Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

ON VOWEL IDENTIFICATION AND PHONOLOGICAL THEORY

Justus C. Roux

This paper[1] focuses on some recent calls for a more substance-based approach to phonetic and phonological descriptions, and presents some concrete examples from Sotho[2] to demonstrate some of the inadequacies found in traditional approaches.

1 Introduction

In the preface of a recent edition of one of the most well-known phonetic journals, viz. Phonetica (1981 : 5 ff.), the editor prof. Klaus Kohler repeats a previous plea for "... a revival of phonetics as speech science, in which scientific method is subordinated to a language perspective rather than to a professional linguistic one, and in which experimental investigations of specific phenomena in individual languages are integrated into a general theoretical discussion, thus shifting the general aspects of human speech back into the focus of our attention." He refers to a theme of great theoretical import in modern phonetics, viz. the temporal aspects of speech production and perception, and states that while the classical segmental approach of phonology and linguistics have had very little concern for the time dimension of speech, a new approach in terms of language categories integrating the time dimension of speech has become a necessity. "And this approach will inevitably result in a revision of phonological theory and thus have important repercussions on present-day linguistic thinking" (1981 : 5).

In an equally recent publication with the title "Phonology in the 1980's", Griffen (1981 : 617 ff.) illustrates a model of phonological description void of the traditional 'segmental' concept. He argues that the segment

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

in phonology is a convenient fiction that has been used by linguists only because of the lack of a more adequate way of representing continuous speech sounds (cf. Goyvaerts 1981 : 20). He bases his model on an existing experimental phonetic model capable of describing speech on a nonsegmental basis. He illustrates how such a model can account for aspects of the Welsh consonant system in a way in which segmental models cannot.

These views, from different domains of linguistic practice, can most probably be seen as a reaction against the atomistic reductionism that has dominated linguistics up to now. It can be seen not only as a call for greater empirical input to phonological theory (cf. the by now well-known views of Ohala 1974, 1979 in this respect), but in essence as a call for a return to 'substance' in phonetic and phonological descriptions. The substance approach is in fact a functional approach as it looks upon the acoustic speech signal as "... the central, publicly observable vehicle of linguistic communication, which is au fond to be regarded as a human artifact, manufactured ad hoc to serve a definite communicative end" (Öhman 1979 : xxi). It furthermore aims at an unprejudiced study of the acoustic speech signal, unmediated by any preconceived theory of segmentation or whatever, in the hope that it "... could give us a rather new and different understanding of language" (Öhman 1979 : xx).

Classical phonetics and phonology on the other hand, have always looked upon the speech utterance as a linear series of discrete articulatory postures. One of the problems inherent to this approach is that it tends to ignore dynamic information (e.g. prosodic and intonational features which are very important in real speech), because this type of information is not compatible with the principles of discreteness,

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

linearity, and articulatory posturality" (cf. Öhman 1979 : xix). Classical phonetics and phonology furthermore lean heavily on the auditory judgements of either the "trained phonetician" or the "ideal speaker-listener" in the collection of primary data.

Referring to aspects of vowel identification in Sotho it will now be demonstrated

(i) how dynamic prosodic information may cause perceptual confusion of vowel qualities;

(ii) how unaided auditory phonetic judgements on segmental and suprasegmental aspects of vowels may lead to totally different results.

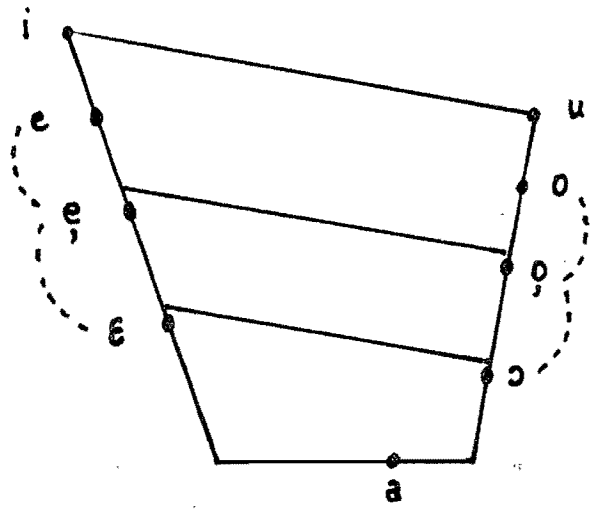
2 The vowel segments of Sotho

Even the most cursory glance at existing views on the vowel segments of Sotho will show that there is a great amount of disagreement among researchers, not only as to how many, but also as to which vowel segments are to be identified in these languages.[3] A specific problem seems to have developed around the raised mid high vowel as posited by proponents of the official view. However, to come to grips with the problem it is necessary to have a quick look at the other relevant views.

The first comprehensive and scientific[4] treatment of the vowel segments of Sotho can be found in the by now historical work of A.N. Tucker (1929). In this work he distinguishes between "seven main vowel sounds" (1929 : 23), viz. i, e, ɛ , a, ɔ , o, u and two "open varieties", viz. e̞ and o̞. These sounds may be represented as follows on the Cardinal Vowel Chart:

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

(1)



It is important to note that Tucker distinguishes seven phonemic forms and two allophones (e̞ and o̞). These allophones are regarded as positional variants of e/o and ɛ/ɔ respectively, i.e. as "lowered/opened" variants of e/o and "raised/closed" variants of ɛ/ɔ .

Doke and Mofokeng (1967 : 1) take the same position as Tucker, i.e. they acknowledge nine different phonetic qualities, but they are not so sure as to the phonemic status of these forms and state: "When applying narrow phonetic principles , ... all [may] be considered as constituting separate phonemes," (1967 : 5).

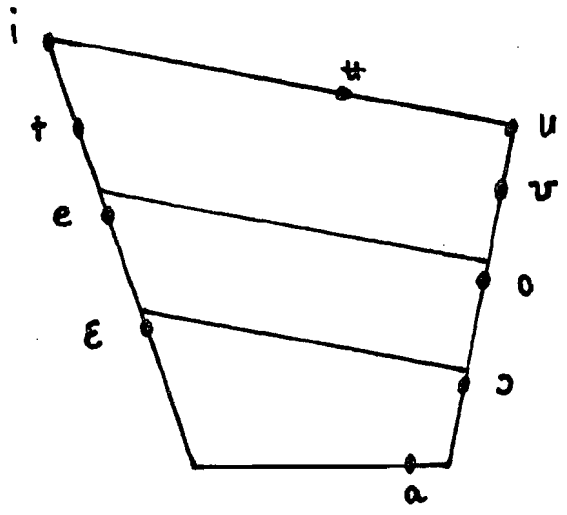
This view of Tucker (1929) was criticized severely by Cole (1955 : xxvii) when he remarked: "Though this is a most useful piece of work, especially as regards the treatment of consonants, Tucker fumbled badly over the Sotho-Tswana vowels, and thus helped to perpetuate the erroneous phonemic grouping which persists in the present (1937) Tswana orthography".

The most interesting view, however, dates back to 1916 when none other than the famous Daniel Jones in collaboration with S.T. Plaatjie claimed

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

that Tswana had ten vowel qualities. He identified four front vowels, five back vowels and one low vowel (cf. Jones and Plaatjie 1916 : xiv, xxiii - xxv):

(2)



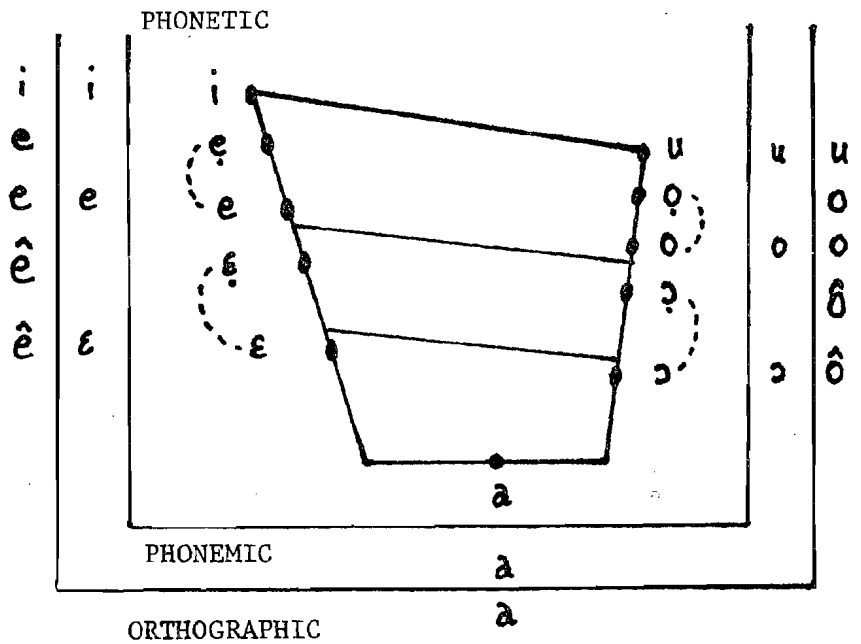
In his analysis Jones states that "... the vowel sound ɨ [is] probably 'non distinctive' in respect to ... u" (1916 : xiv). He furthermore makes no distinction between the other nine vowel segments in terms of their phonemic and/or allophonic status.

Two points are to be noted at this stage. Firstly, it is seen that in both instances only one vowel segment is identified between Cardinal Vowels 1 - 2, and 7 - 8 respectively. Secondly, a maximum of ten phonetic segments are identified in the "phonetically orientated" descriptions.

The most generally accepted view seems to be the following one claiming that Sotho has eleven phonetic vowel segments which may be reduced to seven phonemic forms (cf. Ziervogel 1967 : 117-9; 125-7; 132-4; Cole 1955 : 5):

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

(3)



This "phonemically orientated" description, owing much to Cole (1949 : 124), makes provision for two extra vowel segments in the positions just below Cardinal vowels 1 and 8 respectively. The question, however, is whether these two forms [e̝] and [o̝] do occur phonetically or whether they are artifacts of a phonological argument of the following kind: [ɛ] and [e] stand in phonemic opposition to each other; there are exact rules governing the occurrence of raised [ɛ̝] , i.e. [e̝] (cf. Cole 1955 : 15); raised [e̝] occurs phonetically (the 'older' phonetic observations confirm this); if vowel raising holds true for one member of the phonemic opposition, then, surely, it must hold true for the other member as well; thus there should/must be a phonetic segment [e̝] . (The same argument, naturally, may apply in the case of [o̝] .)

Whatever the origin of this raised mid high vowel, Swanepoel and Lenake (1979) consider it to be phonetically distinct. They present examples of all of the vowel forms in (3) on magnetic tape as part of a language course in Sesotho. The impression gained from the tape is that the forms [e] and [e̝] are in fact different and hence must be regarded as different vowels. A subsequent acoustic (spectrographic) analysis of

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

these segments, however, reveals the following interesting facts:[5)]

1. The spectral qualities of these two segments differ minimally:

	[e]	[ẹ]
Formant 1	240 Hz.	240 Hz.
Formant 2	2025 Hz.	2000 Hz.
Formant 3	3200 Hz.	3200 Hz.

It is common knowledge that the quality of a vowel segment is reflected in the values of its first and second formants (F_1 and F_2) respectively. The results of the analysis only show a difference of approximately 25 Hz. in F_2, that is a difference of 1,23% for this formant. Psycho-acoustic experiments have shown that the difference limen (DL) for a formant frequency is in the order of $\pm$ 3% of that formant frequency (cf. Flanagan 1957 : 533). This suggests that, as far as the perception of different vowel qualities is concerned, no difference should be detected. The claim that these two vowels are different in quality and, hence, entitled to different positions on a vowel chart, thus seems to be false.

2. The respective pitch levels (F_o) of the two segments differ considerably:

	[e]	[ẹ]
F_o	95-77 Hz.	77-90 Hz.

3. The respective amplitude (stress) patterns of the two segments are different:

[e] [ẹ]

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

During the articulation of [e] relatively more stress was exerted on the initial part of the articulation than on the last and, concomitantly, the pitch level decreased from 95 to 77 Hz. (The concomitancy of the two phenomena is coincidental). On the other hand, the opposite seems to have happened during the articulation of [e̝] : more stress was exerted towards the final part of the articulation, whereas the tone level increased at the end of the articulation. One can only speculate as to the psychological effect the latter articulation may have on a listener : the sound gradually "climbs" to a "higher point", it sounds "higher" and thus may be regarded as a "higher" vowel. Hence it probably has a higher tongue position than [e] . If it is taken into account that the just discriminable changes in F_0 are considered to be in the order of 0,3 to 0,5 Hz. (cf. Lehiste 1970 : 64), then the perceptual difference between these two segments is quite real.

The point to be made, then, is that although these two sounds may sound different, they are not different in their respective (static) qualities, but rather in their temporal realization of prosodic features. It is exactly this type of dynamic prosodic information that is lost in the decoding of an acoustic signal into linear series of static discrete segments. A model of description integrating the time dimension of speech would most probably be able to account for this phenomenon.

3 Unaided auditory judgements

In the introduction it was mentioned that classical phonetics and phonology find it fit to rely on the sensory abilities of the "trained phonetician" or "ideal speaker-listener" in the gathering of primary data. In the previous section it was also indicated that perceptual confusions may occur in the judgement of vowel quality (which is determined by the

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

total configuration of the vocal tract) on the one hand, and prosodic features of vowels (which are determined, inter alia, by the tempo of vocal fold vibrations, and by "physical effort", cf. Lehiste 1970 : 106) on the other hand. The question, however, is: how general is this perceptual confusion of quality and tone in, for instance, Sesotho?

Two sets of examples may be cited from Sesotho indicating that researchers have not been very successful or consistent in detecting the differences in their decoding of the acoustic speech signal.

Firstly, consider the problem of so-called unexplainable exceptions to the vowel raising rule in Sotho. The vowel raising rule in Sotho states, roughly, the following:

(4) A mid vowel is raised to a higher position if it is followed by a high vowel in the next syllable.

Hence,

ɛ → ɛ̝ / ___ Ci

e → e̝ / ___ Ci

ɔ → ɔ̝ / ___ Ci

o → o / ___ Ci

This rule is a clear example of an anticipatory assimilation rule. However, many examples are usually presented (cf. Doke and Mofokeng 1967 : 7-8; Cole 1955 : 14-15) of raised forms occurring in environments not conforming to that of the rule. These forms are normally found in demonstrative pronouns, adjectival, relative and quantitative concords, locative adverbial formatives etc. If these forms are marked for their tonal qualities, it appears that approximately 80% of these forms are highly intoned. Hence tsê [ts'ɛ̝] "demonstrative pro-

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

noun class 10 (Sesotho)",may even be [ts'ɛ́]. The question then arises whether this unexplainable raised [ɛ] , i.e. [ɛ̝], is in fact raised in quality or whether it only "sounds" raised due to the high tone it carries.

More explicit examples may be found in dictionaries with tone indications.[6] Compare the following forms from (Du Plessis et al. n.d.) and (Mabille et al. 1974):

(5)	(i) Mabille	-bënya	'glitter, shine'	[bɛ̀ɲà] :	[ɛ̀]
	(ii) Du Plessis	-bênya	'glitter, shine'	[bɛ́ɲà] :	[ɛ́]
	(iii) Mabille	kötsi	'danger, accident'	[k'ɔ̀ ts'ì] :	[ɔ̀]
	(iv) Du Plessis	kôtsi	'danger, accident'	[k'ɔ́ ts'ì] :	[ɔ́]
	(v) Mabille	-lëboha	'to thank'	[lɛ̀bɔ̀hà] :	[ɛ̀]
	(vi) Du Plessis	-lêboha	'to thank'	[lɛ́bɔ̀hà] :	[ɛ́]
	(vii) Mabille	molömo	'mouth'	[mɔ̀lɔ̀mɔ̀] :	[ɔ̀]
	(viii) Du Plessis	molômo	'mouth'	[mɔ̀lɔ́mɔ̀] :	[ɔ́]

These examples show that a specific vowel segment may be identified by one person as a mid low vowel with a high tone and by another as a raised vowel with a low tone. The confusion is self-evident. The fundamental question to be answered here is: at which point does this confusion take place? In other words, in the decoding of the acoustic speech signal, what should the relationship between F_o and F_1, F_2 be for different segments so as to cause these confusions? Put in another way : given two vowel segments, what should the "inner" relationship be between vowel tone (F_o) and vowel quality (F_1, F_2) for each segment, so that these segments may be judged as either perceptually identical, or perceptually distinct? Specific psychoacoustic experimentation could come

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

up with an answer in this respect.[7)]

The point to be made here is that theories relying solely on human sensory acuity in the collection of primary data are bound to run into trouble. The disparity of the data in (5) can only set one thinking as to the consequences it may hold for an adequate phonetic description and for abstract phonological and tonological studies. It must be accepted that there is a point in the process of the decoding of the acoustic speech signal beyond which unaided auditory perception is incapable of taking us. Objective scientific methods must be adopted not only to verify auditory judgements but also to check on the correctness of forms predicted by theories.

4 Conclusion

Results from experimental phonetic tests on vowel perception increasingly show that dynamic information plays an overriding role in determining accurate vowel identification (cf. Strange and Jenkins 1979; Borzone de Manrique 1979, to name but two). Considering this fact as well as the deficiencies immanent in traditional approaches (of which two were discussed in this paper) it seems as if the need to formulate more adequate units of description than those borrowed from phonological theory, i.e. units making explicit reference to the time dimension of speech, is long overdue.

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

NOTES

1. This paper is a revised version of a lecture which was presented to the Study Group of the Department of General Linguistics at the University of Stellenbosch during May 1982. As such, this paper contains preliminary results of a larger research project (cf. Roux, in preparation). The results of phonetic experiments presented here stem from research that was done at the Institut für Kommunikationsforschung und Phonetik in Bonn, West Germany during 1981. Financial assistance from the HSRC and the Alexander von Humboldt Stiftung is hereby greatfully acknowledged.

2. Sotho is a cover term referring to three related indigenous languages found in Southern Africa, viz. Northern Sotho, Tswana and Sesotho.

3. For more specific views in this respect, see Roux 1979, 1980.

4. This work was submitted for a Ph.D. degree at the University of London.

5. No categorial claim is made here as it is obvious that much more data need to be investigated.

6. Dictionaries unfortunately are very often sparsely marked for tonal qualities, especially (Mabille et al. 1974).

7. Cf. Roux in preparation.

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

REFERENCES

Borzonne de Manrique, A.M.

1979 'On the recognition of isolated Spanish vowels' in Current Issues in Phonetic Sciences 9.2, 676 ff.

Cole, D.T.

1949 'Notes on the phonological relationship of Tswana vowels' in African Studies 8.3.

1955 An Introduction to Tswana Grammar. Cape Town: Longmans, Green and Co.

Doke, C.M. and S.M. Mofokeng

1967 Textbook of Southern Sotho Grammar. Cape Town: Longmans.

Du Plessis, J.A., J. Gildenhuys and J. Moilwa

n.d. Bukantswe ya Maleme-pedi : Sesotho-Seafrikanse. Cape Town: Via Afrika.

Flanagan, J.L.

1957 'Estimates of the Maximum Precision Necessary in Quantizing Certain "Dimensions" of Vowel Sounds', in Journal of the Acoustical Society of America, 29.4, 533-4.

Goyvaerts, D.L. (ed.)

1981 Phonology in the 1980's. Ghent, Belgium: E. Story-Scientia.

Griffen, T.D.

1981 'A nonsegmental model for description and analysis' in Goyvaerts (ed.).

Jones, D. and S.T. Plaatjie

1916 A Sechuana Reader. London: The University of London Press Ltd.

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

Kohler, K.J.

1981 Preface, Phonetica 38.1-3, 5-8. Basel, München: S. Karger.

Lehiste, I.

1970 Suprasegmentals. Cambridge, Massachusetts: The M.I.T. Press.

Lenake, J.M. and C.F. Swanepoel

1979 Mathe le leleme 1. Pretoria: University of South Africa.

Lindblom, B.E.F. and S.E.G. Öhman

1979 Frontiers in Speech Communication Research. London: Academic Press.

Mabille, A. and H. Dieterlin

1974 Southern Sotho-English Dictionary. Morija Lesotho: Morija Sesuto Book Depot.

Ohala, J.

1974 'Phonetic Explanation in Phonology', in Papers from the Parasession on Natural Phonology, Bruch et al. (eds.), Chicago Linguistic Society.

1979 'The contribution of acoustic phonetics to phonology' in Lindblom & Öhman (eds.).

Öhman, S.E.G.

1979 'Introduction. Gunnar Fant, Pioneer Student of Human Speech' in Lindblom & Öhman (eds.).

Roux, J.C.

1979 'Eksperimentele tegnieke en die verifikasie van fonetiese data', in Taalfasette 26.2, 1-20.

1980 'The Vowel System of Southern Sotho Revisited', in Proceedings of the Third International African Language Congress, UNISA, 274-288.

in preparation Vokaalkwaliteit, Toon en Konsonantale Invloede in Sotho.

Stellenbosch Papers in Linguistics, Vol. 8, 1982, 79-93 doi: 10.5774/8-0-114

Strange, W. and J.J. Jenkins

1979 'Three Sources of Information in Vowel Identification', Proceedings of the Ninth International Congress of Phonetic Sciences, Copenhagen, 251.

Tucker, A.N.

1929 The Comparative Phonetics of the Suto-Chuana group of Bantu Languages. London: Longmans Green & Co.

Ziervogel, D. (ed.)

1967 Handboek vir die Spraakklanke en Klankveranderinge in die Bantoetale van Suid-Afrika. Pretoria: University of South Africa.