Periphrasis in Ontogeny and Phylogeny

Michael Tsai

May 17, 2001

1 Introduction

Periphrasis may be defined as “the use of longer, multi-word expressions in place of single words”\(^1\) [Haspelmath 2000, p. 654]. The term applies in many linguistic domains, from verb conjugation (with temporal auxiliaries) to comparative forms (with more) to interrogatives (with do). A periphrastic expression is composed of free morphemes, while a monolectic expression consists of a single word that may be composed of several morphemes bound together, e.g. through inflection.

This distinction applies not just to expressions, but also to whole languages: so-called analytic (isolating) languages contain mostly monomorphemic words and rely heavily on syntax to express meaning, synthetic (inflectional) languages rely more on morphology, and polysynthetic languages can express whole sentences as a single string of bound morphemes [O’Grady 1997]. However, these terms are rough classifications, and many languages contain both analytic and synthetic elements, sometimes even in the same paradigm. The balance between the two differs with the area of grammar under consideration, but it may also differ within the same area of grammar over time.

In this paper we will examine the tension between periphrastic and bound forms with respect to two aspects of grammar: tenses and interrogatives. We will examine these from the perspectives of language creation and language change, drawing on evidence from creoles\(^2\) and historical linguistics. Results from language acquisition will link these two disciplines and suggest motivations for the diachronic changes and the patterns that emerge.

---

\(^1\)Strictly speaking, the boundary between periphrasis and discourse is ill-defined, for it depends upon what languages exist and what linguists know about them. Haspelmath notes [Haspelmath 2000, p. 655] that the desirative is not considered periphrastic, because no known language expresses it via a bound form; however, this classing would be subject to change if a language were found that inflected for the desirative. Nevertheless, this definitional fuzziness will not hamper our analysis because we are interested in comparing what we know about languages that do exist, not in proving broad claims about what might be.

\(^2\)We will use “creole” to refer to languages generally classed as such (usually on the basis of their sociohistories) and ignore the question of whether creoles languages constitute a synchronically identifiable class.
2 Periphrasis From Three Branches of Linguistics

2.1 Creolist Perspectives

Seuren and Wekker, in motivating their semantic transparency, write that “it is generally agreed that creole languages have little or no morphology and that they place the burden of semantic expression on syntax and the lexicon” [Seuren 1986, p. 61]. In other words, there were practical reasons for creole languages to rely on periphrasis to a greater extent than other languages do. This enabled listeners to “carry out semantic interpretation with the least possible machinery and with the least possible requirements on language learning” [Seuren 1986, p. 64].

Whinnom says that the creole speaker may be “handicapped by his language” and that creoles represent languages at the start of evolution. Pidgin speakers try to use periphrases to express abstract ideas, but periphrases are broken; creole speaker must “repair” them and “adapt” the language to his needs. Adaptation necessarily involves a shift towards more synthetic structures [Whinnom 1971, pp. 109–110].

In this vein, McWhorter [McWhorter in press, p. 31] claims that “the first language had no affixes” and that inflection emerged “from the grammaticalization, reanalysis, or reinterpretation of material which was not originally inflectional.” He contends [McWhorter 1998, p. 812] that creoles, as young languages, make extensive use of periphrasis and that older languages have drifted over “millennia” towards increased use of bound morphology.

For McWhorter, inflectional morphology (lack of periphrasis) “renders a grammar more complex than another one in most cases” [McWhorter in press, p. 11]. He acknowledges that some old languages have become “simpler,” e.g. through phonetic erosion of affixes, and is aware of exceptions to the trend toward increased inflection. However, he seems to truly regard these as exceptions to the general pattern, thus echoing the typical creolist view that languages tend to develop from analytic to synthetic, in a monotonic fashion away from periphrasis.

2.2 Historical Perspectives

The perspective from historical linguistics is that this progression from analytic forms to synthetic ones is only half of the story. Bazzanella writes that there is a “cyclic pattern of alternating synthetic and analytic structures in modern Romance languages” [Bazzanella 2000, p. 180]. This sentiment is echoed by McMahon [McMahon 1994, p. 324], who says that change is in both directions in different parts of the language at the same time. She stresses that these changes represent neither evolution nor decay. Indeed, Thomason writes that changes are simply the result of “tendencies within the language to change in certain ways as a result of structural imbalances” [Thomason 1988, p. 9].

Haspelmath [Haspelmath 2000] agrees with the creolists that periphrastic expressions can be grammaticalized and eventually become inflections, however he notes that the synthetic forms then become reduced and eventually are replaced by new roundabout expressions. Heine [Heine 1991, p. 13] describes a natural progression where the expression of information (of a particular type) shifts over time from discourse, to syntax, to morphology, to
morphophonemics, and then back to discourse. This cyclic perspective is summarized succinctly by Givón, who states that not only is today’s morphology “yesterday’s syntax,” but “today’s syntax is yesterday’s pragmatic discourse” [Heine 1991, p. 13].

2.3 Acquisition Perspectives

It has long been thought that bound morphology is difficult to learn, perhaps because it is not easy to hear (or see) where the morpheme boundaries are. Affixes certainly pose problems for those learning a language late. In Newport’s American Sign Language study, the proper ASL description of an event required the combination of two inflectional morphemes. “The parents virtually never produced such combinations. Instead, they produced one inflection and represented the other meaning periphrastically (e.g. with an adverbial phrase)” [Newport 1999, p. 169].

Vainikka finds that the ability to analyze bound morphemes declines with age. Adults and L2 learners initially acquire free morphemes, then affixes; while children tend to learn affixes first, then bound morphemes [Vainikka 1995, p. 98]. There is much support for the idea that bound morphemes do not pose a problem for L1 acquisition. For instance, Peters writes: “Learners may produce productive combinations of two bound morphemes earlier than productive combinations of two free morphemes when full sentences of their language tend to consist of single words with many affixes” [Peters 1995, p. 480]. Along these lines, Wexler characterizes children as “little inflection machines,” stating that they have very early knowledge of inflection and are “extremely good at learning [it]” [Wexler 1998, p. 43]. Goldowsky and Newport’s “less is more” hypothesis claims that this difference between early and late acquisition is due to the limited cognitive abilities of children, which give them an advantage over adults in learning inflection [Goldowsky 1993, p. 126].

However, it is far from clear that children have an easy time analyzing bound morphemes, and it may be that bound and free morphemes are equally difficult for learners to segment. After all, both can become insalient when proceeding stress. For instance, Peters mentions a child who analyzed two free morphemes, did you, as a simplex completion marker, didja [Peters 1995, pp. 470]:

\[
\begin{align*}
(1) & \text{ Didja hear car} = \text{ I heard a car} \\
(2) & \text{ Didja throw everything out} = \text{ I threw everything out}
\end{align*}
\]

3 Tenses

Hornstein proposes (following Reichenbach) that human perception of tense can be represented using three points in time [Berwick 2001, p. 5–6]. The event point denotes the time the verb’s action took place, the speech point denotes the time of the utterance, and the reference point denotes some third time. These events may be pairwise ordered or unordered,

---

3 The event involved blowing out multiple candles on two cakes, and this was represented by the verb for blow and morphemes for repetition and dual: [BLOW-REP]-DUAL.

4 On a smaller time scale, the popular MP3 algorithm exploits this human limitation to compress music, with little perceptible loss in quality.
and the sixteen resulting combinatorial possibilities are exactly the tense structures that
humans perceive. English expresses only six of these possibilities with verbal morphology
(some requiring auxiliary verbs) and the remaining ten using adverbs and adjuncts. Some
languages have more synthetic tenses, and Chinese expresses all of them periphrastically.

Heine reports [Heine 1993, p. 68] that cross-linguistically, the progressive, the perfect,
and the future are most often expressed periphrastically, while the past and perfective are
usually expressed using bound morphology, and the imperfect always is.

3.1 Tenses in Diachrony

Miller describes [Miller 1993, p. 14] how the Romance perfect shifted from synthetic to aux-
iliating. The Latin for I have wanted is vol-u-i (want-PERF-1sS), which became the French
passé simple: vou-lu-s. However, during French diachrony the periphrastic passé compose,
which is formed using the auxiliary to have (avoir) or to be (être), gained a perfective sense.
For a while, it was in competition with the passé simple, but now the latter is restricted to
use in literary written materials.

A similar case is presented in [Haspelmath 2000, p. 659], where German’s periphrastic
perfective, formed using haben/sein (to have/to be) plus a participle, lost its perfective sense
and replaced the simple past. Thus, in both these cases the periphrastic form replaced the
bound one.

However, O’Grady et. al. [O’Grady 1997, p. 309] provide an example of change in the
opposite direction. Many English speakers have partially analyzed could have as the bound
coulda, as evidenced by their pronunciation and the fact that they commonly spell it as could
of, which does not make sense grammatically.

We can also see changes in the history of the future tense, which shifted from synthetic to
analytic, and back to synthetic [Miller 1993, p. 14]. The Latin cantabo (first-person singular
future of to sing) became cantare habeo\footnote{Fox confirms that though both Latin forms were present in anti-classical Latin, the form with habeo was based around the implication of duty or necessity. It was not until later that this form took on the future meaning, which eventually gave rise of the synthetic French form [Fox 1971, pp. 17–18].}, and then the French chanterai.

Interestingly, Bybee reports that the French future is shifting back to a periphrastic
form. The future proche is formed using a form of the verb aller (to go) and the infinitive
of the verb. It has expressed the notion of near future for some time, but in Modern
French vais chanter is replacing the synthetic chanterai\footnote{When I learned French as a second language, I was taught the future proche before the synthetic future, because the teacher deemed it easier to learn. Likewise, I was taught the recent past (using venir + de + infinitive) before the (also periphrastic) passé compose, which I learned before the imperative. Although above I cited work showing that free morphemes are easier for L2 learners to analyze, it is not clear to me that there is a difference in formal L2A where learners are provided with the morpheme segmentation. I remember students having trouble remembering that de was needed with venir but not aller, but it seemed easy for everyone to form the synthetic future by affixing conjugations of avoir (though the, often reduced, irregular future stems were hard to remember).} [Bybee 1994, p. 158]. Along these
lines, Bazzanella reports that the Italian morphological future (which presumably followed a
history similar to the French one) is also being replaced in Modern Italian with periphrastic
constructions, in this case involving modals [Bazzanella 2000].
According to Fleischman [Bybee 1994, pp. 158–159], the future often develops by incorporating modal notions such as intention, obligation, desire, movement, and intention. The modals develop and begin to take scope over the whole clause, and thus become grammaticalized as a periphrastic future. This then leaves a gap for new modals to fill, starting the process anew. This explains why the future cycles faster from synthetic to analytic than other tenses, why it is often periphrastic, and why its etymology is often easily identifiable even when it is expressed with bound morphology.

3.2 Acquisition of Tense

Clark reports that when acquiring tense children first learn to make a distinction between the present and the non-present. “In languages where both aspect and tense are marked on the verb, children appear to acquire both forms of inflection at an early age (starting well before two)” [Clark 1998a, p. 382]. However, in English (which only marks tense on the verb), children use adverbials such as yesterday as general tense markers, which can refer to any non-present time—the past or the future.

Clark goes on to say that forms like the present perfect and future are acquired late (ages 4–5 and 3, respectively). However, she implies that this is not related to the way these tenses are expressed in English (with periphrasis instead of bound morphemes), but rather that the acquisition cannot occur until the children learn the conceptual contrasts between these tenses—what they actually mean. As Labelle writes, “the appearance of the plus-que-parfait marks the point when children are able to distinguish between the three times T₀ ‘now’, Tᵣ ‘yesterday’, & Tₑ ‘two days before yesterday’” [Labelle 1994].

Allen [Allen 1995] points to evidence that English-learning children, at least, have a tendency to prefer free morphology. He reports instances of “tense doubling” such as I did broke it where tense is redundantly specified on the verb and on do. Children apparently do not know that the verb itself carries tense, so they insert do as a tense-marking particle. Interestingly, this result was also found for verbs such as walked, indicating that it is not simply a problem with learning irregular forms.

Pierce reports that children learning English acquire inflectional tense markings late, compared to children learning other languages. The first English affix tense acquired is the progressive marker -ing, but Pierce speculates that “it is not represented at syntactic levels as an analyzed stem + affix” [Pierce 1992, p. 70]. However, later she notes that children acquire French and Italian inflections early [Pierce 1992, p. 73].

Bazzanella found that Italian children learned the periphrastic future after the present and imperfect (which are both synthetic). Even though Italian has both morphological and periphrastic futures, children almost always targeted the periphrastic one first. One child in the study did target the morphological future, and he acquired it several months before the other children learned their (periphrastic) future [Bazzanella 2000].

Thus, these results from acquisition are ambiguous. Further, it is difficult to know whether children acquire certain tenses late because of the way the tenses are expressed grammatically, or because they have not yet learned what the tenses mean conceptually (or perhaps some combination of the two).
3.3 Tense in Creoles

Seuren and Wekker summarize the common view of how tense, mood, and aspect are expressed in creoles:

Languages with a complex morphology are generally hard to master for non-native learners...it is not surprising to find that in many creole languages verbal [TMA] are expressed by means of preverbal particles...[unlike] the majority of more advanced languages. [Seuren 1986, pp. 66–67]

Thus, creoles tend to express tenses periphrastically for historical and acquisitional reasons. As Rizzi writes, the creole TMA system “appears to be thoroughly invented in creole languages to make up, so to speak, for the lack of an adequate grammatical specification of the verbal system in the input” [Rizzi 1999, p. 464].

In Haitian Creole we find data such as these where the particle *ap* is used to mark future tense [DeGraff 2000, p. 6]:

(3) Bouki ap konn lesan an
    Bouki FUT know lesson the
    ‘Bouki will know the lesson’

(4) M ap vini (lè m fin manje)
    1sg FUT come when 1sg finish eat
    ‘I will come when I’m done eating’

Similarly, *te* is used for past tense and *a* for irrealis [DeGraff 2001, p. 11]:

(5) Boukinèt te renmen Bouki
    Boukinèt ANT renmen Bouki
    ‘Boukinèt loved Bouki’

(6) Boukinèt a renmen Bouki si ...
    Boukinèt IRR love Bouki if...
    ‘Boukinèt would love Bouki if...’

In the same document, DeGraff reports that regional dialects of modern French similarly mark tenses with particles. The situation in other French creoles is similar. Louisianna Creole uses *ale* (from French *aller*) to mark the future: *m’ale mange* [Verbix]. Mauritian Creole uses *ti* to mark past and *pe* to mark the progressive [McWhorter 1998, p. 804]:

(7) Zot ti pe ale
    they PAST PROG go
    ‘They were going’

Ford’s paper on St. Lucian French Creole [Ford 1996] affirms Bickerton’s analysis [Bickerton 1988] that in creole TMA systems the “unmarked condition is the time initially being talked about, while the marked is any time prior to the original time in focus.” There is a general marker for tense (*tè*), but the actual tense must (apparently) be inferred from adverbials.
Thus, in (8) we see that *ale* takes on the past tense because, from the context, the time in focus is after the person went to the market. In the second clause, the tense marker signals a shift in tense, but the actual tense shifted to must be inferred from the adverbial *ye*. Presumably the future could also be expressed with *té + tomorrow*.

Similarly, St. Lucian Creole contains pre-verbal mood markers:

(9) Mwen kay *ale laplas*
    I MOD go market
    ‘I will go to the market’

(10) Mwen pé *ale laplas denmen*
     I MOD go market tomorrow
     ‘I might go to the market tomorrow (not sure)’

DeGraff, Goodman [Goodman 1964, pp. 79-90], and others have shown that TMA particles in the French creoles have etymons in French periphrastic constructions. For instance, *ap* likely derives from *après*.

### 4 Interrogatives

#### 4.1 Interrogatives in Diachrony

Let us now examine the diachrony of periphrasis in interrogatives, beginning with the case of French. In Old and Middle French, questions could be formed by fronting the wh-word (11) or by simple inversion (12) [Clark 1993, pp. 333–334]:

(11) Que voelt ceste parolle dire?
    what wants this word to-say
    ‘What does this word mean?’

(12) A qui estes vous?
    whose are you

However, Middle French also saw the emergence of questions using complex inversion (13) and *est-ce que* (14) [Clark 1993, p. 336].

(13) Où Jean est-il allé
    where Jean is-he gone

(14) (qu’) est-ce que
    (what) is-it that
Initially, *est-ce que*’s meaning was probably compositional, that is, it would not be considered periphrastic because it followed the same inversion pattern found elsewhere in discourse. However, somewhere between Old French and Modern French, *est-ce que* lost its compositional meaning, as well as agreement with other parts of the sentence. It thus became an interrogative particle [Price 1971, p. 267].

(15) Que est iço que est avenud a Saul?
‘What is this that has befallen Saul?’ (Old French)
‘What has befallen Saul?’ (Modern French)

While *est-ce que* was developing into a periphrastic expression, French interrogatives lost simple inversion [Clark 1993, p. 319]:

(16) Comment fu ceste lettre faitte? (Old French)
how was this letter made

(17) * A Jean pris le livre? (Modern French, lexical)
has Jean taken the book

(18) Où est-il allé (Modern French, clitic)
‘where did he go’

In Modern French, simple inversion is only possible with clitics; according to Clark and Roberts, this is because Old French allowed nominative case assignment to lexical subjects under government, whereas Modern French does not. In any case, the history of French interrogatives displays a shift from using syntax (inversion) to greater use of periphrasis (*est-ce que*)\(^7\).

The situation in English diachrony is similar. As with French, simple inversion was the original way to form questions, but it became ungrammatical [O’Grady 1997, p. 314]:

(19) Speak they the truth? (Old, Middle, Early Modern)

(20) * Speak they the truth? (after 17th Century)

However, as Modern French allows inversion with clitics, Modern English allows it with auxiliaries:

(21) Can they speak the truth? (auxiliary)

\(^7\)Contradicting this trend, Vecchiato presents evidence from modern Québec French (as well as other dialects) that *ti* and *tu* are used as interrogative-marking verbal suffixes [Vecchiato 2000]. Thus, as with tenses, we see change taking place in both directions.

(1) Je suis-tu obligé de manger ma soupe?
I am-INTERR obliged to eat my soup
‘Am I obliged to eat my soup?’
In addition, Modern English admits constructions using periphrastic *do*, which appears in environments similar to French *est-ce-que*:

(22)  Do they speak the truth?  \(\text{(periphrastic)}\)

Kroch reports that periphrastic *do* was first used in the 13th century, in all the environments in which it is now found. Its use gradually increased in all these environments until it became obligatory in the 18th century. Kroch gives many reasons for why *do* might have been favored; among them are that was simpler (requiring less psycholinguistic resources than inversion), that it fit better with other changes in the language such as word order shift, and that it eliminated the following ambiguity [Kroch 1989, p. 149]:

(23)  Which knight saw the King?  \(\text{(Old English, ambiguous)}\)

(24)  a.  Which knight saw the king?  \(\text{(Modern English, unambiguous)}\)
    b.  Which knight did the king see?  \(\text{(Modern English, unambiguous)}\)

### 4.2 Acquisition of Interrogatives

According to Plunkett, the most common type of spoken French question is the periphrastic, which is used more often than inversion, wh-fronting, and wh-in-situ [Plunkett 2000, p. 105]:

(25)  Qu’ est ce que tu veux?  \(\text{(periphrastic, most common)}\)
    what is it that you want

(26)  Que veux tu?  \(\text{(inversive)}\)
    what want you

(27)  Tu veux quoi?  \(\text{(in-situ)}\)
    you want what

Thus, in investigating the acquisition of French interrogatives, we would like to know which type of questions is acquired first and how children analyze *est-ce que*. Plunkett suggests that a simplex analysis of *est-ce que* would render (28) no more difficult to learn than simple fronting of the wh-word [Plunkett 2000, p. 106]. Therefore, if children analyze *est-ce que* as a simplex\(^9\), sentences like (28) should appear before sentences like (26).

(28)  [lequel est-ce que] tu veux?  \(\text{bracketed section as a simplex}\)
    which-one *est-ce que* you want
    ‘Which one do you want?’

On the other hand, a complex analysis of *est-ce que* might look like [Plunkett 2000, p. 107]:

(29)  Lequel\(_k\) est\(_j\) ce \(t_j\) [que tu veux \(t_k\)]

\(^8\)A skeptic might say that *do* is not really periphrastic because it provides obligatory tense support. However, it does seem to fit a generalized definition of periphrasis in that *do* questions need a separate word to mark tense, whereas inversion could form questions without increasing the word-count.

\(^9\)Indeed, when I learned French, I analyzed *est-ce que* as a simplex meaning *do.*
It would therefore not be produced until after children had acquired inversion and embedded clauses. Plunkett’s results show that children acquire est-ce que questions late\textsuperscript{10}, at about the same time as inversion. This supports the complex analysis. However, one child in the study produced questions much earlier than the others. He used only a single form qu’est-ce que and had apparently analyzed it as a chunk. Thus, for him, it was easier to acquire a periphrastic question marker than a syntactic rule.

4.3 Interrogatives in Creoles

Creole interrogatives seem to be simpler than those found in French. DeGraff reports [DeGraff 2000, p. 37] that yes-no questions in Haitian Creole can be indicated with a rising intonation; there is no change in word order. Alternately, the speaker may use a clause-initial question marker èske. This is a simplex that obviously derives from French est-ce que.

\begin{equation}
\text{(Èske) Ou manje diri a?}
\end{equation}

QM 2s eat rice the

‘Did you eat the rice?’

In Mauritian Creole, yes-no questions are also indicated with rising intonation instead of word order change. However, unlike Haitian Creole and French, Mauritian Creole does not have a periphrastic interrogative [Adone 1994, pp. 28–29].

5 Conclusion

At the least, the above discussion and data have shown that the accepted view of many creolists is incomplete. Languages certainly also change in the direction away from bound morphology, towards greater use of periphrasis. However, even though the above examples might suggest this to be the dominant trend, we must not forget that many languages contain forms such as the imperfect that were once periphrastic, but that have remained synthetic for a long time. Thus, we seem to have found evidence that supports the historical linguistic perspective that this sort of change is cyclic.

The results from acquisition are ambiguous. Evidence showed cases where children preferred free morphemes, but also that they are excellent at acquiring bound morphology. Adults, on the other hand, are much better at analyzing free morphemes than bound ones. This may provide an explanation for why the French creoles examined appear to utilize periphrasis to such a great extent. For though children certainly played a role in the creation of creoles, adult pidgin speakers provided the primary source of input.

We must of course note that the creoles examined are a biased sampling; in each case French was the lexifier language. Future work could examine the use of periphrasis in other

\textsuperscript{10}Although Plunkett’s results seem to represent the accepted view, Hulk & Zuckerman found nearly the opposite. Their data [Hulk 2000, pp. 442–443] show that children produce questions using 89% fronting, 6% in-situ, 5% stylistic inversion, and no occurrences of est-ce que. However, these results are suspect because they also found that adults use inversion 62% of the time and inversion in only 5% of the cases, which seems to be the opposite of the accepted distribution in Modern French. They hypothesize that their data are skewed because the subjects were trying to impress their interviewers with complex constructions.
creoles and also in the substrates of the French creoles. The emergent use of periphrasis could be a result of features of the substrates, not of the language creation process. Nevertheless, it is interesting to note the similarities between the French creoles and modern French and its dialects, in the areas of tense and interrogatives.
References


http://www.siu.edu/cwis/departments/cola/ling/reports/stlucian/descript.htm


