

LOAN WORDS AND ABSTRACT PHONOTACTIC CONSTRAINTS*

JONATHAN KAYE

Université du Québec à Montréal

BARBARA NYKIEL

The Pedagogical University of Bydgoszcz

0. In this paper we will discuss several cases of the phonology of loan words. We will show that loan words are adapted phonologically to meet certain constraints on possible words or morphemes imposed on them by the borrowing language. Most significantly, such constraints are defined on the level of lexical representation. To our knowledge, so-called "surface phonetic constraints" (henceforth SPC's) play no role in the phonology of loan words, nor indeed in any other part of phonology. Our examples show that deep phonotactic constraints, constraints not necessarily reflected on the surface, are the drums to which loan words must march. This is not to say that all aspects of loan word phonological behaviour are now understood. Many mysterious facts still remain unexplained. We do claim that substantial aspects of this phenomenon may be understood by assuming that the systematic phonemic level of generative phonology is the level¹ on which the constraints controlling loan word phonological behaviour are defined.

We will first sketch briefly and informally our model for the nativization of loan words. In the following sections we will rediscuss two examples of non-phonetic constraints governing loan word nativization: Northern Algonquin (first discussed in Kaye 1979 and to appear) and Odawa (discussed in Kaye

* We wish to thank Professor J. Fisiak, Dr. E. Gussmann and Dr. J. Rubach for their valuable suggestions and criticism. They are, of course, not responsible for any errors to be found in this article.

¹ Or at least one of the levels. To date we are unaware of other levels on which phonotactic constraints governing the nativization of loan words are defined.

1975a). Finally, we will discuss a similar example from Polish loan word phonology. Assuming an analysis along the lines of Gussmann (1978), we will show that once again non-phonetic phonotactic constraints² are at work in shaping loan words from various sources to Polish phonology. Our conclusion is that loan word evidence repeatedly confirms the existence of deep phonotactic constraints in language whereas evidence from a similar source supporting SPC's seems non-existent.

1. In attempting to formulate a model of loan word nativization we will assume a principle along the lines of the naturalness principle discussed in Postal (1968:54ff.), viz. that loan words are borrowed as they appear in the source language unless there is some reason to modify them. Thus, *alarm* was borrowed into Polish from French more or less intact.³ Accordingly we assume that a phonetic output of the source language is interpreted as the phonetic output of the borrowing language *if this is possible*. This claim is represented in (1).

(1) Source Language Borrowing Language
 [a b c ... x] ⇒ [a' b' c' ... x']⁴

The phonemic status of the sounds in question appears to be largely irrelevant in the borrowing situation. This is hardly surprising in that a detailed knowledge of the source language is not essential in a borrowing situation. Thus, Desano /waka/ [waka] 'cow' has been borrowed from Spanish [βaka] although Spanish [β] is an allophone of /b/ and *baka* would be a perfectly pronounceable Desano word. Similarly, English loans from Japanese such as *hibatchi* and *sashimi* do not reflect the underlying forms of the source language. In Japanese [č] and [š] are allophones of /t/ and /s/, respectively, before high front vowels. Notice once again that *hibati* and *sasimi* are possible English words.

If the underlying representation of the source language is irrelevant in loan phonology,⁵ the same cannot be said for the underlying form of the borrowing language. Indeed, a second strategy for the nativization of loan words is to borrow the form in question as an underlying form and allow it to undergo

² Non-phonetic phonotactic constraints have been termed "counter-factual" constraints by some adherents of concretist phonology. We leave it to the reader to decide just how "counter-factual" these constraints are.

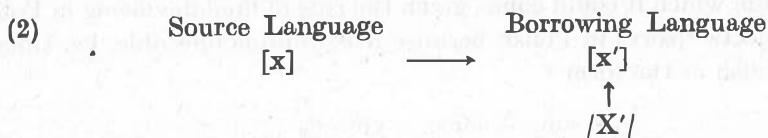
³ We ignore here the fine phonetic details of vowel and consonant quality. Note also that *alarm* is stressed on the initial syllable in conformity with the Polish pattern of penultimate stress.

⁴ The primes of the phonetic output of the borrowing language indicate that the sounds in question are the borrowing language's equivalent of the source language's sounds, e.g. Engl. *tape* > Cdn Fr *tape* where the initial alveolar *t* is dental in its borrowed form.

⁵ We leave aside here the question of loans based on the written language which may correspond to underlying forms in the source language.

the phonological rules to yield a phonetic form which differs from the phonetic form of the source language. This second strategy is used in cases where the phonetic form of the source language is, for one reason or another, not a possible phonetic form in the borrowing language. Looked at another way, there is no possible derivation in the borrowing which would yield the phonetic form of the source language as a possible output.⁶

Thus, English *serv* could not enter Polish as [serv] given the Polish rule of final devoicing. English *serve* did, however, enter Polish as /serv/ which is pronounced [serf] in the nom. sg. Oblique forms such as [servu] gen. sg., [servi] nom. pl., etc. as well as related forms such as *serwować* 'to serve' indicate that /serv/ is indeed the underlying form. Hyman (1970:18) cites the Yoruba word *sísì* 'sixpence' which has been borrowed into Nupe as *šišì*. Note that no Nupe derivation could yield [sísì] as its output although the sequence [...si...] is pronounceable in Nupe. As was the case in Polish, Nupe has used the Yoruba source word for the underlying form of 'sixpence' /sísì/. The rules of Nupe phonology then apply to yield [šišì]. We represent this second strategy in (2)

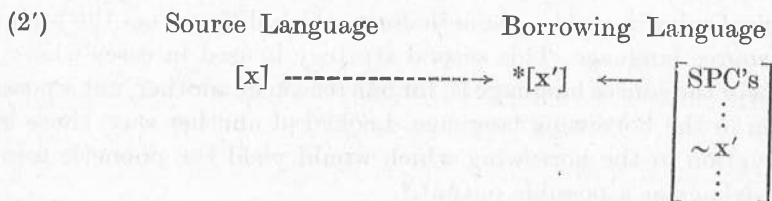


Model (2) is to be interpreted as follows: in cases where a phonetic output corresponding to the source form in uninterpretable, i.e. cannot be the final step in any derivation in the borrowing language, model (2) applies. The phonetic source string, *x*, is interpreted as an underlying representation, *X'*, where *X'* is the underlying form which most closely resembles [x] of the source language and is at the same time a possible underlying form in the borrowing language. For this notion to make any sense at all it is obvious that features of underlying forms must be phonetically interpretable. Such is the case in generative phonology — the theoretical framework of this paper.

Once /X'/ is posited as the underlying form, it undergoes all the phonological rules in the usual manner yielding [x']. It should be noted that [x'] need not be the phonetic sequence that most resembles [x] of the source language. In discussing notions like "possible output" it is necessary to fit in the phonological rules and the phonotactic constraints against which putative underlying forms must be checked, into our model. We will return to this point below.

We have eschewed another possible approach in our model (2), viz. a model using SPC's. A possible model along these lines is shown in (2').

⁶ This notion is quite similar to that of phonological recoverability explored in Kaye (1974, 1975b and 1979). Eliasson's (to appear) idea of directionality is likewise germane to this point.



In (2') we again attempt to incorporate the source string $[x]$ into the borrowing language as $[x']$ (in this example we assume that x' is the closest phonetic equivalent to x). In this model the string x' is rejected as being phonetically impossible, i.e. unpronounceable. This is done by matching the string $[x']$ against the SPC's of the language which include, among other things, a constraint against strings of the form $[x']$. What happens next does not interest us here. What is crucial is that (2') incorporates a different explanation for the failure of certain forms to follow model (1), i.e. for the rejection of $[x']$ in the borrowing language. To return to a concrete example, the string $*[\text{serv}]$ may be rejected because it is underivable in the borrowing language: there is no underlying form from which it could come, given the rule of final devoicing in Polish. Model (2') rejects $*[\text{serv}]$ in Polish because it is unpronounceable, i.e. there is an SPC in Polish of the form

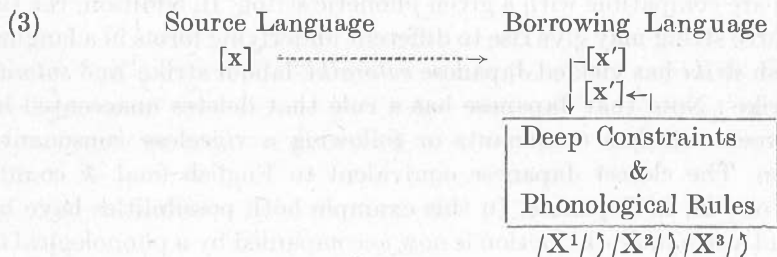
~[-son, +cons, +voiced] # #

We will try to show that the model (2') is not the correct characterization of a native speaker's competence; more particularly that aspect which deals with loan word nativization. SPC's appear to play no role in phonology and are merely the result of phonological rules which have not been opacated by subsequent events (later rules, lexical exceptions, etc.) in the history of the language. Hyman's Yoruba-Nupe example is but the first indication of the non-phonetic nature of loan word nativization: Yoruba *sísi* being reshaped as Nupe *šiši* in spite of the fact that $[si]$ exists in Nupe. In our examples we will show situations where loan words are reshaped to conform to the *deep* phonotactic constraints of a language. What is crucial is that these deep phonotactic constraints do not correspond to SPC's. We thus have the following situation: in cases where a deep constraint happens to be "surface-true" loan words typically conform to this constraint. In cases where an SPC is without a corresponding deep constraint, the SPC has no discernible effect on loan words. Polish has the above mentioned SPC against final voiced obstruents. However, loan words with both underlying voiceless final consonants, *czek* 'cheque', *test* 'test' and underlying voiced consonants, *fez* 'fez', *nerw* 'nerve', exist in Polish.⁷

⁷ These facts have been discussed from a radically different point of view in Shibatani (1973). In arguing for a role for SPC's in loan phonology he presents cases where SPC's correspond to MSC's or WSC's. The fact that loan words conform to those constraints

Finally, in cases where a deep phonotactic constraint exists but not a corresponding SPC, we see that loan words systematically conform to the "abstract" constraint.⁸ This last statement, if correct, shows that deep constraints have a justification in the grammar independent of any considerations of phonetic constraints. The same cannot be said for SPC's. Only in cases where SPC's happen to coincide with deep constraints do they have even the appearance of shaping loan word behaviour. In view of the lack of independent motivation for SPC's we feel warranted in maintaining our point of view that they play no role in phonology. We thus reject the model (2').

We may now refine our model (2) with the addition of a component containing the deep phonotactic constraints as well as the phonological rules.



Model (3) shows a phonetic string from the source language being interpreted as a phonetic string in the borrowing language, *mutatis mutandis*. The string is then processed abductively,⁹ i.e. given the phonetic form and the rules and constraints of the phonology a possible source is sought which would yield [x'] and at the same time conform to the phonotactic constraints of the language. If no such source is found then the source language string is interpreted as an underlying form in the borrowing language, of course still being subject to the same constraints.¹⁰ We do not pretend to have exhausted all aspects of loan word phonology with our model, nor is it clear that the same strategy is

is no argument for SPC's. The crucial cases are those where the SPC's and the deep constraints are in conflict. Shibatani's evidence concerning speaker judgements rejecting nonsense words to which obligatory phonological rules have not been applied, strikes us as dubious at best. The metatheoretical arguments made by Shibatani seem equally suspect. His notion of "plausibility" leads one to reject models such as (2), which have some measure of empirical support, in favour of others which are, as we shall see, demonstrably false.

⁸ To the examples to be discussed below we may add Walker (1975) for French and Morin (1976) for Burushaski, as well as Hyman's treatment of Nupe.

⁹ In the sense given by C.S. Peirce. See Andersen (1973) and Chomsky (1977:87) for discussion of this notion in a linguistic context. It is indeed abduction that underlies the notion of recoverability as discussed in Kaye (1974, 1975b and 1979).

¹⁰ We are, in general, in agreement with the model proposed in Hyman (1970). Some differences do exist, however. Here we assume that languages borrow forms in such a way that they remain phonetically as close as possible to their source. Phonetic source forms are interpreted as underlying forms only when this first strategy fails. Nor is it

always adopted by speakers for all languages. Problems concerning the adaptation of segments that do not exist in the borrowing language are not dealt with here.

The abductive nature of our model entails situations where phonetic strings may have more than one source. In undoing phonological rules it may be the case that more than one putative underlying form is compatible with the phonetic string. This model has been objected to by Shibatani (1973:96) on the basis of informal observations regarding reaction time to nonsense words. We shall see that speakers develop strategies for assigning underlying forms to borrowed phonetic strings. Further, recoverability conditions as described in Kaye (1975b) have precisely the effect of reducing the number of underlying forms which are compatible with a given phonetic string. In addition, the same phonetic source string may give rise to different underlying forms in a language. Thus, English *strike* has yielded Japanese *sutoraiki* 'labour strike' and *sutoraiku* 'baseball strike'. Note that Japanese has a rule that deletes unaccented high vowels between voiceless consonants or following a voiceless consonant in final position. The closest Japanese equivalent to English final *-k* could be either */-ku/* or */-ki/* in Japanese. In this example both possibilities have been exploited and a semantic distinction is now accompanied by a phonological one.

Having described the theoretical model for loan word nativization we now proceed to the discussion of specific cases of loan word treatment.

2. Our first example is from the Northern Algonquin dialect of Lac Simon. This example has already been discussed in Kaye (1979 and to appear). We summarize the facts rather briefly here. The reader is referred to the works cited above for more details. Lac Simon contains a large number of words which display the kinds of alternations shown in (4).

| | | | |
|---------------|-------------|-----------------|-------------|
| (4) pi:goška: | 'it breaks' | ki:bi:goška: | 'it broke' |
| te:sibiwa:gin | 'chair' | nide:sibiwa:gin | 'my chair' |
| ði:ma:n | 'canoe' | oʃi:ma:n | 'his canoe' |
| ka:zo:tim | 'he hides' | niga:zota:n | 'I hide' |
| sigiswa: | 'he smokes' | kizigiswa: | 'you smoke' |
| šo:škose: | 'it slides' | nižo:škose: | 'I slide' |

In (4) it is apparent that an initial voiceless obstruent alternates with a voiced non-initial one. The data in (5) show that we have a case of initial devoicing rather than medial voicing.

clear that the primacy that Hyman accords the borrowed segment as opposed to its context is universally true (Hyman 1970:40). The Desano example discussed above showed Sp ([baka] yielding Desano [waka]. But [β] exists in Desano as an allophone of /w/ before front vowels. Note, however, that the context has not been modified to allow the retention of [β], i.e., we get [waka] and not, say, *[βeka]. Cf. Stampe (1973:68ff.) for a similar though not identical view.

| | | | | |
|--------|------------|--------------------|-----------|-----------------------|
| (5) a. | kitiga:n | 'cultivated field' | kiči- | 'very, great, big...' |
| | a:sibi: | 'net' | pi:sa:dik | 'in height' |
| (5) b. | ni-kin | 'my bone' | o-se:z-in | 'his brother' |
| | ki-šogona: | 'your nose' | ni-ka:t | 'my leg' |

In (5a) we see that voiceless consonants appear freely in morpheme-medial position. The examples in (5b) show that voiceless consonants may also appear in morpheme-initial position following the same personal prefixes that occur in (4). The forms of (5b) are dependent nouns which are obligatorily preceded by a prefix of some sort. Given the data of (4) and (5) plus the fact that initial voiced obstruents never occur phonetically, leads us to posit rule (6).

(6) C → [-voiced] / # # _____

Note that the situation described here seems identical to, say, the final devoicing situation in Polish. The Lac Simon alternation, *sigiswa:* ~ *-zigiswa:* is exactly like the Polish alternation *ješ* ~ *jež* 'hedgehog'. There is, however, no form in Lac Simon that would correspond to Polish *b'egač* ~ *b'egač-* 'runner'. In other words, Lac Simon does not contain initial obstruents which fail to show a voice-voiceless alternation if the opportunity arises. Thus, a hypothetical pair of forms like **či:ko:n* ~ *-či:ko:n* is impossible in Lac Simon. Expressing this point somewhat differently, we may say that all Lac Simon words which alternate (i.e. all those which may appear initially as well as preceded by some pre-element) are underlyingly voiced. The question now remains as to what to do with forms that occur only initially. In Polish uninflectable forms such as *oraz* [oras] 'and' are assumed to have underlying final voiceless consonants. This is because Polish final devoicing neutralizes a contrast. We find final voiced and voiceless consonants (/bez/ vs. /b'egač/) which are neutralized phonetically in word final position. Assuming Kiparsky's alternation condition or the equivalent, we are obliged to posit final /-s/ for *oraz*.

The Lac Simon situation is different. Aside from the forms which are always initial (the Lac Simon equivalents of *oraz*) all initial obstruents are underlyingly voiced. Whether rule (6) is a neutralization rule depends on the status of non-alternating forms like *ka:-* 'relative marker', *ki:sp'in* 'if', *te:di:go* 'much'. If these form, which appear only in word-initial position, are analyzed as having underlying voiceless consonants, then rule (6) is neutralizing. Precisely these non-alternating forms with underlying initial consonants would appear in the immediate input to (6), giving it its neutralizing status. On the other hand, if we assume all such forms are analyzed as undergoing (6), i.e. with underlying initial *voiced* consonants, then rule (6) is non-neutralizing. No underlying voiceless consonants would occur in word-initial position.

What distinguishes these two possible analyses is the existence of a WSC

word structure constraint) excluding initial voiceless stops and fricatives.¹¹ If we assume that *all* initial voiceless consonants are derived from underlying voiced ones via (6), whether they alternate or not, then such a WSC may well still be present in the phonology of Lac Simon. Indeed, loan word behaviour supports such a view as the data in (7) show.

- (7) a. pa:na:n 'banana' (<banane) niba:na:nim 'my banana' /ba:na:n/
 pa:stone: 'American' (<Boston) ka:ba:stone:na:gozič 'he who looks
 like an American' /ba:stone:/
- b. pitik 'potato' (<petac) nibitikim 'my potato' /bitik/
 toma:do:s 'tomato' (<tomatoes) nidoma:do:zim 'my tomato'
 /doma:do:z/
 ko:fi:ke:¹² 'he makes coffee' (<coffee) nigo:fi:ke: 'I make coffee'
 /go:fi:ke:/

The examples in (7a) show loan words borrowed from English and French with initial voiced consonants. According to our model sketched in section 1 of this paper, the forms of (7a) are interpreted as underlying representations since rule (6) makes any interpretation of initial voiced consonants impossible on the phonetic level. The underlying forms undergo rule (6) and are accordingly realized with initial voiceless consonants, *pa:na:n*, *pa:stone:*. These examples have no bearing on the existence of our WSC. They show that rule (6) applies to loan words. The forms of (7b) concern us more directly. These forms contain initial voiceless consonants in their source language. We note that these forms have initial voiceless consonants which alternate with medial voiced consonants. This state of affairs indicates that they have been interpreted as having underlying initial voiced consonants. In initial position the Lac Simon forms correspond phonetically to the forms from the source language. In medial position, as we have noted, voiced consonants appear. Why should this be the case? There is no rule of medial voicing in Lac Simon. Forms such as **nēpitikim*, **nītoma:do:zim*, **nīko:fi:ke:* are perfectly pronounceable. If there is a WSC against underlying word-initial voiceless consonants, then our model can account for the data of (7b). Consider the form *toma:do:s*. The presumed phonetic form from the source language was [tōmaʎuz]. The form was interpreted as [toma:do:s]¹³ in Lac Simon. Where could this form have come from? Assuming the WSC to be present, [toma:do:s] must have come from /doma:do:z/ via rule (6). If this form occurs in non-initial

¹¹ This constraint also prohibits initial consonant clusters. See Kaye (to appear) for details.

¹² This form is taken from the Amos dialect of Northern Algonquin.

¹³ Note that the final [z] is likewise impossible in Lac Simon. A rule of final devoicing also exists in this dialect. *z* is interpreted as the underlying final consonant, at least by some speakers.

position, rule (6) does not apply and the underlying initial *d* surfaces as in [nidoma : do : zim]. By assuming our WSC, in spite of its being ruthlessly opacated by rule (6), we can account for all the data of (7).

At this point our analysis will have become extremely unpalatable to adherents of concrete phonological theories. Alternative analyses are sought for the data just presented. Let us be quite clear as to what is to be explained. Loan words with initial voiceless consonants show up in non-initial position with voiced consonants as in (7b). Our explanation was to posit a WSC forbidding underlying word-initial voiceless consonants and a rule of initial devoicing. To accept our analysis is to accept a WSC which is „abstract” in the sense that it is not *directly* reflected on the surface. To avoid such a conclusion one could posit a morphologically governed rule of medial voicing and a WSC excluding initial voiced obstruents. The prefixless forms of (7b) undergo no pertinent rules and simply reflect their underlying forms. The initial consonants appear voiced in the prefixed forms due to the putative rule of medial voicing. Serious problems are associated with this analysis.¹⁴ Adherents of concrete phonology normally distinguish processes and rules; with processes being exceptionless, surface-true generalizations. Now the rule of medial voicing is anything but surface-true. Word medial voiceless consonants abound, and yet loan words undergo its effects, a property supposedly restricted to processes.¹⁵ Assuming such theoretical wrinkles may be worked out, it is nonetheless odd that advocates of concrete phonology must claim that speakers choose an opaque morphologically conditioned rule over one that is completely transparent and exceptionless. Be that as it may, the concretist argument will be annihilated if we can show that abstract MSC's and WSC's are commonplace. In that case no reason exists for not accepting our analysis presented above. The next two sections will be devoted to illustrating abstract phonotactic constraints influencing loan word behaviour where no alternative analysis involving rule inversion seems available.

3. Odawa,¹⁶ another Algonquian language, provides us with the second example of an abstract WSC. In fact, it is the Odawa version of the WSC discussed in the previous section, a WSC which dates from the Proto-Algonquian period. Odawa has two series of obstruents: a fortis (voiceless-aspirated) series and a lenis (unaspirated) series. Traditionally the lenis series is represented by simple voiceless obstruents: p, t, ç, k, s, š. The fortis series is represented by geminate versions of the lenis series: pp, tt, çç, kk, ss, šš. This

¹⁴ See Kaye (to appear).

¹⁵ Cf. Hooper (1979:2), Stampe (1973:68ff.).

¹⁶ Odawa (or Ottawa) is a member of the Ojibwa dialect complex. It is spoken by some 2,000 speakers on Manitoulin Island in Ontario.

latter series may be assumed to have the status of underlying clusters. We shall make this assumption for the purposes of this article although nothing essential to our argument depends on this. With these preliminaries established, we can state the Odawa WSC.

(8) Odawa WSC — No word may begin with a cluster.

In loan word phonology, English voiceless consonants are interpreted as fortis consonants (e.g. English *p* is interpreted as Odawa *pp*) and English voiced obstruents are interpreted as Odawa lenis consonants (e.g. English *b* is interpreted as Odawa *p*).¹⁷

This constraint is rendered opaque at the phonetic level by a series of rules that (among other things) have the effect of deleting short vowels in initial syllables.¹⁸ Since underlying forms having / $\check{V}CC\dots$ / and / $C\check{V}C\dots$ / occur with normal frequency, we have abundant examples of [CC...]. Note the examples of (9).

| (9) | <i>Underlying Form</i> | <i>Phonetic Form</i> | <i>Gloss</i> |
|-----|------------------------|----------------------|-------------------|
| | akkwe:w | kkwe: | woman |
| | išpatina: | špatna: | it is high |
| | appakiso | ppakso | he throws himself |
| | askanisi | skanse | he is thin |
| | sikime:nh | skimē | mosquito |
| | sanakat | snakat | it is difficult |
| | mittikw | mttik ~pttik | tree |

It remains to be seen if the WSC against initial clusters is still operative in spite of its systematic violation on the surface. Once again loan words provide us with evidence that such is the case. Consider the data in (10).

| | | | | | |
|------|------------|---------------------|--------------------|------------------|---------------------|
| (10) | ppe:n | 'pen' | ntappe:n | 'my pen' | /appe:n/ |
| | kka:ppi: | 'coffee' | ntakka:pi:m | 'my coffee' | /akka:ppi:/ |
| | šta:mp | 'stamp' | ntašta:mpim | 'my stamp' | /ašta:mp/ |
| | ššo:w | 'show' | ntaššo:w | 'my show' | /aššo:w/ |
| | so: | 'zoo' | nso : m | 'my zoo' | /so:/ |
| | mpass | 'bus' | ntampssim | 'my bus' | /ampass/ |
| | mpi:ŋko:hw | 'he has a bingo' | ntompi:ŋko: hiw | 'I have a bingo' | /ompi:nko: hiwi/ |

¹⁷ Odawa represents the more conservative situation with respect to this WSC. The fortis-lenis distinction has been reanalyzed as a voiceless-voiced one in Lac Simon. A cluster analysis for the voiceless stops does not seem warranted for Lac Simon.

¹⁸ The details do not concern us here. See Kaye (1973, 1975a) for details.

In (10) we give the prefixless version of Odawa loan words. Phonetically, they correspond to their source form with the exception of the final two words which have (for reasons that remain obscure) *b-* in English and *mb-* in Odawa. The second column gives the same form preceded by a prefix. In comparing the two forms we arrive at the underlying form shown on the right. The prefix in each case is /n \bar{V} -/, the first person prefix. The following *-t-* is epenthetic, which is inserted to separate a prefix from a following *vowel-initial* morpheme. It follows then that a form like [ntappe : n] must be analyzed as underlying /n \bar{V} -appe : n/ and [ppe : n] is accordingly /appe : n/. Why should this be so? Assuming a more concrete model one would expect English *pen* to yield /ppe : n/ in Odawa. In that case 'my pen' would come out as [mppe : n] from /n \bar{V} -ppe : n/. Such a form is phonetically quite possible: *nttam* 'first', *nkka : t* 'my leg', *nšši : mē* 'my younger sibling'. English *pen* is interpreted as /appe : n/ rather than /ppe : n/ not for phonetic reasons but because /ppe : n/ is not a possible underlying form, given (8). Using the same abductive reasoning illustrated in section 2 and following the model of loan word nativization discussed in section 1, we assume an Odawa speaker interprets English [pⁿen] as [ppe : n]. The problem now is to assign this form an underlying representation. There are a number of possibilities since any short vowel will be syn-copated in an initial syllable. It is hardly surprising that speakers develop strategies for determining underlying representations of ambiguous surface forms. The Odawa solution is to assume the structure /aCC.../ for surface forms of the type [CC...]. Note that these surface forms could have come from / $\left\{ \begin{smallmatrix} a \\ o \\ i \end{smallmatrix} \right\}$ CC.../ or /C $\left\{ \begin{smallmatrix} a \\ o \\ i \end{smallmatrix} \right\}$ C.../. The initial clusters or geminates are never

broken up as per the latter possibility shown above. We can appeal to a general linguistic principle of preserving the integrity of geminates and consonant clusters.¹⁹ It is not clear why Odawa speakers usually select *a* to nativize loan words with initial clusters but let us eliminate one hypothesis immediately. What we are trying to show here is that an abstract WSC (8) is at work in reshaping loan words. We assume that *pen* winds up as /appe : n/ because of this WSC excluding */ppe : n/ as a possible underlying form. Suppose one claims that the *a* that pops up comes from a misanalysis of the English indefinite article *a*. Quite frequently loan words are taken with an accompanying article from the source language, if one exists. Micmac (Algonquian), Hare (Athapascan), Dida (Kru) are but three examples of such languages. Note that in all three cases it is the *definite* article that is taken, not the indefinite one, e.g. Dida *lèkùlòó* 'school' (<French *l'école*). Now if it turns

¹⁹ See Kenstowicz and Pyle (1973) for discussion. Kaye (1975b) shows a general tendency in Odawa to avoid *CC ~ CVC* alternations in native vocabulary, where the *CC* in question is a possible underlying cluster.

out that the *a* appears in the underlying form of loan words is simply a reflection of the fact that *a pen* served as the source for /appe : n/ and not *pen*, then we have no evidence from loan words for the existence of our abstract WSC. In fact, we can show that the initial *a* of these loan words is not a reflex of the English indefinite article. First, note in the examples of (10) that *a* is not always present initially. Indeed, it appears only in those instances where borrowing the source word as such would violate the WSC. Compare the treatment of *show* > /aššo : w/ vs. *zoo* > /so : /. In the former case the initial English voiceless fricative is interpreted as a fortis consonant in Odawa. Fortis consonants are not found in word-initial position and hence a vowel is added to the underlying form. In the case of *zoo*, the English initial voiced fricative becomes a lenis *s* in Odawa. Such consonants may occur initially and hence no further modification is required. In the cases where English *b* has been interpreted as *mp* (for whatever reason) an initial *a* (or *o*) is added in order to satisfy the WSC. Why should Odawa speakers borrow words with an article only in those cases where an abstract WSC would be violated? It makes far more sense to attribute such behaviour directly to the WSC.²⁰

Finally, there are older loan words borrowed from French, none of which display any tendency to include an indefinite article. The initial fricative of the first example is probably from the [z] of the French plural definite article. The final example is a verb and is given only for purposes of comparison.

| | | |
|----------------------------|-------------------------|----------------------------------------------------|
| (11) <i>ša : kana : šš</i> | 'English Canadian' | (<i><les anglais plus pejorative -šš></i>) |
| <i>a : nima :</i> | 'German' | (<i><allemand></i>) |
| <i>mo : tay</i> | 'bottle' | (<i><bouteille></i>) |
| <i>ma : sa : n</i> | 'bazaar' | (<i><bazar></i>) |
| <i>ko : ppe : sse :</i> | 'to confess' (religion) | (<i><confesser></i>) |

We conclude then that our English loan words are not analyzable as coming from English sources including the indefinite article.

We may now ask if the same sort of reanalysis by rule inversion of the sort discussed in the previous section is possible here. The syncope rule subsumed in our analysis is given in (12).

(12) $[+voc, -stress] \rightarrow \emptyset$

What concerns us here is that one of the effects of this rule is to delete initial short vowels. Any attempt to use some sort of rule inversion analysis to avoid an abstract WSC would result in dividing up the lexicon into three arbitrary

²⁰ A possible claim would be that speakers misanalyze nouns (i.e. include the indefinite article in their UR) because of the WSC. This view would be equally supportive of our position although we do not wish to argue for it here. The Odawa version of *bingo* poses problems for all analyses involving the English indefinite article.

classes. To illustrate this point let us suppose that the inverted rule is along the lines of (13).

$$(13) \quad \emptyset \rightarrow \left[\begin{array}{c} i \\ o \\ a \end{array} \right] / \text{prefix}] + \text{--- CC}$$

Following this analysis one would posit that the underlying form of 'pen' is *ppe : n* and that the vowel found in *ntappe : n* is inserted by rule (13).²¹ A problem arises since it is not always the same vowel that appears following the prefix, as is shown in (14).

| (14) | <i>Prefixless Form</i> | <i>Prefixes Form</i> | <i>Abstract UR</i> | <i>Concrete UR</i> |
|-------------|------------------------|----------------------|--------------------|--------------------|
| 'pen' | ppe : n | ntappe : n | /appe : n/ | /ppe : n/ |
| 'say' | kkito | ntikkito | /ikkito/ | /kkito/ |
| 'come from' | nèipa : | ntonèipa : | /onèipa:/ | /nèipa:/ |

Since there is no phonological way of predicting which vowel will appear following the prefix, every noun and verb in the language which has a phonetically initial cluster must be supplied with a diacritic feature to indicate which vowel is to be inserted by (13). But things get worse. Consider the forms in (15).

| (15) | <i>Prefixless Form</i> | <i>Prefixes Form</i> | <i>UR</i> |
|---------------|------------------------|----------------------|-----------------|
| a. 'car/sled' | ta : pa : n | ntota : pa : n | /ota : pa : n/ |
| 'man' | nine | ntinni : m | /ininiw/ |
| b. 'touch' | (ŋki :)ta : nkškwa : | nta : nkškwa : | /ta : nkiškaw/ |
| 'stand up' | na : ni : pwe | nna : ni : paw | /na : ni : paw/ |

In (15a) we see that we find the same $\emptyset \sim$ vowel alternations as occurred in (14). No solution involving an inverted rule like (13) is possible because not all consonant-initial stems in the prefixless form alternate with a vowel occurring after a prefix as (15b) shows. Once again one could distinguish the forms of (15a) from those of (15b) by another diacritic but now we are required to mark every noun and verb in the language with an arbitrary diacritic. If we eschew this analysis for the forms of (15) and retain the vowels of the prefixed forms as part of the UR of the stems, then we need a rule like (12) independently of the data in (14). But rule (12) will automatically account for the data of (14) if one assumes the "abstract" UR's. No arbitrary lexical features are required for this analysis. One might add that if an analysis involving an inverted rule like (13) is invoked, the entire analysis of stress placement (Kaye 1973) falls into disarray. We do not wish to beat a dead horse and we conclude that our analysis involving an abstract WSC is correct.

Now if we have an independently motivated basis for abstract phonotactic

²¹ We pass over the problem of deriving the epenthetic *i* in these forms.

constraints, there is no longer any reason for rejecting such an analysis in the case of Lac Simon. The only conceivable reason for rejecting the completely productive and transparent rule of initial devoicing was because to do so would lead one to posit an abstract WSC. Now we have just shown that abstract WSC's are required regardless of what analysis we choose for Lac Simon. The unmotivated "concrete" morphological rule of medial voicing must be abandoned and we arrive at the abstract analysis once again. We have seen that both Odawa and Lac Simon Algonquin have an abstract WSC. We will now proceed to show that Polish has an abstract *morpheme* structure constraint (MSC).

4. Edmund Gussmann (1978) in his extremely insightful analysis of Polish segmental phonology considered the case of vowel $\sim \emptyset$ alternations of the sort illustrated in (16).

| | | | | |
|------|----------------------------|-----------|----------|--------|
| (16) | (after Gussmann 1978 : 52) | | | |
| | forma | 'form' | foremny | 'adj.' |
| | taśma | 'ribbon' | tasiemka | 'dim.' |
| | wyspa | 'island' | wyseпка | 'dim.' |
| | szużba | 'service' | szużebny | 'adj.' |

Gussmann argues convincingly that the vowel that occurs in the right column and is absent in the left column is part of the UR. The details of the analysis do not concern us here. Suffice it to say that the vowel is deleted unless followed by another deleting vowel. Such deleting vowels are posited for the diminutive suffix, *-ek* (*-ka*, *-ko*), the adjective forming suffix *-ny* and sporadically for the genitive plural of feminines and neuters and the nominative singular of masculine nouns. It is in these contexts then that the underlying vowel will surface, if indeed one is present.

It is the first form of (16) that catches our eye. According to the analysis we have just sketched, the UR of *forma* must be /forEm-a/.²² But *forma* is a loan word and *-rm-* is a consonant cluster in the source language. Why, then has this cluster been broken in the UR of the Polish form? The examples of (17) show that *forma* is by no means an isolated case.

| | | | | |
|---------|--------------|-----------|------------|-------------------------|
| perla | 'pearl' | perel | 'gen. pl.' | (<Lat <i>per(u)la</i>) |
| mebel | 'furniture' | meble | 'nom. pl.' | (<Fr <i>meuble</i>) |
| klub | 'club' | klubek | 'dim.' | (<Fr <i>trèfle</i>) |
| tawerna | 'tavern' | taweremka | 'dim.' | (<It <i>taverna</i>) |
| turnia | 'jail tower' | turemka | 'dim.' | (<Germ <i>Thurm</i>) |
| filtr | 'filter' | filtrerek | 'dim.' | (<Lat <i>filtrum</i>) |
| palma | 'palm' | palemka | 'dim.' | (<Lat <i>palma</i>) |

²² We use *E* to represent the deleting vowel. Gussmann posits the underlying lax vowels *i* and *e* (non-palatalizing and palatalizing, respectively) as its source.

All the above forms show an alternation of the type CC~CeC. Typically, the nominative singular shown in the left hand column contains the cluster. In the right hand column the cluster is broken by *e*. We have used the diminutive form to illustrate the presence of the underlying vowel. A comparison of *mebel* and *treśl* shows the unreliable nature of masculine nom. sg. forms as indicators of the presence of the vowel.

We must now ask the question as to why these loan words have been interpreted with an underlying vowel breaking up the consonant cluster from the source language. Note that in our brief discussion of Japanese loans there was at least a vague sort of surface motivation for such an adaptation. A surface explanation for this phenomenon might appeal to such notions as avoiding awkward consonant clusters and the like. Any such surface explanation is doomed to failure. First of all, Polish can hardly be termed a language which avoids surface consonant clusters. Furthermore, all of the consonant clusters broken up in loan words occur on the surface, frequently in final and/or preconsonantal position (cf. *filtr*, *treśl*). We conclude that the behaviour shown in (17) has nothing to do with surface constraints. We must look elsewhere for an explanation.

Given our experience from the first two examples (Odawa, Lac Simon) discussed above, we would do well to determine the constraints governing the shape of Polish morphemes. A study of the *Indeks a tergo* for possible stem final clusters in Polish nouns is quite revealing. In general, all genders permit stem final clusters ending in an obstruent.

| | | | | |
|------|---------|------------------|-----------|--------|
| (18) | szelest | 'rustling noise' | szelestny | 'adj.' |
| | skurecz | 'cramp' | skureczny | 'adj.' |
| | wilk | 'wolf' | wilczek | 'dim.' |
| | skarpa | 'escarpment' | skarpka | 'dim.' |
| | ospa | 'smallpox' | ospka | 'dim.' |
| | baszta | 'tower' | basztka | 'dim.' |
| | kapusta | 'cabbage' | kapustka | 'dim.' |

We are not claiming that *all* surface clusters ending in an obstruent correspond to underlying clusters. Indeed Gussmann (1978) points out numerous examples of such forms. We are rather trying to show that these clusters indeed occur in UR's and accordingly are not involved in any MSC which would exclude them. A comparison of the forms in (19) illustrates this point.

| | | | | | |
|------|-------|------------|----------|--------|-----------|
| (19) | wyspa | 'island' | wyseпка | 'dim.' | /vysEp-a/ |
| | ospa | 'smallpox' | ospka | 'dim.' | /osp-a/ |
| | walka | 'struggle' | waleczny | 'adj.' | /valEk-a/ |
| | wilk | 'wolf' | wilczek | 'dim.' | /vilk/ |

Loan word behaviour is consistent with this state of affairs. Loan words with such clusters in the source language do not manifest broken clusters in Polish.

| | | | | |
|------|----------|-------------|-----------|------------|
| (20) | park | 'park' | parczek | 'dim.' |
| | plansza | 'board' | planszka | 'dim.' |
| | farba | 'paint' | farbka | 'dim.' |
| | ryksza | 'rickshaw' | ryksz | 'gen. pl.' |
| | reszta | 'remainder' | resztką | 'dim.' |
| | langusta | 'lobster' | langustka | 'dim.' |

Restrictions on stem final clusters seem to involve final sonorants. For all genders it is extremely difficult to find unequivocal²³ cases of integral clusters with a final liquid. As a first approximation we propose the following MSC for Polish:

$$(21) \quad \sim X [+cons] \left[\begin{array}{l} -syll \\ +cons \\ +son \end{array} \right] \alpha]$$

α = masc, fem or neuter nouns

MSC (21) excludes stem final clusters ending in a liquid. Indeed nominative singular feminines of the form *-CL-a*, neuters of the form, *-CL-o* and masculines of the form *-CL* almost invariably display a *-CL- ~ -CEL-* alternation if the opportunity presents itself. Examples are given in (22).

| | | | | |
|------|---------|----------|-----------|------------|
| (22) | wiatr | 'wind' | wiaterek | 'dim.' |
| | wydra | 'otter' | wyderka | 'dim.' |
| | kółdra | 'quilt' | kółder | 'gen. pl.' |
| | pętla | 'noose' | pętelka | 'dim.' |
| | gra | 'game' | gierka | 'dim.' |
| | iskra | 'spark' | iskierka | 'dim.' |
| | dziupła | 'hollow' | dziupelki | 'dim. pl.' |

Masculine nouns in *-CL* in the nominative singular that are not obvious loans are a bit hard to come by since our deleting vowel frequently occurs in this context. Additional examples of masculine nouns in *-CEL* are given below.

| | | | | |
|------|---------|-------------|---------|------------|
| (23) | cukier | 'sugar' | cukru | 'gen. sg.' |
| | węgiel | 'coal' | węgla | 'gen. sg.' |
| | bimber | 'moonshine' | bimbru | 'gen. sg.' |
| | rygiel | 'lock' | rygla | 'gen. sg.' |
| | kafel | 'tile' | kafa | 'gen. sg.' |
| | plaster | 'slice' | plastra | 'gen. sg.' |

²³ In many cases the forms simply do not occur in contexts that would cause an underlying vowel to surface, if one existed.

The status of *l* is somewhat problematic. If *l* is considered to be an underlying liquid then we would assume that no stems could end in underlying *-Cl*. For feminine and neuter nouns this is correct.

| | | | | |
|------|--------|----------------------------|-----------|--------|
| (24) | miotła | 'broom' | miotelka | 'dim.' |
| | kukła | 'puppet' | kukielka | 'dim.' |
| | ćwikła | 'beets and horseradish' | ćwikielka | 'dim.' |
| | mgła | 'fog' | mgielka | 'dim.' |
| | masło | 'butter' | masełko | 'dim.' |

Masculine nouns in [Cl] are rather rare. Very limited data seem to indicate that [Cl] clusters are permitted.

| | | | | |
|------|----------|-------------|------------|-------------|
| (25) | pomysł | 'idea' | pomysłek | 'dim.' |
| | przemysł | 'ingenuity' | przemysłny | 'ingenious' |

However the overwhelming majority of masculine nouns in [Cl] follow the feminine and neuter pattern: *poseł—posła* 'deputy', *supel—supła* 'knot', *Paweł—Pawła* 'Paul' and so on. In the light of the above we will assume that masculine *-Cl* clusters are excluded along with the *-Cl* clusters. Of course this will follow automatically from (21) if we assume that *l* is an underlying liquid. The examples of (25) are then exceptional. Note that the feminine loan word *perła—perel* 'pearl' conforms to (21).

The behaviour of clusters ending in a nasal is most interesting. These clusters are excluded from feminine and neuter nouns, while they are permitted in masculine nouns. We now add these facts to (21) and reformulate our MSC as in (26).

$$(26) \quad \sim X[+cons] \left[\begin{array}{c} -syll \\ +son \\ \langle -nasal \rangle_a \end{array} \right] \alpha$$

condition: if α =masculine, then a .

MSC (26) collapses two constraints. In the case of masculine nouns the material in the angled brackets is included in the constraint. Thus, for masculine nouns only non-nasal sonorants are excluded as the final member of clusters. In the case of feminines and neuters, the material in the angled brackets is not included and the constraint is more general. All non-syllabic sonorants, including nasals, are excluded as final members of clusters. Note the behaviour of feminine and neuter nouns.

| | | | | |
|------|---------|------------------|---------|-----------|
| (27) | sarna | 'deer' | sarenka | 'dim.' |
| | córučna | 'daughter, dim.' | córučen | 'gen.pl.' |
| | cizma | 'shoe' | cizemka | 'dim.' |

| | | | |
|--------|----------|----------|--------|
| wydma | 'dune' | wydemka | 'dim.' |
| taśma | 'tape' | tasiemka | 'dim.' |
| pasmo | 'streak' | pasemko | 'dim.' |
| jarzmo | 'yoke' | jarzemko | 'dim.' |
| ziarno | 'corn' | ziarenko | 'dim.' |

Masculine nouns which both end in *-CN* and may occur in contexts which indicate whether we are dealing with a true cluster or just a secondary cluster created by a deletion rule, are exceedingly rare. One clear example appears in (28). A feminine and a neuter form are also included for purpose of comparison.

| | | | |
|-------------|--------|----------|--------|
| (28) pokarm | 'food' | karmny | 'adj.' |
| sarna | 'deer' | sarenka | 'dim.' |
| ziarno | 'corn' | ziarenko | 'dim.' |

No vowel breaks up the *-rm-* cluster in the form *karmny*. The fact that **karemny* is not the adjectival form of *pokarm* shows that /po-karm/ is indeed the UR. This is not to say that *-CN-* clusters are hard to find in Polish. Unfortunately, they do not occur before our two criterial suffixes *-ek* and *-ny*. In fact, one of the most productive suffixes in Polish is *-izm/-yzm* occurring in hundreds of forms: *szowinizm*, *faszyzm*, *sarkazm*, *komunizm*, *socjalizm*, and so on. Although not of Polish origin, this suffix seems well established in the morphology. It does not appear in contexts that cause an underlying vowel to surface and so we have no way of proving definitively that its underlying form is /*(i/y)zm*/ rather than /*(i/y)zEm*/.

There exist *CN ~ CeN* alternations among masculine nouns as shown in (29).

| | | | |
|------------|---------|------|-----------|
| (29) dzień | 'day' | dnia | 'gen.sg.' |
| pień | 'trunk' | pnia | 'gen.sg.' |
| sen | 'dream' | snu | 'gen.sg.' |

The data in (29) does not contradict our MSC (26). The fact that some surface *-CN-* clusters are broken in masculine nouns does not mean that all *-CN-* clusters must be broken. There are two types of evidence that indicate that our MSC (26) expresses the correct generalization. Feminine loan words receive different treatment from masculine ones.

| | | | | | |
|---------|---------|--------------|------------|--------|------------|
| (30) a. | forma | 'form' | foremny | 'adj.' | *formny |
| | turma | 'jail tower' | turemka | 'dim.' | *turmka |
| | tawerna | 'tavern' | tawerenka | 'dim.' | *tawernka |
| | drachma | 'drachma' | drachemka | 'dim.' | *drachmka |
| b. | alarm | 'alarm' | ?alarmek | 'dim.' | *alaremek |
| | uniform | 'uniform' | ?uniformek | 'dim.' | *uniformek |
| | film | 'film' | ?filmek | 'dim.' | *filemek |
| | szturm | 'attack' | ?szturmek | 'dim.' | *szturemek |

Without exception, feminine loan words in *-CN-* show breaking in the appropriate context. Even where the diminutive in *-ka* is not normally used, speakers seem quite sure that breaking must take place. As for the masculine nouns, none occur spontaneously before *-ek* or *-ny*. Nevertheless, speakers' judgements seem quite clear: if one must use *-ek* to form a diminutive, the integrity of the consonant cluster is maintained. Broken clusters before *-ek* are uniformly and resoundingly rejected.

No appeal to surface phonotactic constraints can explain this behaviour. If speakers are capable of pronouncing *-rnm-* in *karmny*, they are capable of pronouncing **formny*. Yet it is *foremny* that obtains. The diminutive suffixes are somewhat misleading. This suffix surfaces as *-ek* in masculine nouns and as *-ka* in the feminine. One could argue that the syllabic shape of the suffixes is the determining factor in the integrity of consonant clusters. Failure to break in the feminine would create sequences of the form *-CNka*. In the masculine such clusters lead to *-CNek* on the surface. In fact, the feminine shape of *-CeNka* rather than **-CNka* has nothing to do with surface phonotactics. We have only to look at masculine diminutives in oblique cases to prove this point. The gen. sg. forms of masculine diminutives are identical to the feminine nom. sg. —viz. *-ka*. Thus, if, say, *alarmek* is acceptable, then *alarmku* is the gen. sg. of this form. In (31) we present a series of masculine nouns in *-ek*. These may or may not be analyzed as synchronic diminutives. What is important is that the oblique cases illustrate all the surface clusters that a putative SPC would supposedly exclude in order to account for the behaviour of loan words.

| | | | |
|--------------|----------|---------|-----------|
| (31) giermek | 'squire' | giermka | 'gen.sg.' |
| garnek | 'pot' | garnka | 'gen.sg.' |
| kosmek | 'wisp' | kosmka | 'gen.sg.' |
| czosnek | 'garlic' | czosnku | 'gen.sg.' |

The second piece of evidence concerns certain gender shifts from feminine or neuter to derived masculines. We note that the consonant clusters that are broken in the original form are treated as true clusters in the derived masculines. While no MSC forbids masculine stems of the type *-CEN*, the tendency to preserve the integrity of consonant clusters *where possible* discussed above, plus the well-known regularizing influence of derived forms appears to have led to this reshaping.

| | | | | |
|---------|-------------|-------------|-------------|-----------|
| (32) a. | wiosna | 'spring' | wiosenka | 'dim.' |
| | pierwiosnek | 'primrose' | pierwiosnka | 'gen.sg.' |
| b. | biodro | 'hip' | bioder | 'gen.pl.' |
| | nabiodrek | 'hip guard' | nabiodrka | 'gen.sg.' |

The underlying form of the feminine stem for 'spring' /věsEn/ conforms to the MSC (26). The derived masculine has the form /pír-věsn-Ek/ (literally 'first spring (dim.)') and exhibits the final cluster *-sn* which is consistent with our MSC which permits final *-CN* clusters in masculine nouns. (32b) is somewhat strange. The underlying form of the non-derived neuter noun is /bědEr/. The derived form has been reshaped as /na-bědr-Ek/ ('on hip (dim.)'), which is masculine. What is strange is that masculine nouns normally may not end in a *-CL* cluster. Why the constraint is apparently relaxed in this case is not clear. Nonetheless, both examples of (32) show the role of gender with respect to the MSC. It is precisely this correlation which enables us to explain the different treatments of *forma* vs. *alarm*.

To conclude our study of Polish noun stem phonotactics, let us briefly consider feminine stems in *-Cw* clusters. In Polish orthographic *w* is normally pronounced [v]. In some ways it behaves like a glide rather than a fricative. Its role in the phonotactics puts it in a class with the other sonorants. Consider the examples in (33).

| | | | | |
|---------|---------|---------------|-----------|---------------|
| (33) a. | barwa | 'colour' | barwny | 'adj.' |
| | malwa | 'mallow' | malweczka | 'double dim.' |
| | morwa | 'mulberry' | morwka | 'dim.' |
| b. | łyżwa | 'iceskate' | łyżew | 'gen.pl.' |
| | poszwa | 'pillow case' | poszewka | 'dim.' |
| | brzytwa | 'razor' | brzytewka | 'dim.' |
| | sakwa | 'bag' | sakiewka | 'purse' |

The generalization involved here seems to be that clusters in *w* are tolerated if the initial member is a sonorant (33a). Clusters in *w* with an obstruent as the first member are excluded (33b). Going on the supposition that *w* is indeed a glide [-syll, -cons], we can incorporate the above generalization in the final version of the MSC which we give in (34) below.

$$(34) \quad \sim X \left[\begin{array}{l} +\text{cons} \\ < +\text{son} >_b \end{array} \right] \left[\begin{array}{l} +\text{son} \\ -\text{syll} \\ < +\text{cons} >_c \\ < -\text{nasal} >_\alpha \end{array} \right] \alpha$$

Conditions: if α = masculine then *a*.

if *b* then *c*.

Details of this analysis may require eventual modification. These data once again show the non-phonetic nature of loan word behaviour. Loan words in Polish are not adapted because of some phonetic constraint excluding these clusters. Indeed, most clusters that are broken up in their passage from the source language into Polish, exist on the surface. The abstract nature of the above constraint is due to the fact that countless Polish nominals containing the

forbidden clusters never occur in an environment where their putative underlying vowel surfaces. Words like *kopra*, *lepra*, *masakra*, *dromla*, *budowla*, *szmer*, *akr*, *kadr* do not occur in the critical environments. Are they to be analyzed with underlying clusters? If so, there is no MSC (34). The fact that loan words conform to (34) with great regularity indicates that even non-alternating forms are to be analyzed with an underlying vowel breaking up a forbidden cluster.

In this paper we have discussed three cases of loan word adaptation. In all these cases loan words conformed to abstract phonotactic constraints that are flagrantly violated on the surface. It has been repeatedly shown that theories of loan word behaviour based on surface constraints are consistently unable to account for these facts. As we have indicated above, our three examples are by no means the only ones of their kind. Phonotactic constraints that are not reflected on the surface have been termed "counterfactual". What our study reveals is that it is rather the surface phonetic constraints that are counterfactual — counterfactual in the sense that they play no role in linguistic theory.²⁴

REFERENCES

- Andersen, H. 1973. "Abductive and deductive change". *Language* 49. 765—793.
- Chomsky, N. 1977. *Dialogues avec Mitsou Ronat*. Paris: Flammarion.
- Dinnsen, D. (ed.). 1979. *Current approaches to phonological theory*. Bloomington: Indiana University Press.
- Eliasson, S. to appear. "Analytic vs. synthetic aspects of phonological structure". In Goyvaerts, D. L. (ed.). forthcoming.
- Goyvaerts, D. L. (ed.). forthcoming. *Phonology in the 1970's*.
- Gussmann, E. 1978. *Explorations in abstract phonology*. Lublin: Maria Curie-Skłodowska University.
- Hooper, J. 1979. "Substantive principles in natural generative phonology". In Dinnsen, D. (ed.). 1979. 106—125.
- Hyman, L. 1970. "The role of borrowing in the justification of phonological grammars". *Studies in African linguistics* 1. 1—48.
- Kenstowicz, M. and C. Kisseberth. (eds). 1973. *Issues in phonological theory*. The Hague: Mouton.
- Kenstowicz, M. and C. Pyle. 1973. "On the phonological integrity of geminate clusters". In Kenstowicz, M. and C. Kisseberth. (eds). 1973. 27—43.
- Kaye, J. 1973. "Odawa stress and related phenomena". In Piggott, G. and J. Kaye. (eds). 1973. 42—50.
- Kaye, J. 1974. "Opacity and recoverability in phonology". *Canadian journal of linguistics* 19. 134—149.

²⁴ The existence of abstract MSC's and WSC's is not accidental. See Kaye (to appear) for discussion.

- Kaye, J. 1975a. "Contraintes profondes en phonologie: les emprunts". *Cahier de linguistique de l'UQAM* 5. 87—101.
- Kaye, J. 1975b. "A functional explanation for rule ordering in phonology". *CLS 11 (Parasession on functionalism)*. 244—252.
- Kaye, J. 1979. "On the alleged correlation of markedness and rule function". In Dinnsen, D. (ed.). 1979. 272—280.
- Kaye, J. to appear. "Abstractness, recoverability and phonotactic constraints". In Goyvaerts, D. L. (ed.). forthcoming.
- Morin, Y.-Ch. 1976. "Contrainte de structure morphématique en bourouchaski". *Recherches linguistiques à Montréal* 6. 197—204.
- Piggott, G. and J. Kaye. (eds). 1973. *Odawa language project: second report*. Toronto: Centre for Linguistic Studies, University of Toronto.
- Postal, P. 1968. *Aspects of phonological theory*. New York: Harper and Row.
- Shibatani, M. 1973. "The role of surface phonetic constraints in generative phonology". *Language* 49. 87—106.
- Stampe, D. 1973. A dissertation on natural phonology. Unpublished Ph. D. thesis, the Ohio State University.
- Walker, D. 1975. "Contraintes profondes en phonologie française". *Cahier de linguistique de l'UQAM* 5. 77—85.