Transparency and language contact

The case of Haitian Creole, French, and Fongbe

Luisa Seguin
University of Maryland, College Park, USA

When communicating speakers map meaning onto form. It would thus seem obvious for languages to show a one-to-one correspondence between meaning and form, but this is often not the case. This perfect mapping, i.e. transparency, is indeed continuously violated in natural languages, giving rise to zero-to-one, one-to-many, and many-to-one opaque correspondences between meaning and form. However, transparency is a mutating feature, which can be influenced by language contact. In this scenario languages tend to evolve and lose some of their opaque features, becoming more transparent. This study investigates transparency in a very specific contact situation, namely that of a creole, Haitian Creole, and its sub- and superstrate languages, Fongbe and French, within the Functional Discourse Grammar framework. We predict Haitian Creole to be more transparent than French and Fongbe and investigate twenty opacity features, divided into four categories, namely Redundancy (one-to-many), Fusion (many-to-one), Discontinuity (one meaning is split in two or more forms,) and Form-based Form (forms with no semantic counterpart: zero-to-one). The results indeed prove our prediction to be borne out: Haitian Creole only presents five opacity features out of twenty, while French presents nineteen and Fongbe nine. Furthermore, the opacity features of Haitian Creole are also present in the other two languages.

Keywords: Haitian Creole, Fongbe, French, transparency, language contact, Functional Discourse Grammar

1. Introduction

When communicating, speakers map meaning onto form. It would thus be expected for languages to show a one-to-one correspondence between meaning and form. Nevertheless, this is often not the case. This transparency principle is continuously violated in natural languages, giving rise to redundancy and reduced

https://doi.org/10.1075/jpcl.00060.seg
Journal of Pidgin and Creole Languages 35:2 (2020), pp. 218–252. issn 0920-9034 | e-issn 1569-9870

forms for example. To our knowledge no fully transparent language exists, since to some extent they all somehow violate this ideal one-to-one correspondence. Languages, however, can have different degrees of transparency and violate it in different ways. The present study is on transparency in Haitian Creole and its super-, French, and substrate language, Fongbe, within the Functional Discourse Grammar (FDG) framework. Transparency is an important topic and has a great impact on the subfields of theoretical linguistics, language contact, diachronic variation, and language acquisition. This research deals with all of them, but more specifically with the former two and aims at theoretically analysing transparency in a specific language contact context, namely that of a creole.

The paper is organized as follows. Section 2 defines the concept of transparency and its definition within FDG (2.1 and 2.2). The last Subsection (2.3) explains the relation between creoles and transparency and outlines the hypothesis and prediction. Section 3 presents the methodology, introducing and explaining all non transparent features investigated in the study, while Section 4 outlines the results. Section 5 concludes the paper.

2. Transparency

Over the years transparency has been defined in different ways. For this research we will make reference to Hengeveld's (2011a) definition of transparency as a one-to-one correspondence between meaning and form. However, we first need to clarify what transparency is *not* but is often defined as: absolute simplicity, ease of acquisition, and iconicity.

Absolute simplicity is the simplicity of a language system as such, that is the amount of form (surface simplicity) and the levels of embedding (structural simplicity) needed to express any meaning (Miestamo 2006). The more linguistic material is needed and the more layered its structure is, the more complex the language. Examples of language properties that might be defined as signs of absolute simplicity are (Dahl 2004; Langacker 1977; McWhorter 2001):

1. relatively few morphological and phonological rules;
2. a small phoneme inventory;
3. limited degrees of syntactic depth, e.g. little sub-ordination or lack thereof;
4. regularity in linguistic paradigms.

Another type of simplicity is relative simplicity, as defined by Miestamo (2006), or ease of acquisition. According to this definition, the easier it is for learners to acquire a language, the simpler it is (Kusters 2003). Iconicity, on the other hand, is the predictability of the meaning of a certain word from its form (McWhorter

1998). The relation between the form and the meaning of a word is mostly arbitrary, except for onomatopoeia and ideophones. The matter becomes slightly more complicated when looking at compounds and derivations. A compound or derived word, or sign, can be defined iconic if its meaning can be predicted on the basis of the meaning of the single forms (Downing & Stiebels 2012). As the above descriptions suggest, all these concepts have something in common with each other and with transparency, but they should not be confused. Transparency in the current paper is defined as an interface property between two levels, meaning and form, and not an intrinsic property of the language.

Transparency is interesting for different linguistic subfields, namely diachronic change, language acquisition, and language contact. As far as diachronic change is concerned, Hengeveld (2011b) and Seuren & Wekker (1986) have argued that, as they 'are born', languages start out with transparent features that over time tend to become more opaque. These forms that evolve and lose their function, e.g. grammatical gender, have been defined as 'historical junk' and 'male nipples' (Lass 1997: 309). A typological study conducted by Leufkens (2015) shows that, although languages differ greatly in their degrees of transparency, the variation between them is not random. From her data,[1] Leufkens (2015) drew up an implicational hierarchy (1), which mirrors the presence of non transparent features in a language.

(1) nominal expletives, clausal agreement
→
grammatical gender, tense copying
→
suppletion
→
phrasal agreement, irregular stem formation
→
predominant head-marking
→
morphophonologically conditioned stem alternation
→
morphologically and morphophonologically conditioned affix alternation
→
redundant referential marking, phonologically conditioned stem and affix alternation, grammatical relations

The hierarchy is thereby implicational in the sense that the presence of a certain feature in a language implies the presence of all features lower in the hierarchy. For

1. For reasons of space, we will not present examples to motivate the implicational hierarchy and refer the reader to Leufkens (2015).

example, if a given language possesses nominal expletive, then that language will also possess grammatical gender, suppletion and so on. The hierarchy also represents diachronic change: languages always start out transparent (to some degree) and then slowly acquire the opaque features in (1) in a bottom up way.

However, this is the only situation in which a change in this direction can be witnessed. In both L2 acquisition and in language contact the tendency is in fact towards more transparency. Lupyan & Dale (2010) found that languages spoken in esoteric communities have more opaque features, such as fusional morphology, showing that diachronic change can be reinforced by linguistic isolation. Similarly, Trudgill (2011) provides evidence for the idea that the properties of a speaker community influence the complexity of a language variety. The second challenge to opacity concerns learnability. It has been shown that the more opaque a language is, the more difficult it is to acquire (Omar 1973; Slobin 1977; Aksu-Koç & Slobin 1985), especially for L2 learners (Blom et al. 2008). In fact, the more L2 learners a language has, the more it will lean towards losing certain opaque features (Kusters 2003). Slobin (1977) and Lightfoot (1979) argue that there is an 'opacity ceiling', also called Transparency Principle (Lightfoot 1979). The ceiling is a learnability limit: if a feature crosses this limit, it will get lost. An example is the diachronic variation of relative clauses from Old to Modern English. In Old English, relative clauses were introduced by the complementizer *þe* or a demonstrative pronoun. In Middle English, however, relative clauses started to be introduced by Wh-forms followed by the evolved complementizer *þat*. This stage represents the opacity ceiling, after which a reanalysis occurred in the 15th and 16th centuries: relative clauses could be introduced by a complementizer or a Wh-form, but not both (Lightfoot 1979: 313–342).

The opacity ceiling prevents languages from becoming too opaque and therefore a burden to learners. The only possible way for an opaque system to be learned is if enough evidence is provided, like in the case of grammatical gender in German (Audring 2009). The next section will outline the framework used in this research, namely Functional Discourse Grammar, and its importance for the investigation of transparency.

2.1 Functional Discourse Grammar

Functional Discourse Grammar is a linguistic theory developed by Hengeveld & Mackenzie (2008) following up on Dik's (1978) Functional Grammar model. The former, as opposed to the latter, is a structural-functional framework that aims at finding explanations for the structures of human language but that does so in the communicative function of the language. FDG sees the communicative intentions of a speaker as the starting point of the speech act. It is a top-down representation

of linguistic organization, that models the forming of the intention down to the phonetic (or orthographic) form. In FDG a Grammatical Component interacts with three other non grammatical components: The Conceptual, Contextual, and Output Components.

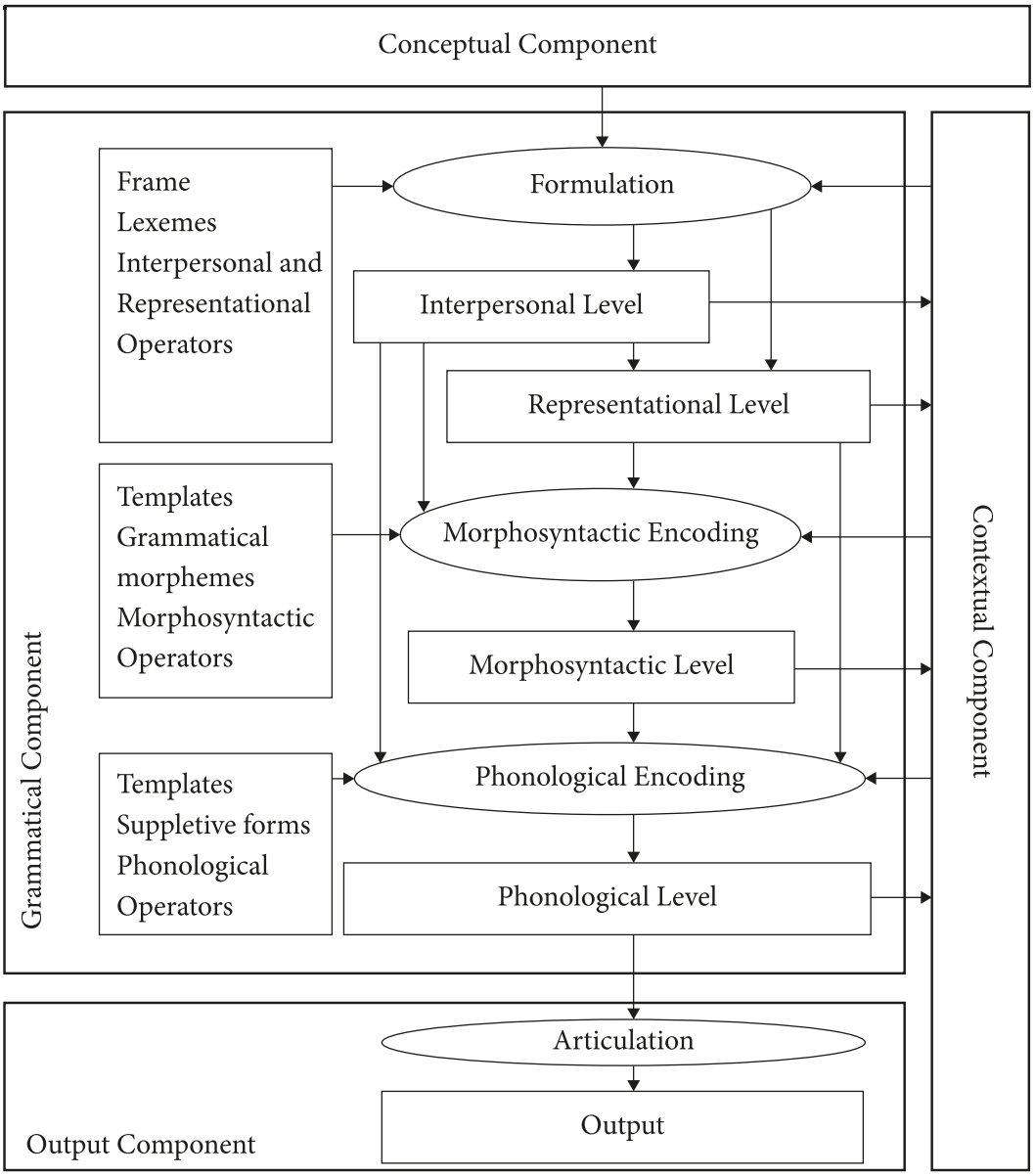

Figure 1. General architecture of FDG (Hengeveld & MacKenzie 2008: 13)

The Conceptual Component is where the intention is formed, the Output Component where the message is articulated, while the Contextual Component contains knowledge about the speech context. The Grammatical Component is subsequently subdivided into four hierarchically ordered levels: a pragmatic level, Interpersonal Level (IL), a semantic level, Representational Level (RL), a Morphosyntactic Level (ML) and a Phonological Level (PL). The speaker's intention

first goes through a Formulation process, which translates it into pragmatic (at the IL) and semantic (at the RL) units, that is primitives. The last transfer is through Phonological Encoding, a process that converts them into phonological units at the PL. During these last two encoding processes, and Individual x from the RL would acquire the form *he*. An example of the whole derivational process will be given in the following Section (2.2). These grammatically complete units are then passed on to the Output Component to be written, signed or pronounced. It is not obligatory for an intention to go through all the processes and levels. In some cases an intention can go directly from the Conceptual Component to the IL and then on to the PL, as in the case of *Ouch!*, which has no semantic or morphosyntactic counterpart. In other words, all four levels are independent but they all interact with each other. Figure 1 is a representation of the general architecture of FDG. Processes are represented by ovals, whereas levels by rectangles.

Every level has its own hierarchical internal structure, but this will not be discussed here, since it is not relevant for the goals of the present research. The following section defines transparency within FDG.

2.2 Transparency in FDG

As outlined above, transparency can be defined as an ideal one-to-one correspondence between meaning and form: all non transparent correspondences can be defined as opaque. However it can still be hard to understand what exactly can be defined as a meaning or form unit. Luckily, FDG can help us make this definition more precise. In this framework, a unit of meaning is a unit, which is a primitive, at either the RL or IL, while a unit of form is a primitive at the two lower levels, ML or PL. This would lead to the following definition (Leufkens 2015: 13):

(a) Transparency is obtained when one unit at one of the upper two levels of linguistic organization (IL, RL) corresponds to one unit at one of the lower two levels of linguistic organization (ML, PL)

It is important to notice once more that transparency is not a property of the level itself, but of the relation between levels, as an interface property. From this perspective this definition is not precise enough. In the Grammatical Component there are four levels, but in (a) only the relations between the two upper levels on the one hand and the two lower ones on the other (IL/RL – ML/PL) are considered, whereas there are also relations between the two 'meaning' and the two 'form' levels themselves. There are in fact six interfaces between the four levels, namely IL-RL, IL-ML, IL-PL, RL-ML, RL- PL, ML-PL and there might be opacity in all six of them. Following from this, transparency can thus be defined as the one-to-one correspondence of linguistic units between the four levels (Leufkens 2015: 13):

(b) Transparency is obtained when one unit at one level of linguistic organization corresponds to one unit at all other levels of organization

This definition may seem less specific than the one in (a) above, but it is more precise as it reflects the complexity of the language interfaces. As outlined in Section 1 above, all languages are somehow opaque and they differ from one another with regard to their degree of transparency. The latter must not be seen as a binary feature though, but as a spectrum: a language can be more or less transparent and its transparency degree may be different for the six different interfaces.

We chose FDG as a framework because of its potential and accuracy in analysing linguistic forms. It allows to break down any form to primitives and, in doing so, allows to better spot mismatches between linguistic levels. Furthermore, FDG is a descriptive framework and so is the study of transparency, making them a perfect fit.

We present here an example of how the framework works. Let's consider the NP *These bananas* (Hengeveld & Mackenzie 2008).

These bananas

a. IL (+id R_I)
b. RL (prox m x_i: [(f_i: /bəˈnɑːnə/$_N$(f_i)) (x_i)$_\Phi$])
c. ML (Np_i:[(Gw_i:this-**pl**(Gw_i))(Nw_i:/bəˈnɑːnə/-**pl**(Nw_i))](Np_i))
d. PL (PP_i: [(PW_i: /**ðiːz**/ (PW_i)) (PW_j: /bəˈnɑːnəz/ (PW_j))] (PP_i))

At the IL (a), the constituent is characterized as being referential (R) and identifiable (+id) by the listener. At the RL (b), it is noted that the quantity of the Individual (x) is more than one (m) and that such Individual (x) has a Property (f), specified by the Nominal (N) Lexeme /bəˈnɑːnə/. Furthermore, the location of the Individual is specified as close to the speaker (prox). At the ML (c), the constituent is defined as a Noun Phrase (Np), in turn composed by a Grammatical Word (Gw), represented here as *this*,[2] and a Nominal Word (Nw). It is interesting to notice how the Operator (m) from the RL is converted into the Plural Operator *pl* twice, on the Grammatical Word and on the Nominal Word, which is an instance of opacity. Finally, the PL (d) consists of a Phonological Phrase (PP), containing two Phonological Words (PW), /ðiːz/ and /bəˈnɑːnəz/. At this stage, the words acquire their appropriate plural forms. For the Nw, the plural is expressed by adding the plural suffix *-s*, while for the determiner by selecting the appropriate form. The realizations of *these* is another example of opacity, called suppletion, which will be explained in details in 3.3.2.

2. At ML there would only be a placeholder for the demonstrative. Here, we used *this* in order to render the representation simpler for the sake of explanation. We refer the reader to (Hengeveld & Mackenzie 2008) for further details.

2.3 Transparency and creoles

As stated by Hall (1966), a creole language is generally considered a natural language that has developed from a pidgin, a simplified language having emerged in a specific contact situation, namely between two groups that do not have a language in common. The pidgin is then nativized by the following generation, which results in a fully developed language, both semantically and grammatically. Although recently it has been heavily questioned whether all creoles start out as a pidgin (Siegel 2008; Blasi et al. 2017), in this paper we follow Hall (1966) in considering creoles as deriving from pidgins.

The most common situation in which creoles are born is the eighteenth and nineteenth centuries' colonization, due to the Atlantic slave trade (Mufwene 2015). All such creoles are based on European languages, such as Dutch, French, English, and Portuguese. The latter are called the superstrate languages, which came into contact with substrate languages spoken by enslaved people, e.g. Fongbe. However, defining creoles is not as easy as it may seem. According to Arends et al. (1995), the distinction between creoles and non creoles is mostly historical, while according to other researchers, such as McWhorter (2011) and Bakker et al. (2011), they also differ in their grammatical structures in a consistent way. For the present study we will follow the definition of Arends et al. (1995), considering a creole as a language that arose in a contact situation during the colonial period. As far as the grammar of such languages is concerned, we will not pursue this topic any further and only focus on transparency.

Contact situations are extremely interesting with regard to transparency and, given the absence of any common language for communication purposes, creole languages are even more interesting in this respect. This kind of situation forces speakers to make their language as clear, simple and understandable as possible (Leufkens 2013). This goal can be reached in two ways. The first is economy: speakers try to use as little material as possible in order to express a certain meaning, which results in the use of reduced forms. The second is intelligibility, that is the need to be understood, which leads to the use of forms that are as easy as possible to perceive, namely more intelligible. This results in the avoidance of reduced forms and, as a consequence, in increased transparency. The choice between the two means is partly based on motivation, e.g. the speaker needs to be quick and efficient in his requests, and partly on the typological distance between the languages. As far as the motivation is concerned, more transparency is expected in those situations in which communication is necessary, such as in trade, and not in those in which both speakers have time to learn a common language in order to interact. Typological distance is also extremely important: the more distant the sub- and superstrate languages are, the more transparent the creole will be. For creoles originating in the colonial period, this is always the

case, therefore both points are important when dealing with such contact situations. Finally, there is another factor that needs to be considered, namely time. As mentioned above, in the absence of contact, languages become more opaque over time. Colonial creoles are young and therefore more transparent than the older languages they originated from. Nevertheless, it has to be noted that a specific type of opacity, namely redundancy, might be very helpful for the perceiver, who in that case would have a double chance to perceive a certain meaning.

A previous study by Leufkens (2013) on the transparency of four creoles (Nubi, Pichi, Sri Lanka Malay and Diu Indo-Portuguese) showed that, even though they all exhibit some opaque structures, they are more transparent than the languages they originated from. She also found no evidence for the existence of so-called form-based forms in creoles, that is, linguistic elements that are not semantically or pragmatically motivated, such as nominal expletives. The languages we investigated are Haitian Creole, its superstrate language French and the substrate Fongbe. Two aspects need to be considered. First, Haitian Creole originated in the colonial period, in the fields and in a context of slavery. Moreover, French and Fongbe are typologically quite distant. The nature and history of the languages and the previous literature lead to the prediction that Haitian Creole is more transparent than its super- and substrate languages French and Fongbe. This prediction is investigated in the current paper. The next section will outline the methodology and describe the features investigated in detail.

3. Methodology

For the present study the research was conducted in a different way for each of the three languages involved. It is important to consider that, ideally, we should have investigated the language varieties spoken at the time Haitian Creole first emerged. However, the lack of data and literature on these varieties has rendered such approach impossible. We thus looked at present day varieties. The French data were discussed with a colleague and native speaker of Parisian French, who agreed to be consulted as an informant. The data from Haitian Creole are based on DeGraff (2001), Lefebvre et al. (1982), Valdman (2015), and Glaude (2012). For Fongbe, on the other hand, reference was made to the dataset created by Leufkens (2015)[3] and based on Lefebvre & Brousseau (2002). Table 1 summarizes all references.

The core method has been to consult the aforementioned sources in order to attest the presence or absence of opaque features from the list presented in the next subsections. For the aim of this study we considered one example of opacity for a

3. The database is available at http://transparency.humanities.uva.nl/.

Table 1. References

Language	References
French	Native speaker consultant
Haitian Creole	DeGraff, Michel. 2001. Kreyòl Ayisien. In John Holm & Peter L. Patrick (eds.), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars*. London: Battlebridge Publications. Glaude, Herby. 2012. *Aspects de la syntaxe de l'Haïtien*. Paris: Anibwé. Lefebvre, Claire, Hélène Magliore-Holly, & Nanie Piou. 1982. *Syntaxe de l'Haïtien*. USA: Karoma Publishers INC. Valdman, A. 2015. *Haitian Creole: Structure, Variation, Status, Origin*. Sheffield: Equinox Publishing.
Fongbe	Lefebvre, Claire & Anne-Marie Brousseau. 2002. *A Grammar of Fongbe*. Berlin: Mouton de Gruyter. Database http://transparency.humanities.uva.nl/

certain feature as sufficient evidence to define that feature as opaque (+ value). If no evidence was found that language has been considered transparent with regard to that feature (– value). Note that we decided to count for opaque rather than transparent features because of our decision to consider one instance as enough evidence for opacity. In case there was no literature available on that particular phenomenon then the value No Data (ND) was assigned. The features have been divided into four subgroups, namely Redundancy, Discontinuity, Fusion, and Form-based Form.

3.1 Redundancy

This category includes all one-to-many relations at different linguistic levels, namely one pragmatic, semantic or morphosyntactic unit that corresponds to more than one semantic, morphosyntactic or phonological units. The redundant opacity features investigated in this research are the following: Clausal Agreement, Cross-Reference, Phrasal Agreement, Plural Concord, Negative Concord, and Tense Copying.

3.1.1 Clausal Agreement and Cross-Reference

Agreement is a morphosyntactic operation in which a semantic or grammatical property of one unit, also called the controller, is expressed on some other unit, which can be defined as the target. Agreement can be found in different contexts, but the phrase is definitely the most canonical (Corbett 2006), followed by the clause, where the predicate agrees with its arguments. In FDG, agreement is seen

as a purely morphosyntactic operation, happening at ML, which copies some features of one unit onto another. The copy is thereby semantically empty. FDG, however, distinguishes Agreement from what it calls Cross-Reference, namely multiple expressions of one semantic unit.

(2) **Je** **vien-s** demain. *Viens demain. Agreement
1SG.NOM come.IND.PRS-1SG tomorrow
'I am coming tomorrow.'

(3) **Io** **veng-o** domani. Vengo domani. Cross-Reference
1SG.NOM come.IND.PRS-1SG tomorrow
'I am coming tomorrow.'

Distinguishing between these two types of agreement is not easy, but Hengeveld & Mackenzie (2008) propose a rule of thumb. If an element can occur on its own, it cannot be a copy, therefore we have Agreement only when both units at ML are obligatory, as in French (2), as opposed to Italian (3).

3.1.2 Phrasal agreement

Phrasal Agreement is the agreement between a noun and its modifiers, determiners and demonstratives. The latter can agree with the former in number, gender, case, and definiteness.

(4) Quell-**a** bambin-**a** biond-**a** rid-e Phrasal Agreement
DEM-SG.F child-SG.F blonde-SG.F laugh.IND.PRS-3.SG
'That blonde girl laughs.'

Italian is extremely opaque in this respect, since it always requires the targets to agree with the controller in both gender and number, as Example (4) shows.

3.1.3 Concord

The term concord refers to situations in which one semantic meaning is expressed by both a morphosyntactic and a lexical unit. There are three types of concord, namely plural, negative, and temporal concord. As far as the former is concerned, number can be expressed lexically, that is by the use of a numeral or a quantifier, and morphologically, by means of a plural morpheme. Some languages, however, use both, as in the English phrase *two cars*, where the plural is expressed twice, namely by *two* and by the morpheme *-s*, which clearly leads to opacity. A similar redundancy can be seen with negation. Negation can be expressed lexically, by means of adverbs (e.g. *never*), quantifiers, and pronouns (e.g. *nobody*), and morphologically, by means of inflection or a free grammatical morpheme. Moreover, in certain languages (like English), there are also other negative elements, that is

Negative Polarity Items (NPI). NPI's, like *anyone*, are not semantically independent, which means that they can only be used in the presence of another negative item, as in *I haven't seen anyone*. Logically the coexistence of two full (that is no NPI's) negative items should result in a positive reading, which is what happens in languages like English (5a). Russian, on the other hand, is a negative concord language: two negations have a negative reading (5b), which is redundant, thereby opaque.

(5) a. Nobody didn't read the book. = Everybody read the book.
b. **Nikto** **ne** chita-l knig-u. Negative Concord
Nobody.NOM NEG read.PFV-PST.SG book-ACC.F
'Nobody read the book.'

Temporal concord is very common and it is the result of the cooccurrence of both a morphosyntactic tense marker and a temporal adverb, as in the sentence *Yesterday I studied chemistry*. As Leufkens (2015) clearly points out, these opacity features are very common cross-linguistically. With regard to transparency, they are extremely similar and therefore will be grouped together under the category Concord.

3.1.4 Tense copying

The last redundancy feature we consider is Tense Copying, a multiple marking of time reference in a main and subordinate clause, also known as *consecutio temporum*. This copying mechanism is a morphosyntactic process that copies the tense value of the tense operator of the main clause to its subordinates. French is opaque in this respect: tense copying is obligatory (6a). Languages like Russian, on the other hand, do not require it and the tense in the embedded clause is a relative tense, that is relative to the time of the utterance (6b):

(6) a. Marie a di-t qu'elle lis-**ait**.
Marie have.IND.PRS-3SG say-PST.PTCP COMP'NOM.SG.F read-IND.PST.3SG
'Maria said that she was reading.'
b. Maria skaza-l-a čto ona chita-et
Maria say.PFV-PST-SG.F COMP NOM.3SG.F read.PVF-PRS.3SG
'Maria said that she was reading.'

Tense copying is opaque, since the tense marking is a morphosyntactic process that is not semantically or pragmatically motivated.

3.2 Discontinuity

Discontinuity is another one-to-many correspondence and groups together all cases in which one pragmatic or semantic unit is split-up into two or more morphological or phonological units, resulting in a non transparent one-to-many relation. The discontinuity features investigated in this research are the following: Extraposition and Extraction, Raising, Circumfixes and Circumpositions, Infixes, and Non-Parallel Alignment.

3.2.1 Extraposition and/or Extraction

Elements that belong together at the IL and RL can sometimes be separated at the ML. This is the result, for example, of a modifier being separated from its head and moved to the right periphery of the sentence, Extraposition, or to the left periphery, Extraction. The former is often resorted to when an element is too complex and, therefore, preferred at the end of the sentence, or when it is focalized, while the latter is usually the result of topicalization. The English examples below show the difference between the two phenomena (Van de Velde 2012: 433). They are both considered opaque.

(7) a. We have several important books about global warming in stock.
b. We have several important books in stock **[about global warming]**.
Extraposition
c. **[About global warming]** we have several important books in stock.
Extraction

3.2.2 Raising

In certain cases, an argument semantically belonging to an embedded sentence can syntactically behave as an argument of the main clause (Davis & Dubinsky 2004). In languages like English raising is restricted to verbs like *seem* and *appear*. Two examples of a sentence before and after raising are reported in (8) below.

(8) a. It seems that the kids are tired. No Raising
b. **The kids** seem (to be) tired. Raising

Raising is clearly opaque, because the predicate and its arguments form a single unit at RL, while in a raising construction they do not do so at ML & PL.

3.2.3 Circumfixes and circumpositions

Circumfixes are affixes, therefore single units at the IL and RL, that are realized in two separate phonological units. Circumpositions, whose only difference is that they are freestanding words rather than affixes, are mostly found in isolating languages. An example of the former is the morphological marker of the past participle

in German, e.g. *ge*-wuss-*t* 'known'. The French negation *ne … pas*, on the other hand, is an example of Circumposition.

3.2.4 Infixes

The opposite of a Circumfix is an Infix, namely an affix which is inserted into a morphological unit. They are not discontinuous but they create discontinuity in the unit they are inserted in, as illustrated by the causative marker *<[(o)ʔ]>* or *<[(o)ʔb]>* in Kharia. The word *botoŋ* 'fear' is made causative (and discontinuous) by infixing the causative marker: *boʔtoŋ* 'scare' (Peterson 2011: 231).

3.2.5 Non-parallel alignment

We define as Non-Parallel Alignment the non parallelism between the ML and the PL. In order to be transparent the relationship should be parallel, but often this is not the case, as the Dutch example below shows:

(9) Ik wou dat hij kwam. Non-Parallel Alignment
NOM.1SG want.PST.1SG COMP NOM.3SG come.PST.3SG
[ik wou] [dat hij] kwam.
/kʊɑu dɑti kʊɑm/
'I wish he would come' Hengeveld & Mackenzie (2008: 18)

In (9) above *ik* and *wou*, that are distinct units at the IL and RL, correspond to a single unit at PL and so do *dat* and *hij*. This non parallelism is clearly opaque.

3.3 Fusion

Besides one-to-many languages also often show many-to-one correspondences. These transparency violations are called Fusion, because two or more units on one level correspond to one single unit at another level. The fusion features investigated in this research are the following: Cumulation of TAME and Case, Morphologically Conditioned Stem Alternation: Suppletion, and Morphologically Conditioned Stem Alternation: Irregular Stem Formation.

3.3.1 Cumulation of TAME and case

Cumulation is the expression of multiple meanings in a single grammatical unit, such as an Affix or a Grammatical Word (Hengeveld 2007), which becomes then a 'portmanteau morph' (Bauer 2003: 19). Cumulation is also known as fusional morphology and as such very common in fusional languages, like Italian, where the morpheme *-o* in *parl-o* 'speak-IND.PRS.PFV.1SG' encodes tense, aspect, person, and number. There are, however, certain semantic categories that are very commonly expressed by portmanteau morphs, also in non-fusional languages. The

first is the cumulation of gender and number, which are very often fused together. The second is case, which is often fused with gender and number, as in *-orum*, the genitive plural morpheme (for nouns) of the second declension in Latin. The last very common category of portmanteau morpheme is that of Tense, Aspect, Mood, and Evidentiality (TAME), as in the Italian form *parl-o*. Cumulation is of course non transparent and these two semantic categories will be of interest in our study.

3.3.2 Morphologically conditioned stem alternation: Suppletion

In order to express grammatical information, languages may resort to affixation, i.e. adding an affix to a stem, or to a change in the stem. There are two possible changes and one of them is called Suppletion. This name refers to a morphological process, during which the marking of specific information requires a stem which is not derivable from other stem forms of the same Lexeme (Bauer 2003: 48; Hengeveld 2007: 39). French is opaque in this respect and shows Suppletion (10): the stem does not only have a lexical meaning but it also encodes aspect and tense (Bauer 2003).

(10) Je **v**-ais j'**all**-ais j'**ir**-ai
NOM.1SG go.PRS-1SG NOM.1SG'go.PST-1SG NOM.1SG'go.FUT-1SG
'I go/ am going.' 'I went.' 'I will go.'

3.3.3 Morphologically conditioned stem alternation: Irregular stem formation

The marking of grammatical information can also be expressed by Irregular Stem Formation, namely a modification of part of the stem. This type of modification, however, is purely morphological and has to be distinct from the morphophonologically driven ones which will be discussed later. According to Bauer (2003) there are four different kinds of Irregular Stem Formation. The first two are vowel and consonant mutation, as in the English paradigm for the verbs *begin-began-begun* and *send-sent-sent*. The third is segmental structure modification, as in *thief-thieve*, where voicing defines if the word refers to an Individual or an action (State-of-Affairs in FDG terminology). Finally, there is suprasegmental modification, namely in the stress pattern, e.g. *INsult* (noun) and *inSULT* (verb). Nevertheless, only irregular modifications, i.e. only applying to some stems but not all, are considered opaque.

3.4 Form-based form

The last subgroup of opacity features consists of zero-to-one correspondences between meaning and form. They are called Form-based Form because these formal units have no counterpart at the pragmatic or semantic levels. The ones investigated in this research are the following: Grammatical Gender, Syntactic Alignment, Nominal Expletives, Influence of Complexity on Word Order or Heavy Shift, Predominantly Head Marking, Morphophonologically Conditioned Stem Alternation, Morphologically Conditioned Affix Alternation, and Conjugation/Declension.

3.4.1 Grammatical gender

Languages tend to divide nouns into classes. There are two kinds of noun classifications known to us: lexical and semantic. An example of the former is the noun classification in Dutch, also called Grammatical Gender. In this language the selection of the common *de* vs. neuter *het* article is lexically defined and has no semantic motivation (Blom et al. 2008). Kikongo, on the other hand, has almost ten noun classes and every class contains nouns that also belong together semantically (Dereau 1995). To class I, for example, only belong humans, to class VII abstract nouns that have no plural, etc. A language is considered opaque if it exhibits Grammatical Gender.

3.4.2 Syntactic alignment

In a clause the arguments can be expressed in different ways, depending on different factors. FDG recognizes three types of alignment, namely pragmatic, semantic, and morphosyntactic. The former, called interpersonal alignment in FDG terminology is typical of Tagalog, where Topic arguments need to be marked by the particle *ang-* and are cross-referenced on the predicate. In both (11a) and (11b) the semantic roles are the same, while the topic functions are different. In (11a) the Topic is the Agent *lalake* 'man', glossed in the predicate as A, whereas in (11b) the Topic is *isda* 'fish', the Undergoer, cross-referenced on the predicate (U).

(11) a. bumilí **ang**-lalake ng-isda sa-tindahan Interpersonal Alignment
PFV.A.buy SPEC.TOP-man OBL-fish LOC-store
'*The man* bought fish at the/a store'

b. binilí ng-lalake **ang**-isda sa-tindahan
PFV.U.buy OBL-man SPEC.TOP-fish LOC-store
'The/a man bought fish at the store.' Bickel (2011: 8–9)

Other languages base their alignment on semantic functions, therefore exhibiting representational alignment as it is called in FDG. There are two kinds of alignment

based on semantic functions. The first type of representational alignment is hierarchical and based on animacy and person. The second one marks their arguments for the categories of Actor, Undergoer, and Location (Hengeveld & Mackenzie 2008), which we find in Acehnese. In the latter arguments are expressed through the use of clitics depending on semantic roles, i.e. *-geuh* for Undergoer and *-geu* for Actor:

(12) gopnyan galak-**geuh** that — Representational Alignment
3.HON happy-3.HON.U very
'He is very happy.' — Durie (1985: 55)

Finally, another group of languages ignores any pragmatic and semantic role of the arguments and presents a purely morphosyntactic alignment, which as opposed to the previous ones is opaque. In FDG this phenomenon has also been called Grammatical Relations or Syntactic Function. It belongs to the Form-based Form category because the marking at the ML has no counterpart at the IL and RL, leading to a zero-to-one relation between levels. English for example shows Grammatical Relations. Example (13) illustrates that the alignment of the arguments in the clause is driven by their syntactic roles. While its semantic role changes through the sentences, Agent in (13a), Theme in (13b), and Patient in (13c), *he* can only occur in preverbal position, because he is the grammatical subject.

(13) a. **He** eats an apple. — Syntactic Alignment
b. **He** falls.
c. **He** was chased by the dog.

3.4.3 Nominal expletives

Nominal Expletives, also known as dummy subjects, are units needed at the ML that do not have any counterpart at the IL and RL and are therefore opaque. Dummy subjects are mostly found in weather and existential predicates and non raised constructions, like in the English sentences *<u>It</u> is snowing*, *<u>There</u> is a dog in the garden*, and *<u>It</u> seems that John is tired*. The verbs *to snow*, *to be* (in the sense of existing), and *to seem* need a placeholder, either because the verbs have a zero-argument structure or because the subject, in this case *a dog* and *John is tired*, do not precede like usual but follow the verb. Om the contrary, other languages, like Fongbe, do not require Nominal Expletives for the expression of the weather and existential predicates:

(14) jì jà
rain fall
'It is raining.' — Lefebvre & Brousseau (2002: 245)

Travis (1984) shows, with an implicational hierarchy, that if language does not make use of nominal expletives in weather predicates, then that language also does not make use of Nominal Expletives in other constructions, such as non raised ones. It follows that if a language has Nominal Expletives, then the latter will show in weather predicates. For this reason, Leufkens (2013, 2015) considered the presence of Nominal Expletives in weather predicates as a litmus test. We however believe that these two types of Nominal Expletives are different in nature and will consider Nominal Expletives in non raised constructions as belonging to another category, that is Discontinuity, because one single semantic unit is split into two units at ML.

3.4.4 Influence of complexity in word order or heavy shift

In FDG the placement of constituents in the sentence is considered to be driven by semantic and pragmatic status. This can, however, be overruled by the complexity of constituent: if a certain constituent is morphosyntactically complex, it can be at the end of the sentence (FDG does not allow for movement, therefore it cannot be said to be moved). The most common instances of Heavy Shift are NP shift, possessive's phrases, and relative clauses, as in the English sentences *Yesterday I saw at work the girl with really long red hair.* The same order would be infelicitous with a non heavy NP, as in ?*Yesterday I saw at work the girl.* This is obviously non transparent, since the position of *the girl with really long red hair* is morphosyntactically motivated.

3.4.5 Predominantly head marking

Grammatical information can be marked by means of affixes, which are head marking, or clitics and free-standing function markers, which are phrase marking. Head marking is opaque, because it is the class or complexity of the host which defines the nature of the affix, which in turn causes a zero-to-one correspondence between the RL and the ML. Clitics and free-standing function markers, however, are transparent because they are not defined by the class of complexity of the host. For this feature, it is impossible to set a plus/minus value, as languages tend to resort to both strategies. We therefore looked at what tendency the language predominantly has.

3.4.6 Morphophonologically conditioned stem alternation

As discussed above in the Fusion subgroup, stems may undergo changes when a morpheme is added to them. Changes may be due to semantic or pragmatic reasons, like in Suppletion, or to pure morphosyntactic reasons, resulting in a zero-to-one relation between RL and ML. In Hungarian, for example, when adding the

imperative morpheme *-s* the final *-t* of the stem of the verb *köt-* 'tie' becomes *-š* as in *köš-s* 'tie!'

3.4.7 Morpho(phono)logically conditioned affix alternation or conjugation and declension

Just like stems, affixes can undergo mutations when being added to certain stems due to morpho(phono)logical reasons. This phenomenon is quite restricted and only applies to certain affixes. An example of this can be found in West Greenlandic, where the replacive affix *-lirtuuq* 'one who likes -ing' is attached to a stem, its initial consonant adapts to the stem-final consonant, e.g. *sin-nirtuuq* 'one who likes sleeping'. (M. Fortescue, personal communication, June 24, 2014, reported in Leufkens 2015). There is another type of affix alternation, which is lexically driven by conjugation or declension classes and only applies to some affixes. We speak of conjugation when the affix mutates based on the class of the verb it attaches to, while we define declension as its nominal equivalent. An example of the latter comes from Latin, in which nouns are arbitrarily, i.e. lexically and not semantically, divided in five classes and every class requires a specific affix paradigm, as illustrated in Table 2 for the second class for the neuter word *bellum* 'war'.

Table 2. The Latin declension for the second noun class

Nominative	bell-**um**	bell-**a**
Genitive	bell-**i**	bell-**orum**
Dative	bell-**o**	bell-**is**
Accusative	bell-**um**	bell-**a**
Vocative	bell-**um**	bell-**a**
Ablative	bell-**o**	bell-**is**

These alternations are clearly opaque, since it is purely phonologically (in some case there may be a thematic vowel indicating the class the word belongs to) or morpho(phono)logically driven and has no semantic motivation.

3.5 Summary of all transparency features

Before starting to analyse the results, we propose a summary of all the transparency features investigated in this study, which are reported in Table 3 with both opaque and transparent values.

Table 3. Summary of all transparency features

Feature	Transparent value	Opaque value
Redundancy		
Clausal Agreement	Absent	Present
Cross-Reference	Absent	Present
Phrasal Agreement	Absent	Present
Concord	Absent	Optional or Obligatory
Tense Copying	Absent	Present
Discontinuity		
Extraposition and Extraction	Absent	Present
Raising	Absent	Present
Circumfixes and Circumpositions	Absent	Present
Infixes	Absent	Present
Non-Parallel Alignment	Absent	Present
Raising Nominal Expletives	Absent	Present
Fusion		
Cumulation of TAME and Case	Absent	Present
Morphologically Conditioned Stem Alternation: Suppletion	Absent	Present
Morphologically Conditioned Stem Alternation: Irregular Stem Formation	Absent	Present
Form-based Form		
Grammatical Gender	Absent	Present
Syntactic Alignment	Absent	Present
Nominal Expletives	Absent	Present
Influence of Complexity in Word Order or Heavy Shift	Absent	Present
Predominantly Head Marking	Mostly phrase marking	Mostly head marking
Morphophonologically Conditioned Stem Alternation	Absent	Present
Morphologically Conditioned Affix Alternation and Conjugation/Declension	Absent	Present

4. Results

In the present section the results are presented, divided by groups of features. Examples are reported only when the three languages differ from each other; the full set of data can also be found at http://transparency.humanities.uva.nl. Before presenting our results, however, an important note is to be mentioned about French, namely the difference between the spoken and the written language. The orthography of French is highly opaque and therefore there may be differences between the two varieties with regard to opacity features too. Since the French we are interested in is clearly the spoken one, we reported examples in which the opacity is noticeable in the spoken language, neglecting those in which the opacity is only visible in the written variety.

4.1 Redundancy

As far as the redundancy features are concerned, Concord seems to be the only one present in all languages, as the Example (a–c) show for plural concord.

(15) a. **Deux** chevaux. French
two horses.PL
'Two horses.'

b. àvɔ̀ vɔ̀vɔ̀ **wè** ɔ́ **lɛ́**. Fongbe
loincloth red two DEF PL
'The two red loincloths.' Lefebvre & Brousseau (2002: 54)

c. M gen **de** pòm **yo** Haitian Creole
NOM.1.SG have two apple PL
'I have two apples.' De Graaff (pers. Comm.)

With respect to the others they all behave differently. Clausal Agreement is present in French and it is obligatory. It cannot be Cross-Reference, because the target cannot be left out (16a). Both Fongbe (16b) and Haitian Creole (16c), on the other hand, do not present either Agreement or Cross-reference.

(16) a. **Nous** mange-**ons** de-s pomme-s. French
NOM.1PL eat.IND.PRS-1PL DET.PART-PL apple-PL
*Mangeons des pommes.
'We eat apples.'

b. kɔ̀kú xò àsíba. Fongbe
Koku hit Asiba
'Koku hits Asiba.' (Lefebvre & Brousseau 2002: 247)

c. M tande Jan vini. Haitian Creole
1SG hear Jean come
'I hear Jean coming.' (Lefebvre 1982: 122)

The same relation holds for Phrasal Agreement. It is obligatory in French, where the noun and its determiners, quantifiers, demonstratives, and adjectives are both marked with gender and number (17a). In the French example, the liaison is visible: the plural marker is spelled out as part of the following words, making the agreement noticeable. Fongbe (17b) and Haitian Creole (17c) do not have Phrasal Agreement:

(17) a. L**es** autre-**s** arbre-**s**. French
DET.PL.F other-PL tree-PL
'The other trees.'
b. ɖìɖè ɖàgbè kɔ̀kú tɔ̀n ɔ́ **lɛ́.** Fongbe
sketch good Koku GEN DEF PL
'Koku's good sketches.' (Lefebvre & Brousseau 2002: 51)
c. M gen **de** pòm **yo** M. Haitian Creole
NOM.1.SG have two apple PL NOM.1.SG
'I have two apples.' De Graaff (pers. comm.)

Finally, tense copying is present in French, as the tense of the subordinate clause in (18) below shows.

(18) Marie **a** **dit** qu'elle dans-**ait**. French
Marie AUX SAY.PST.PRT COMP'3SG.F dance-PST
'Mary has said that she was dancing.'

Unfortunately, no Fongbe examples have been found, but in Gungbe, which is closely related to Fongbe in this respect, there is no such morphosyntactic process. In Example (19) there is a future tense marker in the main clause, but no such marker is present in the embedded clause.

(19) jó kèkè lɔ́ dó, é má nyín mɔ́n súrù **ná** ɖɔ̀ mí Ø hɛ̀n kèkè
leave bike DET at if NEG COP that_way Suru FUT say we NFUT hold bike
émìtɔ̀n gble. Gungbe
his break
'Don't play with the bike, otherwise Suru will say that we caused the bike to break down.' E. Aboh (pers. comm.)

Haitian Creole also does not have any Tense Copying process from the main to the subordinate clause, but resorts to different strategies (Valdman 2015). Anteriority, for example, is expressed by a semantic process, namely repetition of the lexical verb:

(20) **Limyè l** **limyè** lanp lan papiyon an vole. Haitian Creole
light 3SG light lamp DEF butterfly DEF fly
'Since he had turned on the lamp, the butterfly flew away.'
(Lefebvre et al. 1982: 153)

Table 4. Summary of redundancy features language by language

Feature	French	Fongbe	Haitian Creole
Clausal Agreement/Cross-Reference	+	–	–
Phrasal Agreement	+	–	–
Concord	+	+	+
Tense Copying	+	–	–

4.2 Discontinuity

Of the Discontinuity features investigated in this study, one is possible in all three languages, namely Extraposition/Extraction, and one, i.e. Infixes, is not found in any of them.

Raising, Circumfixes and Circumpositions, and Raising Nominal Expletives, on the other hand, are not homogeneously spread. Raising is present in both French and Fongbe. In the former it is very common with verbs like *sembler* 'seem', which can only take one argument, namely a subordinate clause as in (21a), *il* just being a dummy subject. The Undergoer *les élèves* 'the students' is where it semantically belongs, in the embedded clause.

(21) a. Il sembl-e **que les élève-s soi-ent fatigué-s**. French
it seem.PRS-3SG COMP DET.PL.M student-PL be.SBJV-3PL tired-PL.M
'It seems that the students are tired.'

b. **Les élève-s** sembl-ent fatigué-s.
DET.PL.M student-PL seem.PRS-3PL tired-PL.M
'The students seem tired.'

The subject of the embedded sentence can however raise to the left of the main verb, creating semantic discontinuity, as in (21b).

According to Lefebvre & Brousseau (2002), raising is also attested in Fongbe. In (22a) below all arguments are where they belong semantically (and the verb takes a dummy subject *é*) but the Undergoer [j]è in (22b) has risen to the left of the main verb.

(22) a. é hwè **[j]ɛ̀** ɖɔ núsúnû ɔ́ mɛ̀ Fongbe
it lack salt at soup DEF in
'It lacks salt in the soup.'
b. **[j]ɛ̀** hwè ɖɔ núsúnû ɔ́ mɛ̀
salt lack at soup DEF in
'Salt is lacking in the soup.' (Lefebvre & Brousseau 2002: 278)

The other Discontinuity feature in which these three languages differ is Circumposition. The latter is found in French, where the only instance of discontinuity is the negation *ne…pas/jamais/plus* etc., and it is reported in (23).

(23) Je **ne** dor-s **pas/jamais/plus.** French
1SG NEG sleep.INF.PRS-3SG NEG/never/anymore
'I don't sleep/I never sleep/I don't sleep anymore.'

It is, however, important to mention the possible evolution of *ne…pas* from the colonial times to present. Such negative elements undergo a specific chain of changes, also known as Jespersen's Cycle. As Jespersen (1917) showed, negation may go through three stages: a first stage with only one pre-verbal element (*ne*), a second stage in which *ne* is strengthened by another negative element (*pas*) and a third stage in which the first element is lost. French is currently moving towards the third stage: in spoken French *ne* has indeed disappeared. Jespersen (1917) shows that Old French (9–14th century) was at stage I, with only one negative element, Modern French was at stage II, and contemporary French is now moving towards stage III. What needs to be considered here is the timespan between Old French, the 18th century and today. Not having any linguistic records of the use of negation during the colonial time, we cannot make any precise claim, but we can assume French to have been in stage II of the cycle and therefore for the Circumpositions *ne…pas* to exist in the French spoken in the contact context we are interested in. Fongbe also has a Circumposition, namely *ɖò…wɛ̀* to express imperfective aspect (24).

(24) kɔ̀kú **ɖò** àsɔ́n ɔ́ ɖù **wɛ̀**. Fongbe
Koku be.at crab DEF eat IPFV[4]
'Koku is eating the crab.' Lefebvre & Brousseau (2002: 96)

As far as Raising Nominal Expletives are concerned, they are present in both French (21a) and Fongbe (22a).

4. The author glossed wɛ̀ as POST, as they claimed it does not have an independent meaning. However, as the circumposition *ɖò…wɛ̀* expresses imperfective aspect, we decided to gloss it as such.

The last Discontinuity feature we investigated is Non-Parallel Alignment. It is present in both French and Fongbe. In the French Example (25a) below, the plural marker of *beau*, *-x*, forms a phonological unit with the following word, a phenomenon known as Liaison (Gess et al. 2012). A similar phenomenon can be observed in Fongbe (25b,c).

In HC, the preferred syllable structure is CV. Thus, if the vowels *i, ou, e*, and *o* occur before another vowel, they can undergo blending and be replaced by the corresponding semivowels (25d).

(25) a. Les beau-**x** **arbre-s**. ==> le bo **zaʁbʁ** French
DET.PL nice-PL tree-PL
'The nice trees.'

b. ɖɔ́ xó ná è ==> ɖɔ́ xó **nîi** Fongbe
say word to 3SG
'Say a word to him/her.'

c. ɖa e ==> **ɖɛ̀ɛ̀** E. Aboh (pers. comm.)
cook 3SG
'Cook it.'

d. toto ale ==> **totwal** Haitian Creole
all go
'Everybody goes' Valdman (2015: 93)

Table 5. Summary of Discontinuity features language by language

Feature	French	Fongbe	Haitian Creole
Extraposition and Extraction	+	+	+
Raising	+	+	–
Circumfixes and Circumpositions	+	+	–
Non-parallel Alignment	+	+	+
Infixes	–	–	–
Raising Nominal Expletives	+	+	–

4.3 Fusion

This category groups together four features: Cumulation of TAME and Case, Suppletion, and Irregular Stem Formation.

French has Cumulation of TAME, as the verbal forms in (26a) show, and Cumulation of Case with person and number. The latter is however restricted to personal pronouns (26b).

(26) a. Je mang-**e**/ mange-**ais** French
1SG eat-PRS.1SG eat-PST.1SG
'I eat/ate.'
b. **Je** **Me**
1SG.NOM 1SG.ACC/DAT
'I' 'Me'

Fongbe also has Cumulation of TAME, with personal pronouns:

(27) **ná** **màá** Gungbe
FUT.1SG NEG.FUT.1SG E. Aboh (pers. comm.)

The other two Fusion features were only found in French. Suppletion is quite common in irregular verbs, such as *être* 'to be' (28), while the formation of plural in nouns can present Irregular Stem Formation (29).

(28) Je suis/ ét-ais/ ser-ai French
1SG be.PRS.1SG be.PST-1SG/ be.FUT-1SG
'I am/was/will be.'

(29) ciel - cieux fol - foux
sky sky.PL mad mad.PL
'Th sky/skies' 'Mad'

Table 6. Summary of Fusion features language by language

Feature	French	Fongbe	Haitian Creole
Cumulation of TAME and Case	+	+	–
Suppletion	+	–	–
Irregular Stem Formation	+	–	–

4.4 Form-based form

The last category of features investigated in this study is Form-based Form, namely all those forms that do not have any pragmatic or semantic counterpart. The only feature with the same opacity value for all three languages is Syntactic Alignment. For all the others, our languages behave differently.

Grammatical gender has only been found in French, where nouns are lexically assigned either masculine or feminine gender:

(30) Masculine Nouns Feminine Nouns French
le cheval 'the horse' *la grenouille* 'the frog'
l'armoire 'the wardrobe' *la table* 'the table'
le vélo 'the bike' *la voiture* 'the car'

Nominal Expletives are present in both French and Fongbe. In the former, they are quite common, obligatory in weather (31) and existential predicates (32).

(31) **Il** pleu-t. French
3SG.M rain.PRS-3SG
'It rains.'

(32) **Il** y a un chien.
3SG.M there have.PRS.3SG DEF.M.SG dog
'There is a dog.'

In Fongbe, Nominal Expletives are less common, but nonetheless present (33), while in Haitian Creole (34) they do not exist.

(33) jì jà Fongbe
rain fall
'It is raining.' Lefebvre & Brousseau (2002: 245)

(34) Gen yon jè rape ayisyen. Haitian Creole
have DET young rapper haitian
'There is a young Haitian rapper.' (Glaude 2012: 325)

Complexity seems to only have influence on word order in French, where it causes Heavy Shift (35). No data on Fongbe and Haitian Creole was found.

(35) Le cadeau qu'ont tous donn-é à la maître-sse
DET.M.S present COMP'have.PRS.3PL all give-PST.PART to DET.F.S teacher-F.S
les enfant-s de la classe de Marie French
DET.PL kid-PL of DET.F.S class of Marie
'The present that we all kids from Mary's class gave the teacher.'

In French most grammatical information, e.g. gender, number (17a), and TAME (26a), is marked through the use of affixes, while only definiteness (***le*** *cadeau* in (35)) and possession (***de*** *Marie* in (35) above) are marked through free-standing forms, such as determiners and prepositions. Following from this, French can be said to be a Predominantly Head Marking language. Fongbe does not show instances of Head Marking, as exemplified in (17b) re-proposed below as (36) in which both possession (tɔ̀n), definiteness (ɔ́), and number (lɛ́) are marked on the phrase level.

(36) ɖìɖè ɖàgbè kɔ̀kú **tɔ̀n ɔ́ lɛ́.** Fongbe
sketch good Koku GEN DET PL
'Koku's good sketches.' Lefebvre & Brousseau (2002: 51)

Haitian Creole behaves the same in this respect, as the plural marker *yo* in (37) shows.

(37) M wè yonn nan chwal Jan **yo.** Haitian Creole
1SG see DET PREP horse Jean PL
'I have seen one of Jean's horses.' (Glaude 2012: 236)

As far as Morphologically Conditioned Stem Alternations are concerned, French is the only language of our sample to have them. The former can be seen in the plural formation of nouns:

(38) Trav**ail** - Trav**au**-x French
job job-PS
'Job – jobs'

Finally, Morpho(phono)logically Conditioned Affix Alternation was found in all three languages. French shows Conjugation and Declension. In this language verbs belong to three different groups. The first and biggest group (90% of the verbs belong here) contains all regular verbs finishing in *-er* like *aimer* 'to love', the second all those in *-ir* like *finir* 'to finish' and whose present participle finishes in *-issant* like in *finissant* 'finishing', while in the third group we find all the other verbs not belonging in the first or the second. *Aller* 'to go' is an example of verb belonging to such class, even though it ends in *-er* (cfr. Example (10) above).

Table 7. French pronominal declension

	Nominative	**Accusative**	**Dative**
1SG	je	me	me
2SG	tu	te	te
3SG	il, elle, on	le, la	lui
1PL	nous	nous	nous
2PL	vous	vous	vous
3PL	ils, elles	les	leur

Nominal declension does not exist in Modern French, but pronouns do still have a residue of the Latin declension system, as shown in Table 7.

Fongbe and Haitian Creole show Affix Alternation as well. In the former the diminutive suffix *-i* shows vowel assimilation:

(39) àčú-ví => [àčúvú] Fongbe
rat-DIM[5]
'Little rat.' (Lefebvre & Brousseau 2002: 25)

In Haitian Creole, on the other hand, the definite determinant particle *la* changes in relation to the last syllable of the preceding word, as it is shown in (40) and (41).

(40a) C [−nasal] ==> **la** (es. Tab la 'the table') Haitian Creole
V [−nasal] ==> **a** (es. Vwati a 'the car')
V [+nasal] ==> **an** (es. Chen an 'the dog')
V [−nas] + C [+nas], V [+nas] + C [+nas] and V [+nas] + C [−nas] ==> **nan**

(41) Machin **nan**
car DET
'The car.' Glaude (2012: 37)

Table 8. Summary of Form-based Form features language by language

Feature	French	Fongbe	Haitian Creole
Grammatical Gender	+	−	−
Syntactic Alignment	+	+	+
Nominal Expletives	+	−	−
Influence of Complexity in Word Order or Heavy Shift	+	ND	ND
Predominantly Head Marking	+	−	−
Morph. Conditioned Stem Alternation	+	−	−
Morph. Conditioned Affix Alternation and Conjugation/ Declension	+	+	+

4.5 Summary

Before moving to the discussion, we present a summary of our findings. French turns out to be the most opaque language in our sample, with nineteen features out of twenty (only Infixes were not found). Fongbe has nine: Concord, Extraposition and Extraction, Raising, Circumpositions, Non-Parallel Alignment, Cumulation of Case, Raising Nominal Expletives, Syntactic Alignment, and Affix Alternation. In Haitian Creole, on the other hand, only five features were found, namely Concord, Extraposition and Extraction, Syntactic Alignment, Affix Alternation, and Non-Parallel Alignment. The last features is, however, optional. The overall results are summarized in Table 9.

5. Diminutive.

Table 9. Summary of the results

Feature	French	Fongbe	Haitian Creole
Clausal Agreement/Cross-Reference	+	–	–
Phrasal Agreement	+	–	–
Plural and Negative Concord	+	+	+
Tense Copying	+	–	–
Extraposition and Extraction	+	+	+
Raising	+	+	–
Circumfixes and Circumpositions	+	+	–
Infixes	–	–	–
Non-Parallel Alignment	+	+	+
Raising Nominal Expletives	+	+	–
Cumulation of TAME and Case	+	+	–
Suppletion	+	–	–
Irregular Stem Formation	+	–	–
Grammatical Gender	+	–	–
Syntactic Alignment	+	+	+
Nominal Expletives	+	–	–
Heavy Shift	+	ND	ND
Predominantly Head Marking	+	–	–
Stem Alternation	+	–	–
Affix Alternation and Conjugation/Declension	+	+	+

5. Discussion and conclusion

Our results seem to show a very interesting pattern. As outlined in Section 2.3 above, we predicted that Haitian Creole, being a contact language, would be more transparent than its sub- and superstrate languages. The results clearly show that our prediction is borne out. Haitian Creole does not only present fewer opaque features than French and Fongbe, but as Table 9 above shows, all opaque features present in Haitian Creole are also present in both French and Fongbe. Moreover, by comparing the results to the implicational hierarchy proposed by Leufkens (2015) in (1) above and proposed again in Table 10 with our results, it is clear that there is a correspondence. Our results corroborate the hierarchy proposed by Leufkens (2015). French indeed presents both Clausal Agreement,

Nominal Expletives and all other features lower in the hierarchy. Haitian Creole and Fongbe, on the other hand, have Morphologically and Morphophonologically Conditioned Affix Alternation and also the feature underneath them in the hierarchy, Syntactic Alignment (Grammatical Relations in Leufkens 2015).

Table 10. Implicational hierarchy

Features	French	Fongbe	Haitian Creole
Nominal expletives, clausal agreement	+	–	–
Grammatical gender, tense copying	+	–	–
Suppletion	+	–	–
Phrasal agreement, irregular stem formation	+	–	–
Predominant head-marking	+	–	–
Morphophon. conditioned stem alternation	+	–	–
Morphophon. conditioned affix alternation	+	+	+
Grammatical relations	+	+	+

One point needs to be addressed, though. In her study, Leufkens (2015) only looked at Nominal Expletives in weather predicates, while we looked at them in non raised constructions like (22a) above as well. As outlined in Section 3.3.4 above, we believe that they should rather belong to another category, that is Discontinuity. The Nominal Expletives in Table 10 above are in fact the ones found in weather predicates, which we only found in French, and not the ones in non raised constructions, which we also found in Fongbe. This suggests that these two categories do not just belong to different opacity groups, but might also be different in nature and, therefore, being generated at different moments in time. This implicational hierarchy is very important in this respect. As a matter of fact, if the presence of a feature implies the presence of all features under it in the hierarchy, it follows that opaque features should appear in languages in a bottom-up way: first Grammatical relations followed by Morpho(phono)logically conditioned affix alternation and so on. This could provide a very important tool for the study of diachronic variation and language development and therefore needs further investigation.

However, as every study, this also presents some shortcomings. When studying creole languages, certain factors must be taken into account. As every other colonial creole, the one studied here is the result of different languages coming into contact with each other. In this study we only investigated French and Fongbe, because they had the biggest influence on the development of Haitian Creole (Bonenfant 2011). Future research is needed in order to investigate the

other languages that influenced it, which as documented by Bonenfant (2011) are Wolof, Ewe, Portuguese, and Spanish. Ideally another language should be studied as well, namely Taíno, the language spoken in the Caribbean before the European invasion, but since it is extinct since the 16th century not many records are available (Aikhenvald 2012). The other issue with colonial creoles is represented by, first of all, the specific variety of the different languages involved in the contact situation and, secondly, the stage of their development. As far as variety is concerned, those spoken by the colonizers were non standard varieties. This should be considered in a study such as the current one, but as nowadays there are no records available of those vernaculars, it is clearly impossible. The second important issue is represented by the different stages of development of the languages involved in the creole's genesis. These two factors are known as the Founder Principle (Mufwene 1996). When dealing with the genesis of a creole it is indeed important to consider different varieties, including the ones spoken by the colonizers, and the stage of development of the languages involved. This has revealed itself to be difficult in the present study for the aforementioned reasons, with the exception of the French negation.

To conclude, a more in-depth study considering all contact languages Haitian Creole originated from would provide a more specific and detailed picture of the situation and should be considered for further research. Nevertheless, our results are already pretty straightforward. They indeed not only confirm the hypothesis that language contact drives languages towards transparency, but also prove the existence of an implicational hierarchy of opacity features. The latter is very important as it demonstrate that outwardly non related features, such as agreement and grammatical gender, are in fact somehow dependent on each other, i.e. the latter is only present if the former is also present. Finally, such hierarchical dependencies suggest a specific pattern that languages should follow in their evolution towards transparency or opacity.

References

Aikhenvald, Alexandra Y. 2012. *Languages of the Amazon*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593569.001.0001

Aksu-Koç, Ayhan A. & Dan I. Slobin. 1985. Acquisition of Turkish. In Dan I. Slobin (ed), *The crosslinguistic study of language acquisition. Vol. 1. The data*, 839–878. Hillsdale: Lawrence Erlbaum Associates.

Arends, Jacques, Pieter Muysken, & Norval Smith. 1995. *Pidgins and Creoles: An Introduction*. Amsterdam: John Benjamins.

Audring, Jenny. 2009. Gender assignment and gender agreement. *Morphology* 18: 93–116. https://doi.org/10.1007/s11525-009-9124-y

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1). 5–42. https://doi.org/10.1075/jpcl.26.1.02bak

Bauer, Laurie. 2003. *Introducing linguistic morphology* (Second Edition). Edinburgh: Edinburgh University Press.

Bickel, Balthasar. 2011. Grammatical relations typology. In Song, Jae J. (ed.), *The Oxford Handbook Of Linguistic Typology*, 399–444. Oxford: Oxford University Press.

Blasi, D. E., Michaelis, S. M., & Haspelmath, M. 2017. Grammars are robustly transmitted even during the emergence of creole languages. *Nature Human Behaviour* 1(10). 723–729. https://doi.org/10.1038/s41562-017-0192-4

Blom, Elma, Daniela Polišenská, & Fred Weerman. 2008. Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research* 24. 297–331. https://doi.org/10.1177/0267658308090183

Bonenfant, Jacques L. 2011. History of Haitian Creole: from Pidgin to Lingua Franca and English Influence on the Language. *Review of Higher Education and Self-Learning* 4(11). 27–34.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/slcs.71

Davis, William D. & Stanley Dubinsky. 2004. *The Grammar of Raising and Control: A Course in Syntactic Argumentation*. Malden: Blackwell Publishing. https://doi.org/10.1002/9780470755693

DeGraff, Michel. 2001. Kreyòl Ayisien. In John Holm & Peter L. Patrick (eds.), *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars*. London: Battlebridge Pubblications.

Dereau, Léon. 1995. *Cours de Kikongo*. Namur: Wesmael-Charlier.

Dik, Simon C. 1978. *Functional Grammar*. Amsterdam: North-Holland Publishing Company.

Downing, Laura J. & Barbara Stiebels. 2012. Iconicity*. In Jochen Trommer (ed.), *The Morphology and Phonology of Exponence*, 379–426. Oxford: Oxford University Press.

Durie, Mark. 1985. *A grammar of Acehnese: On the basis of a dialect of North Aceh*. Dordrecht-Holland/Cinnaminson-USA: Foris Publications.

Gess, Randall, Chantal Lyche & Trudel Meisenburg. 2012. *Phonological Variation in French, Illustrations from Three Continents*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/silv.11

Glaude, Herby. 2012. *Aspects de la syntaxe de l'Haïtien*. Paris: Anibwé.

Hall, Robert A. 1966. Pidgin and Creoles Languages. Ithaca: Cornell University Press.

Hengeveld, Kees. 2007. Parts-of-speech systems and morphological types, *ACLC Working Papers* 2(1). 31–48.

Hengeveld, Kees. 2011a. Transparency in Functional Discourse Grammar. *Linguistics in Amsterdam* 4. 1–22.

Hengeveld, Kees. 2011b. Epilogue: Degrees of transparency. *Linguistics in Amsterdam* 4(2). 110–114.

Hengeveld, Kees & Lachlan J. Mackenzie. 2008. *Functional Discourse Grammar: a Typologically-based Theory of Language Structure*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199278107.001.0001

Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: Høst.

Kusters, Wouter. 2003. Linguistic complexity: The influence of social change on verbal inflection. PhD dissertation, Leiden University.
Langacker, Ronald. 1977. Syntactic reanalysis. In Charles Li (ed.), *Mechanisms of syntactic change*, 56–139. Austin: University of Texas Press.
Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620928
Lefebvre, Claire, Hélène Magliore-Holly, & Nanie Piou. 1982. *Syntaxe de l'Haïtien*. USA: Karoma Publishers INC.
Lefebvre, Claire & Anne-Marie Brousseau. 2002. *A Grammar of Fongbe*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110880182
Leufkens, Sterre. 2013. The transparency of creoles, *Journal of Pidgin and Creole Languages* 28 (2). 323–362. https://doi.org/10.1075/jpcl.28.2.03leu
Leufkens, Sterre. 2015. *Transparency in Language: A Typological Study*. Utrecht: LOT.
Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
Lupyan, Gary & Rick Dale. 2010. Language structure is partly determined by social structure, *PLoS ONE* 5 (1). 1–10. https://doi.org/10.1371/journal.pone.0008559
McWorther, John H. 1998. Identifying the creole prototype: vindicating a typological class. *Language* 74, No. 4. 788–818. https://doi.org/10.2307/417003
McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5. 125–166. https://doi.org/10.1515/lity.2001.001
Miestamo, Matti. 2006. On the feasibility of complexity metrics, In Kerge Krista & Maria Maren Sepper (eds.), *FinEst Linguistics: Conference Proceeding of The Annual Finnish and Estonian Conference of Linguistics May 6–7 2004, Eesti Filoloogia Osakonna Toimtited* 8, 11–26. Tallinn: TLÜ
Mufwene, Salikoko S. 1996. The founder principle in creole genesis. *Diachronica* 13:1. 83–134. https://doi.org/10.1075/dia.13.1.05muf
Mufwene, Salikoko S. 2015. Pidgin and Creole Languages. *International Encyclopaedia of the Social and Behavioural Sciences* (Second Edition). 133–145. Oxford: Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.52014-5
Omar, Margaret. 1973. *The acquisition of Egyptian Arabic as a native language*. The Hague: Mouton. https://doi.org/10.1515/9783110819335
Peterson, John. 2011. *A grammar of Kharia. A South Munda language*. Leiden: Brill. https://doi.org/10.1163/ej.9789004187207.i-474
Seuren, Pieter & Herman Wekker. 1986. Semantic transparency as a factor in creole genesis. In Muysken, Pieter & Norval Smith (eds.), *Substrata versus Universals in Creole Genesis*. Amsterdam: John Benjamins, 57–70. https://doi.org/10.1075/cll.1.05seu
Siegel, Jeff. 2008. *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.
Slobin, Dan I. 1977. Language change in childhood and history. In John Macnamara (ed.), *Language Learning and Language Thought*. New York: Academic Press. 185–214.
Trudgill, Peter. 2011. *Sociolinguistic typology: social determinants of linguistics complexity*. New York: Oxford University Press.
Valdman, Albert. 2015. *Haitian Creole: Structure, Variation, Status, Origin*. Sheffield: Equinox Publishing.
Velde, Freek van de. 2012. PP extraction and extraposition in Functional Discourse Grammar, *Language Sciences* 34. 433–454. https://doi.org/10.1016/j.langsci.2012.02.007

Address for correspondence

Luisa Seguin
University of Maryland
Department of Linguistics
1401 Marie Mount Hall
College Park, MD 20742-7505
USA
seguinluisa@gmail.com

Publication history

Date received: 6 June 2019
Date revised: 10 January 2020
Date accepted: 8 March 2020