In this paper I address the question what Media Lengua can tell us about the distinction between stems and affixes, the nature of lexical entries, and the relation between the lexicon, syntax, and phonology. It is part of a much larger attempt, coming from a variety of subdisciplines to provide what is sometimes termed external evidence for particular assumptions about the nature of the human language faculty.

Media Lengua is a linguistic variety spoken both as a native and as a second language by acculturated Indian peasants, craftsmen, and construction workers in Central Ecuador (Muysken 1981b). It is essentially Quechua with the vast majority of its stems replaced by Spanish forms. This process of replacement is commonly called "relexification". Examples of Media Lengua utterances are given in (1) through (3):

(1) a. unu fabur- ta pidi- nga- bu bini- xu- ni (Media Lengua)
   one favor AC ask NOM for come PR 1
   'I come to ask a favor.'

b. shuk fabur-da maña-nga-bu shamu-xu-ni (Quechua)

c. vengo para pedir un favor (Spanish)

It is clear that (1a) has resulted from putting the phonological shapes of the words in (1c) into the lexical entries in (1b). Several things should be noted. First, we get an emphatic form of the indefinite article in Media Lengua, unu, rather than Spanish unemphatic un. Second, the Spanish irregular verb form vengo appears in a regularized stem form bini. Third, the Quechua rule voicing accusative case -ta to -da after fabur has not applied in Media Lengua; Quechua dialectological evidence suggests that this is a very recent rule. Fourth, what is peculiar about Media Lengua is not so much that it contains Spanish words (many dialects of Quechua do as well), but that all Quechua words, including all core vocabulary, have been replaced. Fifth, the Spanish forms have been adapted phonologically
to Quechua; mid vowels have been replaced by high vowels. Quechua word order and morphology have been retained.

(2) a. kuyi-buk yirba nuwabi -shka (Media Lengua)
    cavia for grass there is not SD

    'There turns out to be no grass for the caviyas.'

b. kuyi-buk k'iwa illa-shka (Quechua)

c. No hay hierba para los cuyes (Spanish)

Note that the Quechua word *kuyi* appears in the local Spanish as well. The Media Lengua verb maintains the Quechua-specific “sudden discovery tense” marking -shka. The Quechua negative existential verb stem *illa-* has been relexified with a newly formed “frozen” stem *nuwabi-*, derived from Spanish *no* and *haber* ‘have’. The Spanish verb ‘have’ has an impersonal form *hay* which also has existential meaning.

(3) a. yo- ga awa -bi kay -mu -ni (Media Lengua)
    I TO water LO fall CIS 1

    'I come after falling into the water.'

b. ñuka-ga yaku -bi urma-mu -ni (Quechua)

c. vengo despues de caer en el agua (Spanish)

Examples such as (3) show the extent to which Media Lengua follows Quechua verb semantics. Cislocative -mu- can be attached to non-movement verb stems indicating that the subject comes after some action. This possibility exists in both Media Lengua and Quechua.

What examples (1) to (3) show us is that:

(a) Media Lengua is essentially the product of replacing Quechua phonological shapes with Spanish forms, maintaining the rest of the Quechua structure

(b) the Spanish forms chosen have undergone regularization and adaptation to Quechua morphophonology

(c) Media Lengua is conservative in sometimes reflecting earlier stages in Quechua pronunciation. It is not made up on the spot every time it is spoken

(d) The occurrence of Spanish strong alternants, frozen composites, etc. is an indication that the process of vocabulary replacement is not a simple one.

We now turn to the implications of the relexification process operant in Media Lengua for our view of the relation between the lexicon and the other components of the grammar. In the first section I look at the word formation component. In section 2 lexical semantics and syntax are discussed. Section 3 deals with the issue of what external evidence can be in theoretical linguistics. The implications of Media Lengua relexification for our view of the interaction between the lexicon and the phonology lie beyond the scope of this article.
1. Roots versus Affixes, IA versus IP

The first issue on which Media Lengua data may be able to throw some light is the representation of affixes as part of word formation rules or as separate lexical entries. This issue has a long history in structural linguistics (Hockett 1955), where it was discussed in terms of Item-and-Arrangement (IA) versus Item-and-Process (IP). In the IA view, words are constructed from building blocks which all have equal status in the theory of the lexicon, except that some are specified as roots, others as affixes, etc. Consider a Quechua form such as (4):

(4) ri - naku - nga - mi
    go REC-PL 3FU AF

'They will (certainly) go.'

In the IA view, the Quechua lexicon contains, among others, the entries in (5):

(5) ri STEM go
    naku AFFIX reciprocal or conjoint action
    nga AFFIX third person future
    mi CLITIC affirmative

These are then combined by "arrangement" rules to form complete words as in (4). Thus the arrangement rules for Quechua will specify that roots go to the left, clitics to the right, and affixes, in a certain order, in between.

In the IP view, words are constituted mostly out of other words, through word formation processes. In that view, there will be one lexical entry, and two affixation rules, to derive (4):

(6) ri 'go'
    rule 1: V \rightarrow V + naku 'conjoint or reciprocal action'
    rule 2: V \rightarrow V + nga 'third person future form of action'

(To the status of -mi in the IP account we return below).

Within generative morphology, which initially inherited the preoccupation with RULES from the Aspects (Chomsky 1965) and Sound Patterns (Chomsky and Halle 1968) model, the first elaborate treatment was within the IP tradition (Aronoff 1976). Only later do we find a shift towards the IA perspective, in Lieber's (1981) thesis and, most explicitly, in Selkirk (1982). In the latter, word formation is stated in terms of phrase structure rewrite rules. McCarthy's (1981) analysis of Semitic vowel alternations and Marantz's (1982) treatment of reduplication, prima facie the strongest cases for an IP model, in terms of abstract morphemes on separate phonological levels then provided the basis for a wholesale shift to the IA model. This shift was made possible by the modularization of the theory to the extent that the putting together of morphemes (the realm of morphology)
was strictly separated from the way these morphemes are then pronounced together (the realm of phonology).

What evidence does Media Lengua have to bear on this issue? Consider first the way (4) would appear:

(7)  
\[ \text{i - naku - nga - mi} \]
\[ \text{go REC-PL 3FU AF} \]

'They will (certainly) go.'

Notice that only the Quechua stem \text{ri-} has been replaced by \text{i-}. The Quechua affixes remain. A brief glance at the examples in (1) to (3) will give the same result: only roots are relexified, not affixes. In a theory in which roots and affixes are both lexical entries, differing only in that affixes obligatorily have a subcategorization feature, this result is not easily explainable. In a theory, however, in which affixes are introduced by linguistic rules, this is what we would expect.

Can we claim, then, that Media Lengua data provide external evidence favoring an IP account of word formation? I do not think the evidence is so straightforward as that. Several issues need to be considered. (1) Do we not find any cases of affix relexification? (2) Could it not be for independent, perhaps semantic, reasons that affixes are not relexified? (3) What about clitics? If they are not relexified, what does this mean for the IA versus IP controversy? In the following subsections I will take up these issues in turn.

1.1 Are no affixes relexified?

I mentioned above that no affixes were relexified. This is not entirely true. We find the Spanish subordinating gerundial affix \text{-ndu} (originally \text{-ndo}) in Media Lengua, in places where in Quechua the “same subject” subordinator \text{-sha} appears or the “different subject” subordinator \text{-kpi}. Media Lengua shows all three forms. In (8) and (9) I illustrate “same subject” subordination with \text{-ndu} and \text{-sha}, respectively:

(8)  
\[ \text{alla -bi -ga entonces -ga artu terreno propio tini -ndu -ga} \]
\[ \text{there LO TO then TO much land own have SUB TO} \]
\[ \text{riku-ya -na, no?} \]
\[ \text{rich become NOM no} \]

'There one could become rich then, having one’s own land, no?'

(9)  
\[ \text{Isi- munda- ga asta kolera muri- sha bini- xu- ka- ni} \]
\[ \text{this from TO even anger die SUB come PR PA 1} \]

'From this I was coming even dying from anger.'

In (10) and (11) different subject subordinations with \text{-ndu} and \text{-kpi} are illustrated:
In Muysken (1981b) I showed that cases such as (10) are quite rare, but sentences such as (8) frequent in the samples.

It is not evident that -ndu is a counterexample to the claim that no affixes are relexified in Media Lengua because there is increasing evidence for an earlier Amerindian-Spanish contact language, spoken throughout Ecuador, the most striking feature of which was the use of -ndo or -ndu as a generalized verb marker (Muysken 1984). It may well be that Media Lengua was influenced by this contact vernacular and took over the -ndu ending from it in certain adverbial contexts, rather than relexifying directly from Quechua in these cases.

There are several other Spanish affixes that occur in Media Lengua as well: diminutive -iiu/-ila and past participle -do. The first occurs in Quechua as well, and can be seen as a borrowing. The second occurs only in Spanish adjectives that were probably taken over as a whole. Therefore we can maintain the generalization that no affixes were relexified.

1.2 Possible semantic reasons for not relexifying

Even if no affixes are relexified, one might argue that this is not due to structural properties of roots and affixes, but to the fact that it would have been too complicated to find the semantic equivalents of Quechua affixes in Spanish. Thus, one might argue, the fact that the Quechua causative suffix -chi- occurs in Media Lengua as well is due to the absence of a similar causative suffix in Spanish.

This argument is very plausible at first sight, but it has several deficits. First, not all Quechua affixes are without a Spanish equivalent. Thus, Quechua second person singular -nki could have been replaced by Spanish -s, Quechua first person plural -nchi by Spanish -mos. In a dialect of Quechua spoken elsewhere the Quechua agentive suffix -k is often replaced by Spanish-derived -dur (from -dor), but not so in Media Lengua. We could extend this list for other suffixes as well. The point is that semantic considerations may have led to the relexification of some affixes, but not others. They cannot account for the fact that no affixes were relexified.
Second, relexification of lexical items has also involved very considerable semantic adjustments sometimes. Take the Quechua impersonal verb yarixa-n ‘it hungers (someone)’. In Spanish you say yo tengo hambre ‘I have hunger’. Now it would appear to be impossible to relexify here. Still, there is a Media Lengua form: the impersonal verb ambri-naya-n ‘I feel like hunger’, containing the Quechua suffix -naya- ‘having a physical inclination’. Similar cases abound. They demonstrate that semantic factors did not prevent relexification of lexical items, even in complicated cases. We find some Quechua items in Media Lengua, but they generally refer to specific culture items, and are perhaps best thought of as loans. An example is Quechua shutichiy ‘baptism’. Here it is not semantic complexity that has prevented relexification, since there is the perfectly equivalent Spanish bautismo, the term the priest would use anyway.

Similar arguments will be discussed in the next section, where we discuss the relexification of grammatical items or function words. There is no general semantic correspondence between Quechua and Spanish grammatical items. Still they are all relexified. To be fair, there is one grammatical item that is not relexified, and this may well have been partly for semantic reasons: the Quechua verb ka- ‘be’. In Spanish there are two verbs, ser and estar, with a semantically complex distribution. This may have been one of the reasons that Media Lengua has consistently maintained ka-. The possible other reason is morphological: estar is somewhat irregular, and ser suppletive, like English ‘be’.

On the whole, we can safely conclude, semantic reasons do not explain why Quechua affixes were not relexified.

1.3 Clitics in Media Lengua

So far I have established that no affixes have been relexified and that this cannot be for semantic reasons. Does this mean that the relexification process operant in Media Lengua constitutes evidence for an Item and Process (IP) view? I think not, and the reason is that clitics are not relexified either. Quechua has a class of phonologically dependent elements, which do not have affix properties (Muysken 1981a; Lefebvre and Muysken 1988) such as being sensitive to the category of the word to which they are attached (Aronoff 1976), being morphological heads (Williams 1981a), and being attachable only to lexical categories (Aronoff 1976). This class includes the elements in (12):

(7) a. -mi affirmative
    -shi hearsay
    -cha dubitative

b. -chu negation, question
c. -pish too, indefinite
   -tik emphatic
   -ri emphatic
   -ma emphatic
   -ga topic

Even though these elements cannot appear as independent words, they have all the syntactic properties of being separately introduced by the phrase structure rules.

In Media Lengua they are not relexified:

(13) isti sabi- nga- ma-bish
    this know 3FU EMP too
    'he will certainly know'

(14) kital- mi uyari- xu- n
    how AF hear RE 3
    'how it sounds'

(15) nustru- ga alla -wa -bi -mi sinta- nchi
    we TO there DIM LO AF live 1pl
    'we live there'

None of them are. It is hard to establish that this is not for semantic reasons, but the considerations presented in section 1.2 hold here as well. In addition, there is evidence from bilingual rural Spanish that there is an equivalent for -pish 'too, indefinite'. In rural Spanish we sometimes find enclitic -tan (from Spanish también 'too') used in the same way as Quechua and Media Lengua -pish:

(16) onde -tan ha ido
    where too has gone
    'where then has he/she gone?'

(17) yo -tan quiero
    I too want
    'I want it too'

There is no reason why this form was not adopted in Media Lengua.

The reason that clitics were not relexified cannot be the fact that they are part of a word formation rule, since they are not. Rather, this suggests that the lexicon is divided into several parts: a stem lexicon and a clitic lexicon, only the former of which participates in relexification. In that perspective, however, there is no reason not to assume the existence of an affix lexicon as well (Halle 1973; Lieber 1981; Selkirk 1982). Thus the evidence from Media Lengua for the IP approach turns out to be invalid.
There is evidence, of course, for a principled separation between roots on the one hand and affixes and clitics on the other.

We may wonder what the implications of this grouping are. The three classes of elements may be assumed, within the IA perspective, to have the following properties:

(18)  

<table>
<thead>
<tr>
<th></th>
<th>roots</th>
<th>affixes</th>
<th>clitics</th>
</tr>
</thead>
<tbody>
<tr>
<td>phonological dependence</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>subcategorized for a base</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>finite list</td>
<td>-</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

While it is true that there is a finite, though extensive, list of affixes in Quechua (and the same holds for clitics), the list is in no way structured. Except for certain subdomains, such as person and case, it rather appears a loose collection of elements. It is hard to see this difference between roots and affixes and clitics as principled. This leaves us for the moment with phonological dependence as the one distinguishing feature that determines whether relexification has taken place or not.

In the next section we will see that lexical meaning is not what sets roots apart from the rest. A number of relexified stems does not have a lexical meaning.

2. Semantic and syntactic properties of lexical entries

In the lexicalist theory developed in the early seventies, increasing attention was paid to the internal structure of lexical entries, and to the types of information that need to be present in such entries. In Jackendoff (1975) entries of the following type were postulated:

(19) /biliv/ phonological representation
    weak conjugation conjugation class
    +V syntactic features
    [+ ___ NP] subcategorization features
    NP1 BELIEVE NP2 semantic interpretation
    [human] selectional restrictions

Much later work, including Williams (1981b), has tried to derive the properties of lexical items in (19) as much as possible from independent principles. In addition, work of Stowell (1981) and others has been directed at deriving properties of phrase structure rules directly from the lexicon. In this section I will discuss Media Lengua from the perspective of the structure of the lexical entry. In 2.1 I consider lexical versus grammatical meaning, in 2.2 the process of lexicalization and semantic specialization of derived forms, and in 2.3 the directionality of government.

Before going on, I should stress that the very process of relexification confirms the conception of the lexical entry as a loosely organized bundle
of at least three feature clusters: the phonological representation, the morphological characteristics, and the syntactic and semantic properties. This should not be interpreted, however, as directly confirming the notion of dual representations for words that we find in work of Sproat (1987) and others. Media Lengua Spanish-derived roots are the base for Quechua suffixes, and the suffixation process is governed by the morphosyntax of Quechua.

2.1 Lexical and grammatical meaning

Media Lengua relexification takes place on the basis of meaning correspondences. A Spanish stem as close as possible in meaning to the Quechua original is used to replace it. The question now is how closeness in meaning is determined. It turns out we have to distinguish here between lexical meaning and grammatical meaning. The former can be determined by reference to some extralinguistic entity, the former only by reference to the language systems themselves. It turns out that relexification is feasible, with all the difficulties mentioned, for lexical items, but operates in a very incomplete manner for grammatical items. In Muysken (1988) I illustrated this for demonstratives. Here I will try to argue the same point for question words. In (20) the relevant forms are given for Quechua, Spanish and Media Lengua:

<table>
<thead>
<tr>
<th>Quechua</th>
<th>Spanish</th>
<th>Media Lengua</th>
</tr>
</thead>
<tbody>
<tr>
<td>pi</td>
<td>quién</td>
<td>kin</td>
</tr>
<tr>
<td>ima</td>
<td>quéén</td>
<td>ki/inki</td>
</tr>
<tr>
<td>mayxan</td>
<td>cuál</td>
<td>kwal</td>
</tr>
<tr>
<td>mashna</td>
<td>cuánto</td>
<td>kwantu</td>
</tr>
<tr>
<td>a como</td>
<td>a komo</td>
<td>'for how much'</td>
</tr>
<tr>
<td>ima uras</td>
<td>cuándo</td>
<td>kwandu</td>
</tr>
<tr>
<td>qué horas</td>
<td>ki uras</td>
<td>'at what time'</td>
</tr>
<tr>
<td>ima-munda</td>
<td>por qué</td>
<td>purki</td>
</tr>
<tr>
<td>ima-shna</td>
<td>cómo</td>
<td>(purki-munda)</td>
</tr>
<tr>
<td>may+case</td>
<td>(d)ónde</td>
<td>undi+case</td>
</tr>
</tbody>
</table>

These examples show that Media Lengua forms, with the exception of the form for 'where' are modeled on Spanish rather than Quechua patterns. If it had been the reverse, we would have found forms such as *(in)ki-munda* and *(in)ki-shna*, which do not occur at all. The form *purki-munda*, which follows both Quechua and Spanish, occurs only once.

2.2 The process of lexicalization

One of the differences between Quechua and Spanish is that the latter has more separate verbal roots, while the former has fewer roots and tends
to form new verbs through a complex system of verbal derivational suffixes. Thus we have:

(21) Quechua       Spanish
    apa-             llevar     'take'
    apa-mu-          traer      'bring'

Here the cislocative suffix -mu- is added to the original root to create a cislocative verb whose meaning in Spanish (and English) is expressed with a separate verb.

Sometimes the derived meaning is simply compositionally determined, and often the root + affix combination has undergone semantic specialization or drift. We may hypothesize that the lexicalized forms are replaced in Media Lengua by Spanish simplex roots, and the compositional forms by forms directly patterned on the complex Quechua original, or include both a Spanish root with the complex meaning and the Quechua derivational suffix. In Muysken (1981) data are presented which illustrate this pattern for the Quechua verb riku- 'see' and derived forms. Here a few examples are given directly related to the forms in (21). Sentences (22) and (23) illustrate the use of the Spanish-derived forms trayi- and lleba-, respectively:

(22) intonsi lindu radiyu- da trayi- shka
    then nice radio AC bring SD
    'Then it turned out they'd brought a nice radio.'

(23) awa -da ahi -munda lleba- nga zin Ambatu- mun
    water AC there from take 3FU they say Ambato to
    'From there they will take the water to Ambato, they say.'

In (24) an example of a redundant form is given:

(24) grabadora -da trayi- mu -ngi
    tape recorder AC bring CIS 2
    'Bring the tape recorder.'

The interaction between relexification and derivational morphology needs to be studied in much more detail, for different groups of verbs. The main point here was that relexification can be used to study the degree of lexicalization of a particular root + affix combination.

2.3 The lexicon and syntax: the directionality issue

As was briefly mentioned above, research in the early eighties has attempted to drive properties of phrase structure such as order specifications from properties of the lexicon (Stowell, 1981, and later work by various authors). This was done in terms of directionality of government: verbs and
prepositions govern rightward in English, therefore English has [P NP] and [V NP] structures.

We will now see what evidence Media Lengua adduces on this issue. It should not surprise us that Media Lengua shows the same percentage of XV (80%) versus VX (20%) sentences as the Quechua spoken in the same region. Media Lengua verbs are Quechua verbs, for all practical purposes, and they are marked with Quechua affixes. Quechua verbs govern leftward. Thus we expect Media Lengua verbs to govern leftward as well.

With elements of the category P the situation is more complicated. Quechua has a few postpositions, such as k’ipa ‘after’, and shina ‘like’. When these are relexified, we find postpositions in Media Lengua, as with the underlined constituents in (25) and (26), respectively:

(25) miza despuesitu kaza-mu i -naku -ndu-ga, ahi -bi boda
Mass after home to go REC-PL SUB TO there LO feast
da -naku -n, ahi -bi bayla -naku -n
give REC-PL 3 there LO dance REC-PL 3

‘Going home after Mass, there they give a feast and dance.’

(26) asi -lla- di kumu bos, bos kwenta- lla- di gringu
thus DEL EMP like you you like DEL EMP gringo
kunusidu -guna tini-n bastanti miu Rosalina
acquaintance PL have 3 plenty my Rosalina

‘My Rosalina has plenty gringo acquaintances just like you, like you precisely.’

As kumu ‘like’ (derived from Spanish como) (26) shows, however, we also find a few prepositions in Media Lengua. Thus we have entre ‘among’ and asta ‘until’, in (27) and (28), respectively:

(27) nustru- ga entre seys -mi ga-nchi
we TO among six AF be 1pl

‘There are six of us.’

(28) solo isti syera asi, kazi isti Machachi-munda
only this Highlands thus almost this Machachi from
asta Chimborazo Riobamba-gama, barato paga-naku -n
until Chimborazo Riobamaba-until cheap pay REC-PL 3

‘Only here in the Highlands, almost from Machachi here to Chimborazo, to Riobamba do they pay low wages like this.’

The Quechua equivalents of this words are affixes rather than separate elements.
I interpret the contrast between (25)–(26), on the one hand, and (27)–(28), on the other, to support the lexically-based view of directionality. There turns out to be no general \([NP P]\) or \([P NP]\) order in Media Lengua. The P elements that are relexified from Quechua govern leftward, the elements newly introduced from Spanish govern rightward. Similar results, that I will not discuss here, can be obtained from the introduction of Spanish complementizers.

3. Conclusion

This concludes my discussion of some of the evidence that can be gathered from Media Lengua for our conception the organization of the grammar. In what sense can it be viewed as external evidence? It is not external in the sense that it refers to the interaction with a non-linguistic (cognitive) system, which I think would be the proper definition of what external evidence for the organization of the grammar is. It is only external that for the interpretation of the Media Lengua data, two other linguistic systems, Quechua and Spanish, are taken into account. But inasmuch as all recent grammatical research is inherently comparative in nature, this does not make this type of research special.

I will conclude with raising the (open) question of whether we can think of Media Lengua as the result of a kind of word internal code mixing. Schematically this possibility is presented in (29), where the superscripts \(s\) and \(q\) refer to the Spanish, respectively Quechua language index of roots and affixes, and the subscript \(V\) to the categorial identity of an element:

(29)

\[
\begin{align*}
\text{ROOT}_{V^s} & \quad \text{AFFIX}_{V^q} \\
\end{align*}
\]

The affix is the head (Williams 1981a) and determines the language index of the whole word. The advantage of this view is that it relates the possibility for relexification in Quechua to the highly agglutinative character of the language: since there will always be affixes to mark a word as belonging to the language, the root can be derived from a different language. The disadvantage of this view is that it cannot account for cases such as \textit{despuesitu} ‘after’ in (25). This form behaves as a Quechua postposition, even though it is entirely Spanish. We would have to postulate a \(\emptyset\) Quechua case affix on it, marking it as Quechua. I will leave this issue open for further research.
Appendix:
List of abbreviations used in glosses

1, etc. first person etc. marker
AC accusative
AF affirmative
AG agentive
CIS cislocative (movement towards speaker) marker
DEL delimitative
DIM diminutive
EMP emphatic marker
FU future
LO locative
NOM nominalizer
PA past
PL, pl plural
PR progressive aspect
RE reflexive
REC-PL conjoint action: reciprocal or plural
SD sudden discovery tense
SUB adverbial subordinator
TO topic marker

REFERENCES

Aronoff, Mark

Chomsky, Noam

Chomsky, Noam, and Morris Halle

Lefebvre, Claire, and Pieter Muysken

Lieber, Rochelle

McCarthy, John

Marantz, Alec
Muysken, Pieter


Selkirk, Elizabeth


Sproat, Richard


Stowell, Timothy


Williams, Edwin
