THE ROLE OF BORROWING IN THE JUSTIFICATION OF PHONOLOGICAL GRAMMARS[1]

Larry M. Hyman
Department of Linguistics
University of California, Los Angeles

Both in syntax and phonology one of the chief concerns in the process of linguistic analysis is the justification of grammars. If this problem is somewhat more in focus in phonology, it is perhaps because in phonology we are able to more closely approximate a workable solution and indeed argue the merits of alternate proposals to a set of linguistic data, whereas we consider ourselves lucky in syntax if we are able to reach any workable hypothesis at all. Thus it is the case that although no one overlooks the need to justify syntactic analyses, it is within the domain of phonology that questions of this sort have been most deeply discussed and developed.

1. The Need for Empirical Justification

One form of argument used to justify phonological descriptions relies on the so-called 'simplicity' criterion. Two analyses are discussed and one is rejected on the basis that in some well-defined (or not so well-defined) way it is less simple than the more acceptable solution. Only feature counting has been explicitly incorporated into the evaluation metric. Since this often leads to counter-intuitive results or has been unable to provide unique

---

[1]I would like to thank Victoria Fromkin for helping me revise an earlier version of this paper. I am also indebted to Isaac George for many suggestions and bits of information.

adequate solutions, it is not infrequent that we find phonologists justifying their solutions on the basis of a number of varying forms of support: the number of underlying segments, pattern considerations, the number of phonological rules, historical evidence, or the so-called 'naturalness' of various points at issue, this last argument often falling under 'markedness theory' (yet to be developed). Thus, to consider a specific example, in a language such as Igbo, which exhibits a general CVCV pattern structure, one might quibble over whether verb roots (which are otherwise typically CV) such as kwé 'believe', gwá 'tell' and ŋwú 'die' are best analyzed as /CwV/, as in Standard Igbo Orthography, or as /C^wV/, where in both cases C is limited to the velar series, /k/, /g/ and /ŋ/. By recognizing rounded velars (/C^w/) as phonological units we are able to maintain the otherwise exceptionless (ignoring syllabic nasals which possibly have a NV source) CV structure of Igbo. Similarly, in this solution we do not have to account for what would be a strange distributional restriction on /w/ (only after velars), since we take care of this in our underlying segments. However, a solution recognizing /Cw/ requires fewer underlying segments, as it more thoroughly utilizes the independently justified /w/ phoneme.[2] How, if not by mere

[2]The choice of one of these solutions over another may seem trivial. However this choice can have great cognitive significance. If I were to pronounce kwé as ké, would a native speaker tell me that I pronounced a sound wrong (the first solution) or that I left a sound out (the second solution). Since my informants have been greatly influenced by Standard Igbo Orthography (as well as by English) such experimentation was infeasible.

assertion, are we to know which of these criteria is more critical in the construction of a phonology? The problem is that we have no way to interrelate various possible complications in a phonological component. What we need to do is further design our decision-making evaluation metric in such a way as to reflect what is really happening in human language. I strongly believe we should turn our attention toward developing a set of criteria that in a more direct, perhaps more experimental, way can be used to verify the psychological reality of a given proposal. Thus Kiparsky (1968) states: 'what we really need is a window on the form of linguistic competence that is not obscured by factors like performance . . .' (p. 174). He then concludes: "In linguistic change we have precisely such a window." It is through the work of Kiparsky and others that justification for many of the notational conventions of generative phonology has been obtained. Linguistic change is one of the possible considerations open to the phonologist desiring psychological verification.

I am assuming a certain interpretation of the notion 'explanation' (or, if you will, 'explanatory adequacy'). To state that a certain solution is superior to another on the basis of simplicity can be explanatory, if and only if the simplicity metric itself has been empirically justified. As an illustration, the failure of Halle's principle of feature counting to define natural classes has led to markedness theory, which (at least in its current state) fails in just this sense to be explanatory. Let us take an example in French from Schane (1968b, see also 1968a).

Schane discloses two sources for [ã] in French. One derives this low back nasalized vowel from /An/ (capitalized vowels are [+tense]), as well as from underlying /ɛn/. These two sources are seen in the alternations (Schane's examples):

[peizã] 'male peasant' [peizan] 'female peasant'

[žãr] 'genus, kind' [ženerik] 'generic'

In the discussion of these phenomena Schane is concerned about the non-uniqueness of phonological representations (what Lightner has termed 'lexical overlap'). In this regard he states that we do not know whether to derive [vãdr] 'to sell' from /vAndr/ or /vɛndr/. His concern is not unrelated to the problem of justification and explanation. A phonology that is explanatory (descriptively adequate) will permit only <u>one</u> solution per problem, and in so far as we have indeterminate cases, our phonological metatheory has not been amply well defined. To remedy this problem of non-uniqueness, Schane invokes the concept of 'markedness'. He proposes that in cases of non-unique phonological representation we choose the underlying segment that is 'least marked'. So, in accordance with the current theory of markedness which hypothesizes /A/ as less marked than /ɛ/, we recognize the underlying form /vAndr/, despite its orthography <u>vendre</u> and its historical derivation.

Schane's example provides us with a solution, however with its reliance on markedness it is explanatory only to the extent that the particular markedness convention which he uses is itself explanatory. Even within this framework, the proposed theory which Schane

uses does not solve all similar problems. In the same article, for example, Schane talks of a closely related problem: phonetic [ɛ̃] is derivable from both /In/ and /En/ as seen from the examples:

[divɛ̃]	'divine' (m.)	[divin]	'divine' (f.)
[fɛ̃]	'end'	[finir]	'to finish'
[sərɛ̃]	'serene' (m.)	[sərɛn]	'serene' (f.)
	cf. [serenite]		
[plɛ̃]	'full' (m.)	[plɛn]	'full' (f.)
	cf. [plenitüd]		

Given these facts, how are we to decide what should be the underlying representation of <u>scinder</u> [sɛ̃de] 'to divide'? Having recourse to marking conventions, as they now stand, the underlying representation is unquestionably /sInde/, since /I/ is less marked than /E/. A problem arises however when we try to account for certain dialects of French where the distinction between [ɛ̃] and [œ̃] has been lost. Thus we find:

'<u>Standard French</u>'

[œ̃]	'one' (m.)	[ün]	'one' (f.)[3]
[brœ̃]	'brown' (m.)	[brün]	'brown' (f.)

'<u>Parisian Dialects</u>'

[ɛ̃]	'one' (m.)	[ün]	'one' (f.)
[brɛ̃]	'brown' (m.)	[brün]	'brown' (f.)

[3] [ü] is used to represent IPA [y], i.e. a high, front rounded vowel.

In these latter dialects there is then no phonetic distinction between [brɛ̃] (spelled brun) 'brown' (m.) and [brɛ̃] (spelled brin) 'blade of grass'. As soon as we introduce this third possible source of [ɛ̃] our reliance on markedness falters. Chomsky and Halle state (1968) that there is no reason to designate /i/ as more or less marked than /u/, and since [ɛ̃] now is to be derived from /U/ (the normal source of [ü], see Schane, 1968a) as well as /I/, we have no principled way to arrive at a unique solution.[4]

The conclusion is that no matter how nice one's analysis turns out to be, unless there is some empirical justification the possibility still remains that the analysis is a mere formalism without 'psychological reality'. From the Igbo and French examples it is clear that we must look for a 'window' which will help us determine those constraints in the metatheory permitting just the correct solutions. I would like to suggest that borrowing provides one such window. The remainder of this paper will be devoted to determining to just what extent 'foreign sound adaptation' can be utilized to justify phonological grammars. A specific case where Nupe, a Kwa language of Central Nigeria, borrowed from neighboring Yoruba and Hausa will exemplify the discussion to follow.

---

[4] Although a more concrete phonology might recognize /ü/ and therefore scinder would be uniquely marked as /sInde/, the possibility for non-uniqueness is still inherent in the Chomsky and Halle marking conventions.

2. Past Approaches

This paper attempts to view the phenomena of borrowing and of foreign sound perception in its broader sense in the light of generative phonology. Both of these processes will be treated as one. In my dealing with the borrowing phenomenon, I shall be interested only in those cases where the foreign material has been made to conform with the phonological properties of the interpreting language. Since I shall therefore be ignoring cases of what has been traditionally termed 'phonological import' (i.e. cases where the borrowing language has, in violation of its native phonology, adopted new sounds or sequences (see especially Haugen (1950)), we can view 'borrowing' as 'institutionalized foreign sound adaptation'. Similarly, I shall ignore cases where a particularly talented speaker of the interpreting language accurately perceives and reproduces foreign sounds in the process of foreign language acquisition. I shall be interested only in those cases where the forms resulting from both forms of contact (borrowing and foreign language learning) are possible lexical items in the first language. Those lexical items which would require the feature [+foreign], for example, will not be considered. The term 'lexicalization' (which I owe to Kalon Kelley) will be used to refer to how these borrowed forms will be represented in the lexicon.

One of the chief motivations for this study is derived from the following idealized possibility: if we have a theory of borrowing (encompassing the perception of foreign sounds in other

contact situations), then by analyzing the occurring borrowed forms and/or running the necessary tests on foreign sound perception, various aspects of the internalized phonology can be determined. This possibility rests on the proposition that the phonological properties of a language largely determine both the phonological shape and the phonetic realization of a lexicalized loan-word. Not all theorists have accepted this conception of borrowing, as we shall see. The question that I shall be foremost concerned with is: if a foreign word is totally assimilated, what determines its lexicalized shape?

The tacitly (or not so tacitly) assumed explanation that we find in the pre-structuralist literature, and the one we encounter in much of the structuralist literature as well, is generally designated as 'phonetic approximation'. The form this argument usually takes is that speakers of a language, in hearing a foreign sound, replaces that sound with the most closely related phonetic (or perhaps phonemic) unit in its inventory. Thus Hermann Paul writes:

> Um eine fremde Sprache exakt sprechen zu lernen, ist eine Einübung ganz neuer Bewegungsgefühle erforderlich. So lange diese nicht vorgenommen ist, wird der Sprechende immer mit denselben Bewegungsgefühlen operieren, mit denen er seine Muttersprache hervorbringt. Er wird daher in der Regel statt der fremden Laute die nächstverwandten seiner Muttersprache einsetzen und, wo er den Versuch macht Laute, die in derselben nicht vorkommen, zu erzeugen, wird er zunächst fehlgreifen. (p. 394)

According to this interpretation speakers of a language L_1 replace the sounds of a second language L_2 with those sounds closest to them in L_1. In this analysis, then, the nativization of foreign sounds

could and should be explained in terms of physical phonetics. The manner of nativization should be reflected in (and predictable from) the distinctive features and scalar values that make up the phonetic quality of the individual segments. A contrastive analysis of the occurring phones of L_1 and L_2 should then suffice in such a 'theory' to explain how (and demonstrate why) the phones of L_1 will be realized by speakers of L_2, and the phones of L_2 will be realized by speakers of L_1.

That this explanation is at the very least inadequate is seen from the following observations: a Frenchman attempting to render the English sound [θ] will typically realize his efforts as [s]; correspondingly, he will render English [ð] as [z]. This makes sense, according to the argument of phonetic approximation, since [s] is the closest sound French has to [θ], just as [z] in French most closely approximates English [ð]. In terms of phonetic features, the English sounds are [-strident], whereas the French counterparts (substitutes) are [+strident]. Presumably (accepting for these purposes the Chomsky-Halle feature system, 1968) these sounds will agree in all other phonetic feature specifications.[5] This constitutes, then, a physical phonetic explanation of the assimilation of sounds from one language (English) into another (French).

---

[5]It is true that phonetically there are further differences between [s] and [θ] in that the scalar values of certain similar features will differ, e.g. [s] is 'less anterior' than [θ].

However, as soon as we look at the way in which another language L_3 interprets the sounds of L_1, we run into theoretical difficulties. Speakers of Serbo-Croatian, for instance, will tend to realize [θ] as [t] (and not [s]), just as they substitute [d] (not [z]) for [ð]. Once again we can look at these facts as advocates of the theory of phonetic approximation. The English sounds in question differ from their Serbo-Croatian counterparts only in that the former are [+cont], while the latter are [-cont]. We conclude from this that the speakers of Serbo-Croatian replace [θ] and [ð] by those segments closest to them in their language--or do they?

The problem with this approach is that these phenomena cannot be explained in purely phonetic terms. If they could be, then we would expect that French be like Serbocroatian and substitute [t] and [d] for the English sounds; or that Serbocroatian be like French and substitute [s] and [z] for the English non-stridents, since both languages bring equivalent segments to the borrowing task. The only possible chance of salvaging such a theory of phonetic approximation is to claim that [θ] is 'equidistant' in some sense from both [s] and [t] (perhaps in that in both cases the substituted sound differs by only <u>one</u> phonetic feature: [strident] in the French example and [cont] in the case of Serbo-Croatian). But if this were the case then we would expect to find with respect to English [θ], equal numbers of Frenchmen substituting [t] and [s], and equal numbers of Serbo-Croatian speakers substituting [s] and [t]. This doesn't appear to be the case. Thus even if we assume (without

foundation) that [θ] is equidistant from [s] and [t], we still have to explain why one nation chose [s] and the other [t].[6]

Instead of adhering to absolute phonetic criteria we often hear the notion 'phonetic approximation' closely tied in with the more intangeable notion of Sprachgefühl. It appears more realistic to say that a language adopts that sound that is 'felt' to be closest to the prototype. Thus a Frenchman feels that English [θ] most closely resembles his own [s], while a speaker of Serbo-Croatian feels that [θ] most closely resembles his [t]. As soon as we accept this notion of Sprachgefühl we can no longer adhere to a physical phonetic interpretation of foreign sound assimilation. The unequivocal conclusion that we are led to is that foreign sound adaptation is mental in nature. The only way in which we can explain why a sound X from L_2 is realized as sound Y (and not as sound Z) in borrowing language L_1 is by having recourse to the phonological facts of L_1 and the

---

[6] In this discussion I have ignored the possibility that [θ] and [ð] are borrowed as such as a result of the manner in which these sounds are taught in the schools. European languages seem to present this difficulty of interference that often vitiates the otherwise automatic process of foreign sound reproduction (spelling presents still another problem of interference). Whether or not I am justified in dismissing this possibility of interference is a question that can be empirically investigated, I would think. However, I do find it significant at least that in one country (France) the teachers tell the students 'it's like our [s]', while in another (Yugoslavia), apparently, the students are told 'it's like our [t]'. If all we have is the behavior of pedagogues to go by, then there must be an explanation for their varying performances. Someone (and in turn something) has to be responsible for the observed divergences.

phonetic data of L_2. Differing phonological properties are then responsible (at least in part) for different nativization processes. The French and Serbo-Croatian example is far from being resolved. However, if it can be shown that there is no extra-linguistic interference (see footnote 6), then we must look into the individual phonologies (perhaps into the individual feature hierarchies?) for an explanation. But the explanation must be phonological.

This is precisely the change of emphasis that came about with the advent of structuralism (although it is by no means rare to hear one speak of phonetic approximation in cases where no other explanation presents itself). When the phoneme was introduced there was a reinterpretation of the data of borrowing. Bloomfield noted simply: 'In phonetic substitution the speakers replace the foreign sounds by the phonemes of their language' (p. 446). Foreign sounds are not reanalyzed as isolated phenomena, but instead fit into the phonological system (as opposed to the phonetic system) of the borrowing language.

That the notion of phonetic approximation was insufficient to phonemicists is evident in the statement of Haugen:

> Neither the speaker himself nor the linguist who studies his behavior is always certain as to just what sound in his native tongue is most nearly related to the model. Only a complete analysis of the sound system and the sequences in which sounds appear could give us grounds for predicting

> which sounds a speaker would be likely to substitute in each given case. (p. 215)[7]

The example given by Haugen is that the Yaqui Indians reproduce Spanish estufa 'stove' as [ehtúpa] thereby substituting [h] for [s] because the normal 'allophone' of /s/ before /t/ and /k/ is [h].[8] (He doesn't discuss the f/p discrepancy.) Thus as the work of Weinreich best illustrates (1953, 1957), the allophonic distribution of the phonemes in L_1 will greatly influence the manner in which the

---

[7] I would disagree with much of this statement of Haugen's. To a certain extent a speaker of a language can predict how a foreign sound will come into his language, especially if he thoroughly understands what that sound is. How consistent this is, I'm not sure, but it is definitely something we wouldn't want to dismiss. It may well be part of the speaker's tacit knowledge of his language.

[8] Rosario C. Gingràs informs me that certain Caribbean dialects of Spanish contain a rule of the form:

$$s \rightarrow h / \_\_ \left\{ \begin{matrix} C \\ \# \end{matrix} \right\}$$

and furthermore that in some dialects this postvocalic [h] drops out completely. Thus to consider certain English loanwords in Panamanean Spanish:

Eng.	astronaut	→	P.Sp.	ahtronauta/a:tronauta
Eng.	Chase-Manhattan	→	P.Sp.	čehmayxátaŋ/če:mayxátaŋ
Eng.	Astoria	→	P.Sp.	a:toria

Whether or not the Spanish spoken in Yaqui territory (Arizona and Northern Mexico) contains such a rule is uncertain to me, although I will assume that Haugen was aware of this possibility and that it doesn't contain such a rule. In any case, the argument still remains that a presented [st] or [sk] will turn up [ht] and [hk], respectively, in Yaqui, and the reality of the process can at least be seen in the Spanish examples above.

borrowed sounds of L_2 are 'phonemicized', and he makes an attempt to classify the possible types of rephonemicization (overdifferentiation, underdifferentiation, etc.). Polivanov, a Prague functionalist, sums up the phonological relevance to borrowing:

> En entendant un mot inconnu étranger . . . nous tâchons d'y retrouver un complexe de nos représentations phonologiques, de le décomposer en des phonèmes propres à notre langue maternelle, et même en conformité de nos lois de groupement des phonèmes. (p. 80)

With this information available and with the Yaqui example in mind, Haugen should have been led to postulate that sounds are borrowed on the basis of phonemic approximation; that is, a language first finds the closest phoneme that encompasses the phonetic quality of the foreign sound, and then this foreign sound is appropriately phonemicized and subject to the phonological constraints of that phoneme's members. With this type of analysis of borrowing, Haugen could have more readily accounted for the Yaqui example in the following manner (instead of making it appear so exceptional): 1) the perceived [s] is an allophone of /s/ in Yaqui; 2) it is therefore phonemicized as /s/; 3) now that the Spanish [s] is Yaqui /s/ it is subject to the phonological rules of Yaqui and will be realized as [h] before /t/ and /k/, as in native words. Thus the lexicalized form of Spanish estufa is /estúpa/ with a phonetic realization of [ehtúpa]. This would account for why [h] was used, when I assume there was some other element appearing in that context that would have more closely approximated [s] phonetically.

3. A Hypothesis

In the foregoing discussion we dismissed the inadequate theory of phonetic approximation. It was shown that the analysis of borrowing can be fruitful only if the phonological properties of the borrowing language are taken into account. One of the few structuralists to fully recognize the significance of foreign sound perception, its phonological nature, and its role in demonstrating the reality of linguistic descriptions was Harris (1954), who wrote:

> Clearly, certain behaviors of the speakers indicate perception along the lines of the distributional structure, for example, the fact that while people imitate nonlinguistic or foreign-language sounds, they repeat utterances of their own language (i.e. they reproduce the utterance by substituting, for the sounds they heard, the particular corresponding variants which they habitually pronounce . . . . There are also evidences of perception of sounds in terms of their morphophonemic memberships. (pp. 36-37)

The next question to be resolved is whether taxonomic phonemics can provide us with all of the phonological information necessary to account for lexicalization. I shall now present in the remaining sections evidence from Nupe to support my contention that it is only through the apparatus provided by generative phonology that the facts of lexicalization can be truly accounted for.

In the preceding section we rejected the impoverished theory of phonetic approximation in favor of what we termed 'phonemic approximation'. Since others working within the framework of generative phonology (esp. Halle, 1962; Chomsky, 1964; Postal, 1968) have shown the inadequacies of phonemic analysis it is not surprising to find that where the phonological processes of a language escape a

general explanation within the phonemic approach, the facts of lexicalization are also not readily explained. Let us consider a sample case. In Nupe the distribution of stridents is virtually unmanageable by phonemic theory. A rule of strident palatalization derives the appropriate 'allophones' before front vowels as shown in the rule (SR):

1. $\begin{bmatrix} s \\ z \\ ts \\ dz \end{bmatrix} \rightarrow \begin{bmatrix} \check{s} \\ \check{z} \\ \check{c} \\ \check{j} \end{bmatrix} / \_\_\_ \begin{Bmatrix} i \\ e \\ \varepsilon \end{Bmatrix}$

That is, phonemic /sī/ 'to buy' is realized as [šī] (actually, it is further palatalized by a more general palatalization rule (PR) to become [$\check{s}^{y}$ī], though I shall ignore this detail when it is not relevant to the discussion; see footnote 9). The remaining three stridents act accordingly. As expected we obtain the dental series before /u/ and /o/. The problem is that before /a/ we obtain both the dental series and the palatal series as seen in the examples:

2. [čà] 'to begin'[9]

   [tsà] 'to choose'

---

[9]Actually [$\check{c}^{y}$à] would be more accurate phonetically and is crucial for phonological reasons. Whenever the predictable labialization or palatalization from LR/PR (see 25) is not crucial to the example, the unassimilated consonant will be given.

Three discrete level tones are recognized for Nupe: /á/ 'high', /ā/ 'mid' and /à/ 'low'. These are realized identically in the surface phonetics, except for the rising tone [ǎ] to be dealt with below. Falling tones [â] also occur. On the other hand, the Hausa 'high' and 'low' tones are realized differently phonetically (in a complex system of terrace level tonemics) and this accounts for many of the discrepancies between the Hausa and Nupe tones in borrowing. In the cited Hausa and Yoruba forms, the phonological tones are given.

Since we have contrastive pairs such as the above, we are forced within the phonemic approach to recognize <u>eight</u> phonemes: /s/, /z/, /ts/, /dz/, /š/, /ž/, /č/, and /ǰ/. This position is speciously supported by the process of reduplication which is represented by the rule (RED):

3. $$\text{RED} \rightarrow C_1 \begin{bmatrix} +\text{high} \\ \alpha\text{round} \\ \alpha\text{back} \\ \text{MID} \end{bmatrix} / \underline{\quad} \; C_1 \; [\alpha\text{round}]_V$$

where $C_1 = C_1$

That is, the reduplicated vowel is either /i/ or /u/, depending on the roundness of the stem vowel and the reduplicated syllable receives MID tone. This rule is summarized by the following reduplicated forms:

4. /gí/ 'to eat' → [gīgí] 'eating'

/gē/ 'to be good' → [gīgē] 'being good'

/gú/ 'to puncture' → [gūgú] 'puncturing'

/gò/ 'to receive' → [gūgò] 'receiving'

By this rule the normal reduplicated form for a verb of the structure /Ca/ would therefore be [CīCa]. The question is: what happens in the case of the above stridents? Is the reduplicated form of /tsà/ 'to choose' going to be [tsītsà] or [čītsà]? It turns out that the former is the case. Since we derive [čīčà] 'beginning' from /čà/ 'to begin' we find [ts] and [č] contrasting also before /i/. For this reason, according to the phonemicist's approach, we have another

argument for recognizing an underlying dental series and an underlying palatal series of stridents. Similar to the example just given we find that /sá/ 'to cut' reduplicates as [sīsá]. Since /sī/ reduplicates as [šīšī], both /s/ and /š/ must be posited in a phonemic inventory.

Turning now to the problem of lexicalization, how should we expect the sequence [si] to be assimilated into Nupe? We have just recognized above (assuming the phonemic analysis just proposed) both a phoneme /s/ and a phoneme /š/. On the basis of phonemic approximation (and phonetic approximation as well) we would predict that a foreign sequence [si] will be analyzed as [si] in Nupe. This does <u>not</u> however turn out to be so. The normal (apparently exceptionless) nativized version of foreign [si] is [ši] (or, once again, more accurately: [$š^{y}$i]). Thus we observe the manner in which Yoruba 'sixpence' (which Yoruba borrowed from the English word 'six') comes into Nupe:

5. Yoruba sísì → Nupe šíšì

A Nupe who speaks Yoruba (or any other appropriate language) with a Nupe accent will reproduce the [si] sequences of that language as [ši] and the [se] sequences as [še]. Yet the sequence [si] appears

in Nupe as in the form [sīsá] 'cutting'. Thus the theory of phonemic approximation is equally invalid.[10]

In the references cited above, the autonomous phoneme was shown to be a highly questionable unit. In these writings it is dismissed in favor of a more abstract 'morphophoneme' or 'systematic phoneme' which by means of phonological rules is mapped into the phonetic forms of the language. Is it then to be expected that lexicalization takes place on the level of 'systematic phonemic approximation' or by some other related process? At this point I would like to propose, as a natural tendency, the following hypothesis:

6. FOREIGN SOUNDS ARE PERCEIVED IN TERMS OF UNDERLYING FORMS.

As such, they are subject to the phonological constraints of the system--unless the foreign word is to be singled out as an exception (i.e. [+foreign]) and not subject to these constraints. (It is interesting to note that in Nupe most borrowed forms are reinterpreted as non-exceptional Nupe formatives.) Thus a candidate for 'new lexical item' 1) satisfies the MSC's of that language; and 2) has a surface phonetic manifestation that is in accord with the phonological rules of the language.

---

[10]Some might argue that this observation is only natural because [si] occurs only in RED in Nupe. However anyone who tries to use this as a counterargument is in agreement with my position--which is: in order to tell what the lexicalized form will be, we must refer to the borrowing language's phonological properties, and not to its speakers phonetic habits. A Nupe can pronounce [si], but only uncer certain phonological conditions. A phonetic or phonemic analysis fails to perceive this.

Thus if we take a second look at the strident situation in Nupe as proposed in Hyman (1970, forthcoming), all cases of palatal stridents + [a] (e.g. [čà]) are reanalyzed as <u>dental stridents + /ɛ/</u> (e.g. /tsɛ̀/). That is, we propose an abstract entity /ɛ/ that is however realized phonetically as [a] after it has allowed /s/, for example, to palatalize (by means of SR) to [š], and then further palatalize (by means of PR) to [š^{y}]. Similarly, /tsɛ̀/ 'to choose' is realized as [čà] (or [č^{y}à] in more detailed phonetics), and so forth. By recognizing /ɛ/ (and there are other arguments) we are able to account for all palatal stridents as derived from the underlying dental series. Reduplication also represents no problem for generative phonology, since in order to obtain the form [ts$\bar{\text{i}}$tsà] 'choosing' or the form [s$\bar{\text{i}}$sá] 'cutting' we have only to order the strident rule (SR) <u>before</u> RED. Reduplication will then create new sequences of dental strident + [i], but the strident rule will no longer be operative. The complete derivations for 'beginning' and 'choosing' will therefore look something like:

7. /tsɛ̀/ → čɛ̀ → č$\bar{\text{i}}$čɛ̀ → č$^{y}\bar{\text{i}}$č^{y}ɛ̀ → [č$^{y}\bar{\text{i}}$č^{y}à]

SR RED PR ɛ→a

/tsà/ → tsà → ts$\bar{\text{i}}$tsà → [ts$^{y}\bar{\text{i}}$tsà]

SR RED PR

It is only by recognizing a deeper level than the autonomous phonemic level and by conceptualizing the phonological component as a system of rules relating abstract underlying forms to

surface phonetic realizations that borrowing can be coherently handled. Thus, to return to our example (of which there are many more), in the case of Yoruba [sísì] we hypothesize that the Nupes interpret this foreign word as an underlying form--that is, as /sísì/, and then all of the phonological rules are applicable. In other words, a word or formative when borrowed becomes part of the lexicon with a phonological representation which is subject to the morpheme structure conditions and then all of the phonological rules. These new words then consist of phonological strings which are composed of underlying phonological segments already in the language, and not new phonological segments or segments which may be closer to the phonetic output in all cases. (Note how important it is for lexicalization that these two levels, the systematic phonemic and the systematic phonetic, be expressed in the same terms.) Since Yoruba 'sixpence' is 'inserted' into the lexicon (i.e. is lexicalized) at the point where the morpheme structure conditions are in effect and the phonological rules are about to apply, /sísì/ is 'created' early enough to undergo the strident rule. Any other theory would fail to account for this fact. For it is only by recognizing a set of ordered rules that this analysis is possible.

Having demonstrated 1) the phonological nature of foreign sound adaptation; and 2) the necessity of recognizing an independent level of abstract phonological representations and an ordered set of P-rules relating these abstractions to their concrete phonetic realizations, we can now demonstrate the relevance of lexicalization to

the justification of the phonological systems we construct. The 'psychological reality' of the properties postulated in a given phonology is demonstrated in two ways in lexicalization: first, certain morpheme structure conditions and phonological rules are shown to be productive in the process; second, certain rules are shown to interfere with the exact rendering of foreign material. We shall consider both kinds of evidence in the following discussion.

4. Rule Productivity

Although the concept of rule productivity has been in the literature for quite some time, it has not been adequately treated with respect to the borrowing situation. If we return to Haugen's Yaqui example, where the Spanish word estufa was lexicalized with a phonetic realization [ehtúpa], we find that the reason for this s/h discrepancy is the presence of a rule in Yaqui of the form:

8. $s \rightarrow h \; / \; \_\_\_ \left\{ \begin{matrix} t \\ k \end{matrix} \right\}$

whereby underlying /s/ is realized as phonetic [h] before /t/ and /k/. Thus the reality of this rule is attested to by the shape of incoming loan-words. An equivalent verification of the reality of a rule can often be obtained by an analysis of what constitutes a 'foreign accent'. As we see in 9,

9. French [məsyø] → American [məšə]
   Spanish [grasyas] → American [grašəs]

American English speakers, unless they are instructed or are perceptive, will have a tendency to reproduce French [məsyø] as [məšə] and Spanish [grasyas] as [grašəs] because there is a rule in American

English that realizes sequences of /s/ + /y/ as [š]. Thus we say [ɪšuw] and not *[ɪsyuw] and for many Americans the phrase 'I miss you' is realized as [aymɪšuw].

Every morpheme structure condition and phonological rule that I have found necessary to postulate in my work on Nupe, a Kwa language of Central Nigeria, can readily be shown to be productive through borrowed forms. Thus, if we look at 10,

10. MSC: A Nupe morpheme is typically (V)CV(CV)[11]

we find that because of the strict CVCV nature of the Nupe morpheme, the language must somehow deal with the numerous CC clusters in the Hausa words it has borrowed. Epenthetic vowels are created to break up the unacceptable consonant clusters. These facts are represented in 11,

11. $\emptyset \rightarrow i \,/\, C\_\_C$

$\emptyset \rightarrow u \,/\, C\_\_\begin{bmatrix}+\text{labial}\\ C\end{bmatrix}$

$\emptyset \rightarrow a \,/\, C\_\_[h]$

and exemplified in 12,

12. Hausa gàskíːyáː → Nupe gã̀šìkīyā 'truth'

Hausa fúskà → Nupe fúšíkā 'face'

Hausa ƙárƙóː → Nupe kã̄līkō 'to last'

Hausa káskóː → Nupe kã̄šīkō 'shallow pot'

Hausa múlkì → Nupe múlíkī 'authority'

---

[11] This formula ignores syllabic nasals, which possibly have a NV source (and are not too frequent). The initial vowel prefix, when occurring, is obligatorily /e/ before monosyllabic CV roots; otherwise it can also be /a/.

Hausa àlbárkà → Nupe àlùbărík̄ā 'blessing'

Hausa àlwá:šì → Nupe àlùwāšì 'boasting'

Hausa àlbásà → Nupe àlùbăsā 'onion'

Hausa màrhàbá: → Nupe măràhábì 'welcome'

Hausa àlhájì → Nupe ālāhájì 'Mecca pilgrim'[12]

An epenthetic /i/ breaks up two consonants, unless the second is a labial consonant, in which case /u/ occurs. Also, if the second consonant is /h/, the epenthetic vowel is normally /a/. The fact that we find epenthetic vowels being created in the lexicalized forms can only be explained by the reality of the morpheme structure condition of 10. That is, it cannot be explained at the phonetic level for the following reasons.

First, if the process is phonetic in nature (i.e. if we assume that the Hausa words are lexicalized with consonant clusters and a P-rule inserts epenthetic vowels) then there is no explanation why we obtain [u] before labial consonants and [a] before /h/. Sequences of [i] followed by [p], [b], [m] and [h] are seen in the following reduplicated forms:

13. /p^{y}á/ 'to dodge' → [pīp^{y}á] 'dodging'

/bá/ 'to be sour' → [bībá] 'being sour'

/mī/ 'to mould' → [mīmī] 'moulding'

/há/ 'to be hanging' → [hīhá] 'being hanging'

---

[12]Many of these and other words have Arabic equivalents. This is because Hausa (but not Nupe) borrowed these very words directly from Arabic. The choice of vowels and tones demonstrates that these words had to be borrowed from Hausa.

If we attempted to explain the observed epenthetic vowels by reference to the phonetic form of occurring Nupe words, we would overlook the generalization that on the systematic phonemic level there are no sequences of /i/ followed by /p/, /b/, /m/, /w/ or /h/. Instead, two MSC's enumerate the possible sequences in Nupe:

14. a) $[+\text{high}]\ (h)\ [+\text{labial}] \Rightarrow \begin{bmatrix} +\text{round} \\ +\text{back} \end{bmatrix}$ (V)

b) V h a ⇒ a

Thus the question would appear <u>not</u> to be 'what can a Nupe speaker pronounce?' (i.e. the question that would arise if we were to look to the phonetics for an explanation), but rather 'what is a possible underlying representation in the Nupe lexicon?'

The second argument for supposing that epenthesis is introduced as part of the lexicalization is that the 'insertion' of the appropriate vowel <u>precedes</u> all of the phonological rules postulated in Nupe phonology. It would be significant if we could show that the newly created syllables of <u>CiC</u> or <u>CuC</u> or <u>Cah</u> did not satisfy the structural description of a phonological rule, which 'original' sequences of <u>CiC</u>, <u>CuC</u>, and <u>Cah</u> are forced to undergo. As of now I have found no such rule; instead, consider the Nupe forms 'truth', 'face' and 'shallow pot' given in 12 above. In accordance with the phonological rules of Nupe, Hausa [s] has become Nupe [š]. Therefore the epenthetic front high vowel must be introduced into the phonological component at a stage prior to the application of SR (and PR, since it is further modified to [$\check{s}^{y}$]).

The i-insertion must also precede a tone rule (see Isaac George, this volume) which I represent informally as:

15. [HIGH TONE] → [RISING TONE]/[LOW TONE][+voice]___
V C

Thus, to illustrate, we find the following:

16. /èdé/ 'cloth' → [èdě]
/kùlé/ 'bell' → [kùlě]

BUT: /yèkó/ 'road' → [yèkó]

In the form for 'onion' [àlùbăsā] from Hausa àlbásà we note that the high tone of Hausa has been transformed, in accordance with the tone rule just presented, into a rising tone. Other examples are numerous:

17. Hausa àní:yà → Nupe àn̆iyā̃ 'determination'
Hausa dàbá:rà → Nupe dā̃bărā 'skill'
Hausa dìllá:lì → Nupe dìlălì 'agent, broker'
Hausa tàlá:tà → Nupe tàlătà 'Tuesday'
Hausa mùrá:dì → Nupe mùrădī 'desire'

Given, then, the productivity of this tone rule on new forms coming into the language, let us take a second look at the word for 'blessing':

18. Hausa àlbárkà → Nupe àlùbărīkā

Notice from the examples in 12 that the tone of the syllable created by epenthesis is a copy of the phonological tone (with few exceptions) of the preceding syllable. If it is the case that epenthesis is introduced by a P-rule, then the newly created syllable should copy

the phonetic tone of the preceding syllable and not the phonological tone, as here is the case. Given the suggestion that the borrowed forms have been lexicalized in keeping with the MSC's of Nupe, the phonetic tone realizations can be easily explained. The underlying form for 'blessing' is /àlùbárikā/ and the tone rule changes the first high tone to a rising tone. In order to account for the tone on the newly created epenthetic vowel, we must postulate that tone is assigned prior to this tone rule. Thus we have one more rule that must follow epenthesis.

We have thus seen through borrowing the productivity of three morpheme structure conditions (represented in 10 and 14.a, b) and the tone rule represented in 15. It is important to bear in mind, however, that although I make constant reference to attested borrowed forms, all of the processes illustrated are still productive and are observable in situations where Nupes speak a foreign language with a 'Nupe accent'. Thus it is not the case that a handful of borrowed words suffice in themselves to motivate the claims made in this paper, but rather these borrowed cases should be seen as an institutionalized reflex of a more general on-going process, that of perceiving and reproducing foreign sounds and sequences. In other words, both the borrowed words which have been lexicalized and have become part of the Nupe language, and the way Nupes 'repeat' non-Nupe words provide verification of the reality of both the morpheme structure conditions and the phonological rules in a grammar.

As a final illustration of the reality of phonological rules, let us recapitulate the argument put forth in Hyman (forthcoming, 1970). In that discussion /ɔ/ and /ɛ/ were justified on a number of grounds, despite the fact that neither occurs phonetically in the language. After /ɔ/ has labialized the consonant that precedes it, in accordance with the general labialization rule (LR) of Nupe, and after /ɛ/ has palatalized the preceding consonant in accordance with SR and PR (the latter of which collapses with LR to form a general assimilation rule, LR/PR, see 26), /ɔ/ and /ɛ/ merge with /a/ as they surface as [a], by the 'absolute neutralization' (AN) rule:

19. $[+\text{low}]_{V} \rightarrow \begin{bmatrix} +\text{back} \\ -\text{round} \end{bmatrix}$

Thus, to consider three verbs whose underlying vowels contrast, but whose surface vowels are realized identically, we see the following derivations:

20.

/tsɔ́/	→			$ts^{w}ɔ́$	→	[ts^{w}á]	'to watch'
/tsɛ̀/	→	čɛ̀	→	$č^{y}ɛ̀$	→	[$č^{y}$à]	'to begin'
/tsà/	→					[tsà]	'to choose'
		SR		LR/PR		AN	

It is because of the absolute neutralization rule (AN), according to my claim, that when Nupes speak Yoruba (which has [ɔ] and [ɛ]), with a Nupe accent, they replace [Cɔ] with [C^{w}a] and [Cɛ] with [C^{y}a]. It is thus as a result of this process that numerous consistent loanwords can be cited in their lexicalized form:

21. Yoruba kɛ̀kɛ́ → Nupe k^{y}ə̀k^{y}á 'bicycle'

Yoruba ɛ̀gbɛ̀ → Nupe ègb^{w}ə̀ (Yoruba town)

Yoruba tɔ̄rɛ̄ → Nupe t^{w}ār^{y}ā 'to give a gift'

Yoruba kɔ́bɔ̀ → Nupe k^{w}áb^{w}ə̀ 'penny'

Thus, according to this argument, not only are SR, PR and LR productive in lexicalization, but so is AN.

5. Rule Interference

In the Haugen example, the American English example and the Nupe examples in 5, 12, 17 and 21, we found that the phonetic shape of the foreign sound was identical to the distinctive feature content of the underlying form in the borrowing language. Thus in 5 Yoruba phonetic [sísì] is identical to Nupe underlying /sísì/, and it is because of this identity that the borrowed sound is made to undergo the phonological rules of the interpreting language. An important question one might raise is: what happens if the borrowed sound is identical not to an underlying segment in the borrowing language, but to a derived one? In 22,

22. Hausa àlwá:šì → Nupe àlùwāšì 'boasting'

Hausa fú:ší → Nupe fūšī 'anger'

Hausa múná:fúnčì → Nupe mánáfíčī 'hypocrisy'

Hausa mágà:ǰí: → Nupe mágàǰī 'heir'

remembering from 1 that these palatals are derived from dentals, we see that when these segments occur where they occur phonetically in Nupe, they are realized identically, and we can assume, are lexicalized with the appropriate underlying representations, seen in 23:

23. /àlùwāsì/ 'boasting'

/fūsī/ 'anger'

/mánáfítsī/ 'hypocrisy'

/mágàdzī̃/ 'heir'

Consider, however, the examples in 24,

24. Hausa šù:gàbá → Nupe šìgàbā 'leader'

Hausa ǰúmmáʔà → Nupe ǰímâ 'Friday'

where the segments [š] and [ǰ] occur in Hausa in environments where they wouldn't be found in Nupe. Since the forms of 25

25. *sùgàbā *dzúmâ

are possible Nupe words, one must ask why the consonant remained as in Hausa with the vowel altered, rather than the vowel remaining with the consonant altered, as in 25. It must mean that the palatalization of the initial sound has caused the Nupe speaker to perceive the vowel [u] as [i], the corresponding front vowel that would make palatalization possible. Nupe possesses two assimilation rules. The first, SR, was presented in 1. The second is a more general assimilation rule, taking the form:

26. $$[+\text{cons}] \rightarrow \begin{bmatrix} +\text{high} \\ \alpha\text{round} \\ \alpha\text{back} \end{bmatrix} / \underline{\quad} \begin{bmatrix} \alpha\text{round} \\ \alpha\text{back} \end{bmatrix}_{V}$$

Thus this rule (LR/PR) expresses the following information:

27. /gí/ → [gʸí] 'to eat'

/gē/ → [gʸē] 'to be good'

/gú/ → [g^{w}ú] 'to puncture'

/gò/ → [g^{w}ò] 'to receive'

AND: /tɛ́/ → t^{y}ɛ́ → [t^{y}á] 'to be mild'

/tɔ́/ → t^{w}ɔ́ → [t^{w}á] 'to trim'

BUT: /gà/ → [gà] 'to separate'

/tá/ → [tá] 'to tell'

(no assimilation)

Similar to the manner in which Hausa 'leader' comes into Nupe is the treatment of the exceptional Nupe word [g^{w}êgì] 'fifteen'. Since this isolated word must be marked as an exception to the phonological rules of Nupe, it is equivalent to a foreign word that would have to be marked [+foreign], as say, the native Hausa pronunciation of 'leader' with initial [šu] would be, were it to come into Nupe as such. Neither exception is tolerated in Nupe. The exceptional word for 'fifteen' derives from a compound consisting of [gúwō] 'ten' plus another morpheme ending in [egì], which for some reason we cannot recover. Note that this morpheme also has an alternate in free variation:

28. [g^{w}êgì] → [g^{w}ôgì]

We would thus represent the regularized variant as /gôgì/ and the labialization would be automatically obtained from LR. Just as Hausa <u>šù:gàbá</u> is not in keeping with SR (and PR), the exceptional Nupe word for 'fifteen' is an exception to LR. In both cases we find that the vowel changes to agree with the preceding non-vowel. The

exceptional form [$g^{w}\hat{e}g\grave{i}$] is not regularized as *[$g^{y}\hat{e}g\grave{i}$]. Instead it is the labialization that apparently causes the [e] to be perceived as [o]. Corresponding to the data of 'leader' and 'fifteen', then, if a Nupe were to say the Hausa word <u>$\grave{a}k^{w}\acute{i}:y\grave{a}$</u> 'goat' with a Nupe accent (or, equivalently, borrow and lexicalize it), one can predict that it would come out as underlying /$\grave{a}k\acute{u}y\bar{a}$/ with a phonetic realization of [$\grave{a}k^{w}\acute{u}y\bar{a}$]. Thus the labialization of the preceding consonant causes the following [i] to be realized as [u], just as the palatalization in the case of 'leader' caused the [u] to be realized as [i]. Just as we do not encounter the form *[$s\grave{u}g\grave{a}b\bar{a}$], we should not expect to find *[$\grave{a}k\acute{i}y\bar{a}$]. In both cases the vowel changes to agree with the preceding labialized or palatalized consonant, and not vice-versa.

The question immediately poses itself: if the correct form of the palatalization and labialization rule is as stated above in 26, where the consonants are assimilated to the vowels, then why is it that in borrowing and in 'Nupe accents' vowels assimilate to consonants? That is, why do we obtain just the <u>reverse</u> of the process that we postulate? Should the assimilation rules be stated differently, or is there some simple way of accounting for this discrepancy?

Before answering this question, let us consider the question of glides in Nupe, as part of a general process of assimilation in the language. We find the following near-complementary distribution:

29. yi wu

ye wo

ya wa

We do not obtain any of the following:

30. *wi *yu

*we *yo

It appears from these data that in [-low] vowels an archisegment /G/ can be used to represent lexical entries. Thus,

31. /ēGī/ → [ēyī] 'sun'

/ēGé/ → [ēyé] 'eye'

/ēGù/ → [ēwù] 'perfume'

/ēGó/ → [ēwó] 'money'

In each case the glide would be spelled out according to the rule:

32. $\begin{bmatrix} -\text{voc} \\ -\text{cons} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{round} \\ \alpha\text{back} \end{bmatrix} / \underline{\quad} \begin{bmatrix} \alpha\text{round} \\ \alpha\text{back} \\ -\text{low} \\ \text{V} \end{bmatrix}$

We would still, however, need to posit /w/ and /y/ to account for: [ēwā] 'snake' and [ēyā] 'canoe', which of course constitute a minimal pair. In order to extend this glide-spelling rule to cover [+low] vowels, we recognize the following phonological sequences:

33. yi wu

ye wo

yɛ wɔ

Since [h] is obtained only before [a] we are able to complete the system by the addition of /ha/. Thus /h/ can be assumed to be the

only underlying glide and [w] and [y] are derived in the appropriate environments by the glide-spelling (GS) rule as in 34:

34. $h \rightarrow \begin{bmatrix} \alpha\text{round} \\ \alpha\text{back} \end{bmatrix} / \underset{V}{\_\_} \begin{bmatrix} \alpha\text{round} \\ \alpha\text{back} \end{bmatrix}$

This rule may ultimately be collapsed with LR/PR (cf. 26).

From 34 we observe that /hɔ/ underlies [wa] and /hɛ/ underlies [ya]. The arguments for these underlying forms are similar to those presented in my justification of /Cɔ/ and /Cɛ/ in 'How Concrete is Phonology?' First, with reference to the process of reduplication outlined in 3 and 4, we note that the reduplicated form of /wá/ 'to want' is [wūwá] 'wanting' and not *[wīwá], as expected. If the underlying form is /hɔ́/ and not /wá/, as just assumed, then we can easily explain the appearance of [u] in reduplication where we would expect [i]:

35. /hɔ́/ → wɔ́ → wūwɔ́ → [wūwá]
GS RED AN

Unless we recognize this underlying form, we will have to postulate a corrective rule to apply after RED, of the form:

36. i → u / w

This rule would be unnecessarily redundant and unexplanatory, since it is as a result of the MSC's that no occurrences of [wi] (or [we]) ever occur in the language (two exceptions are noted below).

The second argument is one of borrowing and lexicalization. The glide-spelling rule is productive in this process in exactly the same sense as the other rules we have dealt with. Thus (though such

borrowed words are few) we find:

37. Hausa wàllá:hì → Nupe wōláyì 'by God'
Hausa húǰǰà → Nupe wúǰā 'excuse'

In 'by God', [h] has become [y] before [i], and in 'excuse' [h] has become [w] before [u]. The glide-spelling rule is probably also responsible for 38,

38. Hausa màlá:ʔíkà → Nupe mãlǎyíkā 'angel'

where we find glottal stop replaced by [y] before [i]. Although [ʔ] is phonetically closest to [h], we obtain the form for 'angel' as entered above. (Sequences of [hi] do exist in Nupe as noted above.) In a related process, as a result of the non-occurrence of initial vowels except /e/ (and more rarely /a/), we observe:

39. Hausa íri → Nupe yírī 'kind, species'
Hausa í:kò → Nupe yíkō 'power'
Hausa ígí:yà → Nupe yígíyā 'skein of thread'

Again, it is [y] that is obtained before [i].

One fact about the glide-spelling rule, as written, is that before /a/ we do not obtain [w] or [y]. Instead, in this position we find [h] exclusively. The non-occurrence of [w] and [y] before /a/ is apparently responsible for the obtained lexicalized form in:

40. Hausa nàʔám → Nupe nãhám 'yes' (in reply to a summons)

40 represents a recorded pronunciation (Banfield, 1914) of the Hausa word with a Nupe accent. Although /h/ and /a/ frequently conspire together (compare epenthesis in 12 above), when two /a/'s occur

together as in the above example, the glottal stop is often ignored. Thus the actual lexicalized form of Hausa [nà?ám] among most speakers of Nupe, as well as other similar examples, is seen in 41:

41. Hausa nà?ám → Nupe nă 'yes' (in reply to a summons)

Hausa bá?à → Nupe bâ 'defamation'

Hausa ƴàmá?à → Nupe ƴàmâ/ƴèmâ 'crowd'

As a final example, Hausa contains two diphthongs which are represented as /ai/ and /au/. Some attested borrowed words are:

42. Hausa sái → Nupe sā́yí́ 'until'

Hausa râi → Nupe ráyī 'life'

Hausa kâi → Nupe káyī (exclamation of surprise)

cf. Yoruba nâi → Nupe nā́yī 'ninepence'

Hausa bàutá: → Nupe bàwùtā 'slavery'

Thus at the morpheme structure level a glide /h/ is inserted to break up Hausa [ai] and [au]. The glide-spelling rule then derives [y] before [i] and [w] before [u].[13] It is from such evidence that we hypothesize (though with some certainty) that foreign [ahɔ] will be rendered as Nupe [awa] (or [ewa], since [e] is the only initial vowel

[13]There is one slight problem that we must keep in mind in dealing with these Hausa diphthongs: namely, that in many dialects of Hausa (apparently some from which other such words were borrowed into Nupe) they have become simple monophthongs, ai → e, au → o. Thus we note the following pattern:

Hausa	sàráutà	→ Nupe	sàrŏtā	'kingdom'
Hausa	sàlá:mù àláikù	→ Nupe	sālāmù àlēkù	'greetings'
Hausa	láimà	→ Nupe	lémā	'rank'
Hausa	ƴàlá:láinì	→ Nupe	ƴìlālēnì	'book on Mohammedan law'

permitted in bisyllabic words, cf. Hausa <u>àlló:</u> → Nupe <u>èlǒ</u>) and [ahɛ] will be realized as [aya] or [eya]. While I have not been able to test this out experimentally except on one subject (with positive results), and since such sequences are rare--if not nonexistent--in the languages from which Nupe has borrowed, we await empirical justification of these claims. It is clear, however, that such sequences would behave differently from [aha], which comes into the language as such, or which simply becomes [a], as in:

43. Hausa láhàdì → Nupe lâdè 'Sunday'

From the above examples we see that the glide-spelling rule is productive in such cases. The question we are now faced with is: what about those cases where the GV sequence is itself borrowed? As in the case of derived assimilated consonants (see 22 and 23), if the GV sequence is a permissible one in Nupe, the Hausa word is lexicalized with an identical surface glide realization:

44. Hausa wúrí: → Nupe wūrī 'an open space'
Hausa wùyá: → Nupe wūyá 'difficulty'
Hausa másò:yí: → Nupe mā́sòyī 'friend'
Hausa hár → Nupe hárí 'until'

But when sequences such as [wi], [we], [yu] and [yo] occur, are they lexicalized with the glide assimilating to the vowel (as would be predicted from the glide-spelling rule), or with the vowel assimilating to the glide? In order to deal with this possibility, it is not necessary to deal with foreign words coming into Nupe. There are two irregular pronouns in Nupe that are exceptions to the general

restrictions on the distribution of [w] and [y]. These are: [w̌ì] 'him' and [wě] 'you'. Note however that they have 'regular' variants in free distribution: [w̌ǔ] 'him' and [wǒ] 'you'. The irregular forms of these two pronouns arose, most likely, from idiosyncratic historical developments operative on old forms such as *uNi 'him' and *uCe 'you'. Three processes are involved. In the form for 'him', the nasal consonant causes the subsequent vowel to become nasalized (a general rule of Nupe phonology). It (and the C in 'you') then drop and we are left with intermediate forms *ui and *ue. The final step is for /u/ to become [w] and hence to break the general phonetic distribution of glides. The regularized forms represent still a more recent change, which can perhaps be expected to thoroughly wipe out the older forms.

The importance of the foregoing is that LR/PR (26) and GS (34) would seem to predict that the vowel determines the preceding non-vowel, and not vice-versa. We would otherwise expect that the normalizing effect of the Systemzwang should be:

45. w̌ì → *y̌ì

wě → *yě

g^{w}êgì → *g^{y}êgì

šùgàbá → *sùgàbā

I have thus far been unable to discover a significant number of borrowed words violating the permissible Nupe GV sequences. One reason for this is that Hausa (with few exceptions) observes the same

pattern. One such borrowing does occur from Yoruba:

46. Yoruba àkɔ̀wé → Nupe àkʷàwŭ 'clerk'

Although the expected form is *[àkʷàwŏ] there are two possible sources of this discrepancy. The first is that the final vowel in the Yoruba word is often perceived as [i]. And second, there could have been contamination from the Hausa form (also borrowed from Yoruba): àkà:wú:. However, we do not get forms such as *[àkʷàyĕ] or *[àkʷàyĭ]. All of these data then support the following general principle:

47. FOREIGN SEGMENTS EQUIVALENT TO NATIVE SEGMENTS DERIVED BY RULE ARE LEXICALIZED AS THE CORRESPONDING NATIVE UNDERLYING FORMS.

Then the phonological rules are applicative. By this process Hausa wúrí: is lexicalized as /hūrī/ and then the [w] is recreated so-to-speak in the phonetic [wūrī]. In the Hausa word šù:gàbá, however, [š] is a derived segment in Nupe. Since we have hypothesized that foreign sounds are perceived as underlying forms, a Nupe's natural tendency is to 'recreate' the underlying form that would have given us the derived [š]. The Nupe speaker perceives [š]. He 'reasons' that it must have satisfied the structural description of SR (see 1). But clearly the Hausa word violates this rule, since the incoming [š] would have then been derived in the context / ___u. In order to remedy this impulse toward positing underlying /s/ when [š] is heard, the [u] of 'leader' is modified to [i] so that SR can appropriately apply to the then newly entered underlying representation: /sìgàbā/. We therefore need a second part to the above principle:

48. WHEN A FOREIGN SEGMENT APPEARS IN AN ENVIRONMENT IN WHICH THE EQUIVALENT NATIVE DERIVED SEGMENT DOES NOT APPEAR, THEN THE FORM OF THE INCOMING FOREIGN WORD IS MODIFIED SO THAT THE STRUCTURAL DESCRIPTION OF THAT RULE IS MET AND THE SEGMENT IN QUESTION IS THEN DERIVED IN THE APPROPRIATE ENVIRONMENT.

Thus, for Nupe at least, given a rule of the form,

49. A → B / ___ C

if [BD] is borrowed, where [D] does not equal [C], then [D] is modified to [C] so that the structural description of this rule is met and [B] can be appropriately derived. Returning to the form for 'clerk' (see 46), then, [w] is recognized as a derived segment obtained through the glide-spelling rule. One can only suggest that the tendency of the Nupe speaker is to posit underlying /h/ wherever he hears [w]. But the relevant part of the glide-spelling rule that derives [w] says:

50. $h \rightarrow \begin{bmatrix} +\text{round} \\ +\text{back} \end{bmatrix} / \underset{V}{\_\_} \begin{bmatrix} +\text{round} \\ +\text{back} \end{bmatrix}$

The [e] of Yoruba <u>àkɔ̀wé</u> does not however permit us to represent this phonetic form as underlying /àkɔ̀hé/, because by means of another part of the glide-spelling rule, which says,

51. $h \rightarrow \begin{bmatrix} -\text{round} \\ -\text{back} \end{bmatrix} / \underset{V}{\_\_} \begin{bmatrix} -\text{round} \\ -\text{back} \end{bmatrix}$

the derived surface form would be *[àkʷàyě]. In order to obtain [w] from underlying /h/ it is necessary to modify the environment in the incoming word so that it correctly satisfies the structural description of the glide-spelling rule. The final vowel is thereby altered. Thus the rule in question is shown by its <u>interference</u> to be real.

These principles have more general application in the language than is evident from the discussion thus far. For example, consider the related phenomena to be presented. The vowel systems of Hausa and Nupe are:

HAUSA

	i:	i		u	u:	
ai	e:	e		o	o:	au
			a			
			a:			

NUPE

i	ĩ		ũ	u
	e		o	
ɛ	ɛ̃		ɔ̃	ɔ
		ã		
		a		
		(a:)		

Short /a/ in Hausa is realized as [ə]. Thus its feature specifications will be phonetically:

52. $\begin{bmatrix} \text{-low} \\ \text{-round} \\ \text{+back} \\ \text{-nasal} \end{bmatrix}$

Similarly, Nupe nasalized /ã/ is realized as [ə̃]. Its feature specifications will be:

53. $\begin{bmatrix} -\text{low} \\ -\text{round} \\ +\text{back} \\ +\text{nasal} \end{bmatrix}$

Now note the 'normal' way (with few exceptions) for Hausa /a/ ([ə] to be realized in Nupe:

54. Hausa dàgà → Nupe dãgã̀ 'from'
Hausa zárà̀fí: → Nupe zãrã̀fī 'wealth'
Hausa gádó: → Nupe gãdō 'bed'
Hausa gáƒákà → Nupe gãƒãkā 'school bag'
Hausa ƙàfó: → Nupe kã̀fō 'horn'
Hausa ƙárƙó: → Nupe kãlīkō 'to last'
Hausa kábàrí: → Nupe kãbã̀rī 'grave'

The most common way for Hausa [ə] to come into Nupe is thus as a nasalized schwa [ə̃] (/ã/). Compare, however, the following:

55. Hausa wádàrí: → Nupe wódã̀rī 'length of cotton
Hausa wàllá:hì → Nupe wōláyì 'by God'
Hausa wàrkí: → Nupe wòrìkī 'loin-cloth'
Hausa wásàlí: → Nupe wósã̀lī 'vowels in Arabic script'
Hausa wàsíƙà → Nupe wòšíkā 'letter'
Hausa wátàkí:là → Nupe wōtàkīlà 'perhaps'
Hausa àlká:wàlí: → Nupe àlìkáwòlī 'promise'
Hausa ámá:wàlí: → Nupe ámáwòlì 'part of turban'
Hausa yàbó: → Nupe yèbō 'thankfulness'

Sequences of Hausa [wə] come into Nupe as [wo], despite the fact that there are numerous words of the shape [wə̃] in Nupe, as seen in:

56. [ēwə̄] 'pregnancy'

[ēwə̃́] 'falsehood'

[ēwə̀̃] 'lake'

A similar situation is observed in the case of labialized or palatalized consonants followed by schwa in Hausa:

57. Hausa k^{w}àdâyí: → Nupe k^{w}òdə̃̀rī 'covet' (/kòdə̃̀rī/)

Hausa mák^{w}àbčí: → Nupe mák^{w}òčī 'neighbor' (/mákòtsī/)

Hausa ǰámá?à → Nupe ǰémâ 'crowd'

Hausa ǰánàbá: → Nupe ǰénə̃́bā 'cleansing'

Hausa ǰàká: → Nupe ǰìkā 'bag'[14]

Hausa šágàlí: → Nupe šégə̃̀lī 'business affair'

Hausa ǰàlá:láinì → Nupe ǰìlālēnì 'book on Mohammedan law'[14]

Hausa àlǰánnà → Nupe àlìǰěnə̃̄ 'Paradise'

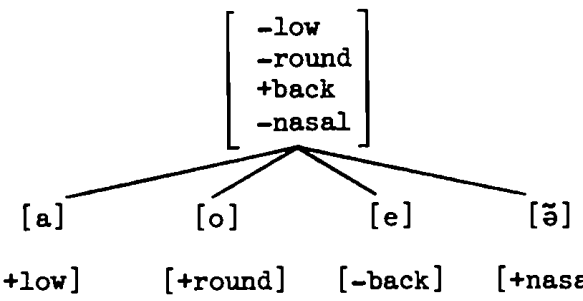
Despite the fact that there are corresponding Nupe morphemes of the structure /Cɔ̃/ and /Cɛ̃/ we do not obtain in borrowed words their respective phonetic realizations [C^{w}ə̃] and [C^{y}ə̃]. The labialization or palatalization of the preceding consonant has blocked the otherwise productive assimilation of Hausa [ə] as Nupe [ə̃], just as the glide-spelling of [w] and [y] had influenced the perception of [ə] in 55. Why should this be?

---

[14] In these three forms [i] is obtained from the Hausa environment / ___ CV: as a result of the de-emphasis of vowels in this position in Hausa.

The answer to this problem has to do with WHY we should obtain a nasalized schwa in 55 in the first place. As is seen in 58

58.



there are FOUR possibilities open for the nativization of foreign schwa in Nupe, each of which differs from schwa by one distinctive feature. Although nasalized schwa is obtained, most linguists I have asked felt that the 'closest' vowel to schwa among these four is [a]. Given the other examples and the principle cited in 48, we can explain the unexpected nasalized schwa. We obtain the nasalized schwa because of the presence of a rule of the form in 59

59. $\underset{A}{\mathbf{a}} \rightarrow \underset{B}{\text{ə}} \; / \; \underset{C}{[\overline{+\text{nasal}}]}$

which raises underlying nasalized /ã/ to [ə̃].[15] Looking at this rule as I have just represented it makes it clear that it is of the form in 49. A schwa is derived only in the environment of simultaneous

---

[15] This rule will ultimately collapse with the absolute neutralization rule (AN) in 19:

$$\underset{V}{\begin{bmatrix} +\text{low} \\ \langle +\text{nasal} \rangle \end{bmatrix}} \rightarrow \begin{bmatrix} +\text{back} \\ -\text{round} \\ \langle -\text{low} \rangle \end{bmatrix}$$

nasalization. Therefore, when this derived vowel quality is heard from Hausa without the simultaneous nasalization, it is nasalized so that the structural description of the rule 59 is met and the schwa quality can be appropriately derived.

The reason why this process breaks down in the presence of glides, labialized or palatalized consonants is that since foreign sounds are treated in terms of underlying forms, the order of the rules affects their phonetic output. We have seen that the rule raising nasalized /ã/ to a nasalized schwa is responsible for Hausa schwa coming in as nasalized schwa. However, the glide-spelling rule, labialization rule and palatalization rules all PRECEDE the nasal raising rule. Thus, when Hausa [wə] reaches the glide-spelling rule, by principle 48, the vowel must be changed to [+round] so that the glide-spelling rule <u>appropriately</u> derives the [w]. Therefore the schwa becomes [+round] or phonetically [o]. Since this has occurred, the schwa never gets far enough into the derivations to become nasalized. The examples in 54 come out nasalized for the sole reason that no rules in the phonology <u>interfere</u> with non-labialized and non-palatalized consonants. Therefore the effect of the low-level nasal raising rule is felt.

6. <u>Summary and Conclusion</u>

In this paper a phonological view of borrowing was supported and it was thus shown that in many respects the manner in which words are borrowed and lexicalized, or equivalently the way

people speak with foreign accents, provides justification for the conditions and rules we propose in a grammar. Three principles were proposed to account for the data of Nupe. Many of the suggestions as to how these processes occur, that is, how foreign items are perceived by speakers, are conjecture. Certain regularities were presented and the hypothesis made would appear to account for what actually occurs. The fact is that these are not random occurrences and until evidence is brought forth to show that these explanations are not plausible, they appear not only to account for the data, but to explain the data as well.

Several questions and problems remain, however. One is that we have no idea how universal these principles may turn out to be. Another is that many languages are for one reason or another prone to borrowing words as exceptions, and do not fully nativize them.

The role of borrowing in the justification of phonological grammars is much less in doubt: one of the criteria that should be considered as part of the metatheory is how foreign words are borrowed. Given this criterion, the transformations that occur between the source-word and the lexicalized-word substantiate the proposed morpheme structure conditions and phonological rules. In other words, the reality of these aspects of the phonology is attested to in just this way.

REFERENCES

Banfield, A. W. (1914) Dictionary of the Nupe Language, Shonga, Northern Nigeria, Vol. 1, Nupe-English.

Bloomfield, Leonard (1933) Language, New York, Henry Holt and Company.

Chomsky, Noam (1964) "Current Issues in Linguistic Theory", in Fodor and Katz, The Structure of Language, Prentice-Hall, Englewood Cliffs, New Jersey.

Chomsky, Noam and Morris Halle (1968) The Sound Pattern of English, Harper and Row, New York.

George, Isaac (1970) "Nupe Tonology", Studies in African Linguistics Vol. 1, No. 1.

Halle, Morris (1962) "Phonology in a Generative Grammar", reprinted in Fodor and Katz The Structure of Language, Prentice-Hall, Englewood Cliffs, New Jersey.

Harris, Zellig (1954) "Distributional Structure", reprinted in Fodor and Katz The Structure of Language, Prentice-Hall, Englewood Cliffs, New Jersey.

Haugen, Einar (1950) "The Analysis of Linguistic Borrowing", Language, V. 26, pp. 210-231.

Hyman, Larry M. (1970) "How Concrete is Phonology?" (forthcoming, Language.)

Kiparsky, Paul (1968) "Linguistic Universals and Linguistic Change", in Bach and Harms, Universals in Linguistic Theory, Holt, Rinehart and Winston, Inc., New York.

Paul, Hermann (1966) Prinzipien der Sprachgeschichte, Tübingen, Max Niemeyer Verlag.

Polivanov, Ergeniy (1931) "La Perception des Sons d'Une Langue Etrangère," TCLP, 4, pp. 79-96.

Postal, Paul (1968) Aspects of Phonological Theory, New York, Harper and Row.

Schane, Sanford A. (1968a) French Phonology and Morphology, Cambridge, Massachusetts, the MIT Press.

Schane, Sanford A. (1968b) "On the Non-Uniqueness of Phonological Representations", Language, 44.4, pp. 709-716.

Stanley, Richard (1967) "Redundancy Rules in Phonology", Language, 43.2, pp. 393-436.

Weinreich, Uriel (1953) Languages in Contact, The Hague, Mouton and Company.

Weinreich, Uriel (1957) "On the Description of Phonic Interference", Word: Vol. 13, No. 1, pp. 1-11.