JL12 (1976) 1–208 Printed in Great Britain

The Duke of York gambit[1]

GEOFFREY K. PULLUM

University College London

(Received 22 July 1975)

A Duke who has not been satisfactorily identified with any historical figure is lampooned in a traditional rhyme (believed to have been directed originally at the King of France) as follows:

The Grand Old Duke of York
He had ten thousand men
He marched them up a great high hill
And he marched them down again.

The implication is that this was incompetent and self-defeating activity on his part. Linguists very frequently seem to give evidence of a tacitly held belief that there is similarly something inept and risible about a linguistic analysis which determines that certain structures are assigned a derivation of the general form A→B→A, that is a derivation in which an underlying representation (or some nonultimate remote representation) is mapped on to an intermediate form distinct from it, and then on to a surface (or other superficial) representation which is identical with the earlier stage. In this paper I shall illustrate this prejudice among linguists from the literature of generative grammar, and discuss the legitimacy of the descriptive strategy to which the prejudice takes exception. Derivations that in some relevant respect have the form A→B→A I shall call *Duke of York derivations*, and the strategy of postulating such a derivation in order to achieve a description of some piece of linguistic data I shall call *the Duke of York gambit*.

The question to which this paper will very largely be addressed is whether the tacit rule of thumb 'don't use a Duke of York derivation if it is possible to avoid it' should be, or could be, promoted to the status of an actual theoretical principle. This would not even be under consideration unless the rule of thumb was

[1] This paper was presented to the Spring Meeting of the Linguistics Association of Great Britain at Nottingham in April 1975. Among the many people I want to thank for their helpful discussion and correspondence about some of its subject matter are Maria Black, Rosella Bonelli, Mike Brame, Ornella Cheshire, Gaberell Drachman, Wolfgang Dressler, Erik Fudge, Dick Hudson, Jim Hurford, Paul Kiparsky, Roger Lass, Steve Parkinson, Neil Smith, Deirdre Wilson, and Arnold Zwicky. There may not be a single one of these, however, who actually agrees with what I say in the paper. I am also grateful to the Department of Linguistics, University of Cambridge, for hospitality and practical help during the visit to Cambridge (July 1975) which gave me the opportunity to prepare this paper. The paper is dedicated to Arnold Zwicky.

quite well entrenched among practising linguists, so I want first to cite a fair number of cases in which appeal to the rule of thumb is openly made. A fairly clear case is provided by Mey (1968), who discusses two possible approaches to a problem of voicing assimilation in Dutch, and then notes:

> The second alternative seems to complicate the description of the particular case studied: in order to derive the correct form with [d], in [brɑndbar], we first have to devoice a /d/ and then voice it again (127).

Mey apparently takes it that a description that shows certain instances of *d* as deriving via a sequence *d*→*t*→*d* is necessarily more complex than one which does not. The other possibility is that he assumes simply that NUMBER OF RULES APPLYING gives a direct measure of complexity.

Brame & Bordelois (1973) are even more explicit. Speaking of Harris's (1969) analysis of vocalic alternations in Spanish verbs, they remark:

> Although this approach does account for the alternations exhibited..., there are several good reasons for rejecting it....First of all, we may note that Harris requires a good many forms to undergo the change *e*→*i*→*e*. While this in itself is not an argument against Harris' approach, such derivations are nevertheless to be viewed with some suspicion (115).

They add to this (142, fn. 28):

> There is not one, but two derivations permitting this kind of juggling.

Rather than descend to this 'juggling', White (1973) introduces an extrinsic ordering constraint into her description of Klamath phonology. In the derivation of the form [keys], she argues:

> ...it is necessary to assume that the rule of n-Deletion is ordered before Vocalization. If this were not the case we would expect to find **kei:s* or be forced to posit a later rule that would turn *i:* back to the original *y*....Since the addition of a rule that would change *i:* back to *y* is unmotivated by other data and requires the segment to make the rather suspicious trip of *y* to *i:* back to *y*, we reject this alternative as patently *ad hoc* and assume that n-Deletion is ordered before Vocalization (159).

One might question whether the addition of a crucial ordering restriction that is unmotivated by additional data is necessarily less *ad hoc* than the addition of a new rule that is unmotivated by additional data; but note that the only further argument White provides is that a Duke of York derivation is 'suspicious'.

Hogg (1973) is more outspoken than this in a passage which again implies a general recommendation that Duke of York derivations should not be employed. In his Part IV, Section 1.2, p. 10 he takes Perlmutter (1970) to task for

> hypothesizing a derivation in which in the first instance emphatic stress is deliberately incorrectly assigned in order to derive [ən] from underlying /wʌn/, after which the stress is reassigned correctly, in order to generate [æn]. This is a theoretically quite illegitimate practice, which has to be rejected.

Hogg's condemnation is forthright, and his objection seems to be to a derivation which goes from an underlying strong form to an intermediate weak form to a surface strong form. If it is not the Duke of York gambit that Hogg is objecting to, it must presumably be the more general strategy of permitting an intermediate stage in a derivation to contain a representation which would be ill-formed if carried through to the surface. On this broader notion, known as the 'false step', see Zwicky (1974), whose conclusions may usefully be compared to my own.

For a final phonological example, let me quote a passage from Chomsky & Halle (1968) which J. R. Hurford brought to my attention, where the authors criticize their own proposed analysis of short *u* in Early Modern English (270; quoted in Hurford (1971: 24)):

> Wallis' grammar must therefore be assumed to contain a rule exempting certain cases of /u/ from Vowel Shift. Since we have formulated the Vowel Shift Rule so that it applies only to round lax vowels, it would be blocked if /u/ were unrounded in the relevant environments. [...] However, [this] rule (27) turns lax /u/ into [i], which is not attested in Wallis' speech. Thus we now need a rule to undo the effects of (27) and reround precisely those instances of /u/ which rule (27) unrounded. This is hardly an attractive solution and indicates that the proposed rules may stand in need of revision.

Hurford (1971), incidentally, agrees entirely that the analysis is unattractive and needs to be revised, thus endorsing again the general prejudice against Duke of York derivations, and Smith (1973) also alludes to the Chomsky and Halle treatment of *u* in a passage describing Duke of York derivations as 'clearly farcical' (33).

The Duke of York gambit is not limited to phonology, however, and nor is the prejudice against it. Berman (1974) discusses McCawley's (1970) suggestion that English might have VSO word order in deep structure, and presents arguments that a grammar postulating such a constituent order would still need to derive surface verb-initial orders (e.g. questions) via a rule of inversion operating on subject-initial structures. She remarks:

> Intuitions, of course, are not an argument. But when, given certain theoretical assumptions, it turns out that every VS sentence must go through a stage of being SV (and every SV sentence must go through a stage of being VS) it seems that those assumptions are missing something (32).

For one further example from syntax of the general prejudice, consider the following passage from Johnson (1974), discussing the hypothesis that the er-

gative construction of languages like Dyirbal is the result of obligatory application of a Passive transformation. Johnson notes that Dyirbal has a transformation, the *ŋay* transformation, which converts ergative NPs into absolutives, and he argues:

> Given the obligatory Passive hypothesis, the derivation of a sentence like
> (78) bayi yara baŋgun dugumbiru balgalŋay
> man–ABS woman–ERG = INSTR hit + *ŋay*
> 'the man hit the woman'
>
> would be schematically:
>
> (79) a. man_S $woman_{DO}$ hit
> b. $woman_S$ man_X hit + Ø ↓ Passive
> c. man_S $woman_X$ hit + Ø + ŋay *ŋay* transformation
>
> The derivation in (79) [where S = Subject, DO = Direct Object, and X = demoted NP lacking a grammatical relation – GKP] entails that the underlying subject, ***man***, is first demoted and then reinstated to the subject position, making the subject demotion aspect of the Passive rule completely vacuous.... The derivation in (79) seems extremely implausible (97–98).

Here it is not a change in the order of constituents that is involved, but a change of grammatical relations such as 'subject of'. What Johnson thinks is 'extremely implausible' is a derivation that runs: subject→nonsubject→subject.

For a prejudice to be shared by such a disparate set of phonologists and syntacticians suggests very strongly, I feel, that a shared belief must exist in a principle that in some way legitimizes the selection of analyses not employing the Duke of York gambit over analyses employing it. I shall now discuss briefly the DESIRABILITY of defending such a principle, and then proceed to examine the FEASIBILITY.

Of the desirability of formulating and maintaining as strong a version as possible of a principle forbidding the use of the Duke of York gambit in linguistic description, there can be little doubt. It continues to be a major embarrassment to transformationalist linguistics that the most explicit attempts at providing a theoretical framework for linguistic description still leave open far too large a number of options in almost every case where a specific subarea of a language is to be described.[2] Any theoretical principle that will decrease this latitude,

[2] I should stress at this point that I recognize the existence of the logical possibility of non-transformational generative models in which there is no concept of DERIVATION (i.e., roughly, no sequences of representations of the same linguistic object in a single representational alphabet) and thus no possibility of any analogue of the Duke of York gambit. Thus Lamb's stratificational grammar would appear not to provide for the possibility of A→B→A derivational sequences since each stratum is defined to include representations in a different alphabet from all other strata; and the same would seem to be true of theories of phonology in which words were listed in the lexicon in their

lessening the huge number of competing analyses the theory would entertain as candidates in a particular case, must be a desirable accretion to the theory. A good example of such a principle is the Alternation Condition of Kiparsky (1968). This defines as inadmissible all phonological descriptions in which morphemes that lack alternant forms have lexical entries that differ from their phonetic representations in any respects other than those dealt with by the automatic rules of allophony in the language. If it is tenable, it reduces enormously the number of conceivable descriptions that the theory defines as permissible, and thus advances considerably the enterprise of finding the description that is most satisfactory (and thus hopefully 'correct').

In the case of the prohibition of the Duke of York gambit, the theory would be substantially strengthened in a very similar way if the prohibition could be sustained. Not only would it cut down the numbers of *a priori* possible analyses in virtually every case of attempted linguistic description, but it would in numerous cases select against exactly those analyses which I believe to be objectionable on independent grounds, adjudicating correctly in favour of superior ones. Among these cases I would cite the following:

1. The analysis of Spanish vocalic alternations in Harris (1969) would be disallowed, but that of Brame & Bordelois (1973), or the modifications of it proposed in Brame & Bordelois (1974) and Dinnsen (1974), would be admitted; I believe the latter two papers exemplify the best approach so far toward an adequate generative account of Spanish phonology (see Pullum, forthcoming).

2. The interpretation suggested by Howard (1973: 74 ff.) of an analysis of Fang-Kuei Li which has a Chipewyan form [θuh+xar] derived /θuh+xar/→ θuh+ɣar→[θuhxar] is clearly wrong, and would be correctly disallowed in favour of Howard's analysis or some other.

3. Brink (1974) suggests that Dutch *zingt*, normally pronounced [ziŋt], undergoes both a *k*-epenthesis (which he motivates by reference to a proposed derivation /koniŋ-tjə/→koniŋ-k-tjə→[koniŋkjə]) and a *k*-deletion (illustrated by /ziŋk-t/→[ziŋt]), so that its derivation is: /ziŋ-t/→ziŋ-k-t→[ziŋt]; I would be inclined to favour an analysis in which the *k* of [koniŋkjə] was derived via assimilation and not by insertion, with the lexical representation of *-tjə* (the diminutive suffix) being something like *C*jə*, where *C** becomes homorganic

---

phonetic shapes, or a theory of syntax which recognized only a surface level. The proposals I know of for theories along these lines seem to me to be demonstrably inadequate, but that is not my topic here. The problem I am considering arises only if one assumes the paradigm of generative grammar or something similar to it in significant respects, it is true. But that does not in any way mean that it is a pseudo-problem. Just as, say, stratificational grammar will pose new paradigm-linked problems of its own ('How many strata are there?'), most questions of interest in science will strictly make sense only in a paradigmatic framework. The hallmark of contentful scientific discussion is the existence of evidence generally accepted as telling in one or another direction, not the existence of questions which can be asked in the absence of all theoretical presuppositions.

with preceding consonants, and I would therefore agree with a principle that disallowed the Duke of York gambit here.

4. Brink also discusses the case of Dutch sequences such as *goed boek* 'good book', pronounced [ɣut buk] in careful speech with the usual final devoicing, but [ɣud buk] in more casual speech, with assimilation across the word boundary. He gives a derivation involving /ɣud#buk/→ɣut#buk→[ɣudbuk]. It would seem preferable, however, to deal with the two speech styles here by reference to the oft-noted observation that casual speech is characterized by reduction in the strength of boundaries. If the word boundary in /ɣud#buk/ were weakened to the status of a morpheme boundary intra-phrasally under casual speech conditions, the structural description of final devoicing would not be met and no Duke of York derivation would be needed here. A constraint against Duke of York derivations would appear to select in quite a reasonable way in this case.

5. Chomsky & Halle (1968) derive the English short [ʌ] from an underlying /u/, and account for the [u] of the class of words that includes *push, pull, bush, bull, full*, etc. by preserving certain instances of /u/ from the lowering process. They achieve the preservation by temporarily converting those instances of /u/ into a high back unrounded vowel *i̇*, so that phonetic [u] derives via /u/→i̇→[u]. This analysis, which is thoroughly scrutinized in Pullum (1974), is completely untenable, as a prohibition of Duke of York derivations would correctly predict.

6. The Chomsky and Halle analysis of *table* and *tabular* similarly seems to me to be a *reductio ad absurdum* of the position that even distant etymological relationships are to be captured in the grammar if rules can be constructed to achieve this. The normal pronunciation [tébjələr] they derive through a sequence involving underlying /tæbulær/ and intermediate *tæbj̄ūlær* and *tæbj̄ūlær*, which runs [−tense]→[+tense]→[−tense]. The variant [tébjūlər] they generate from /tæbulær/ via the steps *tæbj̄ūlær*, *tæbj̄ūwlær*, which involves [+round]→[−round]→[+round]. The whole enterprise seems ill-motivated (while *modular* might be felt by English speakers to contain *module* as a subpart, it is not felt to be closely related to *model* or *mode*, and similarly *tabular* is not felt to relate directly to *table*), and if a constraint against the Duke of York gambit forbade the capturing of such doubtful 'generalizations' this would seem to be a point in its favour.

7. Chomsky and Halle also suggest (213) that the British pronunciation of *cot*, i.e. [kɔt] might undergo the rules that give the American unrounded vowel [a] in [kat], and then undergo an extra re-rounding: /kɔt/→kat→[kɔt]. The complete absence of empirical basis for the claim that British [ɔ] has a presurface *a* stage synchronically makes the prospect of a constraint forbidding such a derivation seem entirely welcome.

What, then, can be said about the FEASIBILITY of adding to phonological theory (and perhaps also to syntactic theory, of which more below) a condition that outlaws absolutely, or attaches great 'cost' to, the Duke of York gambit?

I think it is worth considering first the relation that such a constraint would have to other constraints that have been entertained within the framework of generative grammar. It has been suggested to me, for instance, that Kiparsky's notion of rule opacity seems to relate to the admissibility of Duke of York derivations. Kiparsky's definition of this notion (1971, 621–622) is as follows:

> A rule A→B/C__D is opaque to the extent that there are surface representations of the form
> (i) A in environment C__D or
> (ii) B in environment other than C__D.

The concept of opacity, applied to a phonological rule, is thus essentially that of being 'contradicted on the surface' (as David Stampe has expressed it). It would indeed appear that Duke of York derivations will often involve a rule application that is contradicted on the surface. However, I shall not explore this in detail, but merely remark that even a requirement placed on grammars that no rule application could be opaque (which is an impossibly strong requirement anyway) would not be sufficient to ensure that no Duke of York derivations were admitted. I say that a no-opacity condition would be impossibly strong because it is so clear that late rules can make earlier ones opaque through their effects; consider the Castilian Spanish form *ciudad* [θjuðá] 'city', in which the stress rule that gives final stress on nearly all consonant-final words without inflexional endings has been made opaque through the elision of the final *ð* (cf. *ciudades* [θjuðáðes] 'cities' where the final *ð* shows up). A nonopaque application of stress could easily be obtained if the stress rule applied to the output of the elision, giving *[θjúða] (cf. *casa* [kása] 'house'), but it is the opaque derivation that is the correct one. To show that a Duke of York derivation could occur where there was absolutely no rule opacity of any kind, the simplest way I can devise is to exhibit a phonetically plausible invented case.

Suppose some language contained the forms *káti* 'wallaby', *katínlu* 'wallaby-PLURAL, wallabies', *katenlóma* 'wallaby-PLURAL-ERGATIVE, by wallabies'. Suppose further that examination of the data in this language revealed the functioning of the following four rules in the language:

A. Penultimate vowels are stressed.
B. High vowels become mid in unstressed, closed syllables.
C. Final *n* deletes after mid vowels.
D. Final mid vowels become high.

Katenlóma could then be derived very simply via rule A and rule B in this way:

$$\underset{\qquad\qquad A \qquad\qquad B}{\text{/katin+lo+ma/}\rightarrow\text{katin ló ma}\rightarrow\text{[katenlóma]}}$$

and *katínlu* would undergo rules A and D:

A D
/katin + lo/→katín lo→[katínlu]

but *káti*, assuming simply the underlying form /katin/ that is clearly visible in *katínlu*, would undergo all four rules:

A B C D
/katin/→kátin→káten→káte→[káti]

But notice that although the final vowel of *káti* is derived by means of the Duke of York gambit (*i*→*e*→*i*), none of the forms cited contradict any of the rules on the surface; there is no opacity at all. All and only penultimate vowels are stressed; mid vowels always and only occur in unstressed, closed syllables; *n* never appears finally after a mid vowel; only high vowels appear in final position. There is more to the Duke of York gambit, then, than the permitting of opacity in derivations.

In a very similar way, one might think that the question of abstractness in phonological description was closely bound up with the question of the Duke of York gambit. This is not so in any simple way. It can be shown that a constraint against the gambit would not necessarily militate either for or against abstract description in the sense of Kiparsky (1968).

An excellent example of compatibility between prohibition of the Duke of York gambit and the adoption of a more concrete approach to phonology is provided by the case of linking and intrusive *r* in British English (see Johansson, 1973). Linking *r* is the *r* that appears on a word like *severe* [siviə] when a vocalic-initial morpheme is attached: *severest* [siviərist]. Notice that forms like *severity* seem to indicate that an underlying *r* is present. Intrusive *r* appears in word sandhi even on words like *formula* [fɔ:mjələ], in phrases like *the formula is* [ðəfɔ:mjələriz], despite the fact that related forms show no underlying final *r* in *formula*: cf. *formulaic* [fɔ:mjəlejik]. The four logically possible analyses, as far as status of *r* in underlying representations is concerned, are:

(i) No forms in British (RP) English have underlying final *r*.
(ii) Underlying final *r* is present in contexts where linking *r* appears but not elsewhere.
(iii) Underlying final *r* is present in contexts where intrusive *r* appears but not elsewhere.
(iv) Underlying final *r* is present wherever either linking or intrusive *r* appears.

Of these, (iii) is obviously absurd since it treats *formula* but not *severe* as having a final *r*; no more need be said about it. Of the three that one might seriously entertain, (i) is a highly concrete analysis in which *severe* and *severity* have underlying forms so close to their phonetic representations that there cannot

be said to be a unique morpheme shape that they share, so that to some extent they are treated as phonologically unrelated forms (like, say, *sound* and *sonority*); (ii) is an analysis which saves the principle of unique underlying shapes for morphemes like *sever-* but has a crippling defect in that it misses completely the generalization that linking and intrusive *r* have exactly the same environments (after /əː, ə, aː, ɔː, iə, eə, uə/); and (iv) is the antithesis of (i), postulating underlying final *r* everywhere, including items like *formula* as well as those like *severe*. Johansson (1973) presents arguments against (iv) that I find completely convincing; among them:

(a) Linking *r* alternates with ʔ for some speakers, which would be very hard to explain if it were not epenthetic.

(b) Some dialects do not use linking *r* in the environment...rV_#V, e.g. in *the emperor is*, which would be more easily explicable by means of a blocking condition on an epenthesis rule than by postulating differences in underlying representations.

(c) *Saw it* [sɔːrit] would require underlying /siːr/, or /seːr/ assuming a vowel shift rule, for the representation of the verb *see* if the *r* were not an inserted element. John Wells, in a paper read to the Linguistics Association in 1974, has cited many more grossly counter-intuitive underlying representations that would be necessary: /ʃaːr/ *Shah*, /kæfkər/ *Kafka*, /rɔːr/ *raw*, etc.

(d) Wrong predictions about elision and speech style follow from analysis (iv): *r* deletion would be a consonantal elision rule applying more in slow, careful speech than in fast or casual speech, which is not how elision rules typically behave.

(e) RP speakers extend *r* insertion to foreign languages they are trying to learn, which is a mystery under the underlying *r* hypothesis but a natural consequence of the insertion rule hypothesis.

What this means is that analysis (i) has overwhelming support. It is significant, then, that a constraint against the Duke of York gambit would throw its weight in a correct direction in this case: it is analysis (ii), one of the analyses we would wish to reject, that would be barred, since it would incorporate both an *r* deletion rule (for *severe* etc.) and an *r* insertion rule (for *formula is* etc.) and would therefore determine certain derivations in which both rules applied, as for example in the derivation of *severest* (note that *-est* is preceded by a # boundary in all words except the exceptional *longest*, *strongest*, *youngest*):

/siviər#ist/→siviə#ist→[siviərist]

The condition that no Duke of York derivation can be permitted therefore lends support (in conjunction with much other evidence) to the adoption of a non-abstract solution, namely (i).

Another good example of such a result may be found in the analysis of Brazilian

Portuguese in Brasington (1971), although in this case the question of whether an abstract analysis should be adopted or not is very much more open. An abstract analysis would certainly seek to provide for a unique underlying phonological shape for the morpheme meaning 'lion' which is presumably present in both *leão* [liə̃w̃(ŋ)] 'lion' and *leoa* [lióə] 'lioness'. Brasington offers such an analysis, but shows good reason for postulating both a nasalization rule and a denasalization rule, which results in a derivation of *leoa* that runs [−nasal]→ [+nasal]→[−nasal] as far as the penultimate vowel is concerned:

leão	*leoa*	
lión+i	lión+a	Underlying representations
lión	–	Deletion of class vowel *i* after *n*, *l*, *r*
lián	–	Lowering of *o* before word-final nasals
liã̆n	liõ̆na	Nasalization of vowels before nasals
liã̆w̃n	–	Glide insertion after *a* and *e*
–	liõ̆a	Deletion of nasals between vowels
–	lióa	Denasalization of nonlow vowels before *a*
[liə̃w̃(ŋ)]	[lióə]	Other rules

Native speakers of Brazilian Portuguese that I have asked do feel that *leão* and *leoa* are closely related forms, differing primarily in the gender shown by their endings. Brasington's analysis captures this (and many other aspects of the phonology, particularly with respect to singular/plural relations). A constraint against the Duke of York gambit would apparently force a reanalysis of *leoa* that would be considerably more concrete than this (perhaps one that treated nasalization as a strictly phonetic process and replaced denasalization by a morphological rule of *n*-deletion, as argued for in Parkinson, 1974), and would thus again show solidarity with the anti-abstractness position.

There are instances of solidarity with the other side, however. The dispute over the treatment of Dutch voice assimilation between Mey (1968), Hubers & Kooij (1973), and Mey (1973) illustrates this. The kind of data that are relevant have been mentioned above in connection with Brink's discussion of phrases like *goed boek*: should assimilation be direct and one-step in application so that no assimilation (or only vacuous application of a rule) takes place here, or should the assimilation rule that voices consonants that are followed by voiced stops apply only after another rule of final devoicing that would give [ɣut] as an intermediate stage for *goed*? Mey defends the one-step proposal, while Hubers and Kooij defend the two-process analysis. The data and the arguments are quite complex, and I will not review them in detail here. All I want to point out is that it is the Mey analysis that is abstract in the sense of Kiparsky (1968), for Mey is compelled to hypothesize that in underlying representations all initial fricatives in Dutch are voiceless, and that they become voiced through a general rule of initial fricative voicing. The sporadic instances of voiceless initial fricatives

at the phonetic level, largely in borrowed words, are treated as exceptions to the voicing rule. This violates Kiparsky's Alternation Condition, since nonalternating forms are being given lexical representations that differ significantly from their broad phonetic representations. Yet it is Mey's abstract analysis that dispenses with the Duke of York gambit, and Hubers and Kooij's concrete alternative that invokes it in derivations like /streid#bar/→streit#bar→[streidbar] *strijdbaar*. The issues of whether or not to employ a Duke of York derivation and whether or not to relax Kiparsky's conditions on abstractness of phonological descriptions are therefore independent.

Consideration of one further possible constraint on descriptions, that of prohibiting extrinsic ordering constraints, leads to an even more interesting and (to me) surprising conclusion. It would seem very natural to assume that on the whole it would be analyses which permitted rules to be placed in a fixed order that would provide most readily for Duke of York derivations in any desired situation. One might even question whether any Duke of York derivations at all could result if rules were simply applicable at all points where their structural descriptions were met and no extrinsic sequencing was allowed by the metatheory at all. Closer examination of cases reveals that this assumption is without basis. First, the reader may verify directly that in my invented language above, no statements whatsoever need be made about the order in which the rules must apply in order to generate *káti*, *katínlu*, and *katenlóma*. If a rule is applicable it applies, if it is not it does not, and the phonetic forms are derived when a representation is achieved that does not meet the structural description for any rule. This shows that in principle a Duke of York derivation can easily arise where there is no extrinsic ordering. But second, I have discovered a case in which the very requirement of no extrinsic ordering ensures that a Duke of York analysis must be chosen, given basic assumptions about rules and representations.

The example is from Nootka, and is discussed in Campbell (1973). Campbell takes his data from Sapir and Swadesh's *Nootka grammar and texts*, in which he states there is evidence that in Nootka, *k* is rounded if it follows *o*, but also k^w is unrounded in word-final position. Campbell presents the case as one which is incompatible with the hypothesis of Koutsoudas, Sanders & Noll (1974) that extrinsic ordering is never necessary in grammars. Given a proper understanding of how derivations actually work in the theory being developed by Koutsoudas and his students and colleagues, it can be shown that Campbell is completely wrong in this. In particular, it must be kept in mind that as Ringen (1973) has shown, neither vacuous application of rules nor reapplication of rules in interrupted sequence can be permitted in unordered-rule theories. The situation that Campbell addresses himself to, that of an underlying sequence *ok#* in Nootka, will in fact be correctly described without any extrinsic ordering in the following way:

o k #	Underlying representational string.
↓	ROUNDING applies ($k \rightarrow k^w$/o__), but UNROUNDING would be vacuous here and so must not apply.
o k^w #	
↓	UNROUNDING now applies ($k^w \rightarrow k$/__#) but ROUNDING, which would be vacuous here, does not apply and is discarded.
[o k #]	

The last line constitutes the final output because ROUNDING has been discarded and cannot reapply in this derivation, and UNROUNDING could only vacuously apply and is therefore also discarded. No other rules are applicable, and the derivation terminates. The output, [ok]#, is the correct one. What happens, then, is that the freedom from ordering constraints in this case is itself responsible for determining a Duke of York derivation. It determines, in fact, a Duke of York derivation to which I am unable to see any alternative if the facts are as Campbell asserts and if all constraints on rule application are to be universal ones as argued in Pullum (forthcoming) following Koutsoudas, Sanders & Noll (1974) and many other works.

What I feel has been established so far is that the evidence shows the Duke of York gambit to be an independent issue, cutting across many other issues in phonological description. If it is to be required that no description contain an instance of the Duke of York gambit for any form generated, this will have to be accomplished by an *ad hoc* theoretical principle in the purely nonpejorative sense that the constraint will be set up for this special purpose. Independent constraints on the form of phonological descriptions will in some cases be in agreement with an anti-Duke of York principle and sometimes not.

In order to decide whether to adopt some version of the latter principle and make it part of the theory, we would need to have a sizeable body of unequivocal evidence that supported it. Much study of the forty relevant problem areas that I have so far uncovered convinces me that this is on the whole lacking. All too often the verdict, in crucial cases of clear importance in which masses of different types of relevant evidence are at hand, has to be inconclusive. For instance, in Kiparsky (1973), and Howard (1975) there is a relevant debate about a Sanskrit problem. Kiparsky shows evidence for the superiority of a theory of phonology in which disjunctive application of certain rules is achieved by means of the 'Elsewhere Condition' with the result that a derivation

/ajuṣa+dhwam/ → ajuṣadhuam → [ajuṣadhwam]

is replaced by the straightforward

/ajuṣa+dhwam/ → [ajuṣadhwam]

thus eliminating a Duke of York derivation. Howard, on the other hand, impugns the motivation for the 'Elsewhere Condition' and casts doubt on some of the supporting evidence Kiparsky provides, selecting finally an analysis of the Sanskrit data which does involve a Duke of York derivation. Adjudication between the two would involve absolutely firm decisions in areas which are in fact as open to debate as the Duke of York gambit itself – most notably in the area of phonological interpretation of metrical evidence in ancient texts and constraints on the relations between metrical and phonological structure (the latter being the source of some crucial supportive evidence for Kiparsky). The whole intricate tangle of evidence and counter-evidence represents a scientific puzzle of absorbing interest but will not provide the support for a pillar of theory. It is not untypical in this.

I shall turn, therefore, to some syntactic evidence in the hope that clearer decisions can be made. While no one would think that a result in syntax necessarily entailed a parallel in phonology, there could be heuristic value in knowing whether other areas of grammar were proving compatible or incompatible with a general constraint under consideration in phonology (the case of the extrinsic ordering controversy obviously suggests this).

In syntax it is not hard to exhibit cases where the Duke of York gambit would have satisfying consequences. One instance I would cite is that of the 'disagreement rule' of Abkhaz/Abaza formulated by Anderson (1974). Anderson takes the agreement clitic *y-* which shows up on Abkhaz/Abaza verbs to be an inserted co-referential copy of a subject NP. The fact that it does not appear when the subject immediately precedes the verb he captures by deleting it after insertion when the subject is immediately preverbal; we have, therefore:

$$NP_i - V \rightarrow NP_i - {}_V[y_i + V]_V \rightarrow NP_i - {}_V[\emptyset + V]_V$$

There seems to me to be very good evidence that the remotest underlying structures contain logical variables, the remote structure for a sentence like *Jesus wept* having approximately the form (*Jesus*, *x*)(*wept*, *x*). Much evidence for such a view is given in Keenan (1972), where it is also shown that many languages display surface traces of variables in certain positions, e.g. in verbal morphology. We might therefore reinterpret the Abkhaz/Abaza situation in these terms, with deletion of the underlying variable only if the subject NP immediately precedes, otherwise spelling out of the variable in the verb as a *y-* prefix:

$$(NP, x)\,(x, V) \rightarrow (NP, x)\,(\emptyset, V)$$

If Duke of York derivations of the type $\emptyset \rightarrow A \rightarrow \emptyset$ were forbidden, this analysis would be preferred to Anderson's, and I think that is a correct preference.

It is also easy, however, to exhibit bad predictions that would be made by the Duke of York prohibition. It would appear to predict that an inserted element

like the *there* that shows up as dummy subject in locative/existential sentences in English could not be deleted by a transformation. But this is surely wrong; Conjunction Reduction seems to delete it:

A gorilla is in the kitchen and an alligator is in the bath.
There is a gorilla in the kitchen and there is an alligator in the bath.
There is a gorilla in the kitchen and ~~there is~~ an alligator in the bath.

and the casual speech deletion rules discussed in Schmerling (1973) certainly do:

There's a gorilla in the kitchen.
'S a gorilla in the kitchen.

To object that here a very different process is involved in deleting the inserted item would be to miss the point. If a constraint that something inserted could never be deleted were to have any content or utility, it would be in exactly this sort of case that it should prove itself. An anti-Duke of York condition that said merely that Duke of York analyses would be tolerated in sensible analyses but forbidden where they were vicious, redundant, or inadequate, would be no constraint at all.

I believe I can show more clearly than this that some insertion–deletion analyses in syntax (while not necessarily correct in every detail) are quite reasonable ones to maintain. Consider the following three assumptions:

(a) Extraposition is a cyclic rule that replaces a subject or object clause by a dummy of third person singular neuter form.
(b) Agreement between subject and verb is determined after the cyclic rules have applied.
(c) Every clause has a subject NP through to the end of the cycle on that clause.

Assumption (a) is now widely held, for English at least, and is justified to some extent in Postal (1974) and Jacobson & Neubauer (1974). Assumption (b) can apparently not be held as a universal (Keenan, 1974; Hale, 1975) but is well established for English and a number of other languages. Assumption (c) is valid for English, and almost certainly tenable in the case of numerous other languages despite their lack of superficial pronominal subjects in many sentences; it is only necessary to make the very natural additional assumption that there is a rule of Subject Pronoun Drop which deletes pronominal subjects (optionally but preferredly) if they are not emphatic. Such a rule is postulated and well supported in Perlmutter (1972). I have found no reason to believe that assumptions (a)–(c) may not be tentatively adopted with respect to Italian; indeed, I think each of them can be justified independently for Italian. Consider, then, the derivation they jointly determine for an Italian sentence like *É difficile fare quello* 'It is difficult to do that'.

[PRO fare quello]–COPULA–difficile	Remote structure
[lo]–COPULA–difficile–[PRO fare quello]	Extraposition
[lo]–é–difficile–[PRO fare quello]	Agreement
Ø–é–difficile–[Ø fare quello]	Subj. Pro. Drop

The point to be noted for present purposes is that [lo] 'it' is inserted by the cyclic rule of Extraposition and later deleted (always, since it is logically empty and cannot be emphasized) by Subject Pronoun Drop. I am convinced that this Duke of York analysis is correct. Since Kiparsky & Kiparsky (1970) support has virtually disappeared for the idea that the *it* of Extraposition in English is a deep structure element, and I know of no support for underlying *lo–S* structures in Italian; it would be question-begging to use such structures just to avoid the Duke of York gambit. Keeping the inserted pronoun until the end of the cycle ensures that Agreement can work in a unified way: always by reference to what is the subject NP at the end of the cycle on the clause in question. Making Subject Pronoun Drop a postcyclic rule is fully consistent with what we know about the properties of postcyclic rules (Imperative Subject Deletion, Complementizer Deletion, Preposition Deletion, etc.). Even more significantly, evidence can be found in Italian that the invisible pronoun inserted by Extraposition actually exists, for it can be induced to become visible.

Extraposition in Italian operates with some verbs from object position as well as subject position. This results in an inserted pronoun which cannot get deleted, since Italian has a rule of Subject Pronoun Drop but no rule of Object Pronoun Drop. The inserted pronoun therefore simply moves to become a proclitic on the auxiliary or the main verb in the usual way. Thus from a remote structure like:

io-ho – detto-[di lasciarmi in pace]- ve – tante volte
I have said to leave me in peace to-you many times

(in which the subject of *lasciare* 'leave' has already been deleted by Equi NP Deletion), Extraposition inserts the neuter pronoun *lo* to replace the bracketed object clause and derives:

io-ho – detto- lo – ve – tante volte – [di lasciarmi in pace]
I have said it to-you many times to leave me in peace

from which the normal cliticization rule for object pronouns and the previously mentioned Subject Pronoun Drop derive:

Ve l' ho detto tante volte di lasciarmi in pace
to-you it I-have said many times to leave me in peace
'I have told you many times to leave me in peace.'

Another example is:

Te lo chiedo sul serio di andar via
thee it I-ask seriously to go away
'I ask you seriously to go away.'

This evidence (for which I am indebted to Maria Black) shows that an Extraposition rule seems to exist in Italian which operates on clauses in object position and inserts a dummy object *lo*. In the absence of evidence showing that Extraposition from subject position has to be treated as a distinct process operating in a completely different way (say, by reordering instead of insertion), I take this to support the postulation of an inserted dummy subject in *É difficile fare quello*.

Note also the suggestive fact (pointed out to me by Rosella Bonelli) that in the dialect of Florence the latter sentence generally occurs as *L'é difficile fare quello*, with the dummy subject *lo* that I postulate actually appearing in surface structure.

I conclude that the evidence from Italian indicates that an insertion–deletion analysis in Extraposition sentences is more plausible than any obvious alternatives, which speaks against the idea that Duke of York derivations could be generally outlawed.

My final observation in this paper will be that the same thing seems to be true in the case of another type of possible Duke of York derivation, namely double movements in which a constituent is moved in one direction past a given reference point and then moved in the other direction back past it again.

There are a number of cases, it is true, where independent evidence shows that a prohibition of double movements would correctly rule out bad analyses. A derivation in which a Tough Movement (Nonsubject Raising) rule operated on the output of Extraposition to derive *John is easy to please* from *It is easy* [*PRO please John*], for example, would be one in which *John* was moved first to the right and then back to the left of the predicate *easy:*

[PRO please *John*] be *easy*
It is *easy* [PRO please *John*]
John be *easy* [to please]

This is defined as an impossible derivation by a principle of Perlmutter and Postal known as the Host Limitation Law: a raising rule cannot raise an NP out of a constituent that is not either the subject or the object of its clause. There are good grounds for claiming that the derivation that is correct is:

[PRO please John] be easy → John be easy [to please]

which if it involves movement at all (rather than simply clause membership change in an unordered structure that is later linearized) does not involve double movement.

Again, Zwicky & Zwicky (1973: 929 f.) remove the need for a double move-

ment in dealing with *what for* questions. Lees (1960: 38) had imposed the restriction that NPs cannot be extracted by movement rules from adverbial PPs, in order to block sentences like **What did you leave the parcel at?* (cf. *Where did you leave the parcel?*). This would force a double movement in derivations like this:

Q–you–did–that–*for what*–when you knew it was wrong
for what did you do that when you knew it was wrong
What did you do that *for* when you knew it was wrong?

Zwicky & Zwicky (1973) point out that the rule repositioning the *for* would be very difficult to state since it would have to replace *for* at exactly the point where *for what* came from, which might not be the end of the clause. They also point out that if *what* is regarded as a realization of *wh+something*, a sentence like **What did you leave the parcel at?* can be predicted to be exactly as ill-formed as **You left the parcel at something*, so Lees' constraint is not needed. *What did you do that for...?* can be derived directly from *Q–you–did–that–for wh+something* with no double movement, and should be, as an anti-Duke of York condition would predict.

But another case shows that double movements apparently must be permitted. Grosu (1974) argues that although Ross' 'Left Branch Condition' is an error, it is true that in all languages the leads of complex NPs are frozen (cannot move). This entails that a sentence like *Who do you know who is unfair?* must be derived thus:

Q–you–know–[who [who is unfair]]	Remote structure
[who [who is unfair]] do you know?	Wh-Q Movement
[who] do you know [who is unfair]?	Extraposition from NP

In support of this, note that if we allowed the head of the complex NP [*who* [*who is unfair*]] to move on its own, then from

Q–you–believe–I–told–[who [who is unfair]]–about your plans

we could derive by the application of Wh–Q Movement the sentence

*Who do you believe I told who is unfair about your plans?

If Grosu is right here, we have a case of a sentence which MUST be derived by moving a whole NP to the front of the sentence and then moving part of it back to the end, and a complete ban on such derivations would wrongly forbid this.

The last word on double movements should perhaps be left to Susumu Kuno, who argues in great detail in his (1971) paper 'Position of locatives in existential sentences' for an analysis that could in principle give rise to derivations like this:

A BOOK IS *ON THE TABLE.*	Deepest underlying structure
On the table is a book.	Early preposing of locatives
There is a book *on the table.*	Locative postposing and copying
On the table, there is a book.	Postcyclic preposing of adverbs

Recognizing that this might conceivably boggle a tender mind, Kuno says (377):

> This back-and-forth movement of locatives in the derivation is unusual, to say the least, and is not shared by derivations of other patterns of English. However, there is no a priori reason for rejecting it outright. It might be that this is a fact about the language....

I offer no judgment on the empirical standing of Kuno's analysis of word order in locative/existential sentences, but only call attention to his methodological remark. I find I am forced to agree with him that we have no *a priori* reason for rejecting a derivation such as he sketches. It could indeed just be that the grammar had to be set up that way for an optimal description of the language.

My conclusion, in fact, is that a thorough investigation of the descriptive problems in which proposed Duke of York analyses are embedded does not reveal any basis for a general constraint that would prohibit the Duke of York gambit. This means that all the linguists quoted at the beginning of the paper (Mey, Brame and Bordelois, White, Hogg, Hurford, Chomsky and Halle, Smith, Berman, Johnson, etc.) are in the strictest sense mistaken on a methodological matter. The guiding principle they implicitly appeal to in casting suspicion on Duke of York analysis cannot and does not exist; the rule of thumb they apparently follow when in doubt is like the Free Ride Principle (Zwicky, 1970) in that there are points both for it and against it but it cannot be trusted to select correct analyses over inadequate ones. As with taking 'false steps' (Zwicky, 1974) or 'homing in' on an underlying representation, the strategy of avoiding or of adopting) the Duke of York gambit will be reasonable precisely when the result is a reasonable analysis, and unreasonable precisely when it is not. There is no simple maxim for the descriptive linguist here.

And the point is not merely of methodological concern for linguists. If it is indeed true that the performance of the infant learning to speak and understand his language is something that may be appropriately modelled as a process of constructing and internalizing a fully adequate grammar of the language, we can say that although it could conceivably make the task easier, the child is NOT equipped with a subconscious instruction 'AVOID CONSTRUCTING A GRAMMAR THAT DEFINES DUKE OF YORK DERIVATIONS'. We are reminded again that we have so few clear ideas about how the child IS equipped that as yet we can hardly imagine how he does what he does.

REFERENCES

Anderson, S. R. (1974). On dis-agreement rules. *LIn* **5**. 445–451.
Berman, A. (1974). On the VSO hypothesis. *LIn* **5**. 1–38.

Bierwisch, M. & Heidolph, K. E. (eds.) (1970). *Progress in linguistics: a collection of papers.* The Hague: Mouton.
Brame, M. K. & Bordelois, I. (1973). Vocalic alternations in Spanish. *LIn* **4.** 111–168.
Brame, M. K. & Bordelois, I. (1974). Some controversial questions in Spanish phonology. *LIn* **5.** 282–298.
Brasington, R. W. P. (1971). Noun pluralization in Brazilian Portuguese. *JL* **7.** 151–177.
Brink, D. (1974). On characterizing the natural order of application of phonological rules. *Lingua* **34.** 47–72.
Campbell, L. (1973). Extrinsic order lives. Duplicated, Indiana University Linguistics Club.
Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.
Dinnsen, D. A. (1974). The fallacy of strict sequentiality. Duplicated, Indiana University Linguistics Club.
Grosu, A. (1974). On the nature of the Left Branch Condition. *LIn* **5.** 308–319.
Hale, K. (1975) Counterexamples and explanations in Navajo linguistics: syntax. Navajo Language Review (Washington D.C.: Center for Applied Linguistics) **2.** 29–60.
Harris, J. W. (1969). *Spanish phonology.* Cambridge, Mass. & London: MIT Press.
Hogg, R. (1973). *Determiner and number systems in contemporary English.* Preliminary version. Duplicated, University of Amsterdam and University of Lancaster.
Howard, I. (1973). *A directional theory of rule application.* MIT Ph.D. dissertation. Duplicated, Indiana University Linguistics Club.
Howard, I. (1975). Can the 'Elsewhere Condition' get anywhere? *Lg* **51.** 109–127.
Hubers, G. A. C. & Kooij, J. G. (1973). Voice assimilation in Dutch. *Acta Linguistica Hafniensia* **14.** 25–33.
Hurford, J. R. (1971). The state of phonology. *Linguistics* **71.** 5–41.
Jacobson, P. & Neubauer, P. (1974). Extraposition rules and the cycle. *Berkeley Studies in Syntax and Semantics* **1.** VIII/1–VIII/99.
Johansson, S. (1973). Linking and intrusive /r/ in English: a case for a more concrete phonology. *SL* **27.** 53–68.
Johnson, D. E. (1974). *Toward a theory of relationally based grammar.* Ph.D. thesis, University of Illinois at Urbana-Champaign.
Keenan, E. L. (1972). Semantically based grammar. *LIn* **3.** 413–461.
Keenan, E. L. (1974). A universal definition of 'subject of'. Mimeographed, University of California, Los Angeles.
Kiparsky, P. (1968). How abstract is phonology? Duplicated, Indiana University Linguistics Club.
Kiparsky, P. (1971). Historical linguistics. In Dingwall, W. O. (ed.), *A survey of linguistic science.* College Park, Md.: University of Maryland.
Kiparsky, P. (1973). 'Elswhere' in phonology. In Anderson, S. R. & Kiparsky, P. (eds), *A Festschrift for Morris Halle.* New York: Holt, Rinehart & Winston.
Kiparsky, P. & Kiparsky, C. (1970). Fact. In Bierwisch, M. & Heidolph, K. E. (eds.), *Progress in linguistics: a collection of papers.* 143–173.
Koutsoudas, A., Sanders, G. A. & Noll, C. (1974). On the application of phonological rules. *Lg* **50.** 1–28.
Kuno, S. (1971). Position of locatives in existential sentences. *LIn* **3.** 333–378.
Lees, R. B. (1960). *The grammar of English nominalizations.* (= *IJAL* **26,** 3, Part 2.) Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication 12.
McCawley, J. D. (1970). English as a VSO language. *Lg* **46.** 286–299.
Mey, J. (1968). A case of assimilation in Modern Dutch. *Acta Linguistica Hafniensia* **11.** 123–145.
Mey, J. (1973). Dutch treat: a reply. *Acta Linguistica Hafniensia* **14.** 35–37.
Parkinson, S. (1974). *Nasality in modern Portuguese.* Diploma dissertation, Department of Linguistics, University of Cambridge.
Perlmutter, D. M. (1970). The article in English. In Bierwisch, M. & Heidolph, K. E. (eds.), *Progress in linguistics: a collection of papers.* 233–248.

Perlmutter, D. M. (1972). Evidence for shadow pronouns from French relativization. In Peranteau, P. *et al.* (eds.), *The Chicago which hunt.* Chicago: Chicago Linguistic Society. 73–105.

Postal, P. M. (1974). *On raising.* Cambridge, Mass. & London: MIT Press.

Pullum, G. K. (1974). Sheltering environments and negative contexts: a case against making phonological rules state things that don't happen. *Edinburgh Working Papers in Linguistics* **4.** 31–41.

Pullum, G. K. (forthcoming). Sequential and simultaneous rule application in Spanish phonology. To appear in *Lingua.*

Ringen, C. (1973). Vacuous application, iterative application, reapplication, and the unordered rules hypothesis. Duplicated, Indiana University Linguistics Club.

Ross, J. R. (1966). Relativization in extraposed clauses (a problem which evidence is presented that help is needed to solve). *Mathematical Linguistics and Automatic Translation.* Report No. NSF-17 to the National Science Foundation. Computational Laboratory, Harvard University.

Schmerling, S. (1973). Subjectless sentences and the notion of surface structure. Papers from the Ninth Regional Meeting, Chicago Linguistic Society. 577–586.

Smith, N. V. (1973). *The acquisition of phonology.* Cambridge: Cambridge University Press.

White, R. B. D. (1973). *Klamath phonology.* Studies in Linguistics and Language Learning **12.** Seattle: University of Washington.

Zwicky, A. M. (1970). The free ride principle and two rules of complete assimilation in English. *Papers from the Sixth Regional Meeting,* Chicago Linguistic Society, 579–588.

Zwicky, A. M. (1973). Homing in: on arguing for remote representations. *JL* **10.** 55–69.

Zwicky, A. M. (1974). Taking a false step. *Lg* **50.** 215–224.

Zwicky, A. M. & Zwicky, A. D. (1973). *How come* and *what for.* In Kachru, B. *et al.* (eds.), *Issues in linguistics.* Urbana, Chicago and London: University of Illinois Press.