1997) have reported on. Immigrants naturalistically learning the host language would first acquire content words first because these denote entities, activities, events, and states in the world, which tend to be phonically more substantial than function words and thus more detectable in the speech chain. As a corollary, such learners would also acquire more frequently used forms first before less frequently used forms assuming both are equally detectable. I will return to this below in subsections 4.2. and 4.3. First, I review some of Klein and Perdue's findings and how they apply to the development of Jenny's variety of Spanish.

##### **4.1. Stages of development in naturalistic L2 acquisition**

Klein and Perdue (1992, 1997) report on findings carried out as part of a project named 'Second Language Acquisition by Adult Immigrants' that took place between 1981 and 1988 in five European countries (France, Germany, Great Britain, The Netherlands, and Sweden).<sup>3</sup> In the study, the speech of 40 adult learners representing six native and five target languages was analyzed over a two and one-half year period: native Punjabi and Italian speakers learning English, native Italian and Turkish speakers learning German, native Turkish and Arabic speakers learning Dutch, native Arabic and Spanish speakers learning French, and native Spanish and Finnish speakers learning Swedish. In their data analysis, three developmental stages were identified:

---

**3** There are interesting studies that focus on the L2 acquisition of certain features or constructions by Chinese L1 speakers. The interested reader may consult Jiao & Cuza (2022) and references therein.

Stage 1: Nominal utterance organization (NUO), during which speakers' utterances contain nominal elements.

Stage 2: Infinite utterance organization (IUO), during which speakers' utterances also contain verbs that are not marked for tense or aspectual distinctions.

Stage 3: Finite utterance organization (FUO), during which speakers' utterances contain nominal and verbal elements and there is evidence of tense/aspect marking. At this stage there is also evidence of L1 transfer.

Klein and Perdue (1992, 1997) note that, of the 40 learners whose speech they studied, about two-thirds (27/40) reached the FUO stage, reached the IUO stage, and that up to the IUO stage the development of all participants was found to be similar in that they all developed a relatively stable system to express themselves. This system, which was primarily determined by the interaction of a small number of organizational principles, was largely (though not totally) independent of the specifics of source and target language organization. Moreover, it was simple, versatile, and highly efficient for most communicative purposes. For about one-third of the participants (13/40) in the study, they note that 'acquisition ended on this structural level; some minor variation aside, they only increased their lexical repertoire and learnt to make more fluent use of the BV [basic variety]' (Klein & Perdue 1997: 303). Specifically, the learners developed a versatile, functional basic learner variety governed generally by three constraints: the focus of the utterance appears last (pragmatic), the controller of the utterance appears first (semantic), and the most common syntactic orderings in utterances are: NP1 V NP2, NP1 Cop NP2/AdjP/PP, V NP2 (in presentational utterances only).

Jenny's Spanish does contain substantial evidence that she is at stage 3 of development, However, her tense-aspect marking is not target-like, but innovative, and at times displays transfer. One illustrative example of this is her use of past participle forms to mark perfective aspect.

- 1 a. Mi tío macha-**do** Estados Unido. Tu sabe Yuan Taiwanveni-**do**?  
 my uncle leave-PPART United States you.SG know Yuan Taiwancome-PPART  
 'My uncle came to the United States. You know Yuan, who came from Taiwan?'

|      |    |         |                  |         |     |             |                           |
|------|----|---------|------------------|---------|-----|-------------|---------------------------|
| 1 b. | wǒ | yizhang | likai- <b>le</b> | Meiguo. | Nǐ  | renshi Yuan | Taiwan lai- <b>le</b> ... |
|      | my | uncle   | leave-PERF       | America | you | know Yuan   | Taiwan come-PERF          |

Note how the past participle morpheme *-do* corresponds directly to the Chinese perfective marker *-le*. My interpretation of this is that Jenny came to map the function of Chinese *-le* on to Spanish *-do*. And as we will see below, she also maps the function onto another Spanish form, as well. What this example shows is that Jenny, in making tense-aspect distinctions and transferring functions of Chinese markers into her Spanish, is at stage 3 of development, although her solutions for tense-aspect marking are not target-like.

#### 4.2. Frequency

In the acquisition process, the frequency with which a lexical form appears in discourse impacts how early in the acquisition process it may be acquired. To capture this, Andersen and Shirai (1996) proposed the Distribution Bias Hypothesis (DBH) to account for which forms of a lexical item are acquired first in the acquisition process. In essence, the DBH states that in L2 acquisition the frequency with which, say, a verb form appears in language use affects the order in which it is first acquired. In turn, the DBH is related to the Primacy of Aspect Hypothesis (POA) (Andersen 1993, Andersen & Shirai 1996). The POA states that the lexical aspect of a verb is important in determining in which verb forms a given verb most commonly appear in discourse. For example, stative verbs (such as *saber* ‘know’) and atelic dynamic verbs (such as *andar* ‘walk’) appear more commonly in imperfective forms than do telic punctual verbs such as *nacer* ‘be born’ or *llegar* ‘arrive’, which appear more commonly in perfective forms than they do atelic dynamic and stative verbs. To give an illustrative example, in one of Jenny’s interviews, she talked about a number of events related to her background. The verb form used in the narrative is overwhelming the 3SG present-tense form, independently of person, number, or temporal/aspectual reference. This is the case because of the conjugated forms of any given Spanish verb the 3SG present-tense form is the most frequently occurring in written and spoken Spanish, as shown in Clements (2009, 2014) by the token counts of verb forms from the Davies’ (2002–)

Spanish-language corpus data. There is an exception to this strong pattern though. When relating differences about her parents' backgrounds (place of birth), Jenny used the 3SG preterit form *nació* 'was born' (a telic punctual verb). The immediate question arises: why wasn't the 3SG present-tense form *nace* selected? The Davies' Spanish-language corpus provides a reasonable answer if we assume the DBH and the POA. Of all the present-tense, preterit, imperfect, and infinitive forms of *nacer* found in the corpus, the 3SG preterit form is the most frequent one, with 15% more tokens than the next most frequent form, the 3SG present-tense form (see Table 1). Following the POA and the DBH, the most frequently used form in the corpus is predicted to be the 3SG preterit form, given that *nacer* is a punctual (perfective) verb and that in any paradigm, the most frequent form is almost always the 3SG form. Thus, the prediction that *nació* is the most frequently used form of all of the paradigms of *nacer* is borne out. Although the use of a corpus in this manner is admittedly an imperfect tool to gauge verb form frequency in discourse, it nevertheless makes the correct prediction in the case of the forms of *nacer*, as well as in the large majority of other cases.

**Table 1. Present-tense, preterit, imperfect, and infinitival forms of *nacer* 'to be born' in Davies' Corpus del Español (2002–), listed by token frequency.**

| Person/Number/<br>Tense/Aspect | Form     | Tokens | Percent of Total |
|--------------------------------|----------|--------|------------------|
| 3SG preterit                   | nació    | 1,643  | 35               |
| 3SG present                    | nace     | 975    | 20               |
| INF                            | nacer    | 802    | 17               |
| 3PL present                    | nacen    | 568    | 12               |
| 3PL preterit                   | nacieron | 300    | 6                |
| 1&3SG imperfect                | nacía    | 154    | 3                |
| 3PL imperfect                  | nacían   | 84     | 2                |
| 2SG preterit                   | naciste  | 80     | 2                |
| 1PL present                    | nacemos  | 44     | 1                |

|               |           |       |      |
|---------------|-----------|-------|------|
| 1PL preterit  | nacimos   | 42    | 1    |
| 2SG present   | naces     | 14    | .3   |
| 2PL preterit  | nacisteis | 8     | .2   |
| 1SG present   | nazco     | 6     | .2   |
| 2PL present   | nacéis    | 1     | .1   |
| 2SG imperfect | nacías    | 1     | .1   |
| 1PL imperfect | nacíamos  | 1     | .1   |
| 2PL imperfect | nacíais   | 0     | 0    |
| TOTAL         |           | 4,723 | 100% |

If we assume that in the formation of an immigrant variety, speakers such as Jenny are negotiating a system of communication in which lexical meaning is most important and grammatical meaning is deduced through the situation and the context (i.e. pragmatically, as suggested by Mühlhäusler 1986), speakers will acquire word forms that are most frequently used in discourse and most easily detectable. We can define frequency as the number of times in a given corpus that a certain item or form appears. Following the principle of uniformity (McColl Millar 2007: 284), we assume that the most frequently used word forms in corpora are, roughly speaking, also the most frequently used forms in discourse in a contact situation and the ones that would be the most likely candidates for selection in the formation of an immigrant variety.<sup>4</sup>

### 4.3. Detectability

The notion of detectability is based on two uncontroversial distinctions and one observation. The observation is the ubiquity of CV structure in the world's spoken languages. As for the distinctions, those between stress-bearing vs. non-stress-bearing syllables and between free vs. bound morphemes are relevant here. Thus, for the purposes at hand detectability is defined in

<sup>4</sup> Bybee (1985: 123–27) discusses an additional factor linked to frequency, which she calls lexical connections, that is, shared phonetic material among different forms that make up a given paradigm. Lexical connections will not be discussed here.

relative terms: syllables containing or consisting of a CV structure are more easily detected in the speech chain than those without a CV structure (e.g. V or VC structure), and stressed syllables and free-standing morphemes are more easily detected in the speech chain than unstressed syllables and clitics/affixes, respectively. This is stated in (2).

2. Definition of Detectability
  - a. CV is more detectable than VC, V
  - b. stressed syllables are more detectable than unstressed syllables
  - c. free-standing morphemes are more detectable than clitics and affixes

Based on this definition, I assume that in the conventionalization process of a language-contact variety, the nature and extent of restructuring, and thus the relative importance of frequency and detectability in shaping a newly-emerging language variety, depend on the individual makeup of a given contact situation. At the same time, it must be acknowledged that the adult agents in the acquisition/creation process already know one or more languages and that in creating a language contact variety speakers may introduce into the new language elements from their own language(s), a process I have called transfer, but in the literature it is also known as imposition and interference through shift (van Coetsem 1988, Thomason & Kaufman 1988).

Having discussed stages of development, frequency and detectability, I will set up the discussion of Jenny's speech by showing the frequency and detectability of tense-aspect markers, copulas and pronouns. It turns out that frequency and detectability appear to interact in predicting which forms Jenny has selected in constructing her variety of Spanish.

## **5. Jenny's tense-aspect marking, copular forms, and pronoun system**

Although Jenny's Spanish variety does not resemble any standard-like variety of Spanish in her use of nominal and verbal forms and structure, it does show clear patterns, three of which are in her choice of her tense-aspect markers, her copula marking, and her pronominal system. In Spanish, tense and aspect are marked suffixally. In Jenny's speech, the past participle form occurs



- 4 c. Primero estudio, ahora **ya** casado con una señor.  
 first study now already married-PPART with a-FEM man  
 ‘She studied first; now she’s married to a man.’
- 4 d. **Ya** selado y selado luego yo pensando, me h<sup>o</sup> negocio.  
 already close-PPART and close-PPART then I thinking better business  
 ‘It closed, and after it closed, I thought, better [to launch] another business.’

While the use of *ya* and *-do* in Jenny’s speech are evidence of a developing completive or perfective-marking system, their use as fully conventionalized aspect markers is not apparent.

In Spanish, there are two copulas: *ser* and *estar*. Restricting myself to the infinitive and present-tense forms for the present discussion (the relevant forms), I assume that for Jenny all forms of both copulas were in theory possible candidates for selection as she formed her new linguistic system. Based on frequency counts gleaned from the Davies (2002–) Spanish-language corpus, shown in Tables 2-3, it is clear that the 3SG form is by far the most frequently-occurring form. I have included data from the sixteenth to the twentieth century, in addition to data from the oral portion of the corpus (twentieth century) in order to illustrate that in language use the distribution of the forms of these paradigms is fairly consistent across centuries and genres (written vs. oral).

**Table 2. Distribution of the infinitive and present-tense forms of Spanish *ser* ‘be’ in two periods and two genres.**

| Form      | 16th-18th C.   | 19th-20th C.   | Oral          |
|-----------|----------------|----------------|---------------|
| 3SG es    | 284,568 (62%)  | 257,156 (67%)  | 66,804 (77%)  |
| INF ser   | 81,376 (18%)   | 55,240 (14%)   | 7,351 (8%)    |
| 3PL son   | 67,371 (14%)   | 54,643 (14%)   | 10,239 (12%)  |
| 1SG soy   | 17,548 (4%)    | 10,210 (3.3%)  | 1,569 (1.5%)  |
| 2SG eres  | 6,768 (1.5%)   | 4,019 (1%)     | 337 (.5%)     |
| 1PL somos | 2,445 (.5%)    | 3,113 (.7%)    | 894 (1%)      |
| TOTAL     | 460,076 (100%) | 384,381 (100%) | 87,194 (100%) |

**Table 3. Distribution of the infinitive and present-tense forms of Spanish *estar* 'be' in two periods and two genres.**

| Form               | 16th-18th C.   | 19th-20th C.  | Oral          |
|--------------------|----------------|---------------|---------------|
| 3SG <i>está</i>    | 45,483 (49%)   | 47,014 (49%)  | 12,251 (47%)  |
| 3SG <i>es</i>      | 16,135 (17.5%) | 18,311 (19%)  | 4,614 (18%)   |
| INF <i>estar</i>   | 13,481 (15%)   | 10,581 (11%)  | 2,257 (9%)    |
| 1SG <i>estoy</i>   | 11,089 (12%)   | 9,944 (10%)   | 2,760 (11%)   |
| 1PL <i>estamos</i> | 2,242 (2.5%)   | 6,500 (7%)    | 3,192 (12%)   |
| 2SG <i>estás</i>   | 3,727 (4%)     | 3,523 (4%)    | 856 (3%)      |
| TOTAL              | 92,157 (100%)  | 95,873 (100%) | 25,930 (100%) |

The most frequently occurring form of all those in Tables 2-3 is the 3SG form *es* 'is'. If frequency were the most important criterion, we should find evidence of *es* in Jenny's Spanish. This is, however, not the case. If we take detectability into account the 3SG form *es* is disfavored because it does not contain or consist of a CV structure. Taking both frequency and detectability into consideration, the candidates most likely to be selected by Jenny would be *ser*, *son*, and *está*. Of these, this last one is the form found in Jenny's Spanish almost exclusively. An example is shown in (5).

5. Tú solo *está* allí mucho problema. Quien ayuda tuyo?  
 2SG alone is there much problem. Who helps your?  
 'For you to be there (in Spain) would be a lot of problems. Who would help you?'

However, the most commonly used copula is the null copula. Jenny used the null copula most with adjectival predicates, as in Chinese, but she also occasionally displayed the null copula with nominal or prepositional predicates, which is not found in Chinese. Examples include null copulas with both *ser* and *estar*, shown in (6).

- 6 a. Por ejemplo, ahora yo Ø (estoy) sola. Todavía no Ø (estoy) casado... yo  
 for example now I (am) alone still NEG married I

siempre Ø (estoy) con Yuen, está también hablando, yo no puede Ø (estar) sola.  
 always (am) with Yuen is also talking I NEG can.3SG.PRES (be) alone  
 ‘For example, now I am alone. I’m not yet married. I’m always with Yuen, and  
 am saying [to her that] I cannot be alone.’

6 b. Mamá poquito pénsa. Cabeza más ave. Dise, “tú Ø (eres)  
 Mamá a.little think head more agile says you.FAM.SG (are)

joven puede fuera para trabajo gána dinero.”  
 young can-3SG.PRES out for work earn-3SG.PRES money

‘Mom thought a little. She has a good head. She said, “you are young, you can leave  
 (China) to work and make money.”’

6 c. Papá dise tú no, (eres) una sola mujer,  
 Father says you.FAM.SG NEG (are) an alone woman,

porque fuera Ø (está) muy lejo.”  
 because outside (is) very far

‘Father said, “you can’t [leave China], you’re a single woman, because to leave  
 China, that would be going very far away.’

6 d. Mi papá mamá familia Ø (son) diferente.  
 my father mother family (are) different

‘The families of my mom and dad are different.’

Turning to Jenny’s personal pronoun system, she had several candidates to  
 select from. Here I have divided them into two categories: unbound pronouns  
 (with assigned stress) and bound pronouns (unstressed). The unbound  
 pronouns are shown in Table 4 and the bound pronouns are given in Table 5.

**Table 4. Unbound (free) pronouns in Spanish.**

|          |                         |
|----------|-------------------------|
| 1SG      | yo, mi, mío/-a, conmigo |
| 2SGFAM:  | tú, ti, tuyo/-a         |
| 2SGFORM: | usted                   |

|      |                               |
|------|-------------------------------|
| 3SG: | él, ella, ello, suyo/-a       |
| 1PL: | nosotros, nuestro/-a          |
| 2PL: | vosotros, ustedes, vuestro/-a |
| 3PL: | ellos/-as, suyo/-a            |

**Table 5. Bound Spanish pronouns (clitics).**

|          |                   |     |
|----------|-------------------|-----|
| 1SG      | me                | me  |
| 2SGFAM:  | te                | te  |
| 2SGFORM: | se                | se  |
| 3SG:     | lo, la, se, le    |     |
| 1PL:     | nos               | nos |
| 2PL:     | os                |     |
| 3PL:     | los, las, se, les |     |

Based on frequency of occurrence, all forms listed in Tables 4-5 have a substantial rate of occurrence, the least frequent being the 2PL forms. Thus, all forms would be candidates for selection. However, in terms of detectability, the expectation would be that the unbound forms with their own stress would be more detectable in discourse and thus more likely to be selected by Jenny as she developed her personal pronoun system. In addition, given that Chinese makes no distinction between subject and object pronouns (see Table 6), we would expect that Jenny's personal pronoun system would not distinguish between subject and object pronouns.

**Table 6. Subject and object pronouns in Mandarin Chinese (Li & Thompson 1981: 134).**

| Singular | Plural |
|----------|--------|
| wǒ       | wǒmen  |
| nǐ       | nǐmen  |
| tā       | tāmen  |

Assuming that many of the candidate forms for selection shown in Tables 4-5 were frequent enough in Jenny's interactions with Spanish speakers to be available, the prediction that the more detectable forms is borne out. That is, Jenny's pronoun system, shown in Table 7, exhibits no clitic pronouns. She selected from the candidate pool of unbound, more detectable forms.

**Table 7. Jenny's subject and object pronoun system.**

| Subject pronouns |          | Object pronouns  |                    |
|------------------|----------|------------------|--------------------|
| Singular         | Plural   | Singular         | Plural             |
| yo               | nosotros | mio <sup>6</sup> | nosotros           |
| tú               | -----    | tuyo             | -----              |
| él/ella          | ellos    | él/ella          | ellos <sup>7</sup> |

With regard to the distinction between subject and object pronouns, Jenny has developed a highly innovative system with stress-carrying pronominal forms. For 1SG and 2SG, she has reanalyzed the masculine forms of the unbound possessive pronouns *mio* 'mine', *tuyo* 'yours' as object pronouns. Moreover, she uses the 3SG subject pronouns *él* 'he, him' and *ella* 'she, her', as well as 1PL *nosotros* 'we', as object pronouns. Illustrative examples are given in (7).

7 a. Él no sabe mio.  
 he NEG knows mine  
 'He doesn't know me.'

7 b. yo no sabe el.  
 I NEG knows he  
 'I don't know him.'

**6** Spanish possessives (singular and 3pl) have proclitic forms (*mi(s)* 'my', *tu(s)* 'your', *su(s)* 'his, her, their'), as well as stressed, postplaced forms (*mío(s)*, *mía(s)* 'my', *tuyo(s)*, *tuya(s)* 'your', *suyo(s)*, *suya(s)* 'his, her, their') (cf. Picallo & Rigau 1999). Jenny interprets the former as possessives and has reanalyzed the latter as object pronouns.

**7** In Spanish, *él* 'he, him', *ella* 'she, her', and *nosotros* 'we, us' are used as a subject pronouns, but also after prepositions, as in *con él/ella/nosotros* 'with him/her/us'.

7 c. ¿Quién **ayuda** **tuyo?**  
who helps yours  
'Who would help you?'

7 d. Una amiga **presenta** **nosotros.**  
a friend introduces we  
'A friend introduced us.'

Jenny's Spanish shows signs of transfer from Chinese in the manner in which she sometimes marks the indirect object argument. In (8), she marked the 1SG indirect object pronoun using the unbound form *conmigo* 'with me', the order of which relative to *hablando* she seems to have calqued from her native Chinese *shūshu dui wo shuō* [lit. uncle to 1SG say] 'uncle said to me'.

8. yo dise tio **conmigo hablando**, puede fuera China.  
I says uncle with-me talking can-3SG outside China  
'I said, "uncle was telling me I could leave China."'

In addition to the pronominal system, Jenny has developed an innovative manner of marking indirect objects, using either *para* 'for' or *con* 'with'. Examples are given in (9).

9 a. Un gente presenta una presona **para mi.**  
a people introduce a person for me  
'A person introduced me to a person.'

9 b. yo siempre sale levista, pasa allí, coge,  
I always comes.out magazine goes-by there picks-up  
sube casa, milando, luego bajo, volvé **para ella.**  
goes.up house looking then descend-1SG.PRES return-INF forher  
'I always, when the magazines hit the stand, go by there, pick them up, go up to my place, look at them, then take them down and return them to her.'

- 9 c. Ya **conmigo** dice, “tú no pasa Australia,  
 already with-me says 2SG NEG goes-to<sup>8</sup> Australia  
 con nosotros fue para Malaga, o Almería.”  
 with us went for Malaga or Almería  
 ‘She said to me, “if you don’t go to Australia, you can come with us  
 to Málaga or Almería.”’

In sum, it is apparent that Jenny has developed patterns in marking tense-aspect and in her use of copula and pronouns, and that these patterns can be seen as an emerging grammar she has developed naturalistically in order to communicate with friends and colleagues.

## 6. Concluding remarks

In this article, I have presented features of the naturalistic variety of Spanish created by Jenny, a native speaker of Chinese who immigrated to Spain to work and live. I focused on three areas: her tense-aspect marking, her use of the copulas, and her pronoun system, and the manner in which she marks arguments.

Because of her work situation, Jenny did not pursue formal Spanish lessons. The first several years of her residence in Madrid, she lived and worked with native Chinese speakers in the restaurant industry. Once she obtained employment in another sector, she began to interact with native Spanish speakers daily, but at this point she had already created a variety of Spanish that suited her for purposes of work and interaction with friends and colleagues. I submit that she began to construct her variety of Spanish while at Klein and Perdue’s stage 2 of development and built on the system of that stage to then create a variety in which through transfer and innovation she increasingly was able to mark tense-aspect, use pronouns, and mark case. Interestingly, Jenny’s selection of Spanish *ya* ‘already’ to mark completive

8 In Jenny’s speech, *pasa* seems to be an amalgam of *pasa* ‘goes by’ and *vas* a ‘you.SG go to’. In Spanish, the letter *v* stands for /b/ as there is no labio-dental sound in Castilian. It is common for Chinese native speakers learning an L2 to devoice voiced stops. In this way, Spanish *pasa* and *vas a* came to be realized identically as [pa-sa] by Jenny.

aspect is reminiscent of the tense markers that emerged in some Portuguese-based creoles (see Clements 2009: 140). Her use of the past participle suffix *-do* seems to correlate fairly to some extent with the use of the Chinese perfective marker *-le*. That is, in Jenny's speech two aspect-marking strategies seem to be present: one a reanalysis of an adverb that often occurs preverbally, and the other a mapping between two suffixes that have comparable functions, Chinese *-le* and Spanish *-do*. In her copula use, Jenny's overt copula form *está* is predicted by frequency of occurrence of this form in discourse. Her use of the null copula is accounted for by transfer from Chinese. Her creation of an innovative pronoun system is predicted by detectability of forms in discourse: she selected only forms that carry their own stress. No clitic forms were selected. Finally, Jenny's Spanish is an individual solution of the problem of communication with restricted native-language input. It is also a testament to the creativity she has displayed in constructing a new grammar from Spanish content and function words that has served her in her communicative interactions with friends and colleagues. But it also sheds some light on the various strategies speakers may use to learn subsequent languages naturalistically. 

J. CLANCY CLEMENTS

INDIANA UNIVERSITY – BLOOMINGTON

## References

- Andersen, Roger W. 1993. Four operating principles and input distribution as explanation for underdeveloped and mature morphological systems. *Progression and regression in language*, eds. K. Hyltenstam & A. Viborg. 309–339. Cambridge: Cambridge University Press.
- Andersen, Roger W. & Yasuhiro Shirai 1996. The primacy of aspect in first and second language acquisition: The pidgin-creole connection. *Handbook of second language acquisition*, eds. William C. Ritchie & Tej K. Bhatia. 527–570. San Diego, CA: Academic Press.
- Bakker, Peter 1994. Pidgins. *Pidgins and creoles: an introduction*, eds. Jacques Arends, Pieter Muysken & Norval Smith. 25–39. Amsterdam: Benjamins. <https://doi.org/10.1075/cll.15.07bak>
- Beltrán, Joaquín & Crescen García 2001. The Chinese community. *Multilingualism in Spain*, ed. María Teresa Turell. 282–300. Clevedon UK: Multilingual Matters.
- Blackshire-Belay, Carol 1991. *Language contact: verb morphology in German of foreign workers*. Tübingen: Gunter Narr Verlag.
- Bybee, Joan 1985. *Morphology*. Amsterdam: Benjamins. <https://doi.org/10.1075/tsl.9>
- Clements, J. Clancy 2003a. The tense-aspect system in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition* 25: 245–281. <https://doi.org/10.1017/S0272263103000111>
- Clements, J. Clancy 2003b. On classifying language-contact varieties. *Selected Proceedings of the First Workshop on Spanish Sociolinguistics*, ed. Lotfi Sayahi. 1–10. Somerville MA: Cascadilla Press.
- Clements, J. Clancy 2005. Immigrant speech, creoles, and the ‘basic variety’: a usage-based account of some traits in the Portuguese-based creoles. *Journal of Portuguese Linguistics* 4: 149–165.
- Clements, J. Clancy 2009. *The linguistic legacy of Spanish and Portuguese: Colonial expansion and language change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511576171>
- Clements, J. Clancy 2014. Form selection in contact languages: evidence from some Portuguese- and Spanish-lexified contact varieties. *Portuguese/Spanish Interfaces*, eds. Patrícia Amaral & Ana Maria Carvalho. 377–401. Amsterdam: Benjamins. <https://doi.org/10.1075/ihll.1.20cle>
- Croft, William 2000. *Explaining language change. An evolutionary approach*. London: Longman.
- Davies, Mark 2002–. *Corpus del Español: 100 million words, 1200s–1900s*. Available online at <http://www.corpusdelespanol.org>.
- Jiao, Jian & Alejandro Cuza 2022. Early exposure and input type in the production of Spanish accusative clitics by Chinese-Spanish bilinguals. *Glossa: a journal of general linguistics* 7 (1): 1–30. <https://doi.org/10.16995/glossa.5770>
- Hall, Robert 1966. *Pidgin and creole language*. Ithaca, NY: Cornell University Press.
- Hinnenkamp, Volker 1984. Eye-witnessing pidginization? Structural and sociolinguistic aspects of German and Turkish foreigner talk. *York*

- papers in linguistics: Papers from the York Creole Conference September 24–27, 1983*, eds. Mark Sebba & Loretto Todd. 153–166. York, Canada: University of York.
- Holm, John A. 1988. *Pidgins and creoles: theory and structure* 1. New York: Cambridge University Press.
- Holm, John A. 1989. *Pidgins and creoles: reference survey* 2. New York: Cambridge University Press.
- Klein, Wolfgang & Clive Perdue 1992. *Utterance structure: Developing grammars again*. Amsterdam: Benjamins. <https://doi.org/10.1075/sibil.5>
- Klein, Wolfgang & Clive Perdue 1997. The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13 (4): 301–347. <https://doi.org/10.1191/026765897666879396>
- Li, Charles & Sandra Thompson 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press. <https://doi.org/10.1525/9780520352858>
- Li, Wei (ed.) 2016. *Multilingualism in the Chinese diaspora worldwide: Transnational connections and local social realities*. New York: Routledge. <https://doi.org/10.4324/9781315759371>
- McColl Millar, Robert 2007. *Trask's historical linguistics* (2<sup>nd</sup> edition). London: Hodder Arnold.
- Mühlhäusler, Peter 1986. *Pidgins and creole linguistics*. Oxford: Blackwell.
- Picallo, M. Carme & Gemma Rigau 1999. El posesivo y las relaciones posesivas (ch. 15, vol. 1). *Gramática descriptiva de la lengua española*, eds. Ignacio Bosque & Violeta Demonte. 3 volumes, 973–1023. Madrid: España.
- Schuler, Kathryn Dolores 2017. *The acquisition of productive rules in child and adult language learners*. Unpublished doctoral dissertation. Washington D.C.: Georgetown University.
- Sebba, Mark 1997. *Contact languages: Pidgins and creoles*. New York: St. Martin's Press. <https://doi.org/10.1007/978-1-349-25587-0>
- Thomason, Sarah & Terrence Kaufman 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press. <https://doi.org/10.1525/9780520912793>
- van Coetsem, Frans 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris. <https://doi.org/10.1515/9783110884869>
- Winford, Donald 2003. *An introduction to contact linguistics*. Malden MA: Blackwell.