On segmental complexity

Affricates and patterns of segmental modification in consonant inventories*

Jeroen van de Weijer and Frans Hinskens
University of Leiden / Meertens Institute and Free University Amsterdam

1. Introduction

Affricates occur in over two thirds of the world's languages. They are therefore a common type of segment; as a matter of fact, cross-linguistically they are by far the most common type of complex segments. It is also common for languages to have some type of segmental modification, such as labialization (and other types of secondary articulation), aspiration (and other phonation types) or prenasalization. This paper investigates whether there is a relation between segmental complexity in affricates and in segments with segmental modification. Evidence for such a relation could consist of languages allowing for *both* affricates *and* other types of segmental modification, or languages ruling out both. In the former case, the branching structure that is present in both would be generally permitted, in the latter case it would be generally ruled out. This would amount to a more abstract, structural interpretation of Clements' (2003) notion of "feature economy".

However, it could also be possible — off hand — that languages compensate a lack of affricates by allowing for modified segments. This would argue for a trade-off relationship between complex segments (affricates and perhaps prenasalized stops) and segments with segmental modification. — i.e. between segments with "contouring" complexity and segments with "compounding" complexity.

In Section 2, we describe earlier feature-geometrical approaches to affrication and several types of segmental modification. In Section 3, we investigate the occurrence of both in Maddieson (1984), a large, typologically balanced sample of languages and a statistical investigation of these findings. We will see

Linguistics in the Netherlands 2004 21 (2004), **217–228**.
ISSN 0929–7332 / E-ISSN 1569–9919

that there is usually no trade-off relationship between complex segments and segments with segmental modification. Unlike what might be expected, there is no relation between prenasalized and affricated segments, although both are described as "contour" segments in Sagey (1986).

Section 4 is a brief conclusion. It sums up the findings and places them in an overall framework which links empirical questions to a formal proposal.

2. Affrication and other types of segmental complexity

Traditionally, affricates have been described as complex segments, and, more in particular, as segments which combine features characteristic of stops with features characteristic of fricatives. The following partial representation of affricates is based on Sagey (1986).

(1) *Affricates as contour segments* (Sagey 1986)

o Root node

[-cont] [+cont]

It is a matter of some debate where exactly the feature [cont] is located; in the model presented by Clements & Hume (1995), the structure of affricates is as represented in (2):

(2) *Branching at the Oral Cavity node* (Clements & Hume 1995)

o Root node

o Oral Cavity node

[-cont] [+cont]

We will abstract away from differences such as these.

Representations such as those in (1) and (2) have come under attack in the past few years. For instance, in Hinskens & van de Weijer (2003a), we show on the basis of the large and typologically balanced sample of languages in Maddieson (1984) that affricates behave like stops with respect to the types of segmental modification (henceforth MOD) they take. Other work has also shown that affricates primarily behave like stops in phonological processes and might therefore be represented as such (e.g. Kehrein 2002, Lin, to appear and references cited there).

On the other hand, affricates behave like fricatives with respect to the place of articulation they take: (coronal) affricates typically occur at the same places of articulation as (coronal) fricatives in segmental inventories (van de Weijer 1996).

The present paper is related to this debate, by examining whether the branching relation in (1) falls under the scope of the same constraint that has been argued to be relevant for the type of segmental complexity present in, for instance, secondary articulation. This would represent an extension of Clements' idea of "feature economy" (Clements 2003, following Martinet 1955 and others), one aspect of which is made explicit in (3):

(3) *Prediction 1: Mutual Attraction* (Clements 2003)
A sound S will have a higher than expected frequency in languages that have another sound T bearing one of its features, and vice versa.

Clements (2003) rephrases his Prediction 1 as "if a feature is used once in a system, it will tend to be used again". This paper essentially extends this prediction to a structural (as opposed to featural) level by examining whether languages tend to make use of segmental branching in a range of segments, rather than in just one segment or category of segments. We will rephrase this idea as the hypothesis in (4):

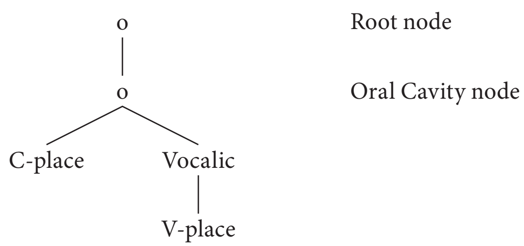
(4) *Hypothesis*
If segmental branching is permitted for a certain segment (or class of segments) in a given language, then it will be permitted at a higher than chance frequency for other segments (or classes of segments) in that language.

If confirmed, then languages allow both affricates and complex segments and if disconfirmed, then there may well be a general tendency for affricates and complex segments to be in a trade-off relationship.

Let us, with this idea in mind, examine what other types of segmental complexity might exist, in order to test this hypothesis. We do not claim that the types of complexity examined here necessarily represent an exhaustive list. For instance, in some theoretical frameworks nasal consonants are also represented as involving some type of branching. In this paper, we will limit our investigation to segment types which involve some articulatory and/or acoustic complexity.

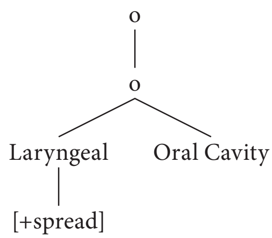
A first obvious candidate is secondary articulation. In Clements & Hume (1995: 287), secondary articulation is represented as follows:

(5) *Secondary articulation* (Clements & Hume 1995)



It is seen that both affricates and segments with secondary articulation involve segmental branching. The same goes for segments with distinctive laryngeal modification, such as aspirated stops (see (6) below).

(6) *Aspirated stops*

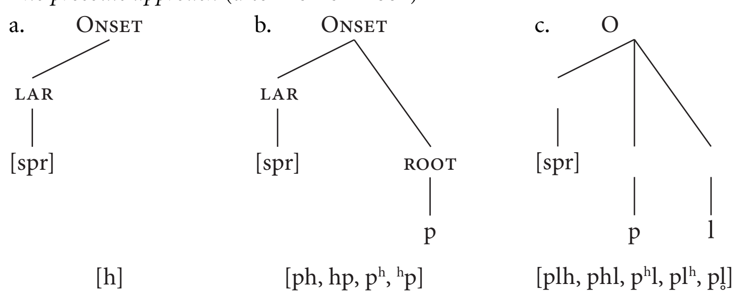


The question is whether affricates and modified segments involve the same relation or whether they are fundamentally different and independent.

One observation to make is that in secondary articulation (and phonatory complexity) there is a "compounding" kind of complexity, in which two nodes (here the C- and V-place nodes) are present (whereas non-complex consonants or vowels only have one). Affricates (and prenasalized stops, see below), on the other hand, have a "contouring" relation of "+" and "−" values of the same feature. If compounding and contouring complexity are not independent, structural constraints may be expected to take both types in their scope.

An issue arises concerning the representation of secondary articulation as in (5). Kehrein (2002) argues that properties such as labialization and palatalization but also aspiration, etc. are not properly described as Root-internal properties. In other words, they do not involve segmental complexity, but rather they should be represented as prosodic properties, i.e. linked to nodes such as Onset, Nucleus, or Coda. Examples of representations that illustrate this idea, taking laryngeal features as the relevant secondary property, are given below (Kehrein 2002: 69).[1]

(7) *The prosodic approach* (after Kehrein 2002)



In individual languages, phonological representations like these are interpreted by the phonetics as one of the realizations given beneath each feature tree; hence, these realizations cannot contrast *phonologically*. For instance, in this proposal, a contrast between pre- and postaspiration is not possible phonemically.

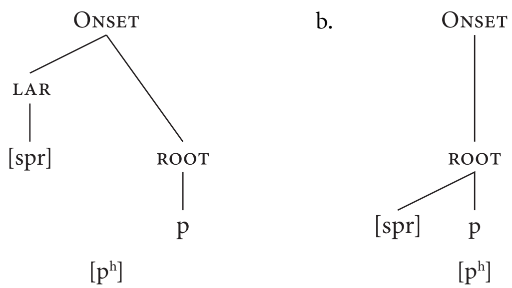
This proposal is very interesting because it makes a large number of very specific predictions, such as the following:

(8)
- contrasts like /ph/ vs. /hp/, /p^h/ vs. /^{h}p/ (etc.) are not possible underlyingly, unlike /p/ vs. /p^h/, etc.
- there can be no conflicting laryngeal specifications within onsets (e.g. */t^{h}n'/, in which /t/ would be aspirated and /n/ would be glottalized)
- assimilation (and neutralization, etc.) of voice (or other laryngeal features) always concern entire prosodic constituents, not individual segments

So far, this proposal has held out well against possible counterevidence. In his thesis, Kehrein looks at a number of possible counterexamples and shows how they can be re-analysed.

Our investigation is to some extent independent of this question, although it assumes, in line with earlier work (Hinskens & van de Weijer 2003a), that there is a close connection between natural class (e.g. velars, or voiceless stops) and segmental modification, maybe even a specific modification type (e.g. labialization, or aspiration, respectively). We prefer to express this by assuming a direct, segmental (not prosodic) relation between both. The difference between Kehrein's and our view can be illustrated with the following two representations of an onset with an aspirated voiceless stop:

(9) a. b.



In (9a), Kehrein's representation of an aspirated stop is repeated. This does not directly express the relation between aspiration and the natural class on which aspiration typically occurs, i.e. that of voiceless stops, while this relation is more directly captured in the more traditional representation of (9b), where both are under the same root node. On the other hand, the fact that aspiration typically occurs in Onsets is more directly captured by (9a). The fact that voiceless stops are preferably in the onset, i.e. a matter of phonotactics or more generally markedness, may reconcile both positions.

The question whether constraints exist that take segment-internal branching in general in its scope could be investigated for a range of other segments which have been described as complex, e.g. prenasalized stops (10) or segments which have a laryngeal MOD type, such as aspirated stops (see (6) above):[2]

(10) *Prenasalized stops*

```
            o
         /     \
[+nasal]        [-nasal]
```

In this paper we examine these types of complexity from a cross-linguistic perspective. Are there languages that ban all types of complexity illustrated so far? This would argue for a maximally general constraint on segmental complexity. Alternatively, if languages ban, say, affricates, do they allow other types of complex segments to compensate, and if so, are these preferably of the secondary articulation (5), laryngeal (6), or the nasal (10) type? We investigate these questions in the next section.

3. Cross-linguistic patterning of affrication and segmental modification

In earlier work (Hinskens & van de Weijer 2003a,b; van de Weijer & Hinskens 2003) we surveyed MOD in the languages of the world, basing ourselves on the data offered in Maddieson (1984). 166 out of the 317 languages in Maddieson's survey have one or several instances of MOD. In all, these 166 languages have 283 instances of MOD. In 55 out of these 283 cases, MOD is completely redundant, in the sense that all segments at issue lack MOD-less counterparts (Hinskens & van de Weijer 2003b).

In (11) we present the types of modification that occur in the languages in the sample. For each single type of modification, we add an abbreviation as well as the diacritic.

(11)

a.	*secondary articulation (supralaryngeal/oral)*		
	labialization	LAB	C^{w}
	palatalization	PAL	C^{j}
	velarization	VEL	C̴
	pharyngealization	PHA	Cˤ
b.	*phonation types (laryngeal)*		
	aspiration	ASP	C^{h}
	preaspiration	PRA	${}^{h}C$
	breathy voice	BRV	C̤
	with breathy release	BRR	$C^{ɦ}$
	laryngealization	LAR	C̰
	ejectivity	EJE	C′
c.	*nasality*		
	prenasalization	PRN	${}^{N}C$

For the present investigation, we included information on affrication for each of the 317 languages in the sample. It turned out that 219 languages (=69%) in the sample had one or more affricates in their segmental inventory. Some language families are particularly rich in affricates (e.g. the Amerindian languages) or particularly poor (e.g. the Australian ones).

In Table 1, we present the main findings for the quantitative patterning of affrication and MOD in the sample.

Table 1. The quantitative patterning of affrication and segmental modification; *n* of languages

		Affricates?		**total**
		yes	**no**	
MOD?	yes	136	30	166
	no	83	68	151
total		219	98	317

The overwhelming majority of languages in the sample ($n=136$) has both affricates and MOD (e.g. Armenian, with both affrication and ejectivity, Sedang with both affrication and prenasalization, and Tlingit, with both affrication and labialization). The second-largest group ($n=83$) is formed by the type of languages that have affricates but lack MOD (e.g. Suena, a New Guinea language, and Andamese, an Indo-Pacific language). There are languages that appear to lack any kind of segmental complexity; of the 98 languages that have no affricates, 68 lack MOD. Kanakuru lacks both affricates and any other type of segmental modification, while Nambakaengo has no affricates but does have labialization and palatalization.

A statistical way of answering the question whether affrication and segmental modification are structurally related or not requires a test of the mutual independence of affrication and MOD. The outcomes of the chi square test ($\chi^2=26.92$ df$=1$ $p<.001$) show that in our sample both are far from independent. On the contrary, there is a highly significant relationship between the two properties.[3] It turns out that languages with affricates in most cases also have MOD and, at the same time, languages without affricates in most cases also lack MOD. So it would appear that the presence or absence of primary complexity goes hand in hand with the presence or absence of secondary complexity, respectively. Thus, there is no evidence for the idea that there is a trade off-relationship between complex segments such as affricates and segments with MOD. On the overall level, our hypothesis in (4) is confirmed by the languages in the sample.

On closer inspection of the figures in Table 1, the situation turns out to be somewhat less straightforward, however. Whereas languages with MOD in most cases also have affricates, most of the languages without MOD nevertheless have affricates. Hence, we face a problem of interpretation: as there is obviously no perfect reciprocity, what is predictable by what?

One way of tackling the problem is a purely quantitative one, i.e. comparing the proportions. While 136/219 = 62.10% of the languages with affricates also have MOD, 68/98 = 69.39% of the languages without affricates also lack MOD. For the complementary dimension, the proportions differ considerably: whereas 136/166 = 81.93% of the languages with MOD have affricates, 83/151 = 54.97% of the languages without MOD nevertheless have affricates. In the dimension 'presence versus absence of MOD', the difference in percentage points is hence considerably larger than in the dimension 'presence versus absence of affricates'. Therefore one might decide that the MOD dimension outweighs the affricate dimension. Cross-linguistically, MOD is thus the better predictor of affrication than the other way around.

Note that our segmental modifications involve a number of properties: both laryngeal, supralaryngeal or oral and nasal properties are lumped together. It might be the case that one of these is an even better predictor of affrication than (some of the) others. To investigate this, we split up the MOD types into laryngeal ones, supralaryngeal ones and nasal ones, and cross-classified them with affrication.

Table 2. The quantitative patterning of affrication and segmental modification, broken down for category; *n* of instances of MOD

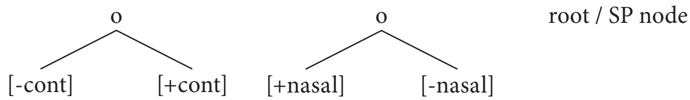
		Affricates?		**total**
		yes	**no**	
	Laryngeal	162	22	184
MOD	Oral	67	12	79
	Nasal	10	9	19
	total	239	43	283

In Table 2, we present the raw figures concerning patterning of affrication and segmental modification, broken down for category. Note that in this table, we deal with numbers of instances of segmental modification, rather than numbers of languages, as in Table 1. Still, both tables refer to the same data.[4]

This pattern of distribution is highly significant ($\chi^2 = 16{,}75$ df = 2 $p < .001$). Obviously, the relationship between MOD category and the presence or absence of affricates is, that the overwhelming majority if languages with MOD also have affricates, with the exception of nasal modification.

Let us look at prenasalization and affrication in some more detail. Unlike other types of segmental complexity, such as aspiration and labialization, which we referred to as "compounding" types of complexity, both prenasalization and affrication were described in terms of "contouring" of a binary distinctive feature in Sagey (1986).

(12) *Affricates and prenasalized stops* (Sagey 1986)



We must note that, just like in the case of affricates, some doubt the status of prenasalized stops as single segments. Downing (to appear), for instance, argues on the basis of a variety of evidence that prenasalized stops in Bantu are more properly described as coda-onset *sequences* of segments, although she explicitly remains uncommitted on the status of prenasalized stops in other language families (see also Herbert 1986).

If the representations in (12) are correct — contrary to Downing's proposal —, i.e. if prenasalized stops are single segments and if prenasalization and affrication are indeed of the same nature, we might expect to find that languages would tend to have both, or ban both. However, as was shown in Table 2 above, of the 19 languages that have prenasalization in our sample (Maddieson 1984), 10 also have affricates, 9 do not.[5] To the extent that this finding can be generalized, we conclude that there is no reason to believe that the incidence of affrication and prenasalization is in any way correlated.

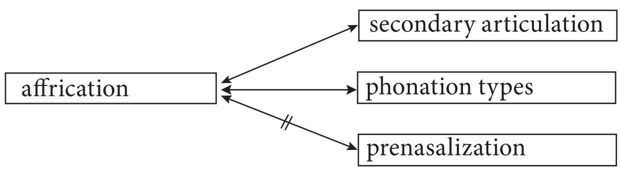
To conclude this section, MOD and affrication are mutually dependent for laryngeal and oral modification, but they are independent for nasal modification (i.e. prenasalization). For MOD of the laryngeal and oral categories, our hypothesis is confirmed, while for MOD of the nasal category it is disconfirmed.

4. Discussion and conclusion

In this study we have seen that affricates on the one hand and oral and laryngeal modification on the other are positively correlated. In our analyses, we have treated segmental modification and affrication as if they were mutually unrelated. This is, of course, an artifact of our strictly synchronic view. It is a well-known fact, that the affricates of modern German, for instance, have developed out of aspirated stops, i.e. out of modified simplex segments. Similarly, palatalization on coronals may well give rise to affricates and have done so in many languages (e.g. in Muinane, as discussed in Hinskens & van de Weijer 2003: 1074). It may not always be simple to keep the diachrony and synchrony apart in cases like these. Ladefoged & Maddieson (1995: 368) expressed the role MOD can play in diachrony as: "Today's secondary articulations may be the primary articulations of the future". This results from the process which Clements (1991) calls 'feature promotion'.

Taking this caveat into consideration, we have looked at four different types of segmental complexity involved in affricates, prenasalized stops, secondary articulation and laryngeal modification, i.e. phonation types. *All* these types of complexity have come under attack in recent years: affricates are "really" stops (Hinskens & van de Weijer 2003a, among much other work), prenasalized stops are "really" clusters (Downing to appear, among other work), and secondary articulation and laryngeal complexity "really" involve prosodic rather than segmental properties (Kehrein 2002). Note that these re-analyses of these types of complexity are directed in different ways: elimination of affrication as a phonological property, clustering and prosodification. It would therefore not be expected that these types of complexity are correlated in any way in segmental inventories. In this paper, we hope to have shown that — although a number of issues should clearly still be resolved — there *are* interesting relations between these types of complexity. These relations can be illustrated by way of the diagram in (13):

(13)



To capture the correspondences expressed in (13), we propose a simple OT constraint that militates against segmental branching in general (cf. also van de Weijer & Hinskens 2003):

(14) *Complex(segment)
No segment-internal branching

For languages that permit both affricates and the various types of segmental modifications, the constraint in (14) is low-ranked in the grammar. For languages that rule out both, the constraint is high-ranked. Languages that permit affricates and rule out segmental modifications, or vice versa, are defined as the focus of investigation: they will have to show whether the constraint in (14) can be maintained in its pure form, or whether, ultimately, it must be reformulated as a family of constraints.

In future work, we will look more closely at the relations between supralaryngeal MOD types and laryngeal MOD types, and between each of these and prenasalization, i.e. the relations between the types of complexity on the right-hand side in (13). It might also be possible to investigate other types of complexity, e.g. that in doubly-articulated stops or in diphthongs.

Notes

* This paper was presented at the TINdag 2003 in Utrecht. We thank the audience, and in particular Norval Smith, for questions and suggestions. We also thank Laura Downing, Wolfgang Kehrein, Marc van Oostendorp and an anonymous reviewer for comments on an earlier version of this paper. The usual disclaimers apply.

1. Note that Kehrein (2002: 65) maintains that his Laryngeal node "consists of" the three laryngeal features, [spread], [voice] and [constricted glottis], although in (7) these features (in this case [spread glottis]) are attached to this node.

2. In this proposal, we might investigate whether voiced stops are also represented as "complex", especially in a theory in which voiceless stops are unmarked and voiced stops have a feature [voice].

3. The chi square test compares the actual mutual distribution of the values of two variables (*in casu* the presence or absence of affrication and the presence or absence of MOD) to chance, i.e. to the independence of the two variables, which is the zero hypothesis. In the zero hypothesis scenario, the values in the cells in the table perfectly reflect the proportion of the row and column totals to the grand total. As the discrepancy between the actual distribution and chance increases, the effect of chance (p) decreases and the dependence between the two variables becomes more probable.

4. The total number of instances of MOD investigated in Hinskens & van de Weijer (2003a) was 281, while in Hinskens & van de Weijer (2003b) it was 280. The different numbers are due to a) an original inaccuracy and b) the fact that in three languages MOD occurs merely on affricates and affricates (which were the object of a separate hypothesis) were excluded from the analyses reported in those contributions.

5. The relatively small number of languages with prenasalization does not necessarily lead us to expect them to show an essentially different pattern of distribution with respect to the co-occurrence with affricates than the oral and laryngeal MOD types. Moreover, the chi square test calibrates for large differences in absolute numbers.

References

Clements, G. N. (1991) Place of articulation in consonants and vowels: a unified theory. *Working Papers of the Cornell Phonetics Laboratory* 5, 77–123.

Clements, G. N. (2003) Feature economy as an organizational principle of sound systems. Paper presented at the First Old World Conference in Phonology, Leiden, January 2003. Ms, University of Paris 3, CNRS.

Clements, G. N. & E. V. Hume (1995) The internal organization of speech sounds. In J. A. Goldsmith (ed.) *The Handbook of Phonological Theory*, 245–306. Basil Blackwell, Oxford.

Downing, L. (to appear) On the ambiguous segmental status of nasals in homorganic NC sequences. In M. van Oostendorp & J. M. van de Weijer (eds.) *The Internal Organization of Phonological Segments*. Mouton de Gruyter, Berlin.

Herbert, R. K. (1986) *Language Universals, Markedness Theory, and Natural Phonetic Processes*. Trends in Linguistics, 25. Mouton de Gruyter, Berlin (revised Doctoral dissertation, Ohio State University, 1977).

Hinskens, F. & J. M. van de Weijer (2003a) Patterns of segmental modification in consonant inventories: A cross-linguistic study. *Linguistics* 41, 1041–1084.

Hinskens, F. & J.M. van de Weijer (2003b) Patterns of segmental modification in consonant inventories: Contrastive vs. redundant systems and phonology vs. phonetics. In P. Fikkert & L. Cornips (eds.) *Linguistics in the Netherlands 2003*, 71–81. John Benjamins, Amsterdam and Philadelphia.

Kehrein, W. (2002) *Phonological Representation and Phonetic Phasing: Affricates and Laryngeals.* Niemeyer, Tübingen.

Lin, Y.-H. (to appear) Piro affricates: Phonological edge effects and phonetic anti-edge effects? In M. van Oostendorp & J.M. van de Weijer (eds.) *The Internal Organization of Phonological Segments.* Mouton de Gruyter, Berlin.

Maddieson, I. (1984) *Patterns of Sounds.* Cambridge University Press, Cambridge.

Martinet, A. (1955) *Économie des changements phonétiques.* Francke, Paris.

Sagey, E.C. (1986) The Representation of Features and Relations in Non-Linear Phonology. Doctoral dissertation, MIT, Cambridge, Massachusetts.

Weijer, J.M. van de (1996) *Segmental Structure and Complex Segments.* Niemeyer, Tübingen (revised Doctoral dissertation, Leiden University, 1994).

Weijer, J.M. van de & F. Hinskens (2003) Patterns of segmental modification in consonant inventories: Contrastive vs. redundant systems. In B. Palek, O. Fujimora & S. Haraguchi (eds.) *Proceedings of LP2002* (Web version). Printed version to be published by Charles University Press, Prague and Meikai University Press, Uruyasu.