



# A multidimensional perspective on the acquisition of subject-verb dependencies by Haitian-Creole speaking children

Isabelle Barrière, Blandine Joseph, Katsiaryna Aharodnik, Sarah Kresh,  
Guetjens Prince Fleurio, Géraldine Legendre, Thierry Nazzi

## ► To cite this version:

Isabelle Barrière, Blandine Joseph, Katsiaryna Aharodnik, Sarah Kresh, Guetjens Prince Fleurio, et al.. A multidimensional perspective on the acquisition of subject-verb dependencies by Haitian-Creole speaking children. Journal of Pidgin and Creole Languages, In press, 10.1075/jpcl.22001.bar . hal-04197080

**HAL Id: hal-04197080**

**<https://hal.science/hal-04197080v1>**

Submitted on 5 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*To appear in Journal of Pidgin and Creole Languages, July 2023*

**A Multidimensional Perspective on the Acquisition of Subject-Verb Dependencies by  
Haitian-Creole Speaking Children: Insights from Comprehension and Production**

Isabelle Barrière<sup>a,b,c</sup>, Blandine Joseph<sup>d</sup>, Katsiaryna Aharodnik<sup>c</sup>, Sarah Kresh<sup>c</sup>, Guetjens Prince  
Fleurio<sup>e</sup>, Géraldine Legendre<sup>f</sup> and Thierry Nazzi<sup>g</sup>

<sup>a</sup>Molloy University, <sup>b</sup>YVY Research Institute/ <sup>c</sup>CUNY Graduate Center/ <sup>d</sup>Public School 298,  
Brooklyn, NY/ <sup>e</sup>Ecole Nationale des Arts, Port-au-Prince, Haiti/ <sup>f</sup>Johns Hopkins University/  
<sup>g</sup>Integrative Neuroscience and Cognition Center, CNRS-Université Paris Cité, UMR 8002

**Abstract**

The present multidimensional study investigates the acquisition of pronominal subject-verb dependencies in Standard Haitian Creole (HC). A corpus analysis confirms that HC subject pronouns are phonological clitics in the target grammar and that their reduction is optional and unpredictable. The comprehension and production of dependencies involving these subject pronouns in 20 preschoolers acquiring HC as their first language were investigated. While the production of third person singular and plural subject pronouns *l(i)* and *y(o)* reveals early mastery of adult constraints on their phonological reductions, the systematic assignments of *l(i)* to singular subjects vs. *y(o)* to plural subjects of the verb in the syntactic dependency emerge later, in both production and comprehension. The few syntactic contexts in which HC-learning children show evidence of comprehension involve full forms, rather than phonological reductions. Possible factors that explain these findings include the relative unpredictability of their forms and the linguistic status of HC pronouns.

Haitian Creole, subject verb dependencies, pronouns, clitics, morphophonology, language acquisition, learnability, paradigm cell complexity, video-matching comprehension task

## 1. Introduction

Mastery of syntax includes the acquisition of syntactic dependencies such as subject-verb dependencies. Their form varies cross-linguistically as the subject may be realized as a lexical DP marked for number, a null subject in pro-drop languages, or a pronominal subject whose syntactic status may be equivalent to that of a lexical subject (e.g. in English) or to an agreement marker, e.g., in Colloquial Hexagonal French that we refer to below as CHFrench (Culbertson 2010; Legendre et al. 2010a). Additionally, the association between the number and person features of a subject and/or of the agreement marking on the verb may be expressed with different degrees of consistency across languages: for instance, in both Mainstream American English (MAE) and Mexican Spanish the verb agreement marking for 3<sup>rd</sup> person singular and plural, respectively, is obligatory while in African American English (AAE) and Dominican Spanish both overt and zero-marking are grammatical. These differences in how subject-verb dependencies are expressed across languages or even across different varieties of the same language have led researchers to investigate how these different forms are acquired in different languages and language varieties. The objective is to understand how the trajectory of the acquisition of subject-verb dependencies is modulated by specific properties of the native language. In this context, the present study explores the acquisition of subject-verb dependencies in an understudied language, Standard Port-au-Prince Haitian Creole (henceforth HC), chosen because it has distinctive properties, including optional resyllabification and variability in the expression of the singular and plural subject pronouns subjects. The present study contributes to disentangling some of the acquisition issues left open by previous research summarized in Tables 1 and 2.

[INSERT TABLE 1]

[INSERT TABLE 2]

The present study of HC subject-verb dependencies is best understood in the context of two main findings that have emerged from the cross-linguistic and cross-dialectal acquisition studies

summarized in Tables 1 and 2. Cross-linguistic studies have revealed that there is one exception to the relatively ‘late’ comprehension of subject-verb dependencies, CHFrench, for which comprehension of pronominal subject-verb agreement has been attested in younger children –aged 2 years and 6 months- than in any other language examined. Notably, the CHFrench verbal contrast for which such early comprehension was obtained involves preverbal subject clitics preceding verbs starting with a vowel, that trigger a process of obligatory resyllabification (associated with liaison). In addition, the syntactic status of these subject clitics is demonstrably that of agreement markers rather than argumental pronouns (Culbertson 2010; Legendre et al. 2010a). In addition, most studies on the impact of variation on the acquisition of the marking of subject-verb agreement (Table 1.2) have revealed that the more variability in the input for a given morpho-syntactic feature (found in varieties such as African American English, Southern White English, White Working Class English), the longer it takes to acquire the various expressions of that feature (but see Molina, Marcilese & Name 2017 for different findings in Brazilian Portuguese). Taken together, these cross-linguistic and cross-dialectal studies have suggested that five distinct factors may impact the age of acquisition of Subject-Verb dependencies: (i) the pre-verbal position and adjacency between the number marker and the verb; (ii) the syntactic status of subject pronouns (argument vs agreement marker); (iii) resyllabification; (iv) semantic transparency (cue reliability) and frequency (cue availability); (v) variation in the expression of number marking and complexity of the paradigm cells.

Part two presents a corpus analysis of HC subject pronouns that considers these five dimensions. Part three explores all the factors that have been proposed to account for the cross-linguistic and cross-dialectal differences in acquisition patterns and demonstrates that the linguistic characteristics of HC subject pronouns make it an ideal language in which to test the respective roles of the five acquisition factors. Comprehensive acquisition studies require an analysis of the nature and distribution of the relevant constructions (part 2) combined with the investigation of children’s production and comprehension of these same constructions (Part 3). This approach adopted in various studies (e.g. Miller 2019 and references cited therein) has been labelled ‘multidimensional’ (Barrière et al. 2016). This multidimensional approach is the one adopted in the present study.

## **2. Nature, Expressions and Distribution of Haitian Creole Subject Pronouns**

One of the key issues in the analyses of subject pronouns lies in the distinction between strong subject pronouns and weak subject pronouns or clitics. Strong pronouns are assigned the same argument status as full DPs. In contrast, weak subject pronouns or clitics may exhibit different syntactic properties. They can either be phonologically reduced but still have the syntactic status of DPs, in which case they are phonological clitics. Or they operate as heads of functional categories and they function as agreement markers or inflectional affixes, in which case they are syntactic clitics. A language with preverbal syntactic subject clitics is de facto a null-subject or pro-drop language in which agreement is simply preverbal (see Culbertson 2010 and Legendre et al. 2010a for CHFrench and DeGraff 1993, for HC). In the acquisition literature, developmental patterns partly depend on whether the language is pro-drop (Valian 1991) and differ for full and weak pronouns: in the early stages of acquisition of CHFrench, weak pronouns are systematically tied to finite verbal forms while strong pronouns are not (e.g., Pierce 2012; Jakubowicz & Rigaud 1997; Legendre et al. 2010a).

The age at which subject-verb dependencies expressed as suffixal subject-verb agreement are used productively has also been found to depend on the morphological richness of subject-verb agreement paradigms (e.g., Collins 1995; Xanthos et al. 2011). In languages that distinctly mark different numbers and persons (e.g., Greek, Hebrew, Italian, Spanish), productive use is observed at a younger age than in languages that exhibit fewer subject-verb person and number agreement markers (e.g., English that has a single overt marker- 3<sup>rd</sup> person singular -s- for most verbs) (Table 1).

Most verbs in HC are not overtly marked for subject-verb agreement and do not change form (Valdman 2015). Person and number of the subject are expressed through the use of lexical and pronominal DPs. HC has a set of full (unreduced) subject pronouns (e.g., *mwen- 1<sup>st</sup> Person SG*, *li- 3<sup>rd</sup> person SG*) and so-called phonologically weak/reduced/clitic subject pronouns (e.g. *m- 1<sup>st</sup> person SG*, *l-3<sup>rd</sup> person SG*). Previous analyses of HC weak subject pronouns have given rise to two hypotheses. According to DeGraff (1993), they are best analyzed as syntactic clitics (i.e., agreement markers) which would make HC a pro-drop language. In contrast, Déprez (1994), Cadely (1995, 1997), Baptista (1995) and Lefebvre (1998) have argued that they are best analyzed syntactically as their full pronoun counterparts (i.e., argument DPs) but they may undergo phonological reduction. However, these analyses were proposed on the basis of adult

grammaticality judgments, rather than the ‘casual speech’ (Labov 1966) that constitutes young children’s HC input. The need to establish the status of these clitics in casual speech is reminiscent of the French case, where recent studies on CHFrench, relying on analyses of casual speech typical of what children are exposed to (e.g. Culbertson 2010; Legendre et al. 2010a), have challenged traditional analyses of weak/clitic subject pronouns that previously relied exclusively on adult grammaticality judgments in Standard Hexagonal French (e.g. Kayne 1977).

In contrast to previous linguistic studies of HC subject pronouns, Barrière et al. (2016b) conducted a corpus analysis. The speech samples constituting the corpus included 574 utterances produced by native and fluent speakers of HC in the context of their participation to a radio show and 330 utterances produced by two native speakers of HC addressing their child in their home environments. The regional variety of the speakers (Standard Port-au-Prince variety) was identified by the third and fifth authors of these articles, both native speakers of HC, based on the speakers’ accents, the authors’ familiarities with the radio show guests and a background questionnaire for the child’s parents. All seven adult speakers were raised and educated in Haiti and immigrated to the US after they became adults. They reside in an area of Brooklyn recently officially named ‘Little Haiti’ in which a high proportion of speakers of HC reside and socialize using HC, both in informal (e.g., local stores) and more formal (e.g. churches, cultural events organized by community-based not-for-profits) contexts. In other words, their pattern of acquisition (combined with their literacy skills in HC) and use is not the one that is typically associated with either incomplete acquisition characteristic of Heritage Language learners and/or language attrition (Monéreau-Merry 2017).

Both the radio and the parent corpora were transcribed using the CHAT (MacWhinney, 2000) format and the latest orthographic conventions of HC, with one exception. The transcription uses apostrophes to indicate which host words phonologically reduced words attach to (which the latest conventions have ruled against; see section 3.1 of Akademi Kreyòl Ayisyen, 2017 p5). It is useful for our purposes to retain apostrophes to indicate instances of encliticization and procliticization in the spoken language. The full paradigm of HC subject pronouns is presented in Table 3 below.

[INSERT TABLE 3]

Several important features characterize HC pronouns. Subject, direct object and indirect object pronouns share the same forms and enter in the formation of reflexive pronouns and possessive determiners (see Table 3 for the full paradigm). While subject pronouns occur in preverbal positions, object pronouns occur in post-verbal positions (see examples (1) and (2), with the indirect object preceding the direct object pronoun (see example (2)).

(1) **Yo** wè **yo**  
**SPn PL3 V DOPn PL3**  
 ‘They see them.’

(2) **Ou** te di **m’** **li**  
**SPn-SG2 TMA V IOPn-SG1 DOPn-SG3**  
 ‘You said it to-me.’

As Table 3 shows, HC subject pronouns can undergo phonological reduction. These reductions have raised debates regarding the syntactic status of HC pronouns, i.e., whether they should be treated as syntactic clitics (agreement markers) or phonological clitics (phonologically reduced pronouns). Barrière et al. (2016b) provide arguments illustrated with relevant examples extracted from the corpora for an analyses of HC pronouns as syntactic strong pronouns – sharing the same syntactic status as argument DPs- rather than weak or clitic pronouns (functioning as agreement markers) in that their behavior is the same as that of argument DP, in several respects. First, HC subject pronouns can be conjoined with other DPs, as in example (3). Secondly, they can be separated from their verbal hosts with adverbs, as in example (4). Finally, HC subject pronouns display an additional property relevant to their syntactic analysis as argument DPs (rather than agreement markers): when both a subject pronoun and a lexical subject DP referring to the same entity co-occur, there is an obligatory prosodic pause between the two, as in example (5), which results in an instance of dislocation rather than clitic-doubling (co-occurrence within a single prosodic phrase, as frequently found in CHFrench; Culbertson 2010). (5) is the only instance of subject DP dislocation identified in the corpora.

(3) **Nou** avèk Jean-Beliard-Lucien bò tab la  
**SPn PL2 CONJ N PREP N Def-Det SG**

‘We and Jean-Beliard-Lucien (are) at the table.’

(4) **Nou** toujou jwe

SPn PL2 ADV V

‘We always play.’

(5) e nati a, li genyen tout sa 'l bezwen

CONJ N Def-Det-SG, SPn SG3 V QUANT DemPN Spn-SG3 V

yo pou 'l fonksyone

DOPn-PL3 CONJ SPn-SG3 V

‘And nature, **it** has everything **it** needs them so that **it** functions’

‘And nature has everything it needs to function.’

Example (5) involves the use of two reduced 3<sup>rd</sup> person singular enclitics: the second and the fourth pronoun of the sentence. Phonologically, the two pronouns combine with elements that precede them, although they are not in a relation of subject-verb syntactic dependency with these. In the case of *sa'l*, the pronoun /li/ is reduced to the consonant /-l/ and combines with the object (expressed by the demonstrative pronoun *sa* ‘that’) of the verb of the main clause whereas at the syntactic level, it constitutes the subject of the verb *bezwen* ‘need’, which follows it. In the same way, the last pronoun /li/ combines with the conjunction *pou'l* ‘for that’ while syntactically /l/ is the subject of the verb *fonksyone* ‘operate’, which follows it (see further discussion and additional examples in Barrière et al. 2016 b).

For the purpose of the present study on possible effects of contractions on acquisition, a quantitative analysis of the proportions of full versus reduced subject pronouns was also conducted for each of the two corpora. In Port-au-Prince/Standard HC, reductions involve the exclusive use of the onset consonants and the deletion of the vowels. The overall results are presented in Table 4. Across corpora, different proportions of reduced forms characterize pronouns. For instance, in both the Radio and Parents’ corpora, most occurrences of the first person singular pronouns are reduced (80% of the time by the contributors to the radio program and 100% of the time by both parents). Valdman (2015) proposes that reduced *mwen* consists of syllabic /m/, which would explain the fact



that *mwen* optionally reduces before words that start with both consonants and vowels, providing a broader range of contexts for phonological reductions, leading to the observed high frequency of reductions of *mwen*. Note however that neither the analysis provided by Valdman (2015), nor the present analysis (due to the limited quality of our audio-recordings) provides perceptual or acoustic evidence for this statement. In contrast, *nou* ‘you-PL/we’ is less frequently reduced- fewer than a third of the time it is used. Moreover, third person plural *yo* ‘they’ is produced in its full form most of the time in the Radio Corpus. However, general statements regarding percentages of phonological reductions can only be made with caution, given that some pronouns rarely occurred in the corpora (for example, *n(ou)* ‘you-PL/we’ is produced 113 times in the radio corpus whereas it is only produced three times by the mother and never by the father in the parents’ corpus; see Table 4 for details).

[INSERT TABLE 4]

Despite these limitations, occurrences in the corpus reflect most of the phonological constraints described by Cadely (1993), with one exception: the phonological reduction of *li* occurs in contexts when it either precedes a word starting with a vowel- giving rise to a proclitic (as in example (6) below)- or when it follows a word ending in a vowel as in example (5) above – giving rise to an enclitic. Although Cadely (1993) mentions that procliticization can occur in the context of a verb that starts with a consonant, such reductions are not judged grammatical by the 2<sup>nd</sup> and 5<sup>th</sup> authors, both native speakers of Port-au-Prince HC, and further analyses of the corpora below explore this issue.

(6)	men	<u>l'</u>	ap	di'	m
	CONJ	SPn-SG3	TMA	V	IOPn-SG1
	‘But	he	will	say	to-me.’

In contrast to *li*, reduction of *yo* only occurs when it precedes a word that starts with a vowel giving rise to a proclitic, as in example (7) – a preceding word ending with a vowel by itself cannot give

rise to a reduction of *yo*, hence the grammaticality of (8.a) and the ungrammaticality of (8.b) (See Valdman, 2015, for a different distribution reflecting different constraints in Cap-Haitian Creole).

- (7) y' \_\_\_\_\_ ap pale de premie yo  
 SPn-PL3 TMA V Prep QUANT DefDetPL

‘They are speaking of/talking about the first ones.’

- (8) a. sa vle di yo te deja la nan premye pati lavi  
 DemPn TMA V SPn-PL3 TMA ADV Prep QUANT N N  
 This want(s) to-say they PAST (were) already in the first phases/cycle life  
 ‘This is to say/ means that they were in the first phases of.’

- (8) b. \* sa vle di'y te deja la deja nan premye pati lavi  
 ‘want(s) to-say they PAST (were) already in the first phases/cycle life.’

Finally, given our specific interest in 3rd person singular and plural subject pronouns for the acquisition study of subject-verb dependencies, an analysis of the context in which the full and reduced forms of *li* and *yo* occur was conducted. Table 5 presents the various phonological contexts in which *li* is expressed through a full form /li/, a proclitic /l’-/, or an enclitic /-‘l/. Given the existence of contexts that may foster either enclitics or proclitics, for which determining the proclitic or enclitic status of the reduction was difficult, such cases were placed in an additional category. Table 5 presents the various phonological contexts in which *yo* is expressed through a full form /jo/ or a proclitic /j-/ in the Radio Corpus; Subject Pronoun *yo* was not produced in the Parents’ Corpus.

[INSERT TABLE 5]

[INSERT TABLE 6]

Although we have to remain cautious given the small size of our corpus, several findings emerged from the analyses. The first one is that even in these samples of casual speech, neither *li* nor *yo* is systematically reduced in phonological contexts that allow reductions. For *li*, the results confirm that reduction only occurs when *li* either follows a word that ends with a vowel (enclitic) or when it precedes a word that starts with a vowel (proclitic): shaded areas in the table indicate contexts in which the reduction is not predicted according to the grammaticality judgments of the 2<sup>nd</sup> and 5<sup>th</sup> authors (both native speakers of HC). Moreover, *li* seems more often reduced in the context of enclitics (62% in the Radio Corpus and 100% in the Parent Corpus) than proclitics (50% in both the Radio and Parent Corpus). Finally, when two factors -the preceding and following words ending and starting with vowels, respectively- can lead to reduction (e.g. see column 4 of Table 5), *li* is more likely to be reduced than in a context when only one factor (either the preceding or the following word) allows for reduction.

For *yo*, as predicted by Cadely (1993, 1997), phonological reductions that give rise to phonological proclitics only occur when the plural subject pronoun precedes a vowel-initial word (e.g. a Tense-Mood-Aspect (TMA) marker, a verb). Thus, not only is *yo* less frequent than *li* in the corpus, it is much more often realized as a full pronoun than as a proclitic. And like *li*, in environments where it can be reduced, it gives rise to a proclitic only half of the time.

Further corpus analyses were conducted to assess the degree of semantic transparency for *li* and *yo*, a factor that has been proposed to impact acquisition (see fourth factor presented in section 2). With respect to semantic transparency, 3<sup>rd</sup> person pronouns exhibit complex characteristics in HC. Their marking of person and number across different syntactic functions (i.e. subject, direct and indirect object) remain the same (see Table 3). However, HC word order is very rigid so it is unlikely that syncretism across different functions impedes their interpretation as subject or object. We examine full pronouns first. While post-nominal possessive determiners in HC indicate the number of the possessor (i.e. *li* refers to ‘her’, ‘his’ or ‘its’), not the number of the possessed entity (e.g., his/her child/children), the post nominal determiner *yo* (following *li*) is used to convey the latter’s plurality; thus, *timoun li* refers to ‘her/his/its child’ while *timoun li yo* refers to ‘her/his/its children’.

As for *yo*, it does not systematically denote plurality of the possessed entity: while it is both the 3<sup>rd</sup> person plural subject pronoun and an optional plural definite determiner, it is also the 3<sup>rd</sup> person plural possessive determiner, in which case plurality pertains to the possessor, resulting in possible ambiguity. Thus, *timoun yo* is ambiguous between referring to ‘the children’ or ‘their child’ or ‘their children’ and can optionally be disambiguated with singular (not plural) determiners. In this context it seems that both preverbal full pronouns *li* and *yo* are relatively transparent – though less so than CHFrensh /z-/ (Barrière et al. et al., 2016) but more so than the English suffix -s.

Reduced forms exhibit a very different pattern. Many word types have /l-/ as an onset, including nouns (numerous HC nouns borrowed from CHFrensh have incorporated the CHFrensh definite articles *l’*, *le* or *la* as their onset and therefore start with /l-/) and verbs, as well as other lexical categories and function words: Vilsaint (1991) bilingual HC-English dictionary lists about 400 words. In contrast, /j-/ occurs as an onset in only a handful of words: according to Vilsaint (1991) fourteen words in addition to the 3<sup>rd</sup> person plural subject and object pronoun *yo*, the possessive determiner *yo* and the reflexive pronoun *yo-menm* ‘themselves’. These include the adverb *yè* ‘yesterday’, two verbs, five nouns, the pre-nominal quantifier *yon* /jõ/ ‘one’, and the cardinal number *yonn/youn* /jun/ that also refers to ‘one’ in the context of counting.

With respect to token frequencies (see Table 7), the Radio Corpus contains a total of 221 words that start with /l-/. Thirty-one percent of these words refer to 3<sup>rd</sup> person singular and the full form /li/ (16%) exhibits higher semantic transparency than the reduced form /l-/ (5%).

[INSERT TABLE7]

A total of 75 words ending in /-l/ were identified in the Radio Corpus. For 54% percent of these words, the coda /-l/ is part of their root. For 32%, /-l/ constitutes reductions of 3<sup>rd</sup> person singular object pronoun and for 14% the reduced form of the 3<sup>rd</sup> person subject pronouns. In other words, 46% of words that end in /-l/ indicate 3<sup>rd</sup> person singular but only a third of these constitute subjects. The pattern for *yo*//*j*o/ and its reduced form /j-/ is presented in the third row of Table 7. Fifty-nine percent of instances of *yo* indicate plural. Words starting with /j-/ are six times more likely to refer

to plural (60%) than singular (less than 10%). Based on these figures it seems that both *yo* and /j-/ are twice more likely to refer to plural than *li* and *l-* to singular and that /-l/ is more likely to refer to singular out of all words that end in /-l/. Words with /-l/ as a coda are also more likely to refer to object pronouns and they are not frequent. Thus, /-l/ is a cue with relatively high semantic transparency but low availability, in contrast to /l-/ that has lower semantic transparency but is more frequent. The qualitative and quantitative analyses of the syntactic distribution and phonological realizations of HC pronouns facilitate the examination of specific factors with respect to both the age at which they are likely to be acquired, the order of acquisition of singular subject *li* and *yo*, and the order of acquisition of full and reduced forms. Each prediction is outlined below.

### **3. Morphosyntactic, morphophonological and distributional acquisition factors**

In this section, we review the different factors on the basis of which different hypotheses have been proposed to account for the cross-linguistic and cross-dialectal findings summarized in Tables 1.1 and 1.2 in the introduction. For each factor, we argue that HC subject pronouns exhibit unique properties that will enable us to test them in a novel way.

#### **First Factor: Pre-verbal Position and Adjacency between the Number Marker and the Verb**

##### *Cross-linguistic perspectives*

In five of the eight languages mentioned in Table 1, number marking is expressed through a verbal suffix. In none of these languages, children understand subject singular and plural number marking before they reach age 3. The dependency on which evidence of comprehension of 3<sup>rd</sup> person singular and plural marking in CHFrench is found in children aged 2;6 (Legendre et al. 2014) involves a pre-verbal clitic pronoun independently analyzed as an agreement marker resulting from syntactic change (Culbertson 2010; Legendre et al. 2010a). The preverbal position of the marker might thus facilitate its earlier comprehension, especially given that word onsets have been found to be more clearly perceived than codas (Swingley 2008; von Holzen Nishibayashi & Nazzi, 2018; but see Nazzi & Bertoncini, 2009; Nishibayashi & Nazzi, 2016). However, this positional factor alone cannot account for the fact that the preverbal position of number agreement markers in Xhosa (a Bantu language, see examples of stimuli in Table 1) does not facilitate comprehension: only the plural marker is understood at age 4. One difference between the subject-verb agreement systems

of CHFrench and Xhosa pertains to the distance between the preverbal subject agreement marker and the verb. Many bound morphemes (marking Tense, Aspect, Object Agreement, etc.) typically occur between them in Xhosa, while in CHFrench only a small set of functional categories, including object clitics and the (optional) negative clitic *ne* may occur between the preverbal subject clitic pronoun qua agreement marker and the verb. If there is a facilitating effect of preverbal position at all it may be modulated by the distance between the dependent/agreeing elements (see Santelmann & Jusczyk 1998, for inhibiting effects of number of syllables between the auxiliaries and the verb forms in English-learning toddlers' preferences) or the nature of intervening elements (i.e. type of lexical or functional categories, Höhle et al. 2006).

#### *Possible implications for the acquisition of HC*

If the preverbal position of the expression of the subject facilitates comprehension of the subject-verb dependency, the within-language prediction is that all expressions of HC subject pronouns, full and reduced (phonological clitics), will be comprehended at the same age. And the cross-linguistic prediction is that they should all be comprehended at around age 2;6- the youngest age at which CHFrench-speaking children demonstrate evidence of subject-verb comprehension expressed through preverbal liaison (Barrière et al 2016). If the distance between number marking and the verb moderates the positive contribution of this factor (as the findings from Xhosa suggest), HC-learning children may not understand 3<sup>rd</sup> person singular and plural subject-verb dependencies as early as CHFrench-learning children. As mentioned in the corpus analyses, many elements can appear between the subject clitics and the verb in HC, including conjoined DPs, Adverbs, Negative Markers and Tense-Mood-Aspect markers, and the phonological distance that applies to the enclitic /-l/ sometimes involves being part of a prosodic unit outside the relevant subject-verb dependency.

### **Second Factor: Syntactic vs. phonological status and distribution of subject pronouns**

#### *Cross-linguistic perspectives*

In languages that have both full subject pronouns and clitic subject pronouns (commonly referred to as weak pronouns), clitics tend to be systematically associated with finiteness earlier in children's production than DP subjects, including full pronouns (e.g., Pierce 1992; Jacobowicz & Rigaud 1997; Culbertson 2010; Legendre et al. 2010a). Among the languages listed in Table 1, CHFrench is the only language in which subject clitics have the syntactic status of agreement (in terms of tree

structure they are functional heads rather than DP specifiers of a TP; see Culbertson, 2010, for details). One of the possible explanations for the earlier comprehension of subject-verb dependencies in CHFrench compared to other languages investigated so far (e.g., Czech, English, Farsi, German, Greek, Spanish) therefore lies in the syntactic clitic (or agreement) status of the weak subject pronouns that were tested. The few studies that have investigated the acquisition of the mapping between pronouns and number features have demonstrated that in languages such as English in which full pronouns indicate Case, Person and Number, Number is the last feature acquired. In a picture matching task testing the comprehension of both subject (*he, she, they*) and object (*him, her, them*) pronouns (Scholes 1981), number comprehension was not found at age three, but at age four, despite the fact that the stimuli included two number cues- the subject pronoun and the auxiliary *is*-SG or *are*-PL; and up until age seven, children made significantly more errors on Number than on Case or Gender.

#### *Possible implications for the acquisition of HC*

Our corpus analyses demonstrate that syntactically HC subject pronouns behave like full DP arguments. For instance, they can be conjoined with DPs and separated from their verbal hosts by adverbs. And their co-occurrence with a DP constitutes instances of dislocations rather than clitic doubling. However, all HC subject pronouns can undergo phonological reduction.

The argumental analysis of HC pronouns in turn leads to the prediction that they should play the same role in subject-verb dependencies as lexical DPs, with no facilitatory effects expected. Compared to null-subject Spanish in which number distinctions were exclusively expressed as a verbal suffix in the stimuli used by Pérez-Leroux (2005) and English in which number on the subject was neutralized by using *s*-initial verb stimuli by Johnson et al. (2005 see Table 1 and 2), HC can also express the subject-verb dependency with a single distinction, albeit on the pronoun in conjunction with an invariable verb. Such a single marker of the dependency is not necessarily expected to be comprehended earlier than those manifested as verbal number markers in other languages, i.e., not earlier than 3;6- 4 years of age- the earliest ages obtained for English (Barrière et al. 2019) and Spanish (Hsin et al. 2021) agreement markers. In English, the assignment of number to subject pronouns does not occur before age four, including in the context of stimuli that contain two number cues – the pronoun (*he/she* versus *they*) and the auxiliary BE overtly marked for number (*is/are*) (Scholes 1981).

The phonological realizations of HC subject pronouns as full or reduced pronouns should neither facilitate nor hinder the acquisition process since all forms enjoy the same syntactic status. So no difference should be found between the ages at which HC learners understand full and reduced forms of 3<sup>rd</sup> person singular and plural subject pronouns.

### **Third Factor: Resyllabification**

#### *Cross-linguistic perspectives*

An alternative explanation for the early comprehension of CHFrench subject clitic pronouns used with verbs starting with a vowel lies in the resyllabification associated with liaison that disrupts lexical boundaries (Legendre et al. 2014) and (counter-intuitively) would make children develop an early awareness of the variation between singular and plural number that consists of a single consonant, e.g., the distinction between /i.la.riv/ (SG) versus /i(l).za.riv/ (PL), especially in combination with high cue reliability (fourth factor, presented below). In CHFrench, such liaison is obligatory in the context of verbs starting with vowels. Although phonological reductions of subject pronouns are possible in HC with verbs that start with vowels (e.g., Cadely 1993) and such contraction gives rise to resyllabification, the distribution of these reductions in Port-au-Prince HC has not been empirically investigated. An important issue is whether resyllabification itself facilitates acquisition in CHFrench and/or whether facilitation is tied to its systematic occurrence in specific phonological contexts. Study 1 provided an analysis the distribution of phonological reductions of HC subject pronouns and addresses the issue of the systematicity- or lack – of resyllabification in HC and Study 2 explores the effect of this feature on acquisition.

#### *Possible implications for the acquisition of HC*

The quantitative analyses of the distribution of full versus reduced pronominal forms (Table 3) cast light on an important distributional feature of HC subject pronouns. Their phonological reduction is optional: for instance, both *li* and *yo* can either be realized as full forms or reduced when preceding a vowel-initial word. Like in CHFrench, HC subject pronouns enjoy a preverbal position. Contra Saint-Martin (2005), they are otherwise quite different by virtue of their distinct syntactic status discussed above, plus the fact that they do not involve an obligatory *liaison* process that systematically triggers resyllabification in the presence of vowel-initial verbs as in CHFrench, which affects the course of acquisition of subject-verb agreement (Legendre et al. 2014).



In HC, resyllabification is tied to elision and is only optional because phonological reduction is optional. Resyllabification occurs in at least two different ways. Like in CHFrensh, it can give rise to phonological proclitics when *li*-SG or *yo*-PL are reduced to /l-/ and /j-/, respectively, in which case they attach to either the verb or other elements within their prosodic phrase, including Tense-Mood-Aspect (TMA) markers. If resyllabification within the prosodic phrase that includes the verb with which they entertain a syntactic dependency facilitates the acquisition of HC subject pronouns, it is predicted that a) the phonological proclitics will be understood earlier than full pronouns and enclitics; and b) they may be understood around age 2;6, the earliest age at which CHFrensh-learning children understand 3<sup>rd</sup> person singular and plural subject clitics.

If, on the other hand, the positive effect of resyllabification is insensitive to the relation between the reduced form and the prosodic constituent/unit in which it occurs, enclitics – that apply to 3<sup>rd</sup> person singular /li/ reduced to /-l/ when following a vowel-final word – should be understood at the same age as proclitics and before full pronouns.

Finally, if it is the obligatoriness of the CHFrensh liaison that facilitates the learning of the subject-verb dependency in CHFrensh at age 2;6, the optional resyllabification process that HC subject pronouns undergo may not facilitate acquisition, in which case we would not expect HC pronouns to be acquired before subject pronouns in other languages in which such pronouns do not undergo reduction (e.g. English), that is from age 4

#### **Fourth factor: Semantic transparency and frequency**

##### *Cross-linguistic perspectives*

Another factor that is thought to contribute to the earliest age of comprehension of CHFrensh and to a certain extent Greek versus English and Spanish agreement markers pertains to the semantic transparency or cue reliability of the agreement cue (e.g. Legendre et al. 2014; Barrière et al. 2016 for CHFrensh; Gonzalez-Gomez et al. 2017 for Mexican Spanish; Kenanidis et al. 2021, for Greek). In CHFrensh, the phoneme /z-/ is very rare in onset position of lexical words; it is typically the product of resyllabification, marking plural and most frequently 3<sup>rd</sup> person subject plural in verbal contexts (between 75% and 95% of the time; Barrière et al. 2016). Its semantic transparency is thus high despite its relatively low frequency or cue availability (Barrière et al. 2016). In Greek, while the singular agreement marker is relatively frequent, it exhibits low semantic transparency

since in addition to 3<sup>rd</sup> person singular verbal suffix ‘-i’ is also a nominal marker associated with both singular and plural numbers for different nouns with different genders and cases (Kenanidis et al. 2021). While the Greek 3<sup>rd</sup> person singular plural markers are much less frequent, they all exhibit high cue reliability (Kenanidis et al. 2021). In contrast to both CHFrench and Greek, the consonant /-n/ in Spanish is not infrequently found at the end of singular and plural nouns (26% of words ending in -n) in addition to marking plural in verbs (66% of words ending in -n) (Gonzalez-Gomez et al., 2017). In English, suffixal -s marking 3<sup>rd</sup> person singular on verbs is much less frequent than the possessive marker -s, which in turn is less frequent than nominal plural -s, that occurs about 100% more than 3<sup>rd</sup> person singular -s in the input in the Brown (1973) corpus (Larsen-Freeman, 1976). Section 3 explores the semantic transparency and cue availability of HC subject pronouns and Section four investigates the effects of these features on the acquisition of 3<sup>rd</sup> person singular and plural subject pronouns that differ in these respects.

#### *Possible implications for the acquisition of HC*

The quantitative analyses of semantic transparency and frequency of *li* and *yo* have revealed an interesting contrast: while *l(i)* exhibits higher frequency (cue availability) than *y(o)*, it is less systematically associated with singular than *y(o)* with plural, hence it has lower cue semantic transparency (cue reliability).

If semantic transparency alone contributes to early comprehension, it is expected that 3<sup>rd</sup> person singular *yo* and its reduced form /j-/ are mapped onto plural around the same age as enclitic /-l/ is mapped onto singular, and earlier than full pronoun *li* and proclitic /l-/ are mapped onto the singular. If semantic transparency plays a role in combination with frequency, then /jo/ and /j-/ are expected to be associated with plural earlier than all the expressions of *li* will be interpreted as a singular.

From a cross-linguistic perspective, the distinct expressions of the 3<sup>rd</sup> person plural in HC are less reliable than the marker of the same number and person involved in liaison in CHFrench but more so than the Spanish 3<sup>rd</sup> person plural verbal marker and much more so than the English 3<sup>rd</sup> person singular verbal marker. So, if semantic transparency is a facilitating factor in acquisition, HC subject pronouns are expected to be acquired later than CHFrench subject clitics (understood at age 2;6) but earlier than the Spanish and English agreement markers that are respectively understood at age 3 and 4.

## **Fifth factor: Variation in the expression of number marking and complexity of the paradigm cells**

### *Cross-linguistic and cross-dialectal perspectives*

The studies that have explored the impact of input variation in acquisition (see Table 2) have led to two different conceptualizations of the issue. Following Labov (1972) on adult representations in AAE, Newkirk-Turner and Green (2016) have proposed that overt regular 3<sup>rd</sup> person singular verbal -s is not part of the grammar of 3- to 6-year-old speakers of AAE, and that occasional occurrences of 3<sup>rd</sup> person singular -s in their speech production does not reflect their ‘core grammar’. In contrast, studies on the acquisition of variable nominal and verbal number markers in different varieties of English and Spanish have led to the conclusion that acquisition takes longer in the presence of more variable forms – e.g., a null and an overt marker mapping onto the same number (Miller 2012, 2019; Miller & Schmitt 2010, 2013). This latter hypothesis helps explain why children aged 3 to 5;5 years exposed to different varieties of English perform differently in both a comprehension and a production task (Barrière et al. 2019). While preschoolers acquiring 3<sup>rd</sup> person in the variety in which there is no variation master it when it occurs in a range of syntactic constructions, their age peers acquiring dialects in which it varies understand it in fewer syntactic contexts, which suggests that consistency/variation of 3<sup>rd</sup> person forms impacts the speed and context-(in)dependence of its acquisition and processing. Similar results have been obtained for both nominal and verbal plural markers in different varieties of Spanish. Nominal plural markers are acquired later in Chilean Spanish in which they are optional than in Mexican Spanish, in which they are obligatory (Miller 2019, and references cited therein). Verbal plural markers are understood earlier and in more syntactic contexts in Mexican Spanish in which they are obligatory and always expressed with the same phonological forms (Gonzalez-Gomez et al. 2017; Hsin et al. 2021) than in Dominican Spanish in which they are optional and the overt forms exhibit different phonological forms (Poline, Aharodnik & Barrière 2019).

Variation results in different degrees of morphological complexity and theories of morphological complexity themselves provide further insights that have implications for acquisition. Many early analyses of the outcomes of language contact including the characterization of ‘creole’ languages have paid much attention to the reduction of the expressions of morphological features that have led to increased case syncretism and regularization of morphological paradigms

(e.g., McWhorter 1998, 2005). From this ‘traditional’ morphological perspective, creole languages have been characterized as lacking morphological complexity and richness (McWhorter 1998, 2005), which in turn would increase their degree of learnability (McWhorter 2005). However, such assessment of morphological complexity based on syncretism (‘complexity of exponence’, Anderson (2015) or ‘enumerative complexity’, Ackerman & Malouf (2013)) is a unidimensional one that considers relations between cells but fails to consider the complexity within single cells, i.e., the number of forms they contain and their degree of predictability (Ackerman, Blevins & Malouf 2009; Ackerman & Malouf 2013; Bonami & Luis 2013). This more recent conceptualization of complexity has been used to argue that language contact has, for example, led to an increase of complexity in nominative case marking across two generations of speakers of Gurindji Kriol (Meakins & Wilmoth 2019).

Extending this approach to language acquisition, the co-existence of the null form  $V-\emptyset$  and the  $V-s$  in children who speak varieties of English that differ from MAE would result in a more complex cell than that which MAE learners have to acquire (limited to  $V-s$ ). The presence of two possible expressions in non-MAE would likely hinder their comprehension and production, especially if the expression of a given morpheme is not constrained by linguistic features such as the aspect of the verb (Newkirk-Turner & Green 2016). Similarly, overt, varying and null forms mapping onto the same nominal and verbal number in Dominican and Chilean Spanish take longer to acquire than the consistently overt number markers in Mexican Spanish (e.g. Miller & Schmitt 2010, 2012; Miller 2013; Poline et al. 2019). In contrast, although CHFrench singular and plural subject clitics are realized differently when they precede vowel-initial verbs, auxiliaries or object clitics, the expression of the cell in the CHFrench paradigm is systematically constrained by the phonological environment. The process of liaison is obligatory for subject clitics preceding vowel-initial auxiliaries, verbs, or object clitics, resulting in variation across forms that is predictable.

#### *Possible implications for the acquisition of HC*

Both *li* and *yo* vary with respect to the degree of complexity of their paradigm cells. In HC, singular subject pronoun *li* is more complex than plural *yo* in that it gives rise to three different forms (full form, proclitic and enclitic) which, in addition, are constrained by three aspects of the phonological context. These include presence or absence of vowels in the coda of the preceding word and in the onset of the following word. In contrast, *yo* gives rise to two forms (full form and

proclitic) that depend on whether the following word starts with a vowel. If degree of complexity of the paradigm cells has an impact on the course of acquisition, 3<sup>rd</sup> person singular subject pronoun *l(i)* should be acquired later than plural *y(o)*.

From a cross-linguistic perspective, the variability of *li* and *yo* differs from that of the non-MAE and Spanish number markers mentioned above. The latter involve the selection between two forms – overt or null – for the 3<sup>rd</sup> person singular marker in non-MAE and for nominal plural in Dominican and Chilean Spanish. In contrast, the HC pronominal variation in *li* and *yo* involves the selection of different phonological forms that are constrained by the phonological environments in which they occur, but the forms are not entirely predictable. Three dimensions are typically held to contribute to a paradigm's degree of complexity: the number of forms in the cell (two for non-MAE, and Chilean Spanish nominal number markers, and for *yo* in HC); the number of factors in the linguistic environment that must be taken into account to select the form (at least three aspects of the phonological environment for *l(i)* and two for *y(o)*); and the predictability of the form selection – relatively low for both *li* and *yo* given that in contexts in which they can be reduced they do so between 50% and 62% of the time.

The acquisition study presented below was designed to experimentally test these within-language, cross-dialectal and cross-linguistic factors.

#### **4. Experimental Study: Acquisition of HC 3<sup>rd</sup> Person Singular and Plural Subject-Verb Dependencies**

##### **4.1 Introduction**

A paradox characterizes the linguistic literature on creole languages. On the one hand, linguistic analyses that contrast them with non-creole natural languages (e.g., Bickerton 1984; McWhorter 2005) emphasize their higher degree of learnability than other natural languages. At the same time, very few empirical studies have documented the acquisition of creole languages. Those that do fall into three categories.

The first category seeks to provide evidence in favor of specific hypotheses with respect to distinctive characteristics of creoles, such as the study of Adone (1994) on the acquisition of Mauritius Creole, which tests Bickerton's Aspect First hypothesis. The second category reflects a

negative view of the bilingual learning experiences of creole speakers as a deficient learner, such as the study by Noël (2014) on the acquisition of Reunion Creole and CHFrensh. These two types of study differ in that in the first, the focus is on the unfolding of individual abilities and the contribution of specific linguistic features to the acquisition process. In contrast, in the second the emphasis is on the unequal sociolinguistic status of the two languages and its contribution to the developing linguistic abilities of children. However, both share the assumption that creoles constitute an exceptional class of natural languages categorized by a cluster of distinctive linguistic features or tied to a specific socio-cultural context that determine their learner's (mis)acquisition path. Like DeGraff (2005a) and Fon Sing (2017), we reject the assumption that creole languages, including HC, constitute a separate category of natural languages. Like Henri, Stump & Tribout (2020), we assume that the degree of complexity of creoles cannot be merely related to the apparent complex features of their lexifiers but in examining their own characteristics (e.g., appealing to concepts such as predictability and consistency explored in Sections 2 and 3 above), and that the degree of complexity is likely to impact both acquisition and processing. Also, from a developmental psycholinguistic perspective, it seems circular to determine a-priori that the features that characterize such languages facilitate acquisition given that the nature of facilitating factors has not yet been established and their discovery constitutes one of the main objectives of the developmental psycholinguistic enterprise (e.g., Lust 2012; Erickson & Thiessen 2015).

A third category of studies has emerged more recently: they investigate whether acquisition principles that apply to non-creole languages also characterize the acquisition of creole languages. This approach is the one adopted by Pratas & Hyams (2009) on (Portuguese-based) Capeverdean Creole and De Lissier, Durrleman, Rizzi & Shlonsky (2016) on (English-based) Jamaican Creole: both investigated whether young learners of these languages go through a Root Infinitive stage- an early stage during which children optionally produce non-adult-like simple sentences lacking tense and/or agreement. Although they differ with respect to the languages under investigation and their methodological approaches (relying exclusively on production data for the Jamaican Creole study while involving a combination of production and comprehension data for the Capeverdean Creole study), they both treat creole languages as unexceptional languages in relation to other natural languages.

We adopt this third approach in our investigation of the acquisition of subject-verb dependencies in HC, a language that has been the focus of very few acquisition studies, despite the

fact that it is the creole language with the highest number of speakers (Barrière & Monéreau-Merry 2012). It is the native language of all Haitians raised in Haiti and it is spoken by a substantial number of speakers in the US: it is the second Language-Other-than-English in Miami (Buchanan & Beaulieu 2010) and the fourth one in New York (Barrière & Monéreau-Merry 2012). Studies on the acquisition of HC have started exploring phonological development in preschoolers (Archer, Champion, Tyrone & Walters 2018), the acquisition of discourse skills (Champion, McCabe & Colinet 2002/2003), the effects of educational context in which HC-speaking school-aged children are raised (Dejean 2010; Roumain 2013; Jean-Pierre 2016; Bentolila 2019) and the acquisition of reading skills (DeGraff 2017), but to the best of our knowledge none of them has documented the development of morphosyntactic abilities in young children acquiring HC. The present study aims at starting to address this gap by experimentally investigating the production and comprehension of 3<sup>rd</sup> person subject pronouns by toddlers and preschoolers acquiring HC as their first language.

## 4.2 Method

### *Participants*

Twenty monolingual HC native speakers (18 girls), between the ages of 2;5 and 5;10 and enrolled at the Fansbada Académie preschool located in Port-au-Prince, participated in the study. The sex imbalance reflects the preschool population during the recruitment period. Similarly, the age range was relatively wide and included two children before the ages of 2;6; all the other children were between 3;5 and 5;9, one to three months apart from one another. This age range 2;6-5;9 was selected based on the findings that emerge from acquisition studies on different languages (Tables 1.1), however we recognize that it would have been optimal to also recruit children between 2;6 and 3;5 in order to obtain a more comprehensive developmental picture. The fact that the older participants were close to six years is not surprising: this is the age at which primary school starts in Haiti. In contrast to higher levels of education, preschool is not regulated by the Ministry of Education in Haiti, so not all parents send their children to preschool. Many preschools consist of relatively small structures with classes including many different ages and the Fonsbada Académie is no exception.

Although the participants were not administered formal assessments, they were considered typically developing by their educators and families, based on a background information questionnaire. All children used HC at home and in the preschool environment. Although French

words are introduced in the preschool, based on the classroom observation of the second author who also interacted with children before the data collection, at the time of testing none of the children spoke CHFrench, except for a single child who produced two words in CHFrench in the elicitation task. One child produced a few Spanish words. The highest education level of the parents was completion of high school. All data collection sessions were held in a room assigned to the second author for the purpose of this project in the preschool. Children were tested individually in either one or two sessions and all sessions were audio-recorded on an Olympus DM-720 digital recorder.

### *Experimental Procedures*

The data collected for this project are part of a broader cross-linguistic study on the acquisition of subject-verb dependencies by monolingual and bilingual children. The protocol included a video-elicitation task and a video-matching comprehension task. The order of tasks was counter-balanced across the participants.

#### *Task 1: Video Elicitation*

##### *Visual Stimuli*

Sixteen sets of videos depicting eight different types of actions and states similar to those used in Legendre et al. (2010b) were used but these stimuli were adapted to specific HC transitive and intransitive verbs familiar to young children, based on questionnaire data collected on HC-speaking children in Haiti and NYC (Barrière, Joseph & Guetjens 2014). They included: *li* ‘read’, *fèmen* ‘close’, *souri* ‘smile’, *domi* ‘sleep’, *trape* ‘catch’, *bati* ‘build’, *kouri* ‘run’, *leve* ‘lift’, *achte* ‘buy’, *ouvri* ‘open’, *antre* ‘enter’, *ekri* ‘write’, *anbrase* ‘kiss’, *abiye* ‘dress’, *aplodi* ‘applaud’, *ale* ‘go’. More information on the motivation behind the selection of the verbs is given in the section describing the audio/verbal stimuli in the comprehension task below.

The boys in the videos were chosen to reflect the ethnicity of the majority of Haitian people. In each of the videos two boys appear, and for the singular action, one of the boys performs the action while the other stands still, and for the plural form both boys perform the action together. For the videos depicting transitive verbs, different unfamiliar objects were used in both the singular and the plural videos, similarly to Legendre et al. (2010b) for the same two reasons: to avoid an effect of (un)familiarity between specific verbs and nouns used in the object DPs (Valian, Prasada & Scarpa



2006) and to avoid a collective interpretation of both screens together and ensure that the singular and plural versions were distinct but that the object could not facilitate the selection of the mapping between the singular and plural subject pronouns and the scene enacted in the video (Kouider et al., 2006). Each video lasted six seconds.

[INSERT FIGURE 1]

### *Procedure*

This task was developed for the purpose of this study and its objective was to elicit the production of 3<sup>rd</sup> person singular and plural subject pronouns, in a context that would enable us to determine whether *li*-SG and *yo*-PL are mapped onto the appropriate number of instigators in the event or state described by the verbs depicted in the videos. At the beginning of the testing session, the experimenter (a native speaker of HC) explained to the child that they were about to see some videos, their task was to watch each of them attentively and describe what they saw. The instructions given to the child were *Di 'm sa 'w panse sa ye la nan videyo sa?* ‘Tell me, what do you think happens in this video?’. In line with the elicitation tasks administered previously in other languages (i.e., relevant sections of the DELV screener, Seymour 2005, in English and adaptations to other languages, using still pictures), the order of the visual stimuli used to elicit specific verbs was the same across participants, with the singular and plural videos alternating on the left and right sides of the screen. While administering the task, the experimenter was sitting next to the child and the videos were displayed on a thirteen inch wide and nine inch high Asus laptop.

### *Task 2: Video Comprehension*

#### *Video Stimuli*

The visual stimuli were the same as those used for the video elicitation task and described above.

#### *Verbal/Audio Stimuli*

Given that the objective of the present study was to evaluate the potential contributions of the morpho-phonological characteristics of HC subject pronouns in the acquisition process, five distinct conditions were tested (see Table 8). Three involved testing comprehension of the same vowel-

initial verbs in different contexts. The first one used the full forms of *li* and *yo* occurring with vowel-initial verbs (C1 in Table 8), while the second one used the reduced forms of *li* as /l-/ and *yo* as /j/, a condition involving proclitics and fostering resyllabification (C2). The third condition used the enclitic /-l/ that follows the very common conjunction *e* ‘and’ (C3). In this condition, the enclitic /-l/ indicating singular is contrasted with full *yo*, as *yo* cannot be reduced and attached to a preceding vowel-final word in the standard Port-au-Prince variety.

The last two conditions tested comprehension of the same consonant-initial verbs either in a context of full pronouns (C4) or of the enclitic /-l/ (again in contrast with full pronoun *yo*, for the same reason as above) (C5).

[INSERT TABLE 8]

Given the small number of verbs that start with vowels in HC, especially when taking into account those that are likely to be familiar to young children (Barrière et al. 2014), each condition included a combination of transitive and intransitive verbs. As in Legendre et al. (2010b) and Barrière et al. (2016a), the object DPs used with transitive verbs consisted of nonce words that comply with the phonotactics of HC: they were bisyllabic like many nouns familiar to young HC-learning children (Barrière et al. 2014), and they were made up combining the first and last syllables of different existing words listed among words familiar to young HC-speaking children (e.g. *bane*, *jaka*) (Barrière et al. 2014).

### *Procedure*

Each child was tested individually in a quiet space in their preschool. Children were seated in front of an Asus laptop displaying the visual stimuli. The experimenter was seated behind the child and next to the laptop. First, the child was told that some images would appear on the screen and that they would be asked to touch one of the images. Each child was tested in the five different experimental conditions described above.

Each session began with four training trials, consisting of two images of familiar objects (a house, a car, a cat, a dog, a book, a fork, an apple and a leaf) presented in silence on each side of the screen separately, followed by their simultaneous presentation on both sides. After 12 seconds

of visual presentation, the live experimenter said, for example: *Eske ou we kay la? Montre'm kibo ou we kay la? Touche kibo ou we kay la?* 'Do you see the house? Point where the house is, where is the house?' The child had to select one picture by pointing or touching the picture on the left or the right side of the screen.

After four training trials, a 3-second-long video of two boys waving was presented (the same boys that appear in the test videos) while the experimenter said: *Kounye a, ou pral gade kèk videyo ak de ti gason k'ap jwe. Gade timoun yo !* "Now you will see videos of boy(s) who are playing. Watch the children!" The test phase consisted of eight trials. Each trial started with one video presented in silence twice on the left side of the screen. After the first video disappeared, a second video appeared in silence on the right side of the screen; it was played twice and then disappeared. Both videos depicted the same action, with one video representing the 'singular' form of the action and the other representing the 'plural' form (see Fig. 1). Then, both videos (plural and singular) were displayed simultaneously while the live experimenter said, for example: *Montre'm kibo li souri? Touche kibo li souri?* 'Show me where he smiles, point to where he smiles'. In both the training and experimental phases, the color of the screen background changed to green after the child selected the image, regardless of the accuracy of their response. At the end of each test trial, the same eye-catching video used after the training trials (see above) was displayed on the side which played the *matching* video to keep the children interested in the task, following Kouider et al. (2006), Legendre et al. (2010b) and Barrière et al. (2019).

For half of the test trials, the verbal stimulus, produced by the live experimenter, corresponded to the singular video, while it corresponded to the plural video for the other half of the trials. The side on which the matching video was presented was counterbalanced within participants (hence for the singular trials, two out of four correct responses were on the left side, the other two being on the right side; same for plural trials). Each condition included four verbs used with singular pronouns and four verbs used with plural pronouns. Within each condition, the singular and plural versions of the same verbs alternated across participants and the order of the stimuli was randomized across participants.

#### 4.3 Analyses and Results

### *Transcription, Analyses and Results of the Video-Elicitation task*

All samples were transcribed by the second author. The same transcription conventions as for the adult corpus described above were used. They were checked by two other native speakers of HC, including the fifth author of this article. Apart from orthographic errors in fewer than ten utterances, the transcribers agreed on all but one utterance that was discussed and corrected once an agreement was reached.

Analyses of the CHAT transcriptions were conducted using CLAN (Joseph 2017, MacWhinney 2000). Table 9 presents quantitative analyses conducted with the basic functions of CLAN including lexical density and Type-Token ratio (TTR) with FREQ. Much variability was observed with respect to the number of utterances produced by individual children, ranging between 4 and 50 utterances, with 16 children producing samples of more than 20 utterances. Given this characteristic of our production data combined with the fact the percentage of verbs produced constitutes a standard measure in language acquisition and, more importantly, that the construction targeted in the present study involves the use of verbs, percentages of utterances that contain verbs (tokens) in each sample and percentages of verbs out of total number of words in each sample were also calculated (see Joseph, 2017 for details). No significant correlation (Pearson correlation) was observed between the age (in months) and percentage of utterances that contain verbs ( $r(20)=0.37$ ,  $p=.11$ ). In contrast, a significant positive correlation (Pearson correlation) was observed between the age (in months) and percentage of verbs out of total number of words ( $r(20)=0.67$ ,  $p=.001$ ).

[INSERT TABLE 9]

The next step involves analyzing the distribution of the different forms of the 3<sup>rd</sup> person subject pronouns (full and reduced) and whether participants' productions of *li* were systematically associated with the singular and *yo* with the plural. All instances of *li* realized as full pronouns, proclitics /l-/ and enclitics /-l/ were identified. If a child hesitated and repeated the pronoun more than once with the same verb or the same combination of Tense-Mood-Aspect marker and verb in the same utterance, it was only counted once. The same principles were applied for full *yo* and its proclitic /j-/counterpart. The results are presented in Table 10. As expected, the samples that do not include verbs (HA19, HA20, HA17, HA10, Ha7) do not contain any subject pronoun. Other participants vary in the numbers of subject pronouns they produced, for different reasons. The

instructions that were used to elicit the utterances were meant to avoid modeling and thereby priming of the production of specific subject pronouns. A weakness in that elicitation strategy is that it gave the participants the opportunity to produce different types of utterances including non-target/non-adult-like subject-less utterances and adult-like/target-like subject DPs, including nouns, quantifiers, or subject pronouns. So, while participants HA19, HA20, HA17 and HA10 produced short utterances made up of mostly nouns and greetings that do not function as subjects in adult utterances, HA8 produced all verbs (that make up 42% of the words they used) with subjects, including 34 with the quantifier *youn* (see example (9)) and only one with the reduced form of *yo* (see example (10)).

(9) youn ekri youn pa ekri (HA8)

QUANT V QUANT NEG V

‘One write one not write.’

‘One writes the other doesn’t

(10) y’ ap domi (HA8)

SubjPn 3PL TMA V

He is sleeping

[INSERT TABLE 10]

The fact that children’s production of subject DPs did not systematically involve the use of subject pronouns may explain why no correlation was identified between the participants’ age in months and the number of singular ( $r(19) = 0.19$ ,  $p = .425$ ) and plural ( $r(19) = 0.24$ ,  $p = .309$ ) subject pronouns they produced. Occurrences of the 3<sup>rd</sup> person pronouns were identified in fifteen of the samples, with twelve children producing both the singular and the plural, two producing only the singular (HA15, HA5) and one only the plural (HA8). All subject pronouns were produced in the context of utterances that were either made up of single verbs or of combinations of the Tense-Mood-Aspect markers *ap* or *pral* followed by a verb. Hence, based on the results of the data

analyses presented in Table 10, HC-speaking preschoolers produce both singular and plural 3<sup>rd</sup> person subject pronouns around the time they reach 3;6. For the singular, eight participants produced both full pronouns *li* and the proclitic /l-/, three only produced full forms and another three only produced the reduced form. In contrast, only two participants produced the enclitic /-l/ and both of them also produced full *li* and proclitic /l-/. For the plural, eight participants produced both full pronouns *yo* and the proclitic /j-/, two only produced the full forms and another three only produced the proclitic. In summary, it appears that both full forms and proclitics emerge at the same time in production. As for enclitics, they either emerge later since they are only produced by two children-aged 3;6;11 and 4;4;6, or there are too few relevant contexts to draw a conclusion at this time.

More importantly, the phonological contexts in which the reduced forms were produced reflect the constraints applied by adult HC speakers that were identified in Part 2 (corpus analyses). The examples below illustrate the use of full *li* with consonant-initial verbs (example (11)) and vowel-initial verbs (example (12)).

(11) **li** chita (HA3)  
**he** sit(s)

(12) **li** ouvri'l (HA3)  
**he** open(s) it

Like in the adult corpus, instances of pro-clitic /l-/ were only found when it preceded a vowel-initial verb (example (13)) or a vowel-initial TMA marker (example 14).

(13) **l'**ale ankò (HA16)  
**he** go(es) again.

(14) **l'**ap souri (HA3)  
**he** ap-TMA smile(s)

**he** is smiling

Occurrences of the enclitic /-l/ followed vowel-final words (example (15)). Occurrences of full *yo* were found before consonant-initial verbs (example (16)) and vowel-initial verbs (example (17)). Instances of proclitic /j-/ were identified before vowel-initial verbs (18) or vowel-initial TMA markers (19).

(15) pou'l mete sou li (HA4)

So/For **he** put(s) under bed

(16) yo frape na pòt la (HA3)

**they** hit/knock on door the

**they** knock on the door

(17) yo ouvè radyo a (HA12)

**they** open radio a

**they** switch/turn on a radio

(18) y'ale (HA6)

**they** go

(19) y'ap ri (HA4)

**they** ap-TMA laugh

With respect to the mapping between *li/l-/l* versus *yo/y-* and singular versus plural 3<sup>rd</sup> person subject pronouns, two mapping scores were calculated. After the transcriptions were completed and

double-checked, the second author/experimenter re-listened to each utterance produced by the child and placed it in the context of the video the child was watching when they produced the utterances containing subject pronouns. The singular mapping score was calculated based on the percentages of time the child was producing *li/l/-l* referring to a single boy in the singular video. The plural mapping score considered the proportion of times *yo/y* was used in order to refer to the two boys exhibiting the same state or performing the same action in the plural video. It is important to note that from a semantic perspective, the depictions of two boys involved in the same state or event can be described by making reference to both boys or focusing only on one. Hence using *li/l/-l* to describe a plural video does not necessarily qualify as an error per se. However, the fact that there is an age effect in the mapping scores (see below) indicates that this factor is likely to have played a minimal role in the results.

This task enabled us to determine the mapping between the use of a specific pronoun and the number of instigators referred to by the children in the visual display in a more systematic and reliable way than in spontaneous speech production data. Indeed, most analyses of spontaneous language samples (mentioned in Table 1) have not systematically considered the context in which the children produce singular and plural pronouns or agreement marker. Hence the consistency of the relation between the actual number of referents mentioned by the child and the linguistic forms they use in order to refer to them could not be fully examined.

The results reveal that very few children systematically mapped *li/l/-l* to singular and *yo/y*- to plural 100% of the time: only the oldest child (69 months, or 5 years and 9 months) did so for both singular and plural subject pronouns. Two sets of quantitative analyses revealed that semantic representations of the number assigned to *l(i)* and *y(o)* are still developing in this age range. First, a 2-way ANOVA with age as a between participant factor (younger N=10 versus older N=10, median split, cut-off age: 52 months, or 4 years and 4 months) and number as a within participants factor (singular versus plural) was conducted for the production mapping scores as a dependent variable. The ANOVA revealed a marginal effect for age at  $\alpha = .10$  ( $F(1,18) = 3.065$ ;  $p = .089$ ), no significant effects by number ( $F(1,18) = .06$ ;  $p = .808$ ) effects, nor a significant age x number interaction ( $F(1,18) = .258$ ,  $p = .615$ ). However, given trends observed in Figure 2, we conducted planned comparisons evaluating the effect of age separately for singular and plural. They revealed no significant difference between age groups in singular mapping scores (younger,  $M = 32.0$ ,  $SD =$



36.53; older,  $M = 45.7$ ,  $SD = 29.63$ ;  $F(1,18) = .848$ ,  $p = .369$ . Neither was there a significant difference between age groups in plural mapping scores (younger,  $M = 29.1$ ,  $SD = 34.81$ ; older,  $M = 54.0$ ,  $SD = 37.9$ ;  $F(1,18) = 2.34$ ,  $p = .143$ ), suggesting no significant difference in number-appropriate production of 3<sup>rd</sup> person subject pronouns between the younger and the older participants. Figure 2 shows the difference in age groups in singular and plural production mapping scores.

[INSERT FIGURE 2]

Second, correlation analyses were computed. Participants' age (in months) correlated with the proportions of the time they use *li* to refer to the singular ( $r(18) = 0.50$ ,  $p = .025$ ) and *yo* to refer to the plural ( $r(18) = 0.54$ ,  $p = .015$ ). The relationship between the singular and plural mapping scores was also examined, and the correlation ( $r(18) = 0.36$ ,  $p = .12$ ) failed to reach significance. These correlation results combined with marginal difference between the youngest and oldest age group suggest that the uses of *li* to refer to singular and *yo* to refer to a plural subject develop between 29 months and 69 months but not at the same speed.

To summarize, HC-speaking children start producing both full and reduced forms of the 3<sup>rd</sup> person singular and plural subject pronouns around 3.5 years of age, and the distribution of their full and reduced forms reflects adult morpho-phonological principles. This relatively early mastery of morpho-phonological constraints, contrasts with their non-target like assignment of number to singular *li* and plural *yo* in production, which develops until they reach five years and nine months. It therefore seems particularly relevant to explore the meaning that the same children assign to these pronouns in the context of an experimental comprehension task using the same visual stimuli.

### *Analyses and results of the Video Comprehension Task*

#### *Accuracy analyses*

Out of 20 children, one -HAC3- was not administered all the conditions since they were distracted and fidgety. Their performance in the conditions they were administered revealed a clear

(100%) left side bias. For these two reasons, all the analyses reported below do not include the performance of this participant.

The results of the comprehension experiment were examined in several ways. We again split the participants into a younger and older age group, using the same median split as for the production analyses, which resulted in a younger group of 10 children and an older group of 9 children. The statistical analyses were conducted for accuracy (e.g., Legendre et al., 2010b) and sensitivity scores (e.g., Johnson et al. 2005), presented separately below.

[INSERT FIGURE 3]

Accuracy was calculated considering the correct number of responses over the total number of responses for singular and plural number separately. To examine whether the overall accuracy (all five conditions collapsed) was affected by age and number, a 2-way analysis of variance (ANOVA) was conducted with age as between participant factor (younger versus older) and number as within participant factor (singular versus plural; see Figure 3). There was a significant effect of number ( $F(1, 34) = 4.747, p = .036$ ) with plural number mean accuracy at .553 and singular mean accuracy at .457. There was no effect of age ( $F(1, 34) = 2.944, p = .157$ ) with older mean accuracy at .537 and younger mean accuracy at .473 suggesting that the older group performed somewhat better than younger participants. The interaction of age and number was not significant ( $F(1, 34) = .121, p = .730$ ). This model exhibited an adjusted  $R^2$  of .169 indicating that 16.9% of the variability in collapsed accuracy is explained by this model.

Additional analyses were conducted to explore comprehension effects in the different conditions, and explore two factors that might impact performance: the type of onset of the verb (consonant versus vowel), and whether it is transitive or intransitive.

[INSERT FIGURE 4]

No effect of condition was observed for accuracy scores when adding this factor to the previous ANOVA (possibly due to variability inherent in developmental studies, requiring larger samples). Since the objective of the study was to examine whether and to what extent different conditions involving full and reduced pronouns facilitated or hindered children's comprehension, accuracy

was tested against chance (chance being 50%- given binary experimental conditions) for all conditions (See Figure 4). Plural accuracy for condition C5 that includes the conjunction *e* ‘and’ combined with full plural pronouns followed by a consonant-initial verb ( $t(13) = 2.828, p = .007$ ) was the only condition in which performance was statistically significantly above chance.

Second, to investigate whether accuracy is determined by the onset of the verb, we conducted a 2-way ANOVA with onset type (vowel versus consonant) and number (singular versus plural (see Figure 5). In total, for the plural stimuli, there were 44 vowel ( $M = .506, SD = .068$ ) and 27 consonant ( $M = .620, SD = .234$ ) onset trials; for the singular stimuli, there were also 44 vowel ( $M = .500, SD = .087$ ) and 27 consonant ( $M = .426, SD = .086$ ) onset conditions. There was a significant effect of number, ( $F(1, 141) = 6.627, p = .012$ ), no effect of onset type, ( $F(1, 141) = 0.269, p = .605$ ), and a significant interaction between number and onset type ( $F(1, 141) = 5.807, p = .017$ ). Adjusted R-Squared indicates a poor model explaining only 4.9% of the variance in verb type and number.

[INSERT FIGURE 5]

Third, the results were also analysed in terms of the verb type (transitive versus intransitive), since the use of pseudo-nouns in the object phrase may have impeded comprehension (based on findings in a Spanish study, Gonzalez-Gomez et al., 2017). A 2-way ANOVA with verb type (intransitive versus transitive) and number (singular versus plural; see Figure 6) revealed an effect of number ( $F(1, 296) = 5.02, p = .026$ ), no effect of the verb type, ( $F(1, 296) = 0.14, p = .709$ ), and a significant interaction between number and verb type, ( $F(1, 296) = 6.83, p = .009$ ). This was due to the fact that, for the intransitive condition, accuracy was significantly higher for plural ( $M = .59, SD = .32$ ) than singular ( $M = .42, SD = .30$ ),  $t(146) = 3.41, p < .001$ ; performance in that intransitive plural condition (involving the full plural pronoun *yo*) was significantly above chance ( $t(74) = 2.48, p = .007$ ). In contrast, for the transitive condition, accuracy was not significantly different between plural ( $M = .51, SD = .29$ ) and singular ( $M = .42, SD = .32$ ),  $t(74) = -.27, p = .395$ . Adjusted R-Squared indicates a poor model explaining only 2.9% of the variance in verb type and number.

[INSERT FIGURE 6]

### *Sensitivity Analyses*

For sensitivity, following Johnson et al. (2005) and Barrière et al. (2016), the ratio of correct responses (hits) over the total of correct responses and false alarms was calculated for each participant for plural and singular number videos separately<sup>1</sup>. The sensitivity measure eliminates a potential bias towards a certain type of visual stimuli (e.g., number bias) in an experiment with a binary choice. A two-way ANOVA with age and number failed to reveal any significant effect or interaction (number:  $F(1, 34) = .004, p = .950$ ; age group:  $F(1, 34) = 1.405, p = .244$ ; age x number:  $F(1, 34) = .032, p = .859$ ). Adjusted R-Squared = .041 indicates a poor model explaining only 4.1% of the variance in verb type and number.

Yet, as in the accuracy analyses, we conducted sensitivity analyses to explore performance in the different conditions. First, performance failed to reach significance in any of the conditions for either singular or plural.

Second, a 2-way ANOVA with onset type (consonant versus vowel) and number (singular versus plural) revealed no effect of number ( $F(1, 296) = .020, p = .889$ ), no effect of verb type, ( $F(1, 296) = .901, p = .343$ ), and no significant interaction between number and verb type, ( $F(1, 296) = .098, p = .754$ ). This was due to the fact that, for the consonant condition, sensitivity was not significantly different for the plural ( $M = .52, SD = .019$ ) and the singular ( $M = .53, SD = .013$ ) condition. Also, for the vowel condition, sensitivity was not significantly different between plural ( $M = .509, SD = .015$ ) and singular ( $M = .506, SD = .015$ ). Adjusted R-Squared indicates a poor model explaining only 0.2% of the variance in verb type and number.

Third, a 2-way ANOVA with verb type (intransitive versus transitive) and number (singular versus plural) revealed no effect of number ( $F(1, 296) = .191, p = .664$ ), no effect of the verb type, ( $F(1, 296) = 1.243, p = .268$ ), and no significant interaction between number and verb type, ( $F(1, 296) = 1.964, p = .165$ ). This was due to the fact that, for the intransitive condition, sensitivity was not significantly different for plural ( $M = .52, SD = .029$ ) and singular ( $M = .54, SD = .03$ ) condition. Also, for the transitive condition, accuracy was not significantly different for plural ( $M = .53, SD = .04$ ) and singular ( $M = .49, SD = .12$ ). Adjusted R-Squared indicates a poor model explaining only 0.5% of the variance in verb type and number. Singular sensitivity to intransitive verbs for consonant-initial verbs with full pronoun condition (C4) was significantly above chance ( $t(21) = 2.28, p = .032$ ).

---

<sup>1</sup> Singular (Plural) Sensitivity =  $\frac{\text{SG (PL) hits}}{\text{SG (PL) hits} + \text{SG (PL) false alarms}}$

In summary, between the ages of 2;5 and 4;4, there is no evidence that HC-speaking children assign adult-like meaning to the 3<sup>rd</sup> person singular and plural subject pronouns. Older children are better able to do so and the conditions that positively contribute to their adult-like comprehension include full forms of the pronouns, consonant-initial verbs, intransitives and the combination of the conjunction *e* ‘and’ and the pronoun. We will consider which of the five factors proposed in the introduction best account for these within-language results in the general discussion, taking into account findings that have emerged from other languages.

#### *Analyses of the Relation between Comprehension and Production Data*

The relationship between production mapping scores for singular and plural and comprehension accuracy and sensitivity scores was investigated. Accuracy and sensitivity scores for only best performing conditions with consonant-initial verbs (C4 and C5, see Figure 6) were included in the analyses. For singular accuracy, no significant correlation was identified between the singular mapping production score and singular comprehension accuracy ( $r(19) = -.012, p = .962$ ). For singular sensitivity, the correlation between singular mapping score and singular comprehension scores did not reach significance ( $r(19) = 0.149, p = 0.542$ ). For plural accuracy, no significant correlation was identified between the plural mapping production score and plural comprehension accuracy ( $r(19) = 0.23, p = .353$ ). For plural sensitivity, the correlation between plural mapping score and plural comprehension scores did not reach significance ( $r(19) = 0.186, p = 0.447$ ).

These results – lack of correlation between mapping score in production and comprehension – combined with correlations with age for production (mapping score) and comprehension scores (e.g. accuracy for the plural condition) suggest that both comprehension and production of these pronouns become more-adult-like over time, but we found no evidence that they do so in a related way.

## **5. General Discussion**

The contribution of this study to a better understanding of the status and distribution of HC subject pronouns and of their acquisition are highlighted first. This is followed by a discussion of how our

findings relate to results emerging from studies on other languages and how they enable us to disentangle the possible effect (or lack) of the five factors discussed in Section 3. The limitations of this study and possible follow-ups are also mentioned.

#### *Status and Distribution of HC Subject Pronouns in HC Adult Grammar*

Part two discussed the syntactic status, phonological realizations, and distribution of HC subject clitics, on the basis of analyses of two spoken corpora of casual speech. The syntactic distribution provided evidence for an argumental, DP-like status of HC subject pronouns, including when they are phonologically reduced, leading to the conclusion that they are phonological rather than syntactic clitics (agreement markers).

The phonological reductions shed light on the complex phonological constraints that may trigger phonological reductions and reflect the principles outlined by Cadely (1993) with respect to the encliticization of /li/ and the procliticization of /jo/ but not the procliticization of /li/: according to Cadely (1995) procliticization can occur when /li/ precedes a consonant-initial verb, contra grammaticality judgments from the 2<sup>nd</sup> and 5<sup>th</sup> authors who are native speakers of HC. No instance of proclitic /l-/ was found before consonant-initial verbs in the two corpora. The phonology of HC subject pronouns involves two processes that distinguish HC from the other languages that have been the focus of acquisition of subject-verb dependencies. First, the process of resyllabification triggered by the reduction of pronouns to a single consonant cliticizing to the preceding or following word is optional and is not systematically used by adults, including in the context of casual speech. In addition, when reductions occur they give rise to both phonological proclitics and enclitics. Additionally, enclitics involve elements that are external to the subject-verb dependency and the prosodic units they are associated with. Second, with respect to the various dimensions that enter into the measurement of the complexity of paradigm cells, the singular subject pronoun *l(i)*'s cell is more complex than plural *y(o)*'s cell. With respect to the systematicity of the mapping between *li* and singular and *yo* and plural, *yo* was also found to be less frequent but more semantically transparent than *li*.

#### *Production*

The results of the Video Elicitation task collected on 20 native HC-learning children between the ages of 27 months (2;5) and 69 months (5;9) reveal that at least from age 3;5, HC-speaking children

use the full and contracted forms of the 3<sup>rd</sup> person singular and plural subject pronouns and apply adult-like constraints when they produce the reduced forms. The production of the 3<sup>rd</sup> person subject singular and plural involves the alveolar /l-/ and the palatal glide /j-/ that are both acquired by age 3 in HC (Archer et al. 2017). The fact that the earliest productions of the reduced forms of subject pronouns reflect mature morpho-phonological constraints is also not surprising in light of what is known about the early acquisition of resyllabification processes in other languages, such as CHFrench liaison in nominal (Buerkin-Pontrelli et al. 2017) and in verbal contexts (Dugua et al., 2017).

The mastery of morpho-phonological constraints exhibited in our production data stands in contrast with the finding that at the same age, semantic associations of *l(i)* with singular and *y(o)* with plural in this production task are not systematic. It is difficult to compare these results to those obtained on the production of subject-verb dependencies in other languages given that other studies (see Table 1) tend to rely on spontaneous speech production samples that make it difficult to examine the actual mapping between the number marked on pronouns and/or agreement markers and the actual number of the instigator in the event or state. However, our results confirm the view that children's production of specific structures does not systematically imply that they assign adult-like representations to these constructions (e.g., Kenny & Wolfe 1972; Smolík, Filip & Bláhová, Veronika 2017) and that the mapping between specific forms and number improves with age. The issue of the mapping between number and specific subject pronouns is further discussed below, in light of the comprehension results.

### *Comprehension*

The results of the comprehension task reveal that HC-learning children do not demonstrate any evidence of comprehension of 3<sup>rd</sup> person subject singular and plural pronouns until after age 4;4. The few conditions in which their sensitivity or accuracy is above chance systematically involve full pronouns, combined with other factors such as consonant-initial verbs (Figure 5), presence of a conjunction and/or intransitive verbs (Figure 6). We explore further below the possible factors that may account for this set of results and how they inform the cross-linguistic and within-language hypotheses listed in the introduction.

In our analyses, we could only find some evidence of comprehension with intransitive verbs (plural for accuracy, singular for sensitivity), not transitive ones. In the experimental stimuli,

transitive verbs were presented with nonce nouns, for reasons explained above, and for ultimate comparison with results on CHFrensh acquisition from Legendre et al. (2010b) and Barrière et al. (2016 a). However, a previous study on Spanish using the same procedures demonstrated evidence of comprehension of the suffixal agreement marker on the same set of verbs when the transitive construction included the real word *objeto* ‘object/thing’ as object DP, but not when a nonce noun was used instead (Gonzalez-Gomez et al. 2017). This suggests that the facilitating effect of intransitive verbs in the present study may stem from the fact that they did not involve the use of nonce nouns which appear to increase the processing load and impede the interpretation of the singular versus plural markers in ‘older’ children (the Mexican Spanish ones, and our HC children), compared to the younger CHFrensh ones (2;6), who showed evidence of comprehension.

Additional HC results further highlight the importance of the format of the experiment and the components of the verbal stimuli. While stimuli using the full pronoun *yo* were presented in four different conditions (C1, C3, C4 and C5, see Table 7), *yo* was better understood when it was contrasted with the enclitic /l-/ (C5), despite the fact that it was used in the same form and with the same verbs as in C1, in which it was contrasted with the full pronoun *li*. The only way that the plural construction involving the use of *yo* in C5 differs from the other conditions that also involve the use of the full plural pronouns is that the subject pronoun was preceded by the conjunction *e* ‘and’, which appears to have facilitated the plural interpretation of the stimuli.

On the basis of our results, it is not possible to determine whether the difficulty that HC-speaking children experience with vowel-initial verbs comes from the extremely low frequency of these verbs in HC, both with respect to types and tokens, and/or from resyllabification that both phonological proclitics and enclitics may trigger, and/or from the optionality per se of resyllabification; it could also result from a combination of these factors. For comparison’s sake, vowel-initial verbs also exhibit low frequency in CHFrensh but the process of resyllabification triggered by liaison is obligatory, and the combinations of 3<sup>rd</sup> person singular or plural subject and vowel-initial verbs are understood much earlier- at 30 months of age (2;6) (Legendre et al. 2010b; Barrière et al. 2016). A priori, an obligatory rule is simpler than an optional one (which suggests either a further constraint or randomness). More studies comparing the effect of obligatory versus optional rules are needed to better understand this important distinction for language acquisition. This issue is further discussed in the context of the fifth factor below.



The present HC study is part of a broader cross-linguistic acquisition project that seeks to evaluate a number of factors previously invoked as having a potential effect on the course of morpho-syntactic acquisition. With this cross-linguistic objective in mind, we now turn to the five factors outlined in the introduction from the perspective of the results we obtained on HC.

### **First factor: Pre-verbal Position and Adjacency between the Number Marker and the Verb**

First, the preverbal position of pronouns does not seem to facilitate acquisition, given that in both HC (present study) and Xhosa (Gxilishe et al., 2009) where person and number are marked pre-verbally, such marking is not understood earlier than suffixal agreement in Greek (Kenanidis et al. 2021), Mexican Spanish (Gonzalez-Gomez et al. 2017; Hsin et al. 2021) and English (Barrière et al. 2019), for which evidence of comprehension is obtained before age 4. If the preverbal position is moderated by the adjacency between the number marker and the verb, it may explain why HC learning children take longer to acquire number marking on subject-verb dependencies than CHFrench-speaking children: although the stimuli in our comprehension task involved dependencies between adjacent elements, HC allows for a broader range of lexical and functional categories to occur between subject pronouns and verbs than CHFrench does between the subject clitic and the verb. Another dimension of the distance between the number marker and the verb that may impact acquisition pertains to the fact that in HC in general, and also in our experimental stimuli, the enclitic /-l/ was more distant from the verb in that it encliticized to the conjunction at the beginning of the sentence and therefore formed a prosodic unit with that conjunction rather than with the following verb with which it entertains a syntactic dependency. The enclitic /-l/ was produced by few children, who were also found to produce other forms of the 3<sup>rd</sup> person SG pronoun *li*. Moreover, it was not understood by the children in any of the conditions in which it appears (C3, following the conjunction *e* ‘and’ and preceding a verb that starts with a vowel and C5, following the same conjunction and preceding a verb that start with a consonant; see Figure 5 for accuracy and summary of sensitivity analyses). This suggests that distance measured not only with respect to number of syllables (Santelman & Jusczyk 1998), but also in relation to prosodic units, may hinder comprehension. Future studies could further explore this issue.

### **Second factor: Syntactic Status**

The syntactic status of HC subject pronouns as phonologically reduced argumental DPs distinguishes them syntactically from CHFrensh subject clitics analyzed as agreement markers (Culbertson 2010; Legendre et al. 2010a, for CHFrensh and Déprez 1994 for HC). In CHFrensh, strong pronouns are not tied to finiteness as systematically as syntactic clitics in the early stages of grammatical development. In English, which only has strong/argumental subject pronouns, the assignment of the relevant number feature to these strong pronouns does not emerge until age 4 (Scholes 1981). Syntactic clitics are found to occupy a different place in the hierarchical structure of sentences making them closer to agreement markers than to independent lexical subjects. Part two (corpus analyses) demonstrated that despite their phonological reductions, HC subject pronouns are not syntactic clitics: they behave like full pronouns. If syntactic cliticization has a facilitating effect on acquisition, it is not surprising that HC subject pronouns are acquired later than CHFrensh subject clitics and around the same age as English subject pronouns.

### **Third factor: Resyllabification**

The resyllabification process associated with (optional) elision in HC does not seem to constitute a facilitating factor. The few syntactic contexts in which some evidence of comprehension was found (i.e., accuracy above chance for the full plural pronoun *yo* that follows the conjunction *e* ‘and’ and precedes transitive verbs, as well as sensitivity above chance when the full singular pronoun *li* precedes a consonant-initial intransitive verb) involve full but not reduced pronouns, combined with other facilitating factors. So no facilitating effect of the phonological reduction and the resyllabification that it entails was found. Therefore, resyllabification may be a facilitating factor only when reductions are systematic and/or if the subject clitic exhibits high semantic transparency, like in CHFrensh. Another possibility that requires further exploration pertains to the frequency of (optional) resyllabification. Indeed, the adult corpus analyses showed that for 3<sup>rd</sup> person pronouns, contexts in which reduction can occur did not often trigger resyllabification. Future studies on the acquisition of HC could explore this issue in at least two different ways. First, across the different subject pronouns, the proportions of reductions vary across pronouns marked for different persons and numbers so the effects of the frequency of the reductions of specific pronouns on the acquisition of Standard (Port-Au-Prince) HC could be explored. Second, Cap Haitian HC allows for different reductions in a broader range of contexts, including for 3<sup>rd</sup> person singular and plural subject

pronouns. They are also optional but may differ in frequency, compared to Standard HC. A study comparing the acquisition of Standard and Cap Haitian HC may enable us to disentangle the effects of frequency versus obligatoriness of the phonological reductions on acquisition. The potential effect of the variability between phonologically full and reduced forms of HC subject pronouns on acquisition is further discussed below in the section on the fifth factor devoted to paradigm cell complexity.

#### **Fourth factor: Semantic transparency and frequency**

As mentioned in the discussion of the results of Study 1, the quantitative analyses of semantic transparency and frequency of *li* and *yo* have revealed an interesting difference: while *l(i)* exhibits higher frequency, it is less systematically associated with singular than *y(o)* with plural, hence it has lower semantic transparency.

First, the accuracy results do not clearly point to a bias towards the plural since accuracy above chance for the HC 3<sup>rd</sup> person plural pronouns was obtained only when other facilitating factors were present. These facilitating factors included the full realization of the pronoun, the presence of the conjunction *e* ‘and’ and verbs that are intransitive and/or start with consonants (Figures 5 and 6). So the semantic transparency of *yo* does not seem to facilitate children’s comprehension. From a cross-linguistic perspective semantic transparency does not explain either why *yo* is acquired later than the plural subject clitic *ils* in CHFrench, that also exhibits high semantic transparency and in which evidence of comprehension is found in 2;6 year-old children (Barrière et al. 2016 a).

With respect to the 3<sup>rd</sup> person singular pronoun *li*, sensitivity results provide evidence of comprehension only when the condition tested involves facilitating factors such as full expression of the pronoun and use of consonant-initial intransitive verb (Figures 6 and 7); it is therefore not clear that the higher cue availability (frequency) of *li* itself contributes to comprehension. Finally, despite the fact that enclitic /-l/ is almost as transparent as full subject /li/ and three times more transparent than proclitic /l-/, it is not understood earlier.

These two sets of results on the limited contexts in which *yo* and *li* are understood by young children do not enable us to disentangle the respective effects of semantic transparency and frequency. These results also suggest that other factors such as the argumental status of subject

pronouns and the optional reductions that trigger resyllabification mentioned above are likely to play a role in the acquisition of HC subject pronouns, as well as the complexity of the paradigm cell- a final factor that is discussed below.

### **Fifth factor: Variation in the Expression of Number Marking and Complexity of the Paradigm Cells**

The fifth factor that has the potential of explaining various aspects of our results is the degree of complexity of the paradigm cells. In particular, the cell of *li* contains more forms (both the enclitic /-l/ and the proclitic /l-/ in addition to the full form /li/) than that of *yo* (that manifests itself either as a full pronoun /jo/ or a proclitic /j-/ in a very limited set of contexts sometimes when it precedes a consonant-initial word). As a result, the selection of the form of /li/ involves the consideration of more factors in its phonological environment - both elements that precede and follow it whereas for *yo* only elements that follow matter to its realization. Although 3<sup>rd</sup> person singular *li* is more frequent than *yo*, its realization is less predictable when one considers the number of factors involved in its realization. Yet, both are unpredictable in the sense that even in contexts when they can be reduced, the contracted forms are used in 50% to 80% of the time (Tables 4, 5 and 6).

It is also interesting to note that enclitics seem to emerge later in production than full forms and proclitics. The processing (retrieval in production and understanding in comprehension) of enclitic /-l/ is a complex operation: while it involves the relatively simple consideration of whether the preceding word coda is a vowel, it is not restricted to the syntactic domain of the dependency – enclitic /-l/ agglutinates to any element that precedes the utterance in which the verb with which *li* entertains a relation of syntactic dependency occurs. From a comprehension perspective, both proclitic and enclitic processes involving /li/ in the input give rise to ambiguous utterances, as mentioned by Cadely (1995, 12). The examples below are adapted from Cadely (1995, 12): the /l/ can either constitute the subject of the second clause (20b) or object of the first clause (20c).

(20.a) Jak di l achte liv la

(20.b) Jak<sub>i</sub> says/tells (that) he<sub>i/j</sub> buys/has bought the book

(20.c) Jak<sub>i</sub> tells him<sub>j</sub> to buy the book

Although the experimental stimuli to which our participants were exposed to in the video matching comprehension task do not contain ambiguous examples exhibiting the structure of example (20), children are likely to be exposed to such constructions. When they are, their interpretation involves assigning the grammatical functions subject versus object to the pronoun in order to identify with which verbs it entertains the relevant syntactic dependency. This process may delay acquisition. Based on what is known about the sensitivity to prosodic cues by infants and toddlers in the acquisition of morphosyntax (e.g., Nazzi et al. , 2000; Soderstrom et al. 2003) and the fact that phonological encliticization involves elements beyond syntactic dependencies, it seems that tracking this form constitutes a processing challenge that adds to the complexity of the paradigm cell of *li*. The combination of factors that characterize the complexity of the paradigm cell may help explain why few enclitics are found in the speech production samples and why the only form of *li* that seems to be comprehended in a few syntactic contexts involves its full realization, and neither the enclitic (that exhibits higher semantic transparency, Table 7) nor the proclitic (that exhibits lower semantic transparency, Table 7).

It is difficult to compare HC *li* and *yo* to other variable markers of the singular/plural distinction in other languages studied so far. Most studies have involved either relatively unpredictable selections between overt and null nominal and verbal suffixes (e.g., non-MAE, Mexican and Chilean Spanish) or predictable number-marking liaison in CHFrench. However, the acquisition pattern of HC *li* and *yo* and the age range during which systematic mapping between form and meaning emerges in comprehension are relatively close to the age range in which evidence of comprehension is found in other languages listed in Table 1 (such as Brazilian Portuguese, Czech, non-MAE, Farsi and Xhosa). With respect to production results, it is difficult to compare our findings with those published in the literature given that most studies have relied on spontaneous speech production data or constrained elicitation task in formats that impede analyses of the mapping between form and meaning.

To the best of our knowledge, only two other studies have compared elicited production and experimental comprehension results of subject-verb dependencies in similar constructions. They both involve sentence completion tasks that foster the use of the verbal suffix and model the syntactic structure with the relevant verb to foster the use of the number marker; both document convergence between comprehension and production (Barrière et al. 2019 and Hsin et al. 2021). It

is unclear at this point whether the differences between the results of these studies and the present study come from the format of the elicitation task: in our study the children were not exposed to the relevant verb, and the subject pronoun + verb constructions were not modeled.

Most acquisition studies do not combine detailed qualitative and quantitative analyses of corpora that enable the examination of the constructions tested in experimental protocols. Often important information relevant to one or more of the factors investigated here is lacking for other languages and makes it difficult to assess the potential role of the different factors cross-linguistically, using our multidimensional methodology.

Overall, our appeal to the concept of paradigm cell complexity to account for the findings that emerged from this study (and possibly for cross-dialectal acquisition findings in Spanish and English agreement markers) raises an interesting issue regarding the nature of complexity in natural languages and its role in the acquisition process- an issue that has been discussed in the development and testing of computational models of acquisition (e.g., Erickson & Thiessen 2015). Variability in input in the expression of number in non-MAE, Spanish, and HC raises the issue of how the co-existence of different forms has to be distinguished from the presence of noise in the input. This is especially relevant when considering that speech perception studies have demonstrated that infants exposed to an artificial language rely on predictability of adjacent elements when extracting syllables (Aslin et al. 1998). A recent artificial language study revealed a facilitating effect of affix type frequency, a slow emergence of the consideration of phonological constraints, and worse performance on phonologically unpredictable than predictable forms (Nevat, Ullman, Eviatar & Bitan 2018). Our study extends the examination of the role of such (un)predictability to the L1 acquisition of a natural language by focusing on variable HC pronominal forms that exhibit a high degree of unpredictability within a single variety.

### *Comprehension and Production*

The present study is one of the few that examines both comprehension and production of subject-verb dependencies by the same participants (see also Miller 2012; Verhagen & Blom 2014; Barrière et al. 2019; Hsin et al. 2021). To the best of our knowledge, it is the first one that combines Video Elicitation and Video Comprehension tasks and enables us to investigate the comprehension and production of the same 3<sup>rd</sup> person subject singular versus plural pronouns in the context of the same animated visual stimuli. Our combined results do not go against a possible gap between

comprehension and production that has been documented for different other languages (e.g., English, German, Spanish and Xhosa), but it provides us with unique insights into those aspects of expressive language skills that are adult-like versus those that are protracted or emerge at the same time as comprehension. More specifically, the full and reduced forms of HC 3<sup>rd</sup> person subject pronouns reflect the adult morpho-phonological constraints *before* they are systematically mapped onto their corresponding number, both in comprehension and production, bringing support to the view that *‘meaning is hard but form is easy’* (Naigles 2002, 157).

### *Limitations*

Both the corpus study and the acquisition study exhibit limitations. A larger corpus of spontaneous speech with more occurrences of pronouns would enable us to further explore whether the pragmatic context impacts the realization of full versus reduced forms of subject pronouns: contra Cadely (1995), Hilton (2000) claims that it does for object clitics. Both authors rely exclusively on grammaticality judgments, so it would be interesting to determine whether full forms of subject pronouns are more likely to occur in the context of new information and reduced forms for old information, for instance. With respect to acquisition, it would have been best to have the children pass developmental screenings like participants in many studies in the languages listed in Table 1, to ensure that none of them had cognitive, hearing or vision issues. Unfortunately, this was not possible for two distinct reasons: to the best of our knowledge no developmental screening has been adapted to Haitian preschoolers and there is a dearth of personnel qualified to conduct developmental screening and cognitive evaluations. It would also have been best to recruit a higher number of participants with more distinct age groups and for both sexes to be equally represented. Given the results on Mexican Spanish (Gonzalez Gomez et al. 2017), it would have been better to test transitive verbs with a generic noun such as “the object”, rather than pseudo-nouns. However, the fact that HC-speaking children’s age of acquisition of number marking in subject verb-dependencies did not drastically differ from those reported in other languages suggests that these limitations are unlikely to have had a major impact on our results. The same applies to the cultural adequacy of the experimental paradigm. Compared to children tested in other contexts, the HC-speaking children who participated in our study are much less exposed to animated videos than their age peers in the Western Countries in which most of the studies listed in Table 1.1 took place. However, the facts that a) we did not have more instances of biases toward one side or plural or

singular videos than in other populations tested with the same video-comprehension paradigms and b) positive correlations between age and percentages of verbs produced, comprehension and production results suggest it is unlikely that the performance of the children was negatively impacted by the experimental paradigm.

### *Follow-ups*

At least two different follow-ups would help further address issues raised in the present study. First, a parallel study in Cap Haitian would enable us to establish the status and distribution of subject pronouns in Cap Haitian Creole and the effects of their features that differ from Standard Port-au-Prince HC (e.g., broader range of reduction, use of vowels rather than consonant in the reduction process), examining the distribution of full and reduced forms in more details than Valdman et al. (2015), with the aim of determining the frequency of the contractions as well their status with respect to frequency, semantic transparency and predictability. Second, from a theoretical perspective, it would be useful to further examine effects of paradigm cell complexity on the acquisition in different languages.

### *Contribution*

Despite its limitations, the present multidimensional study on the acquisition of dependencies involving subject pronouns in HC suggests that a number of factors impact the acquisition of number marking involved in subject-verb dependencies. Longer distance measured prosodically, higher degree of paradigm cell complexity, and argumental status of subject pronouns seem to impede acquisition, while the roles of semantic transparency and frequency are still unclear. Moreover, cross-linguistic comparisons suggest that the agreement marker status of number marking and obligatoriness of phonological processes seem to facilitate acquisition. Finally, although the present study does not directly address the issue of learnability of syncretic morphemes, it provides some indirect evidence against the assertion that syncretism has a facilitating effect in acquisition – a statement that is often made with respect to creole languages (e.g., McWhorter 2005): the HC 3<sup>rd</sup> person subject pronouns that exhibit syncretism with object pronouns and post nominal possessive determiners are not acquired earlier than their English counterparts that do not exhibit such syncretism (e.g. he-SUBJ vs him-OBJ; she-SUBJ vs her-OBJ; they-SUBJ vs them-OBJ). Finally, when the nature of complexity and its role in facilitating or hindering acquisition has been examined in language acquisition in general and more specifically



in inflectional morphology including the expression of number marking, the focus has been on the facilitating role of morphological richness (e.g., Phillips 1995; Xanthos et al., 2011) in reference to the size of morphological paradigms. The results of the present study suggest that the structure of paradigm cell constitutes another dimension of language complexity that may hinder the interpretation of grammatical features and that the role of this factor must be considered in acquisition. Our findings also emphasize the importance of examining different aspects of language complexity, including in creole languages (Henri et al. 2020), and their actual impact on language acquisition.

### **Acknowledgments**

Thanks to the children, parents and teachers of the Fonsbeda Académie who participated in the study and to Pamela Watters for assistance with the statistical analyses.

### **Funding Information**

The conduct of this study and the writing of this article were made possible thanks to NSF OISE/IRES and Linguistics BCS award # 1952501 and NSF BCS award # 1548147 awarded to Isabelle Barrière; NSF BCS # 125828 awarded to Géraldine Legendre and LABEX EFL (ANR-10-LABX-0083) awarded to Thierry Nazzi.

### **References**

- AKA (Akademi Kreyòl Ayisyen) .2017. Premye rezolisyon sou òtograf lang kreyòl ayisyen. Available at:  
[http://akademikreyol.net/dokiman/rezolisyon\\_otograf\\_kreyol\\_ayisyen\\_06012017115710.pdf](http://akademikreyol.net/dokiman/rezolisyon_otograf_kreyol_ayisyen_06012017115710.pdf). Retrieved on 12/12/2017.
- Ackerman, Farrell, Blevins, James P. & Malouf, Robert. 2009. Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In J. Blevins & J.Blevins (Eds.) *Analogy in Grammar: Form and Acquisition*. 54-82. Oxford: Oxford University Press. .

- Ackerman Farrell & Malouf, Robert. 2013. Morphological organization: the low conditional entropy conjecture, *Language*, 89 (3), 429- 465.
- Anderson, Stephen R. 2015 Dimensions of Morphological Complexity. In M. Baerman, D. Brwon, & G. Corbett (Eds.) *Understanding and Measuring Morphological Complexity*. 11-26. Oxford: Oxford University Press.
- Archer, Justine Champion, Tempii, Tyrone, Martha E. & Walters, Susan. 2018. Phonological Development of Monolingual Haitian Creole-speaking Preschool Children. *Communication Disorders Quarterly*, 39 (3), 426-437.
- Aslin, Richard N., Saffran, Jenny R., & Newport, Elissa L. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9 (4), 321-324.
- Baptista, Marlyse. 1995. On the nature of pro-drop in Capoverdean Creole. *Harvard Working Papers in Linguistics*, 5, 3- 18.
- Barrière, Isabelle, Joseph, Blandine, & Fleurio, Guetjens Prince. 2014. *Developing and Piloting the first language assessment tool for Haitian-Creole speaking toddlers*. Paper presented at the 20th Biennial Conference of the Society for Caribbean Linguistics, Palm Beach, Aruba.
- Barrière, Isabelle. 2015. New York Haitian Creole Corpus. Unpublished.
- Barrière, Isabelle, Goyet, Louise, Kresh, Sarah, Legendre, Géraldine, & Nazzi, Thierry. 2016a. Uncovering productive morphosyntax in French-learning toddlers: a multidimensional methodology perspective. *Journal of Child Language*, 43(5), 1131–1157.
- Barrière, Isabelle, Legendre, Géraldine, Joseph, Blandine, Kresh, Sarah, Fleurio, Guetjens Prince, Nazzi, Thierry. 2016b. Le statut linguistique des pronoms du créole haïtien et du français : une étude de corpus. *Actes du 5<sup>ème</sup> Congrès Mondial de Linguistique Française 2016*. [http://www.shsconferences.org/articles/shsconf/pdf/2016/05/shsconf\\_cmlf2016\\_14006.pdf](http://www.shsconferences.org/articles/shsconf/pdf/2016/05/shsconf_cmlf2016_14006.pdf)
- Barrière, Isabelle, Kresh, Sarah, Aharodnik, Katsyarina, Legendre., Géraldine & Nazzi, Thierry. 2019. The comprehension of 3<sup>rd</sup> Person Singular –s by NYC English-speaking Preschoolers. In Matthew Rispoli & Tania Ionin (eds) *Three Streams of Generative Language Acquisition Research*, 7-33. Amsterdam: John Benjamins.

- Bickerton, Derek 1984. The language bioprogram hypothesis. *The Behavioral and Brain Sciences*, 7, 173-221.
- Bonami, Olivier & Luís, Ana R. 2013. *A Morphologist Perspective on Creole Complexity. Rethinking Creole Morphology*. Paper presented at CIL 19, Genève.
- Brandt-Kobe, Oda-Christina & Höhle, Barbara. 2010. What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua*, 120, 1910–1925.
- Buerkin-Pontrelli, Angelica, Culbertson, Jennifer, Legendre, Géraldine & Nazzi, Thierry. 2017. Competing Models of Liaison Acquisition: Evidence from Corpus and Experimental Data. *Language*, 93 (1), 189-219.
- Cadely, Jean Robert. 1995. Elision et Agglutination en Créole Haïtien: Le cas des pronoms Personnels. *Etudes Créoles*, 18(1), 9-38.
- Cadely, Jean Robert. 1997. Prosodie et clicitisation en Créole Haïtien: Le métissage des Langues et des représentations. *Etudes Créoles*, 20(1), 5-114.
- Champion, Tempii, McCabe, Allyssa & Colinet, Yvonne. 2002/2003. The whole world could hear: the structure of Haitian-American children's narratives. *Imagination, Cognition and Personality*, 22 (4), 381- 400.
- Cleveland, Leslie H. & Janna B. Oetting. 2013. Children's marking of verbal –s by nonmainstream English dialect and clinical status. *American Journal of Speech Language Pathology*, 22 (4) 604-614.
- Childers, Jane B., Echols, Catherine H., & Tomasello, Michael, Fernandez, Ana María. 2000. Experimental investigations of children's understanding and use of verb morphology: Spanish-and English-speaking 2 1/2-and 3-year-old children. In Margareta Almgen, Andoni Bareña, María-José Ezeizabarrena, Itziar Idiazabal, & Brian MacWhinney (eds.), *Research on Language Acquisition: Proceedings of the International Association for the Study of Child Language*, 104–127. Somerville, MA, USA: Cascadilla Press.
- Clahsen, Haald, Aveledo, Fraibet, & Roca, Iggy. 2002. The development of regular and irregular verb inflections in Spanish child language. *Journal of Child Language*, 29(3), 599– 622.
- Culbertson, Jennifer. 2010. Convergent evidence for categorial change in French: from subject clitic to agreement marker. *Language*, 86 (1), 885-132.

- Déprez, Viviane. 1994. Haitian Creole: A Pro-Drop Language? *Journal of Pidgin et Creole Languages*, 9(1), 1-24.
- DeGraff, Michel F. 1993. Is Haitian Creole a Pro-Drop Language? In Francis Byrne & John Holm (eds.) *The Atlantic meet Pacific: A global view of Pidginization et Creolization*. 71-90. Amsterdam: John Benjamins.
- DeGraff, Michel F. 2005a. Linguists' most dangerous myth: The fallacy of Creole Exceptionalism. *Language in Society*, 34, 533-591.
- DeGraff, Michel F. 2005b. Morphology and word order: creolization and beyond. In G. Cinque & R.S. Kayne (eds) *The oxford handbook of comparative syntax* (pp. 293-273). Oxford, UK: OUP.
- DeGraff, Michel F. 2017. Mother-tongue books in Haiti: The power of Kreyòl in learning to read and reading to learn. *Prospects*, 47, 1-30 doi: <http://dx.doi.org/10.1007/s11125-017-9389-6>.
- Dejean, Yve. 2010. Creole and Education in Haiti. In A.K. Spears & C. Berotte Joseph (eds) *The Haitian Creole Language: History, Structure, Use and Education*. NY: Lexington Books.
- De Villiers, Jill G. & De Villiers, Peter. 1973. A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, 2, 267-278.
- De Villiers, Jill G. & Johnson, Valerie E. 2007. The information in third person /s/: acquisition across dialects of American English. *Journal of Child Language*, 34, 133-158.
- Dugua, Céline; Nardy, Aurélie, Liégeois, Loïc, Chevrot, Jean Pierre, Chabanal, Damien. 2017. L'acquisition des liaisons après les clitiques préverbaux est-elle spécifique? Apport d'une experimentation à grande échelle. *Journal of French Language Studies*, 27, 73-86.
- Doukas, Thomas & Marinis, Theodore. 2012. The acquisition of person and number morphology within the verbal domain in early Greek. *University of Reading. Language Studies Working Papers*, 4, 15-25.
- Erickson, Lucy C.. & Thiessen, Erik D. 2015. Statistical learning of language: Theory, validity ad predictions of a statistical learning account of language acquisition. *Developmental Review*, 37, 66-108.
- Fon Sing, Guillaume. 2017. Creoles are not typologically distinct from non-Creoles. *Language Ecology*, 1 (1), 44-74.

Gxilishe, Sandile, Smouse, Mantoa Rose, Xhalisa, Thabisa, & de Villiers, Jill.

2009. Children's insensitivity to information from the target of agreement: the case of Xhosa. In Jean Crawford, Koihi Otaki & Masahiko Takashi (eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America* (pp. 46–53). Somerville, MA, USA: Cascadia Press.

Henri, Fabiola, Stump, Gregory & Tribout, Delphone. 2020. Derivation and the morphological complexity of three French-based creoles. In P. Arkadiev and F. Gardani (eds.), *The complexities of morphology* (105-135). Oxford University Press.

Hilton, Dimitri. 2000. Pronominal Clitics in Haitian Creole: Phonological or Syntactic? In Vicki Carstens & Frederick Parkinson (eds.) *Advances in African Linguistics*. Trenton, NJ: SUNY Press. Pp 51-59.

Höhle, Barbara, Schmitz, Michael, Santelmann, Lynn M., Weissenborn, Jürgen. 2006. The recognition of Discontinuous Verbal Dependencies by German 19-Month-Olds: Evidence for Lexical and Structural Influences on Children's Early Processing Capacities. *Language Learning and Development*, 2 (4), 277-300.

Hsin, Lisa, Gonzalez-Gomez, Nayeli, Barrière, Isabelle, Nazzi, Thierry, Legendre, Géraldine. 2021. Converging Evidence of Underlying Competence: Comprehension and Production in the Acquisition of Spanish Subject-Verb Agreement. *Journal of Child Language*, 49 (5) 851-868.

Jean-Pierre, Marky .2016. *Language and Learning in a post-Colonial Context: a Critical Ethnography Study in Schools in Haiti*. New York: Routledge.

Johnson, Valerie E., de Villiers, Jill, & Seymour, Harry .2005. Agreement without understanding? The case of third person singular /s/. *First Language*, 25, 317–317.

Joseph, Blandine .2017. *Developing a Tagger adapted to the study of Haitian Child Language*. Unpublished Master Thesis Long Island University.

Kayne, Richard S.. (1977) *Syntaxe du français: le cycle transformationnel*. (traduit de l'américain par Pierre Attal). Paris: Seuil. First edition (1975): *French Syntax: the transformational cycle*. Cambridge, Ms: MIT Press.

Kenanidis, Panagiotis, Chondrogianni, Vicky, Legendre, Géraldine & Culbertson, Jennifer. 2021. Cue reliability, salience and early comprehension of agreement : evidence from Greek. *Journal of Child Language*, 48, 815-833.

- Kouider, Sid, Halberda, Justin, Wood, Justin, & Carey, Susann. 2006. Acquisition of English Number Marking: The Singular-Plural Distinction. *Language Learning and Development*, 2, 1–25.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
- Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar: The case of Haitian Creole*. Cambridge, UK: Cambridge University Press.
- Legendre, Géraldine, Culbertson, Jennifer, Barrière, Isabelle, Nazzi, Thierry & Goyet, Louise. 2010a. Experimental et empirical evidence for the status et acquisition of sujet clitics et agreement marking in adult et child Spoken French. In Vincenç Torrens, Linda Escobar, Anna Gavarrò, Juncal Gutiérrez (Eds.) *Movement et Clitics: Adult and Child Grammar*. Newcastle: Cambridge Scholars Publishing. 333-360.
- Legendre, Géraldine, Barrière, Isabelle, Goyet, Louise, & Nazzi, Thierry. 2010b. Comprehension of infrequent subject-verb agreement forms: Evidence from French-learning children. *Child Development*, 81, 1859–1875.
- Legendre, Géraldine, Culbertson, Jennifer, Zaroukian, Erin, Hsin, Lisa, Barrière, Isabelle, & Nazzi, Thierry. 2014. Is children's comprehension of subject–verb agreement universally late? Comparative evidence from French, English, and Spanish. *Lingua*, 144, 21–39.
- Lust, Barbara C. 2012. Tracking Universals Requires a Grammatical Mapping Paradigm. Kleanthese K. Grohmann, Aljona Shelkova & Dionysos Zoumpalidis (eds.) *Linguists of Tomorrow: Selected Papers from the First Cyprus Postgraduate Conference in Theoretical and Applied Linguistics*. 105-130. Cambridge Scholars Publishing.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McWhorter, John H. 2005. *Defining Creole*. Oxford, UK: Oxford University Press.
- Meakins, Felicity & Wilmoth, Sasha. 2020. Overabundance resulting from language contact: Complex cell-mates in Gurindji Kriol. Peter Arkadiev & Francesco Gardani (eds.) *Morphological Complexity* (136-162). Oxford University Press.
- Miller, Karen. 2012. Not all children agree: Acquisition of agreement when the input is variable. *Language Learning and Development* 8 (3) 255-277.

- Miller, Karen. 2013. What Sarah reveals about non-agreeing don't and theories of root infinitives. *Language Acquisition*, 20, 305-324.
- Miller, Karen. 2013. Acquisition of variable rules: /s/-lenition in the speech of Chilean Spanish-speaking children and their caregivers. *Language variation and change*, 25, 311-340.
- Miller, Karen & Schmitt, Cristina. 2010. Effects of variable input in the acquisition of plural in two dialects of Spanish. *Lingua*, 120, 1178-1193.
- Miller, Karen & Shmitt, Cristina. 2012. Variable Input and the Acquisition of Plural Morphology, *Language Acquisition*, 19 (3), 223-261.
- Molina, Daniele, Marcilese, Mercedes & Name, Cristina. 2017. Ora está, ora não está: Input Variável e Aquisição Da Flexão Verbal De 3ª Pessoa Do Plural NO PB, *Matraga*, 24 (41), 288- 309.
- Monéreau-Merry, Marie-Michelle. 2017. "Marie Manje la Pom Nan." *Examining the Cognitive Process of Restructuring and Advantage Selection in the Definite Article System in Haitian Creole among U.S. Born Heritage Speakers of Haitian Creole*. Unpublished Doctoral Dissertation, CUNY Graduate Center.
- Naigles, Letitia (2002) Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, 86 (2), 157-199.
- Nazzi Thierry & Bertoncini, Josiane. 2009. Phonetic specificity in early lexical acquisition: New evidence from consonants in coda positions. *Language and Speech*. 52, 463–480.
- Nazzi, Thierry, Kemler Nelson, Deborah G., Jusczyk, Peter W. & Jusczyk, Ann Marie. 2000. Six-month-olds' detection of clauses in continuous speech: Effects of prosodic well-formedness. *Infancy*, 1, 123-147.
- Nevat, Michael, Ullman, Michael T., Eviatar, Zohar & Bitan, Tali. 2018. The role of distributional factors in learning and generalizing affixal plural inflection: An artificial language study. *Language, Cognition & Neuroscience*, 33 (9), 1184- 1204.
- Newkirk-Turner, Brandi L. & Green, Lisa. 2016. Third person singular –s and event marking in child African American English. *Linguistic Variation*, 16 (1), 103-130.
- Nishibayashi, Léo-Lyuki & Nazzi, T. 2016. Vowels then consonants: emergence of a consonant bias in early word form segmentation. *Cognition*, 155, 188-203.
- Noël, Audrey. 2014. La théorie du macrosystème interlectal et l'acquisition langagière en

context plurilingue: réflexion sur la situation réunionnaise. *Actes du Congrès de Linguistique Française*, accessible at: <http://docplayer.fr/59516587-4e-Congrès-Mondial-de-Linguistique-Francaise.html>

Pérez-Leroux, Ana T. 2005. Number problems in children. In C. Gurski (Ed.), *Proceedings of the 2005 Canadian Linguistic Association Annual Conference* (p. 12). Retrieved from <http://ling.uwo.ca/publications/CLA-ACL/CLA-ACL2005.htm>.

Poline, Yerania, Ng, Sally, Aharodnik, Katsiaryna & Barrière, Isabelle. 2019. *Effects of variation and methodology in the acquisition of Subject-Verb agreement by Dominican Spanish-speaking preschoolers*. Council of Undergraduate Research- 2019 Research Experience for Undergraduates Symposium. Marriott Hotel, Arlington, VA.

Phillips, Colin. 1995. Syntax age two: cross-linguistic differences. *MIT working papers in Linguistics*, 26, 37-93.

Rastegar, Zahra, Shirazi, Haghighi, & Sadighi, Firooz. 2012. An amazing conundrum in children's comprehension and production of verb inflection. *World Applied Sciences Journal*, 18(8), 1095-1101.

Roumain, Maryse Noël. 2013. *L'enfant haïtien et le bilinguisme*. Xlibris.

Saint Martin, Weston. 2005. *Les Formes des Pronoms Personnels de l'Haïtien et leur place en comparaison avec celles du français*. Unpublished Master Thesis, University of Georgia.

Scholes, Robert J. 1981. Developmental Comprehension of Third Person Personal Pronouns in English. *Language and Speech*, 24 (1), 91-98.

Smolík, Filip & Bláhová, Veronika. 2017. Comprehension of verb number morphemes in Czech children: Singular and plural show different relations to age and vocabulary. *First Language*, 37 (1), 42-57.

Soderstrom, Melanie, Seidl, Amanda, Kemler Nelson, Deborah G., Jusczyk, Peter W. 2003. The Prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49, 249-267.

Sundara, Megha, Demuth, Katherine, & Kuhl, Patricia K. 2011. Sentence-position effects on children's perception and production of English 3rd person singular –s. *Journal of Speech, Language, and Hearing Research*, 54, 55–71.

Swingley, Daniel. 2008. Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, 60 (2), 252-269.



- Valdman, Albert. 2015. *Haitian Creole: Structure, Variation, Status, Origin*. Bristol, CT: Equinox.
- Valdman, A., Villeneuve, Anne-José & Siegel, Jason F. 2015. On the influence of the standard norm of Haitian Creole on the Cap Haïtien dialect: evidence from sociolinguistic variation in the third person singular pronoun. *Journal of Pidgin and Creole Languages*, 30 (1), 1-43.
- Valian, Virginia, Prasada, Sandeep, & Scarpa, Jodi. 2006. Direct object predictability: effects on young children's imitation of sentences. *Journal of Child Language*, 33, 247–269.
- Varlakosta, Spyridoula, Vainikka, Anne & Rohrbacher, Bernhard. 1998. Functional projections, markedness and “root infinitives” in early child Greek. *The Linguistic Review*, 15 (2-3), 197-208.
- Verhagen, Josie, & Blom, Elma. 2014. Asymmetries in the acquisition of subject–verb agreement in Dutch: Evidence from comprehension and production. *First language*, 34(4), 315-335.
- Vieira, Maria Cristina.P. 2006. *A emergência do padrão flexional de 3ª pessoa do plural na aquisição de PB como L1*. Master Thesis, Federal University of Rio de Janeiro (in Portuguese).
- Von Holzen, Katie, Nishibayashi, Leo-Lyuki, & Nazzi, Thierry. 2018. Consonant and vowel processing in word form segmentation: an infant ERP study. *Brain Sciences*, 8(2), 24.
- Xanthos, Aaris, Laaha, Sabine, Gillis, Steven, Stephany, Ursula, Aksu-Koç, Christofidou, A., Gagarina, Natalia, Hrzica, Gordana, Ketrez, F.Nihan, Kilani-Sloch, Marianne, Korecky-Kröll, Katharina, Kovačević, Melita, Laalo, Klaus, Palmovic, Marijan, Pfeiler, Barbara, Voiekova, Maria D. & Dressler, Wolfgang U. 2011. On the role of morphological richness in the early development of noun and verb inflection. *First Language*, 3 (4), 461-479.

## TABLES

**Table 1** Summary of recent cross-linguistic studies on the Acquisition of Subject-Verb Dependencies involving Agreement Markers and Pronominal Subjects.

	Production	Comprehension	Examples
<b>Brazilian Portuguese</b>	Variable 3 <sup>rd</sup> person plural: Emergence at age 3; productive adult-like use between age 4 and 5 (Vieira, 2006)	5-6: PL>SG both pro-drop and full DP subject conditions, preference for PL stimuli even in SG condition <i>Molina, Marcilese &amp; Name (2017)</i>	2 Conditions : full subject DPs (2 cues) and pro-drop (1 cue) <i>(A criança) comeu doce</i> <i>(As crianças) comeram doce</i>
<b>Czech</b>	—	3-4;9 SG understood between 4-4;9 <i>Smolík &amp; Bláhová, V (2017)</i>	4 Conditions: no subject intransitive and transitive; lexical subject; lexical object. <i>Tady běží/běhají</i> <i>Here runs. SG/run.PL</i>
<b>Dutch</b>	2-3: Accuracy on 3PL and 3SG > comprehension <i>Verhagen &amp; Blom (2014)</i>	2-3: Above chance on 3PL, not 3SG <i>Verhagen &amp; Blom (2014)</i>	<i>Ze aait/aaien het hondje</i> <i>She/they strokes.SG/stroke.PL the dog</i>
<b>(Mainstream American) English</b>	From 2;6 3SG in 90% of obligatory contexts <i>e.g. Brown (1973) De Villiers &amp; De Villiers (1973) Lahey et al., (1992) Miller &amp; Chapman (1981), Radford (1997)</i> From age 3 3SG in 70% of obligatory contexts and from age 4 90% (Rescorla & Roberts, 2002)	5-6 (Johnson et al., 2005)  4-5 Number-marking on subject pronouns understood (above chance), not earlier; more errors on Number than Case or Person until age 7 (Scholes, 1981)	<i>the duckØ swims in the pond</i> <i>the ducks swimØ in the pond</i>  <i>S/he is touching someone</i> <i>They are touching someone</i>
	3;1-5;5 3SG in 70% of obligatory contexts in constrained elicitation task <i>(Barrière et al., 2019)</i>	3;1-5;5 One regular verbal agreement marker cue in sentence final or medial position (Barrière et al, 2019)	<i>the boyØ skips (in the hall)/(freely)</i> <i>the boys skipØ (in the hall)/(freely)</i>
<b>Farsi Persian</b>	4-6: less than 5% agreement errors (including in pro-drop and non-pro-drop constructions) <i>Rastegar et al (2012)</i>	4-6: Non pro-drop SG & PL, pro-drop SG 6: pro-drop SG & PL <i>Rastegar et al (2012)</i>	Pro-drop: <i>Hendoone mikhore</i> <i>Water melon eats.PRES-SG</i> <i>(He) eats watermelon</i>
<b>CHFRENCH</b>	1;8- 3 Subject Clitics/Weak pns + Finite verbs vs DP/Strong pns + Finite and Non-finite <i>Legendre et al. (2010, 2014)</i>	2;6 SG=PL- above chance <u>Real and nonce verbs</u> <i>Legendre et al. (2010)</i> <i>Barrière et al. (2016)</i>	<i>/i.lā.bras/SG vs /i.zā.bras/PL</i> <i>/ilaruv /SG vs /izaruv /PL</i>

<b>Greek</b>	Before 3: emergence, and overuse of 3 <sup>rd</sup> person SG suffix ‘-i’ 3: Productive and accurate use of agreement markers; SG>PL e.g. <i>Doukas &amp; Marinis, (2012)</i> <i>Valakosta et al., (1998)</i>	2: at chance 3 year olds: above chance for both SG and PL, transitive verbs with nonce object <i>Kenanidis et al., (2021)</i>	Pro-drop+ Nonce Object: <i>Δέν-i-SG to káfo</i> <i>Δέν-un-PL to káfo</i>
<b>Mexican Spanish</b>	1;7-4;7 100% <i>Clahsen et al. (2002)</i>  3;6-5;7 Elicited production 88% accuracy; correlated with age; similar performance on SG and PL <i>Hsin et al (2021)</i>	3;5 – 4;2 : PL and SG above chance <i>Gonzalez-Gomez et al. (2017)</i>  3;6- 5;7 PL and SG above chance; no effect of number or position; correlations with age and production performance <i>Hsin et al (2021)</i>	<i>Nada∅ en el charco</i> <i>Nadan en el charco</i> <i>Agarra∅ el objeto</i> <i>Agarran el objeto</i>  <i>Salta∅ en la banca</i> <i>Saltan en la banca</i> <i>En la banca salta∅</i> <i>En la banca saltan</i>
<b>Xhosa</b>	2;0-3;3 >80% <i>Gxilishe et al., 2007</i>	4-5 SG , above chance > PL- at chance <i>Gxilishe et al., 2009</i>	<i>U-nukisa/ oo-nukisa amablomu</i> <i>U-mama u-ya-wa-nlcenkceshela</i>

**Table 2** Summary of studies examining the effects of variation on the acquisition of 3<sup>rd</sup> person Subject-Verb Agreement on children acquiring different varieties of English (adapted from Barrière et al. 2019)

Refs.	Varieties*	Age	Sources of data	Examples of Structures	Results
<b>Johnson et al. (2005)</b>	MAE	3-6	Comprehension	<i>The duck∅ swims in the pond</i> vs <i>the ducks swim∅ in the pond</i>	Evidence of comprehension of b V-s as a sg marker and V-∅ as a plural marker only in MAE aged 6
<b>Johnson (2005)</b>	AAE	4-6		<i>The penguin dresses</i>	Only MAE learners treat -s as generic tense marker and verbal suffix
<b>DeVilliers &amp; Johnson (2007)</b>	MAE & AAE	5-7	Grammaticality judgments		
<b>Miller (2012)</b>	WMCE & WWCE	3;7-5;4	Comprehension & Production	<i>Where does the duck swim</i> (Target)/ <i>the bird live</i> (Control)	Production rate of <i>does</i> WMCE > WWCE For both Target & Control comprehension of WMCE > WWCE
<b>Cleveland &amp; Oetting (2013)</b>	SWE & AAE	6	Production	Irregular 3SG and regular V-s	Rates of V-s/∅: SWE > AAE
<b>Newkirk &amp; Green (2016)</b>	AAE	3-6	Production	Rates of -s/∅ and <i>is/be</i>	Rate of <i>is</i> > -s High rates of <i>be</i> (vs <i>is</i> ) and ∅ (vs <i>s</i> ) across aspectual contexts, high in retelling than constrained elicitation (imitation)
<b>Barrière et al. (2019)</b>	MAE, SMV, STV		Comprehension & Production	Regular V-s (and irregular <i>do/does, were/was</i> )  <i>The boy∅ sleeps (deeply/in the bed)</i>	<u>Production rates</u> of irregular and regular 3 <sup>rd</sup> person singular SVA in constrained elicitation task (DELV, Seymour, 2005): MAE > SMV > STV <u>Comprehension</u> of Regular V-s: MAE understands in three different syntactic contexts; SMV and STV only in one context

\*AAE= African American English, MAE= Mainstream American English, SMV= Some variation; variety that slightly differs from MAE, STV= Strong variation; variety that greatly differs from MAE, SWE= Southern White English, WMCE= White Middle Class English, WWCE= White Working Class English.

**Table 3.** Paradigm of Personal and Possessive Pronouns and Possessive Determiners in Haitian Creole\*

	Person	Subject	Direct Object	Indirect Object	Reflexive	Possessive Pronoun	Possessive Determiners
		SPn	DOPn	IOPn	REF	POSPn	POSDet
<b>SG</b>	1 <sup>st</sup>	mwen/m	mwen/m	mwen/m	mwen menm	pa mwen/m /pan'm	mwen/m
	2 <sup>nd</sup>	ou/w	ou/w	ou/w	ou menm	pa ou/w	ou/w
	3 <sup>rd</sup>	li/l/-l	li/-l	li/l /i	li menm	pa li/-l/l-	li/l/i
<b>PL</b>	1 <sup>st</sup>	nou/n	nou/n	nou/n	nou menm	pa nou/n	nou/n
	2 <sup>nd</sup>						
	3 <sup>rd</sup>	yo/y	yo/y	yo/y	yo menm	pa yo/y	yo/y

\*This table presents the full and contracted pronouns used in HC, the Standard variety used in Port-au-Prince. For variation in phonological contractions used in the Cap Haitian variety, see Valdman , 2015.

**Table 4.** Distribution of the phonological realizations of subject pronouns in the two corpora

	CORPUS 1	CORPUS 2	
	HC RADIO PROGRAM	MOTHER	FATHER
<b>TOTAL Subject Pronouns</b>	343 (100%)	47 (100%)	14 (100%)
<b>1<sup>st</sup> Person Singular MWEN</b>	74	2	6
Full form MWEN	15 (20%)	0 (0%)	0 (0%)
Reduced form M	59 (80%)	2 (100%)	6 (100%)
<b>2nd Person Singular OU</b>	103	41	4
Full Form OU	61 (59%)	26 (63%)	4 (100%)
Reduced form W	42 (41%)	15 (37%)	0 (0%)
<b>3<sup>rd</sup> Person Singular LI</b>	38	1	4
Full form LI	22 (58%)	0 (0%)	3 (75%)
Reduced form L	16 (42%)	1 (100%)	1 (25%)
<b>1<sup>st</sup>/2nd Person Plural NOU</b>	113	3	0
Full form NOU	76 (67%)	3 (100%)	0 (0%)
Reduced form N	37 (33%)	0 (0%)	0 (0%)
<b>3<sup>rd</sup> Person Plural YO</b>	15	0	0
Full form YO	14 (93%)	0(0%)	0 (0%)
Reduced form Y	1 (7%)	0 (0%)	0 (0%)

**Table 5.** Distribution of 3<sup>rd</sup> person singular full and phonological proclitic and enclitics in both corpora

1.Full form phonological contexts					2. Proclitic phonological contexts				3.Enclitic contexts		4.Enclitic/Proc litic contexts	
V coda	Beginning of Phrase <i>li</i> - $C_{\text{onset}}$		$C$ - <i>li</i> - $C_{\text{onset}}$		Beginning of Phrase <i>l(i)</i> - $V_{\text{onset}}$		$C_{\text{coda}}$ - <i>l(i)</i> - $V_{\text{onset}}$		$V_{\text{coda}}$ - <i>l(i)</i> - $C_{\text{onset}}$		$V_{\text{coda}}$ - <i>l(i)</i> - $V_{\text{onset}}$	
	Full	*Re- duced	Full	*Re- duced	Full	Pro- clitic	Full	Pro- clitic	Full	Enclitic	Full	Re- duced
n	16	0	2	0	2	2	1	1	5	9	1	6
%	100	0	100	0	50	50	50	50	36	64	15	85

C= Consonant; V= Vowel; \*Contexts in which reductions are ungrammatical

**Table 6.** Distribution of 3<sup>rd</sup> person plural full and phonological proclitic in the Radio Corpus

1.Full form phonological contexts					2.Proclitic phonological contexts			
V coda	Beginning of Phrase		C/V <sub>coda</sub> -y(o)-C <sub>onset</sub>		Beginning of Phrase		C/V <sub>coda</sub> y(o)-V <sub>onset</sub>	
	y(o)-C <sub>onset</sub>				y(o)-V <sub>onset</sub>			
	Full	*Reduced	Full	*Reduced	Full	Proclitic	Full	Proclitic
n	2	0	11	0	0	0	1	1
%	100	0	100	0	0	0	50	50

C= Consonant; V= Vowel: \*Contexts in which reductions are ungrammatical



**Table 7:** Frequency (Cue Availability) and Semantic Transparency (Cue Reliability) of full and reduced *li-SG* and full and reduced *yo-PL* in the Radio Corpus

	# of tokens	Singular	Plural	Undefined (for number)
SG full and proclitic /l(i)-/	221 words /l-/	31%	0%	67%
		Full Subject Pns 16%		
		Proclitic Subject Pns 5%		
		Post Verbal Object Pns & Post		
		Nominal Dets 10%		
SG enclitic /-l/	75 word /-l/	46%	0%	54% (part of the root of the word)
		Subject Pns 14%		
		Object Pns 32%		
PL full and proclitic /j(o)-/	107 words /j-/	9.8%	59%	31.2%
		Post nominal Poss. Det* 1.8%	Full Subject Pns 17%	
		SG Quantifier 8%	Proclitic Subject Pns 4%	
			Object Pns 30%	
			Definite Det.** 8%	

\* Two instances of the post-nominal possessive determiners (1.8% of words starting with /j-/) were identified with the words *fwa* ‘faith’ and *aparisyon* ‘appearance’; \*\* In all these post-nominal positions, the contextual information favored a plural definite determiner interpretation rather than a possessive determiner ambiguous interpretation.

**Table 8.** Examples of Experimental Stimuli used in the Comprehension Task that tested the comprehension of Subject Pronouns in different phonological (full, proclitic and enclitic) contexts

Condition	Singular	Plural
C1 Full pronoun with verbs that start with vowels	<i>Li ekri</i> ‘s/he writes’	<i>Yo ekri</i> ‘they write’
C2 Proclitic with verbs that start with vowels	<i>L’ekri</i> ‘s/he writes’	<i>Y’ekri</i> ‘they write’
C3 Enclitic (SG is reduced, PL is full) with verbs that start with vowels	<i>E’l ekri</i> ‘And s/he writes’	<i>E yo ekri</i> ‘and they write’
C4 Full pronoun with verbs that start with consonant	<i>Li souri</i> ‘s/he smiles’	<i>Yo souri</i> ‘they smile’
C5 Enclitic (SG is reduced, PL is full) with verbs that start with consonant	<i>E’l souri</i> ‘And s/he smiles’	<i>E yo souri</i> ‘And they smile’

**Table 9.** Results of basic lexical (type token ratio) and grammatical analyses of the transcripts of the Video Elicitation Task (from youngest to oldest children)

Child ID	Age in months	Age in year; months. days	Lexical Analyses			Basic Language Analyses		
			#types	#tokens	TTR	#utterances	#words	% of Verbs
HA19	29.19	2;5.6	4	8	0.50	8	8	0
HA20	29.88	2;5.27	3	10	0.33	10	10	0
HA17	41.39	3;5.12	3	4	0.75	3	4	33.33
HA15	42.37	3;6.11	14	48	0.29	22	44	27.27
HA10	42.64	3;6.19	3	4	0.75	4	4	0
HA16	43.46	3;7.14	42	114	0.36	38	105	34.21
HA6	44.12	3;8.3	27	63	0.42	50	58	22
HA11	50.59	4;2.17	12	45	0.26	23	45	13.04
HA2	51.45	4;3.12	56	133	0.42	26	125	57.69
HA1	51.51	4;3.14	37	88	0.42	29	75	31.034
HA4	52.2	4;4.6	44	115	0.38	32	106	40.62
HA5	53.16	4;5.4	37	74	0.50	41	72	12.19
HA18	54.54	4;6.15	64	215	0.29	39	204	46.15
HA8	57.92	4; 9.27	20	115	0.17	21	115	42.85
HA14	61.24	5;1.5	23	62	0.37	21	53	47.82
HA9	61.8	5;1.22	29	71	0.40	23	59	46.15
HA3	64.33	5;4.10	65	145	0.44	26	120	65.38
HA7	64.5	5;4.15	6	27	0.22	26	27	0
HA12	66.08	5; 6.1	68	165	0.41	27	152	77.77
HA13	69.33	5;9.9	49	117	0.41	26	114	65.38
Totals			557	1506		469	1386	

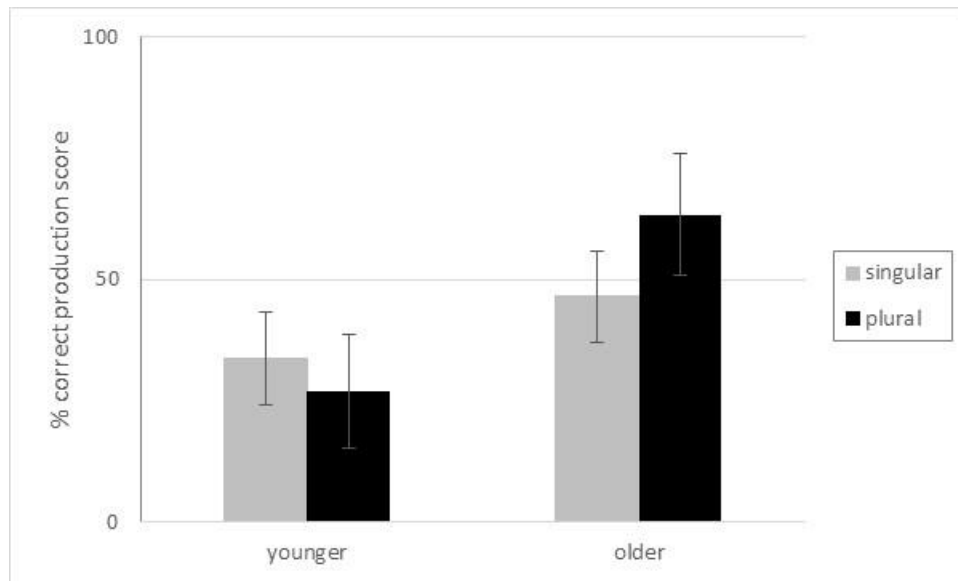
**Table 10.** Distribution of the full and reduced forms of 3<sup>rd</sup> person singular and plural subjects and their mapping onto singular and plural subjects

Part.	Age in months	% of verbs	3 <sup>rd</sup> Person Singular <i>li</i>			% Mapping Score SG	3 <sup>rd</sup> Person Plural <i>yo</i>		
			Full form <i>li</i>	Proclitic <i>l-</i>	Enclitic <i>-l</i>		Full form <i>yo</i>	Pro-clitic <i>y-</i>	% Mapping Score PL
HA19	29.19	0	0	0	0	0	0	0	0
HA20	29.88	0	0	0	0	0	0	0	0
HA17	41.39	33.33	0	0	0	0	0	0	0
HA15	42.37	27.27	19	2	0	<b>48</b>	0	0	<b>0</b>
HA10	42.64	0	0	0	0	0	0	0	0
HA16	43.46	34.21	1	2	2	<b>60</b>	12	0	<b>42</b>
HA6	44.12	22	1	0	0	<b>0</b>	0	4	<b>100</b>
HA11	50.59	13.04	0	2	0	<b>50</b>	17	0	<b>41</b>
HA2	51.45	57.69	2	0	0	<b>100</b>	19	2	<b>47</b>
HA1	51.51	31.034	6	2	0	<b>62</b>	13	1	<b>61</b>
HA4	52.2	40.62	5	14	6	<b>35</b>	0	3	<b>33</b>
HA5	53.16	12.19	0	2	0	<b>50</b>	0	0	<b>0</b>
HA18	54.54	46.15	2	2	0	<b>50</b>	12	10	<b>50</b>
HA8	57.92	42.85	0	0	0	<b>0</b>	0	1	<b>100</b>
HA14	61.24	47.82	0	3	0	<b>67</b>	10	6	<b>44</b>
HA9	61.8	46.15	10	9	0	<b>58</b>	1	1	<b>100</b>
HA3	64.33	65.38	8	13	0	<b>47</b>	2	1	<b>66</b>
HA7	64.5	0	0	0	0	0	0	0	0
HA12	66.08	77.77	1	1	0	<b>50</b>	8	9	<b>47</b>
HA13	69.33	65.38	2	0	0	<b>100</b>	1	1	<b>100</b>
Totals			55	52	8		94	38	

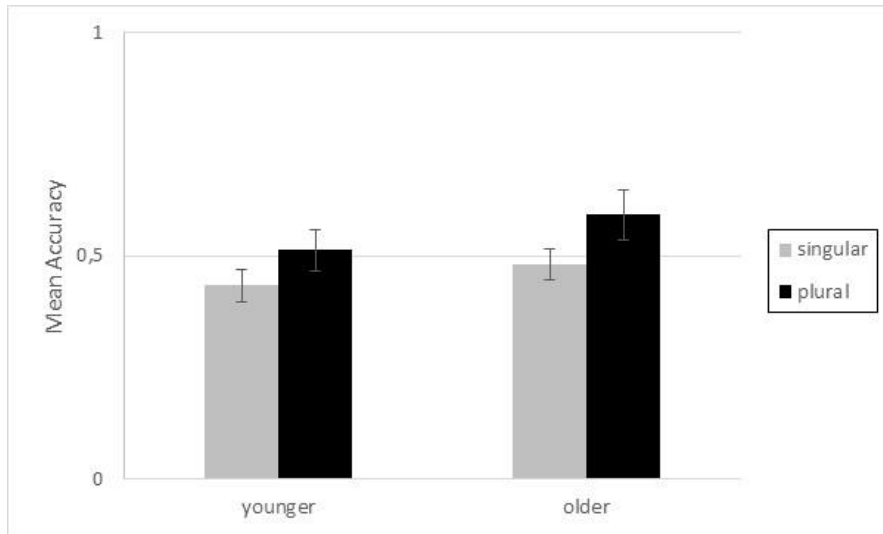
## FIGURES



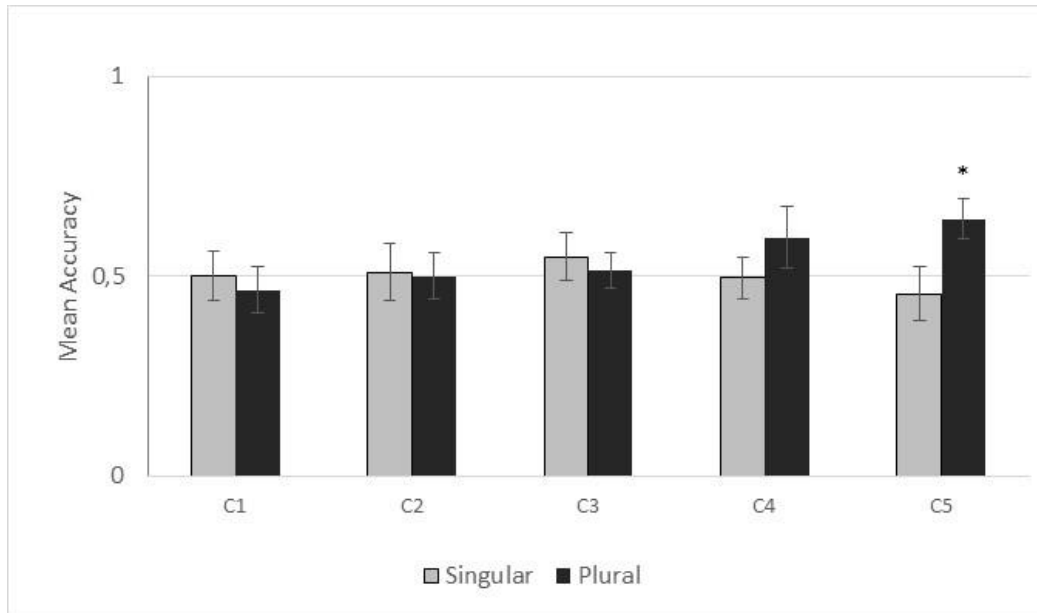
**Figure 1.** Still pictures of Singular (left) and Plural (right) versions of the verb *achte* ‘buy’.



**Figure 2:** Mean correct production mapping scores (and SEs), broken down by Age Group and Number

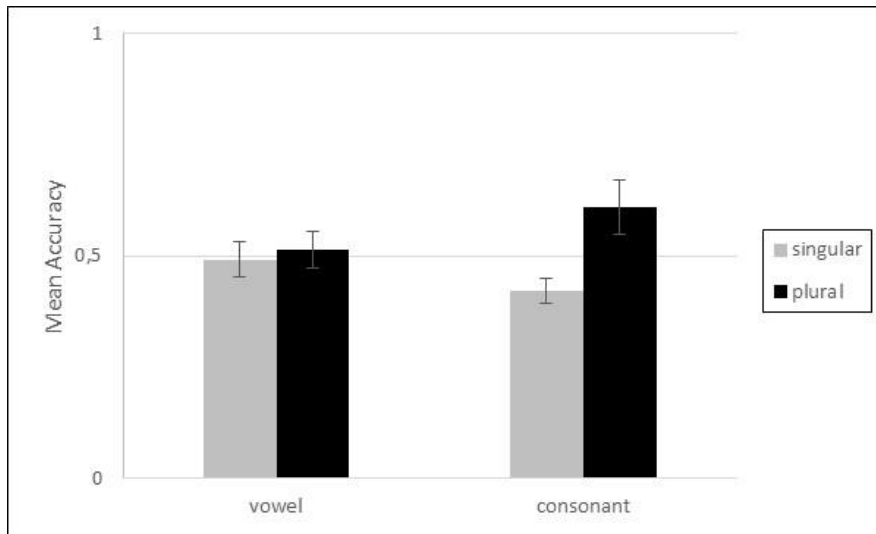


**Figure 3:** Singular and Plural Accuracy percentages across conditions and age groups

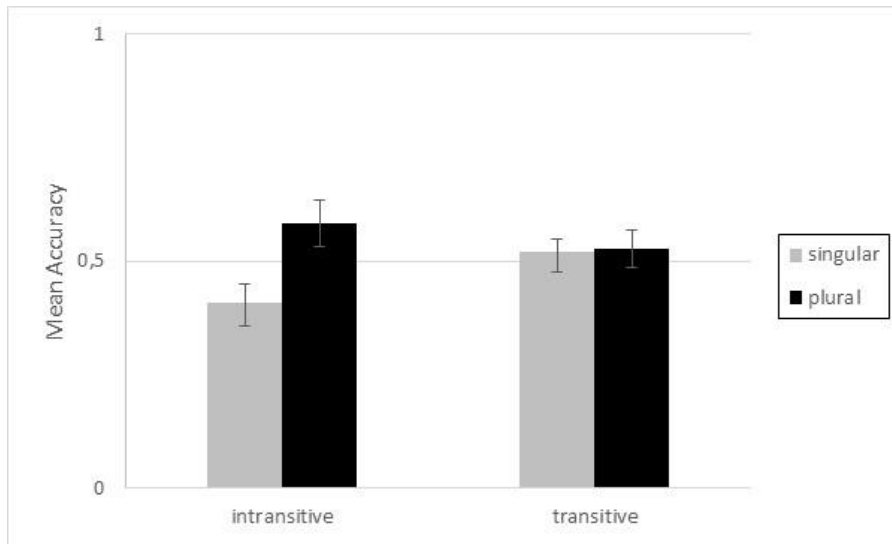


**Figure 4.** Accuracy (and SEs) per condition: C1 Full Singular and Plural Pronouns with verbs that start with vowels; C2 Singular and Plural Proclitics with verbs that start with a vowel; C3 Singular Enclitic and Plural Full Pronoun preceded by conjunction *e* ‘and’ followed by verb that starts with a vowel; C4 Full Singular and Plural Pronouns with verbs that start with Consonant; C5 Singular Enclitic and Plural Full Pronoun preceded by conjunction *e* ‘and’ followed by verb that starts with a consonant.





**Figure 5:** Mean Accuracy (and SEs), broken down by Onset type (Vowel versus Consonant) and Number



**Figure 6:** Mean Accuracy (and SEs), broken down by Verb type (Intransitive versus Transitive) and Number