

Péters, H. (2019). Asymmetric self-repair in gender attribution. In Liesbeth Degand, Gaëtanelle Gilquin, Laurence Meurant & Anne Catherine Simon (eds.) *Fluency and Disfluency across Languages and Language Varieties*. Corpora and Language in Use – Proceedings 4, Louvain-la-Neuve: Presses universitaires de Louvain, 261-281.

Asymmetric self-repair in gender attribution

Hugues Péters

UNSW, Sydney

Abstract

This paper discusses the phenomenon of morphological self-repairs focussing on the determiner system, produced in spontaneous speech, using data from a longitudinal corpus of Jamaican adult second language (L2) learners of French. Our primary concern is to describe – and account for – asymmetries in gender attribution errors and self-correction of the gender of singular definite determiners. The account is grounded in the theories of Distributed Morphology and the Minimalist Program. It is argued that the asymmetries observed reflect the morphological feature configuration assigned by the learners to the determiner system.

1. Introduction

The acquisition of gender attribution is an important topic of L2 research (Prévost & Paradis 2004) that reputedly raises acquisition difficulties for adult learners, particularly those whose native language lacks grammatical gender, such as the case of native speakers of Jamaican English and Jamaican Creole.¹ This paper proposes a new approach by examining oral production from

¹ I am grateful for the comments received from the audience at the (Dis)fluency conference in Louvain-la-Neuve 15-17 February 2017 and the useful comments received from two anonymous referees that have greatly improved the quality of the paper. Thanks to Marcus Taft for advice on statistics. All remaining errors are mine.

a population of Jamaican second language (L2) learners of French and by the examination of their self-correction behaviour in spontaneous speech. More precisely, this chapter discusses the phenomenon of morphological self-repairs focussing on the determiner system. Our primary concern is to describe – and account for – unexpected asymmetries in self-correction of the gender of singular definite determiners, part of which is called “non-systematic agreement due to monitoring” processes in Dewaele & Véronique (2001), that is, inconsistencies in gender agreement within the noun phrase caused by self-correction. The study is based on careful analysis of empirical data from a longitudinal oral learner corpus of 10 Jamaican learners of French at the University of the West Indies (UWI), Mona (Péters 2014, 2017). It is therefore intended to partially bridge the gap that separates corpus methodology from theoretical syntax, in the form of the now current coupling of the Minimalist Program (MP) of inquiry developed by Noam Chomsky (1995, 2001) with the Distributed Morphology (DM) approach to the lexicon (Halle & Marantz 1993).

Here is an example, from our corpus showing the phenomenon of self-repair:

- (1) oui &-euh je ne pense que les Israelis@s a le [//] la [/] <la droit> [//] le droit &-euh [1821_Conv]
- ‘yes (filler) I neg think that the Israelis have the_{+masc} the_{+fem} the_{+fem} right the_{+masc} right (filler) [...]

In (1), we observe that L2 learners use fillers (&-euh) and various forms of retracing such as repetitions (noted [/]) and repairs (noted [//]) (see Hilton 2014, Péters 2017a for a detailed analysis of repair disfluency). In usual studies on the attribution and agreement of gender, only the right-most nominal phrase *le droit* (the right) is considered and judged target-like. One can immediately observe that such approach ignores the processing information included in the complex sequence of repair *le la la droit le droit*. In a case of self-repair, such as *le [//] la*, the original form (Reparandum) can be said to represent the unconscious internal grammar of the learner, and the repairing form (Reparans) to represent the result of monitoring processes that require conscious attentional control from the speaker.

Overall, this research aims to provide an empirical argument that systematic repair disfluency reveals the underlying feature composition characterizing the internal grammar constructed by learners when acquiring an L2. A similar argument has been made by Levelt (1989) about native speakers’ monitoring. Just like native speakers, L2 learners produce disfluencies during their spontaneous oral performance, and if their interlanguage has a systematic grammar constrained by the faculty of language, we hypothesize that the patterns of dis-

fluencies produced depend on the state of the unconscious L2 interlanguage as much as on the state of the conscious Monitor (Krashen 1982, Kormos 1999). Discovering what is repaired and how it is repaired provides a window on both their implicit and explicit grammatical knowledge: the morpho-syntactic distinctions they implicitly make, as well as the explicit rules they employ to monitor their speech. Such quantitative and qualitative research can have benefits not only for a better understanding of the conscious monitoring processes and unconscious grammatical computations, but also to inform the pedagogy of French as a second or foreign language (Granger *et al.* 2002, Whong *et al.* 2013).

In the following sections, I will present a theory of the morphosyntax of gender and self-repairs in French. After a brief review of previous studies on gender attribution and agreement in L2 French, I will detail the methodology of this research, the results of the investigation, and a brief discussion of the results.

2. Theory

2.1. The morphosyntax of Gender in French

As is well known, French distinguishes two genders: masculine and feminine. Gender is a lexical feature of the noun that must be learned for each noun with a focus on the choice of the preceding determiner (Carroll 1989). Noun gender is generally not deducible from the meaning of the word (except with some gendered animate beings), and, even though the morphophonological endings provide some clues towards assigning gender, such combination of semantic and formal principles makes the French gender attribution system opaque (Corbett 1991). This opacity makes the task of acquiring gender particularly difficult for learners of L2 French as it relies principally on the observation of agreement patterns between nouns and their surrounding linguistic context (see Prévost & Paradis (2004), Prévost (2009), for an overview of research on gender in L1 and L2). Across Romance languages, however, the masculine gender is the default gender for nouns: indeed, there are more masculine nouns than feminine ones, masculine is the gender of loan words: *le lavabo* (the sink) and category shifts: *le je-ne-sais-quoi* (the je ne sais quoi), it is traditionally used as a generic to include the feminine: *les étudiants* (male and female students) (Harris 1991, Wise 1997, Hulk & Tellier 1999). All nominals in argument positions must be modified by a determiner. Modifiers (determiners and adjectives) agree in gender with the noun, but gender opposition in determiners only exists in the singular (*le* vs. *la* (the)).

Considering that the features of Definiteness, Gender, and Number are binary (+/-), Svenonius (2012) proposes the following featural composition of determiners in French:

- (2) a. *un* (a) < D [-def, -fem, -pl] > ↔ /œ̃/
- b. *une* (a) < D [-def, +fem, -pl] > ↔ /yn/
- c. *le* (the) < D [+def, -pl] > ↔ /lə/
- d. *la* (the) < D [+def, +fem, -pl] > ↔ /la/
- e. *les* (the) < D [+def, +pl] > ↔ /le(z)/
- f. *des* (some) < D [-def, +pl] > ↔ /de(z)/

Crucially for our analysis, the ‘masculine’ singular definite form *le* in (2.c) remains unmarked for Gender. As shown in Table 1, the feminine determiner *la* - used exclusively with feminine nouns - is the only definite form marked for [+fem]. In all other singular contexts - with masculine nouns, but also with feminine nouns when the definite determiner is immediately followed by a vowel in contexts of elision - the default ‘genderless’ form *le* is used. The *le* form must be ‘genderless’ or its use would cause an agreement mismatch when modifying a feminine noun in elision contexts as in *l’ancienne table*.² As Svenonius (2012) argues, *le* is the only candidate for elision before vowel (*le* > *l’*), as such elision phenomenon only occurs with schwas in French (*me, te, se, ne, que, le, ce, jusque*).

Definite determiners	Masculine noun	Feminine noun
+ a consonant	LE bureau (the desk)	LA table (the table) [+fem]
+ a vowel (with elision)	LE > l’ancien bureau (the old desk)	LE > l’ancienne table (the old table)

Table 1. Singular definite determiners

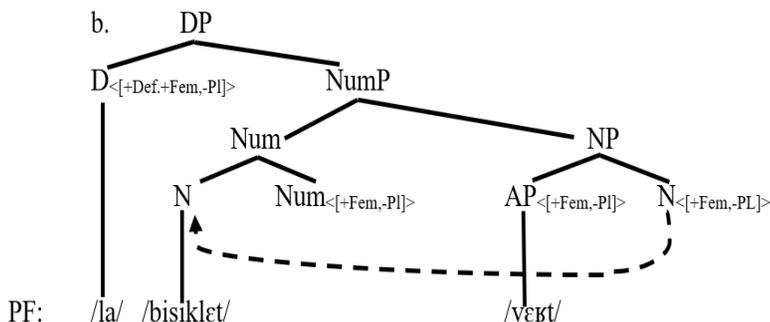
Within the Minimalist Program, a sentence is constructed from the bottom-up by the operation ‘Merge’ applying to lexical items. Lexical items consist exclusively of formal and semantic feature-bundles (devoid of phonological features). At given points in the derivation (known as ‘phases’), the syntactic derivation splits thanks to the operation ‘Spell Out’: one branch is sent towards the interface of ‘Phonetic Form’ (PF) to receive a phonetic ‘interpre-

² The ‘masculine’ form is also the default choice for the singular possessive determiners: *mon/ton/son* (my, your, his/her/its) used before vowels: *mon ancienne table* (my old table) vs. *ma table* (my table).

tation', while the other branch continues its derivation towards the interface of 'Logical Form' (LF). Coupling the Minimalist Program with the theory of Distributed Morphology as is customary since Chomsky (2001), it is assumed that all phonetic features are introduced post-syntactically: Vocabulary Items endowed with a phonetic shape are inserted 'late' in the derivation, at PF (Halle & Marantz 1993).

Simplifying current representations of the syntax of nominals, three layers can be distinguished: the Nominal Phrase (NP), inherently endowed with Formal Features (FF) of Gender and Number, constitutes the lexical core carrying semantic features such as 'artefact', 'mode of transport', etc.; the Number Phrase (NumP) mediates agreement in gender and number of modifiers with the noun; the Determiner Phrase (DP) encodes features of definiteness and specificity, and agrees with the noun. The FFs of NumP are said to be 'strong' in French, and to overtly attract the noun for agreement (while agreement is covert in English), as shown in (3.a) by the position of adnominal adjectives compared to the English translation.

- (3) a. La bicyclette verte.
 The_{+def,+fem,-pl} bicycle_{+fem,-pl} green_{+fem,-pl}
 'The green bicycle.'



As the DP constitutes a phase, the syntactic structure already built is sent to PF for phonetic interpretation: the valued features present in the syntactic derivation are matched with the most adequate Vocabulary Items. In Distributed Morphology, Vocabulary Items (VI) of functional categories compete for insertion at PF: so, all forms in (2.a-f) are competing for insertion in the structure to become the morphophonological realization of D, in (3.b). Forms with unmatched features are immediately eliminated: as D in our example is syntactically specified as [+Definite] and [-Plural], all indefinite and plural forms in (2) are discarded because they would cause an agreement mismatch leading to

ungrammaticality. VIs, however, can remain ‘underspecified’: for instance, the plural determiner *les* is not specified for Gender, yet it can match the Formal Features (FF) of both Masculine or Feminine Plural Determiners. Similarly, in our example, both forms *la* (specified as [+fem]) and *le* (unmarked for gender) are suitable for insertion in D as none would cause an agreement mismatch with the FFs of the noun. In the normal course of events, when several candidates are suitable for insertion, a competition rule selects the one that has the most matching features: so, even though *le* [+def, -pl] constitutes a proper subset of the required FFs, it would not normally be inserted in (3), because there is a better candidate available *la* [+def, +fem, -pl] with more matching features. This explains why failure of gender agreement is unusual in native speaker grammars. However, as White *et al.* (2004: 128) observe, a relaxation of the competition rule may occur under communicative pressure even with native speakers, and with L2 learners.

I hypothesize that the variation in strength of the cognitive link between a given noun and the Gender node (Prodeau 2005), and the communicative pressure to insert a determiner in all argument positions (White *et al.* 2004) will lead L2 learners to resort to the default ‘genderless’ determiner *le*, as a kind of filler to salvage the structure and pursue the communication, as the default does not cause the syntactic derivation to crash. Inversely, I hypothesize that learners will be conservative with respect to the insertion of the gendered determiner *la*, inserting it in D only when they are certain that it corresponds to the Gender of the noun. Therefore, the default form is expected to be overgenerated and, consequently, have a lower accuracy rate than the ‘gendered’ one, used only with feminine nouns. We expect our data to show an asymmetry in the accuracy of gender agreement leading to increased non-target use with feminine nouns. This asymmetry reflects the acquisition of the gender specification of determiners.

2.2. The morphosyntax of repair in French

Just like native speakers, L2 learners’ spontaneous speech contains pauses, fillers, and various forms of retracing, as well as strategies to cope with communicative gaps (Götz 2013). Per current models of L2 speech production adapted from Levelt’s (1989) model, the abundance of disfluencies in non-native speech is accounted for by a lack of automatization of the processes of lexical retrieval and syntactic production (Kormos 2006). For learners, therefore, utterance production involves the conscious arrangement of individual lexemes into syntactic units, a laborious process that rapidly overloads the working memory. As suggested by Temple (2000), self-repairs of L2 learners differ not only

quantitatively, but also qualitatively from those of native speakers: the former, with variations depending on their proficiency, on the nature of the task, or the formality of the situation, make many corrections of syntax and morphology, while the latter focus on meaning, lexis or register (see Pallaud *et al.* (2013) for a morphological analysis of retracing in French). I will define repair as immediate correction of a single morphological or phonological feature. Van Hest (1996) and Kormos (1998) provide taxonomies of repairs in L2 speech production. As suggested by Levelt (1989) for native speakers, I assume that the underlying grammar of the learners will influence the types of repairs they produce.

Investigating self-repair by native French speakers, Fornel & Marandin (1996) adopt Levelt's (1983) claim of a syntactic well-formedness constraint on self-repairs in L1, but they argue against theories that equate (constraints on) repairs to (those of) coordination (Blanche-Benveniste 1987, Levelt 1983, 1989, Candea 2000, for an evaluation). They propose the *principe du bord droit* (the principle of the right-hand side). This principle assumes that the architecture of the syntactic tree of an utterance predicts the types of repairs that a native speaker can produce: the repairing form (Reparans) in a repair structure, must be a valid constituent of the original (O) interrupted syntactic plan (Reparandum), and can only attach on the right-hand side of the tree diagram to some structure already built, that is, the Reparans can attach to any node of O if it is the rightmost one.

Without going into details, here is one example of an interrupted utterance, in (4), with its tree diagram representation and the well-formed repairs in (5), translating the original terminology from the authors (Fornel & Marandin 1996: 194):

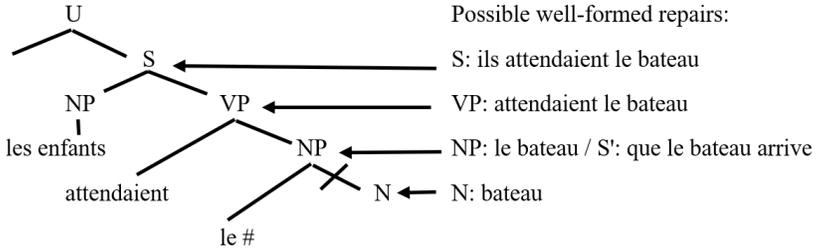
(4) Les enfants attendaient le #.³

'The children were waiting for the #'

The utterance (U) can potentially be repaired at the level of the Sentence (S): *ils attendaient le bateau* (they were waiting for the boat), the Verb Phrase (VP): *attendaient le bateau* (were waiting for the boat), the Noun Phrase (NP): *le bateau* (the boat), or by inserting a complement clause (S') of the verb: *que le bateau arrive* (that the boat arrive/for the boat to arrive), or even the missing Noun (N), preceded by an interruption: *bateau* (boat).

³ The symbol # marks an interruption.

(5)

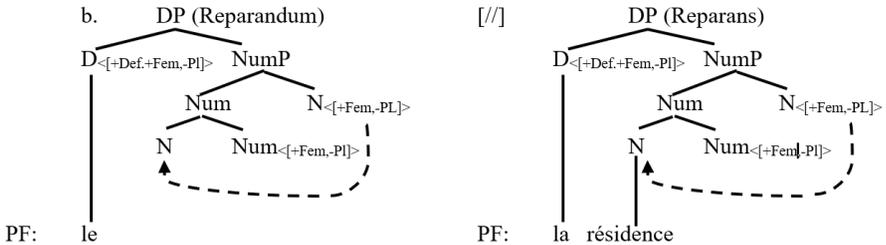


In sequences of successive repairs, each new repair is hooked on the immediately preceding state of the repaired syntactic tree diagram.⁴ I adopt the philosophy of this proposal, adapted to the current syntactic framework (derivational, bottom-up, late insertion) to deal with repairs produced by L2 French learners.

Limiting ourselves exclusively to repairs focussing on a modification of the Gender feature of determiners, and within our morphosyntactic framework characterized by late insertion of Vocabulary Items, the Reparandum is hypothesized to have the underlying syntactic structure of a full DP whether the noun is phonetically realized or not. So, the following repair sequence in (6), extracted from our learner corpus data, will be represented as in (6.b), that is, will have the same syntactic structure as a repair with an overt noun of the type: *<le question>* [//] *la question* [1222_Ex] (the_{+masc} question_{+fem} the_{+fem} question_{+fem}).

(6) a. sur le [//] la résidence [1231_Conv : line 57]

‘in the_{+masc} the_{+fem} residence_{+fem}’



This hypothesis is, of course, speculative, as it would be impossible to prove that the missing noun the speaker had in mind in the Reparandum was the same as the one uttered in the Reparans (*la résidence*). I however consider that, with

⁴ An evaluator points out that, following Shriberg (1999), it is not true that each new repair is hooked on the preceding one. It is important to analyse each example of repair in the flow of its phonetic and conversational context to eliminate cases of false starts rather than immediate correction of a morphosyntactic feature.

a qualitative analysis of the whole context of the retracing sequence, it is a reasonable working hypothesis to make to observe the learners' behaviour. So, the original DP produced by the computational system contains a phonetically empty noun, yet entirely specified in syntax with Formal Features associated with the noun lexeme and stable semantic characterisation. In fact, the learner 'knows' the gender of the noun as she has correctly produced the DP *la résidence* a few lines before (1.33) within the same conversation. But, under communicative pressure, she reverted to the default definite determiner. The learner immediately notices the error and halts the production of the noun. She then initiates a self-correction mechanism with a repaired DP, attached to the right-hand side of the already completed derivation. It is important to know, that modulo a relaxation of the competition of functional items, both DPs converge in syntax proper, even though only the second one conforms to the standard of accuracy in the target language and receives a complete phonetic interpretation at PF. The default definite determiner seems to constitute a form of filler, a dummy morpheme, introduced in the structure to fill in a gap in the structure, retaining only the feature of definiteness, essential for semantic interpretation.

Therefore, when experiencing difficulties in the process of retrieving the Vocabulary Item of a nominal lexeme or remembering the gender, I hypothesize that learners will tentatively revert to the default singular definite determiner *le*, eventually repairing it with the gendered one if need be, when becoming aware of the mistake. So, again there will be an asymmetry: whether the noun in the Reparandum is overtly pronounced or not, as a typical repair sequence will go from the default to the gendered determiner as opposed to the reverse direction. Such asymmetry presupposes that the learners have the correct featural configuration of determiners.

- (7) Default Determiner *le* (+ NP) [//] Gendered Determiner *la* + NP.

3. Earlier research

Bartning (2000) in a study of Swedish learners, and Dewaele & Véronique (2001) in a study of Flemish learners, have shown that singular feminine definite determiners are used more accurately than masculine ones with advanced level learners: most errors of gender attribution to a noun will be of a masculine determiner with a feminine noun, and not vice versa. They also found that indefinite determiners are acquired later than definite ones, and that gender attribution with indefinites seems to be random. Granfeldt (2004) proposes that the definite form *le* (the) and the indefinite form *une* (a) are used as default at

the initial stages of L2 acquisition. He wonders whether *le* cliticizes to the noun or remains an independent phonetic form in the interlanguage of learners.

In their study, Dewaele & Véronique (2001) offer a detailed description of distinct types of errors. They distinguish (1) “Non-systematic gender agreement due to monitoring” (Dewaele & Véronique 2001: 287) which describes the conscious discovery by the learner of a gender error, followed by immediate self-repair: such repair can either occur ‘covertly’, that is, before the actual production of the noun: *le la politique* (the_{+masc} the_{+fem} politics_{+fem}), or ‘overtly’, that is, after the production of the noun: *le mentalité est très bonne* (the_{+masc} mentality_{+fem} is very good_{+fem}) and accounts for 7.3% of errors in their corpus; and (2) “Apparent gender errors” also referred to as “Tip of the Tongue state” (Vigliocco *et al.* 1997): *il y a une, un centre medical* (there is a_{+fem} a_{+masc} medical centre). They convincingly argue that, in the latter case, the speaker was probably looking for the noun *clinique* (clinic_{+fem}), but unable to retrieve it, halts the production process and activates an entirely new lemma: *centre* (center_{+masc}): “What looks like a gender error might therefore only be an aborted attempt at retrieving a particular lexeme” (Dewaele & Véronique 2001: 288). These cases represent 4.2% of the total number of gender errors in their corpus. In the present investigation, this distinction is crucial as I am only interested in the first type that constitutes repair of the gender of the determiner while the semantic and formal features of the noun remain constant.

4. Methodology

The data originates from the UWI learner corpus: a longitudinal corpus of oral productions by 10 Jamaican learners of French (Péters 2014, 2017). 67 interviews took place with students over four successive semesters of the B.A. at the University of the West Indies, Mona: from Year 1 semester 2 (noted as level 12) to Year 3 semester 1 (noted as level 31): these were either formal interviews taking place during end-of-semester oral exams (noted *_Ex*), or informal one-on-one interviews with the investigator (noted *_Conv*), consisting of various activities (conversation, cartoon story retelling, elicitation tasks, and role play). This corpus of more than 50,000 words has been transcribed orthographically, annotated, and encoded using the CHILDES conventions (MacWhinney 2000). Each line of learner’s speech is coupled with a morphological decomposition line, pruned of fillers and retracing, and linked with the corresponding audio-file in electronic format. The corpus is now partially available in Talkbank (Péters 2017b).

Each participant, all early bilingual natives of Jamaican Creole and Jamaican English in a diglossic context, was assigned a two-number code. They had varied educational backgrounds in French: from four years (17, 18, 38) to six years of L2 French in high school (14, 20), while others had studied French through an intensive one-year programme (08, 12, 16, 33), or a mix of both (31). None had been immersed in a French speaking country. All (except 17, 18) had studied Spanish in high school. The measure of Vocabulary Diversity (VocD), calculated at level 21 and level 31 illustrates this variation in proficiency levels among learners (although other measures of grammatical accuracy and complexity would be necessary to give a more precise description).⁵

Learners	08	12	14	16	17	18	20	31	33	38
Level21	60.55	63.95	66.67	48.65	66.94	36.01	60.28	40.65	67.46	52.99
Level31	66.17	63.57	67.13	47.03	82.22	55.17	80.14	46.75	68.85	61.25

*Table 2. Measure of lexical diversity (VocD)
calculated on the morphological tier*

In Table 2, we notice that learners 16, 18, 31 systematically have the lowest rates of lexical diversity at both levels, and learners 17 and 20 reach the highest levels of lexical diversity at level 31, with everyone else in between.

For the present research, I have first investigated gender agreement of singular definite determiners with the noun. Both target and non-target forms were collected from all learners at the highest level (at level 3.1) in structures of “D[+def] (+Adj) +N (+Adj)”. Repaired forms have been omitted and only forms in which the determiner is followed by a consonant have been included (so, the elided *l’* is omitted). Fused forms *au* (at the), *du* (of the) have also been omitted (Embick 2007). I have given the token frequency of target-like use over total use. To evaluate whether gender accuracy is determined by the choice of determiner, a t-test was administered, the probability value (p-value) has been given and the 5% level of significance considered.

For the investigation of self-repair structures, all forms of self-repair (noted as [//] in the corpus) focussed exclusively on the correction of the Gender of the determiner have been collected from all learners, at all levels, because of the low number. However, after careful examination of the data on repair in the context of the whole conversation, I have excluded from the count any instance of “Apparent gender error” (Dewaele & Véronique 2001) such as:

⁵ VocD is a statistical program used in CLAN to measure the richness of vocabulary (McKee *et al.* 2000).

- (8) il [/] il (.) assied sur la plage &*INV:hmhm sur le [/] le [//] (.) la plage, oui, &euh dans la serviette or@s sa serviette de bain [1831_Conv]

‘he he (pause) sits on the_{+fem} beach_{+fem} (back channel) on the_{+masc} the_{+masc} (pause) the beach, yes, (filler) in the_{+fem} towel_{+fem} or (in English) his bath towel.’

In a cartoon story retelling activity, the participant is describing the scene happening on a beach, of a person sunbathing sitting on his towel (*une serviette de bain*) in the sand. The learners, probably unable to retrieve the proper lexeme for ‘sand’ (*le sable*) in the first attempt, marks a short pause, and reverts to the safe word, *la plage* (the beach) she had previously uttered within the same utterance, and immediately after shifts to the target expression *la serviette* (the towel). Therefore, (8) should not be included as a case of self-repair, but more accurately as a case of interruption (with pause) and restart anew.

Finally, I have included data from all learners at all levels on self-repair structures in which a singular definite determiner *le* or *la* is replaced by the elided *l’* when immediately followed by a vowel. However, I have excluded from the count sequences like *la [//] l’ attention* (the attention) because it is impossible to know whether *la* is actually a form of the determiner or a form of elision followed by an interrupted word: *l’ a(ttention)*.

5. The acquisition of gender with determiners: accuracy and self-repair

5.1. Accuracy of the singular definite determiner

Data from the UWI corpus at level 31 confirms previous generalisations found in Bartning (2000) and Dewaele & Véronique (2001): as shown in Table 3, with singular definite determiners (*le*, *la*) in structures of “D [+def] (+Adj) +N (+Adj)”, feminine gender attribution by the determiner to the noun is 93.98% accurate as opposed to an 82.58% accuracy for masculine gender attribution. If we take a 90% accuracy rate as indicative of acquisition, all learners, except 14 and 33, are at that level for feminine determiners (with 3 learners showing 100% accuracy), while only one learners, 12, reaches a 90% accuracy rate for masculine determiners. Note that there are more instances of the feminine determiner (216 tokens) than the masculine one (201 tokens). A t-test shows the high significance of Gender in determiner choice ($t(9)=3.65$, $p=0.0053$).

Learners	08	12	14	16	17	18	20	31	33	38	Total
LA: target forms/total forms	17/18	25/26	24/27	9/10	24/26	29/31	18/18	18/18	26/29	13/13	203/216
Percentage of correct forms	94.44	96.15	88.89	90	92.31	93.55	100	100	89.65	100	93.98
LE: target forms/total forms	14/17	15/16	16/20	20/23	24/28	11/13	8/9	18/27	13/15	27/33	166/201
Percentage of correct forms	82.35	93.75	80	86.96	85.71	84.61	88.89	66.67	86.67	81.82	82.58

Table 3. Accuracy in Gender attribution of singular definite determiners + consonant, at level 31

This asymmetry cannot be explained simply by saying that learners would somehow be more knowledgeable of the gender of masculine nouns than feminine ones. The hypothesis that the asymmetry depends on the morphological properties of the unmarked form *le* overused with feminine nouns in case of doubt by the learners provides a natural explanation to the data.

5.2. Self-repair of the singular definite determiner

Our data on self-repair sheds new light on the issue of noun gender in L2 French. The first set of data, presented in Table 4, shows that in cases of self-repairs focussed on the gender of the determiner, self-corrections of the type: *le* [//] *la* are observed more frequently (23 instances) than repairs of the type: *la* [//] *le* (4 cases, and see below for an explanation of their use), such that sequences of the type: *le* [//] *la résidence* (the the_{+fem} residence_{+fem}) or <*le question*> [//] *la question* (the_{+masc} question_{+fem} the_{+fem} question_{+fem}) are almost six times more frequent than the type: *la* [//] *le cours* (the_{+fem} the_{+masc} course_{+masc}) or <*la droit*> [//] *le droit* (the_{+fem} right_{+masc} the_{+masc} right_{+masc}). Furthermore, if the former *le* to *la* repair is produced in some measure by seven learners, the latter type *la* to *le* is only produced by three learners: the lowest proficiency ones (16, 18), as measured by VocD, and by 38. I would like to propose that the *la* to *le* alternation not only violates the idea that the gendered determiner is to be used sparingly, but also should be considered ungrammatical in the L2 interlanguage because it creates a feature mismatch, and we will explain the apparent exceptions.⁶

⁶ When the learner has a difficulty in retrieving the Vocabulary Item associated with a lexical noun, the strategy employed by learners is usually to repeat the feminine determiner: *la* [//] *la présence* (the_{+fem} the presence_{+fem}).

Learners	08	12	14	16	17	18	20	31	33	38	Total
LE (+NP) [/]											
LA +NP :											
*[-Fem] N	-	-	-	-	-	*3	-	-	-	-	*3 M
[+Fem] N	6	4	-	1	2	4	2	-	-	1	20 F
Total (<i>le la</i>)	6	4	0	1	2	7	2	0	0	1	23
LA (+NP) [/]											
LE +NP :											
[-Fem] N	-	-	-	1	-	2	-	-	-	1	4 M
*+[Fem] N	-	-	-	-	-	-	-	-	-	-	0 F
Total (<i>la le</i>)	0	0	0	1	0	2	0	0	0	1	4

Table 4. Self-repair of definite articles, per learner at all levels

Note that the repair, initiated by the conscious monitoring system, almost always produces a target-like correction. As seen in Table 4, there are however three instances of inaccurate repairs produced by one learner, 18 (that is, three cases of incorrect repair of a form that was originally target-like):

- (9) a. **le** [//] **la** [/] <**la droit**> [...] [1821_Conv]
 ‘the_{+masc} the_{+fem} the_{+fem} right_{+masc}’
 b. **le** [//] **la développement** social [1822_Ex]
 ‘the_{+masc} the_{+fem} development_{+masc} social’
 c. **le** [//] **la ministère** de santé [1831_Conv]
 ‘the_{+masc} the_{+fem} ministry_{+masc} of health’

This time, the derivation must be considered from the point of view of the internal individual grammar of the learner inserting what she mistakenly believes to be the accurate gender attribution.

Let us now examine, in detail, the four exceptional cases of self-correction from the feminine to the masculine determiner found in the corpus.

- (10) a. <parce_que le> [//] parce_que **la** [//] **le cours** [...] [1622_Ex]
 ‘because the_{+masc} because the_{+fem} the_{+masc} course [...]’
 b. **le** [//] **la** [/] <**la droit**> [//] **le droit** &-euh [1821_Conv] (9.a)
 ‘the_{+masc} the_{+fem} the_{+fem} right_{+masc} the_{+masc} right_{+masc} (filler)’
 c. (.) &euh (.) **la** [/] **la** [//] <**le tourisme**> [//] &euh (..) tourisme
 [...] [1812_Ex]

- ‘(pause filler pause) the_{+fem} the_{+fem} the tourism (filler pause) tourism [...]’
- d. j’aime **la** [/] <**la travail**> [//] **le** [/] **le** [/] **le travail** [...] [3831_ Conv]
- ‘I like the_{+fem} the_{+fem} work_{+masc} the_{+masc} the_{+masc} the_{+masc} work_{+masc} [...]’

For (10.a-b), the whole genealogy of the hesitation leading to the unexpected “*la le* + NP” repair reveals that it is, in fact, part of a larger back-and-forth between gendered determiners (*le, la, le...*) happening before the actual production of the lexical item, that is, before Vocabulary Insertion of the noun at PF. In these examples, the learners, unsure of the gender of the noun, are in the process of mentally testing hypotheses, and they start the series of repairs with the default determiner, which, in fact, is akin to the ‘normal’ *le to la* case at the initial stage.⁷

The sentences (10.c-d) represent violations of the tendency towards *le to la* repairs, as they indeed manifest examples of an incorrect singular feminine definite *la* introduced first, repeated once, and then repaired into *le* by the Monitoring process. Here, as the noun is ultimately attributed the target-like masculine gender, the feminine determiner inserted in the syntactic structure should cause an agreement mismatch with the formal features of the noun leading to ungrammaticality: apart from considering these examples as pure performance mistakes, it seems that the learners are testing hypotheses on the Gender node to be associated with the noun and make the determiner agree back and forth with the noun accordingly within the syntactic structure.

It is important to note, however, that sentence (10.c) contains the added difficulty for learner 18 of finding the means to express genericity. The correct DP *le tourisme* (the tourism) is indeed subsequently replaced by the bare generic form: *tourisme* (tourism). This structure is probably an instance of interference from Jamaican English/Jamaican Creole to L2 French as the very English-like pronunciation of the word *tourisme* would suggest. A few lines earlier within the same conversation, the students had already produced non-target like generic bare DPs, this time with no attempt to introduce a determiner, as a word-for-word direct translation from their native languages:

- (11) mais (.) tourisme est nécessaire pour Jamaïque [1831_Ex]
 ‘but (pause) tourism is necessary for Jamaica’

⁷ There is probably a confusion between: *le cours* and *la course* and between *le droit* and *la droite* often observed in learners’ data.

(11) is ungrammatical as bare DPs are not grammatical in argument positions in French. The added cognitive load of trying to express genericity, manifested in (10.c) in the use of fillers and silent pauses, might simply have distracted the learners from the expression of gender.

5.3. Self-repair of the singular definite determiner to the elided form *l'*

A second unexpected and hidden asymmetry, which indirectly confirms the first one, and would have remained mysterious without the hypothesis that the ‘masculine’ determiner is underspecified for Gender, is the asymmetry observed in repairs from the articles *le* and *la* to the elided form *l'* by cliticisation onto the following word starting with a vowel.

From definite to elided	+ Masculine noun	+ Feminine noun	Total
LE [//] L' + vowel	20	12	32
LA [//] L' + vowel	0	5	5
Total	20	17	37

Table 5. Self-repair of singular definite determiners (*le*, *la*) into an elided form *l'*

When observing Table 5 and considering whether the gender of the original choice of determiner was target-like or not, based on the gender of the noun, we notice twenty cases of self-repairs of *le* into an implicit elided *le* followed by a masculine noun, as in (12.a), and twelve cases of *le* repaired into (what we assume to be) an implicit elided *le* followed by a feminine noun, as in (12.b). Again, the default determiner *le* is overused and introduced with either masculine or feminine nouns before its replacement by the elided form *l'* + vowel. Interestingly, if there are five cases of self-repairs of *la* into an implicit elided *le* followed by a feminine noun, as in (12.c), there is not one instance of a repair by cliticization of *la* to a masculine noun in the corpus, that is, the *la* form is used sparingly by the learner and introduced only when she is certain that the noun is feminine.

- (12) a. *le* [//] l' + Noun_{+masc} : *le* [//] l'espagnol (the_{+masc} Spanish-language_{+masc})
- b. *le* [//] l' + Noun_{+fem} : *le* [//] l'ancienne école (the_{+masc} former_{+fem} school_{+fem})
- c. *la*_{+fem} [//] l' + Noun_{+fem} : *la* [//] l'université (the university_{+fem})

Even if that masculine determiner is not actually pronounced (as it is cliticised into *l'*), it is again an example of the ungrammaticality of the *la* to *le* repair.

6. Discussion

In this investigation, we have attempted to fruitfully bridge the gap between the methods of corpus linguistics and the conjoined morphosyntactic theories of the Minimalist Program and Distributed Morphology, interpreting the data with the help of independent morphosyntactic principles. We therefore have revisited the well-known gender imbalance in accuracy between masculine and feminine definite determiners by advanced L2 learners and added data on non-systematic agreement phenomena due to monitoring in the form of self-repair of the gender of singular definite articles.

We have confirmed the results of previous studies on the gender accuracy of singular definite determiner based on original data from a little studied population of learners: Jamaican learners of L2 French: the feminine determiner is significantly more accurate than the masculine determiner for advanced learners. We have shown that retracing phenomena present in the spontaneous speech of L2 learners open a window on the implicit featural composition of their internal grammar, and therefore reveals their interlanguage competence, as well as on the current state of their self-monitoring capacity with regards to noun gender attribution. We assume a syntactic well-formedness constraint on self-repairs in L2 like in L1.

The evaluation of two types of self-repair asymmetries *le* to *la* vs. *la* to *le* with overt or covert nouns, and the absence of *la* to *l'* + masculine noun has confirmed our hypothesis that the form *le* is the overgeneralized underspecified default singular definite determiner. As such, it may be inserted under the communicative pressure to produce a French-like DP as most learners are aware of the ungrammaticality of bare DPs in French in most contexts. It is important to note that the insertion of underspecified *le* bearing only the essential feature of definiteness does not prevent the derivation of the DP to converge, thanks to a relaxation of the morphosyntactic rule of competition regulating the insertion at PF of Vocabulary Items with phonetic shape. Even native speakers, it seems, have recourse to such default determiner under communicative pressure. The similar behaviour between intermediate to advanced learners and native speakers suggests that they may implicitly share the same morphological specification of the definite determiner system. Data has confirmed the hypothesis that the *la* to *le* (*l'*) self-repair should be considered ungrammatical in the learners' L2 interlanguage. The apparent few exceptions to this rule have been accounted for as instances during which the learners are testing and shifting the gender of the noun, or, attributed to a performance mistake caused by the added cognitive pressure linked with the expression of genericity. Finally, from the point

of view of disfluency studies, one could think of this default *le* inserted under communicative pressure as a kind of filler, allowing time for the learner to retrieve the correctly gender-specified determiners and nouns.

It must be recognized that the investigation is highly speculative especially in some cases of covert nouns presupposing an underlying abstract syntactic reconstruction of the *Reparandum* as a full DP. Whatever the explanation of these repairs, the fact of the asymmetry in the use of determiners remains. My hypothesis is that the same explanation for the asymmetry in the accuracy of use of singular determiners, their featural composition, can account for the asymmetry in repair phenomena with overt or covert nouns. Careful attention has been paid to remove obvious cases of ‘apparent gender errors’ / ‘Tip of the Tongue states’ from the count. Of course, this demonstration is only a first examination of the phenomenon, based on a single set of speakers, all bilingual natives of Jamaican Creole and Jamaican English. Considering the small sample size, it is important not to over-interpret the findings, and it would deserve to be tested on spontaneous oral productions by other L2 learners of French with different background. Finally, this paper has focussed on the definite determiner and the scope of the research should be extended to cover accuracy and repair phenomena with other definite (possessive and demonstrative) and indefinite determiners, as well as the interaction with adjective agreement within the DP. Finally, cultural factors such as the importance of the oral tradition in the use of Jamaican Creole, essentially a spoken language, might also play a role, as well as the mode of language instruction in Jamaica.

References

- Bartning, I. (2000). Gender agreement in L2 French: Pre-advanced vs. advanced learners. *Studia Linguistica* 54, 225-237.
- Blanche-Benveniste, Cl. (1987). Syntaxe, choix de lexique, et lieux de bafouillage. *DRLAV* 36-37, 123-157.
- Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané*. Thèse de doctorat, Université de Paris III.
- Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning* 39, 535-594.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA/London, England: MIT Press.

- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (ed.) *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1-52.
- Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Dewaele, J. M & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition* 4(3), 275-297.
- Embick, D. (2007). Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis* 33(3-4), 2-35.
- Fornel, M. de & Marandin, J.-M. (1996). L'analyse grammaticale des auto-réparations. *Le Gré des Langues* 10, 8-68.
- Götz, S. (2013). *Fluency in Native and Nonnative English Speech*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Granfeldt, J. (2004). Domaines syntaxiques et acquisition du français langue étrangère. L'exemple du DP. *AILE* 21, 47-85.
- Granger, S., Hung, J. & Petch-Tyson, S. (eds.) (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, Amsterdam: John Benjamins Publishing Company.
- Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.) *The View from Building 20: Essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111-176.
- Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry* 22, 27-62.
- Hilton, H. (2014). Oral fluency and spoken proficiency: considerations for research and testing. In P. Leclerq, A. Edmands & H. Hilton (eds.) *Measuring L2 Proficiency: Perspectives from SLA*. Bristol: Multilingual Matters, 27-51.
- Hulk, A. & Tellier, C. (1999). Conflictual agreement in Romance nominals. In J.-M. Authier, B. E. Bullock & L.A. Reed (eds.) *Formal Perspectives on Romance Linguistics*. Amsterdam: John Benjamins Publishing Company.
- Kormos, J. (1998). A new psycholinguistic taxonomy of self-repair in L2: A qualitative analysis with retrospection. *Even Yearbook, ELTESEAS Working Papers in Linguistics* 3, 43-68.
- Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning* 49(2), 303-342.

- Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Mahwah: Lawrence Erlbaum Associates.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition* 33, 41-103.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McKee, G., Malvern, D. & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing* 15(3), 323-337.
- Pallaud, B., Rauzy, S. & Blache, P. (2013). Auto-interruptions et disfluences en français parlé dans quatre corpus du CID. *Travaux Interdisciplinaires sur la Parole et le Langage* 29, <https://journals.openedition.org/tipa/995> (last accessed on July 1st, 2018).
- Péters, H. (2014). Étude sur la complémentation infinitive des verbes de déplacement sur la base d'un corpus de productions orales d'apprenants jamaïcains de FLE et suggestions pédagogiques. In L. de Serres, F. Ghillebaert, P.-A. Mather & A. Bosch (eds.) *Aspects Culturels, Linguistiques et Didactiques de l'Enseignement-Apprentissage du Français à un Public Non-Francophone*. Québec: Association Internationale des Études Québécoises, 62-84.
- Péters, H. (2017a). Comportements d'autocorrection et d'hésitation manifestés par des apprenants de FLE. *Bulletin Suisse de Linguistique Appliquée, Numéro spécial, Tome 2*, 133-145.
- Péters, H. (2017b). *The UWI French L2 Corpus*. Talkbank doi:10.21415/T5G975.
- Prévost, P. & Paradis, J. (eds.) (2004). *The Acquisition of French in Different Contexts. Focus on Functional categories*. Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Prévost, P. (2009). *The Acquisition of French. The Development of Inflectional Morphology and Syntax in L1 Acquisition, Bilingualism, and L2 Acquisition*. Amsterdam/Philadelphia, John Benjamins Publishing Company.

- Prodeau, M. (2005). Gender and number in French L2: Can we find out more about the constraints on production in L2? In J.-M. Dewaele (ed.) *Focus on French as a Foreign Language*. Clevedon, UK: Multilingual Matters, 114-134.
- Shriberg, E. E. (1999). Phonetic consequences of speech disfluency. In *Proceedings of the 14th International Congress on Phonetic Science, San Francisco*, 619-622.
- Svenonius, P. (2012). Spanning. Manuscript, CASTL, University of Tromsø. http://ling.auf.net/lingbuzz/001501/current.pdf?_s=gPASKtgCazPxcwFk (last accessed on June 28, 2017).
- Temple, L. (2000). Second language learner speech production. *Studia Linguistica* 54(2), 288-297.
- van Hest, E. (1996). *Self-Repair in L1 and L2 Production*. Tilburg: Tilburg University Press.
- Vigliocco, G., Antonini, T. & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science* 8, 314-317.
- White, L., Valenzuela, E., Kozłowska-MacGregor, M. & Leung, Y-K. I. (2004). Gender and number agreement in non-native Spanish. *Applied Psycholinguistics* 25, 105-133.
- Whong, M., Gil, K.-H. & Marsden, H. (eds.) (2013). *Universal Grammar and the Second Language Classroom*. Berlin: Springer.
- Wise, H. (1997). *The Vocabulary of Modern French. Origin, Structure and Function*. London: Routledge.