DOCUMENT RESUME

ED 350 841 FL 020 624

AUTHOR Ogden, Richard
TITLE Parametric Interpretation in Yorktalk.
PUB DATE Jun 92
NOTE 21p.; In: Harlow, S. J. and Warner, A. R., Eds. York Papers in Linguistics, 16; see FL 020 620.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Computer Software; *Language Patterns; *Language Research; Linguistic Theory; *Phonology; Research Methodology; *Speech Synthesizers
IDENTIFIERS *Parametric Analysis; *YorkTalk

ABSTRACT

The method of parametric interpretation used in the computer program "Yorktalk," software that creates synthetic parameter files from phonological representations of speech, is explained. First, the design of the program is described, and the concept of "exponency" in prosodic analysis is explained as it is applied in the program. Then a hypothetical example of parametric analysis of a syllable is presented. Interpretation of nuclei, codas, and onsets are detailed, and diagrams are provided. The software's treatment of laterals and vowel allophony is given special attention. A 14-item bibliography is included. (MSE)

Parametric Interpretation in Yorktalk*

Richard Ogden

University of York

PARAMETRIC INTERPRETATION IN YORKTALK*

Richard Ogden

University of York

1. Introduction

In this paper we aim to show how parametric interpretation is carried out in the YorkTalk speech generation system (Coleman 1990, Coleman & Local 1987, Local 1989). We do not aim to give a complete description of the system, and in particular we have left out much which may be of interest concerning the temporal interpretation of syllabies. Also we have not made any attempt to relate YorkTalk to its theoretical background, with the exception of a cursory mention of J R Firth's paper 'Sounds and Prosodies'. We welcome comments on our work, and are glad to demonstrate the system to those interested.

2. A quick overview

YorkTalk is a computer program which creates synthesis parameter files from phonological representations which are structured directed acyclic graphs with features distributed over them, as in the diagram below:

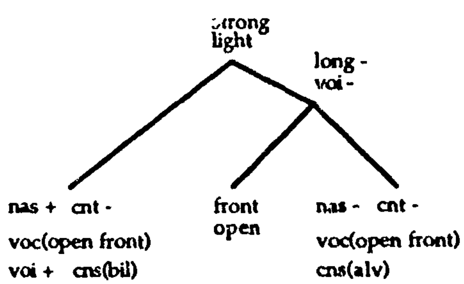

Fig 1: Partial phonological representation for "mat"

* This work is sponsored by British Telecom. Without John Local, John Coleman, Adrian Simpson and John Kelly, the system described here would not exist. Thanks to John Local for comments on earlier versions of this paper.

York Papers in Linguistics 16 (1992) 81-99

The graphs are produced by a parser of English words with a grammar of English syllable structure, metrical and lexical structure. These graphs need to be interpreted in order for them to be 'made audible'. Their interpretation has to stated explicitly. Temporal relations between the constituents of the graphs are worked out in a part of the program that does not concern us here, called 't_interpret'. The resulting structures have features and timings associated with them. In p_interpret (the function which assigns a parametric interpretation to the phonological representations) the relevant parameters, which are all Klatt formant synthesiser parameters, are assigned to these structures. Note that in this context, Start and End, which appear in the diagram below and throughout this description of YorkTalk, are reference points. They do not imply that any given parameter which expones a particular phonological category starts or ends at the time values to which Start and End are instantiated.

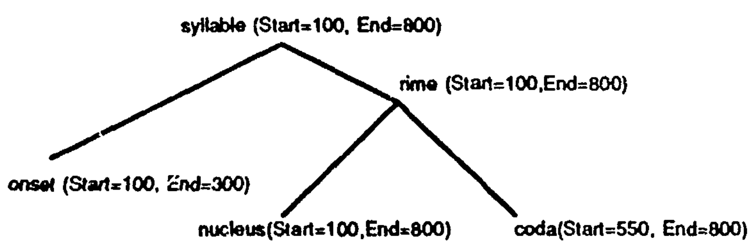

Fig 2: A temporally interpreted graph which can be p_interpreted (features have not been marked, only timings).

p_interpret goes about its work head-first. So the order of interpretation is: nucleus, coda, rime, onset, syllable.

The objects that can be interpreted are any feature or bundle of features at any node in the graph being interpreted. This means that where generalisations can be made the exponency statements can be made to match with just that bundle of features at just that node, and ignore any other feature information. For instance all the relations which can be represented by the partial description (cnt (-) nas (+) _

4

_)[1] might need to have the same exponency statements for some parameters. On the other hand, the parameter statements can be very specific, eg they might relate only to the structure `(cnt(-) nas(-) str(+) voi(+))`.

Parametric interpretation is, just like temporal interpretation, arbitrary (in Saussure's sense) but systematic, compositional and consistent. To put this in less abstract terms: any given bundle of features at a given place in the structure can (indeed *must*) have only one possible interpretation.

Because the synthetic parameters constitute a compositional interpretation of fragments of structure, parameter values cannot be altered (although they *can* be overlaid by something else). An example might be 'stops' in English. Their exponents include plosion and aspiration when they occur in simple onsets in English, and plosion but no aspiration when they occur in onset clusters with friction. Declarative interpretation, which serves as a constraint on YorkTalk, does not allow us to generate first a burst with aspiration and then remove the aspiration in order to achieve the unaspirated stop. The plosion must be generated without aspiration in the first place. This is not a problem since it is *nodes* and *structures* which are interpreted, in other words the *onset cluster* is interpreted rather than an onset 'stop', ie a terminal node containing a particular kind of featural information.

Before working through an example of p_interpretation, we will explain the construct of 'exponency' and consider how it is implemented in YorkTalk.

3. 'Exponency'

Phonological structures and features are associated with phonetic 'exponents', the term used by the prosodic analysts (Firth 1937, among others) for the 'real-world' manifestations of the interpretation of phonological structures. The units of phonology cannot be pronounced - they are abstract and describe structural relations within the language;

[1] YorkTalk is written in Prolog. Initial capital letters stand for variables, initial small letters stand for constants. Underscores (_) stand for unnamed variables. The partial phonological representation given here constitutes an example of a formal instantiation of underspecification. (See eg Gazdar & Mellish 1989)

but their presence is manifested by phonetic 'exponents'.[2] 'Exponency statements' in YorkTalk link the abstract 'silent' phonology with the noisy 'real-world' phonetic material which is speech. Exponency statements make the phonological description audible. They have to be stated explicitly though because phonological features have no inherent interpretation. For example, a feature [nas] need not refer to the position of the velum (which is the case in, eg, autosegmental phonology), although it could; it could also refer to much more. In other words, the relation of phonological feature to phonetic exponent is not one-to-one. One feature may have more than one phonetic exponent.

p_interpret calls all the 'exponency' statements for all the parameters in the Klatt synthesiser; there is an exponency statement relating to each parameter we use. There are statements called 'av_exponency', 'f1_exponency', and so on.

The Form of Exponency Statements

Exponency statements are of the following form: they have a list of features with which they will match, and a list of ordered pairs of the form <Time, Value>. The Start, End and Duration of the constituent are passed to the exponency statements.

The Time Field of Exponency Statements

The value of Time is worked out with reference to Start, End and Duration. Nuclei are 'relatively timed', which means that all times are defined in relation to the syllable Duration, so that any particular acoustic event is timed to occur at relatively the same place in relation to the whole syllable. Below is an example of a possible timing statement:

```
(Start, Val el)
(Start+(x% cf Duration),Value2)
(Start+(y% of Duration),Value3)
(End,Value4)
```

[2]Phonetic exponents can also include the systematic *absence* of a particular feature as well as its presence. (cf Rob[illegible]s 1957:90)

Of course, different values for Start, End and Duration will result in different time values, but the internal temporal structure (the 'timing') of an interpreted constituent with this exponency will always be the same. This allows us to have one exponency statement which will apply in many circumstances, regardless of what the values of the Start and End are, and it also means that it is unnecessary to have different statements for temporally compressed syllables. The claim we are making is that the internal temporal structure of acoustic phonetic parameters is consistent, regardless of actual duration at 'run-time'.

The Value Field of Exponency Statements

The Value field is the second member of the <Time, Value> ordered pair. Values might be 'hard numbers' eg values obtained from instrumental observation of natural speech, or they might be calculated in relation to other values; another important source is refinement of the synthesis through impressionistic listening. In the nucleus, all the values are 'hard numbers'. In his papers on speech synthesis, Dennis Klatt (eg Klatt 1987) usually gives what we have called 'hard numbers'.

Parametric exponents are looked up from a database of exponents on the basis of the phonological representation. In the nuclei, for instance, all the things whose second part is `grv(_), height(close), rnd(_)` statement for f1_exponency for the second part of the nucleus. (Such generalisation is easily achieved with unification[3]). In this way, the parametric interpretation of nuclei is compositional: /iy, ey, oy, ay, uw, iw, ow, aw/[4] all have something in common, which is that phonologically they are all part of a class of V units known as 'closing diphthongs' and are all described as `height(close)` in their second part and have the head feature[5]

[3]YorkTalk is written in Prolog, which makes extensive use of unification. (See eg Shieber 1986)

[4]The phonemic representation is used here only for convenience. It should be clear by now that such phonemic representations are not used anywhere in the system

[5]Our structured representations make extensive use of heads and head features. Heads are given a special status in parametric interpretation, in that heads are always interpreted first.

`long(+)`; what is more, they share part of their phonetic interpretation, which is what makes the interpretation compositional.

Compositionality serves as a strong constraint on interpretation. Since we want to make as many generalisations as possible (and necessary) in our statements of phonetic exponency, we do not want to proliferate statements whenever we can avoid it. As indicated earlier, the interpretation is also arbitrary; any value whatsoever (within the limits of the Klatt synthesiser) could be put in for the f1_exponency in the case above. But the results must sound like good English, and so compositionality and our ears function as strong constraints on the values and the timings we allow. Where generalisations do not result in natural-sounding synthesis, we prefer to have many more exponency statements whose applicability is more confined.

4. A Hypothetical Worked Example

To bring the above sections together, we will provide an imaginary example. We will work through the interpretation of an imaginary syllable (whose identity is irrelevant to our purpose) and show how the parametric interpretation is assigned to the sort of structure drawn in Figs 1 & 2.

The first step in p_interpret(Syllable) is to p_interpret the *head* of the syllable, which is the rime, and its head is the nucleus. So the first thing to be p_interpreted is the nucleus.

Parametric Interpretation of Nuclei

Imagine we wish to interpret a structure with the following featural description:

```
((grv(a), height(x), rnd(a)),
 (grv(b), height(y), rnd(b))).
```

The f2_exponency statement might be:

```
f2_exponency (nucleus ((_, height(x), _), (grv(b),
                height(y),          _)),
  (Start,1500)
  (Start+(x% of Duration),1500)
  (Start+(y% of Duration),1000)
  (End,1000), (Start=400,End=900,Duration=End-Start) ))
```

Below (Fig 3) is a diagram of the resulting formant shape.

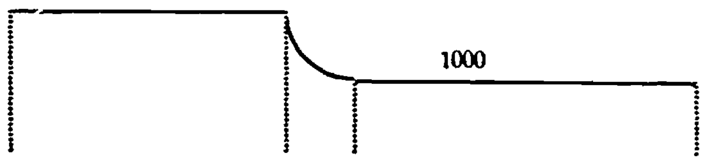

Fig 3: The formant shape resulting from interpreting an 'exponency statement'

An interpolation function is used to join the times and values; where two time points have the same value, a straight line is drawn. Where two adjacent time points (as expressed in the exponency statement) have different values, a smoothed curve interpolation is used to join them. This can be clearly seen in Fig 3. Parametric interpretation goes on in an identical way for all the parameters implicated, so the example of f2 can easily be generalised to other parameters.

Times and values are of equal importance in phonetic description, although traditional phonetics handles time badly. In YorkTalk, the internal temporal structure is crucial to interpretation. Recall that the method we are describing here is compositional and declarative; so not even times can be altered once they are instantiated. The traditional segmental approach to synthesis is to treat values as primary and sit timing on top; in YorkTalk timing and values have to work together simultaneously.

Parametric Interpretation of Codas

Codas are interpreted in much the same way as nuclei. In other words, all the preceding descriptions of interpolation, the structure of the exponency statements and so on apply to codas as well as nuclei. The main point we shall illustrate here is the implementation of *overlaying*.

Recall from Fig 2 that the coda End is that same as the syllable End, and that the temporal domain of the coda falls within that of the syllable, the rime and the nucleus. The result of this was that the coda exponents would be *overlaid* on the nucleus, ie *coproduced* with the nucleus.

Another principle of the Time field of the exponency statements needs to be described here, which is the use of named variables to pick up values. Let us continue with the syllable whose nucleus we have just interpreted. The timing of the syllable is (Start=400, End=900); the timing of the coda might be (Start=750, End=900). Let us say the coda contains the representation for a dark liquid. We will not bother with features here, but use the name Liquid for convenience. The coda exponency statement for f2 might look this:

```
f2_exponency(coda(Liquid)
   (Start-(a% of Dur), Value1)
   (Start+(a% of Dur), Value2)
   (Start+(b% of Dur), Value3)
   (Start+(c% of Dur), Value4)
   (End, Value4), (Start=750, End=900)),
   Value2 is Locus + (Const1 * (Value1 - Locus))
   Value3 is Locus + (Const2 * (Value1 - Locus))
   Value4 is Locus + (Const3 * (Value1 - Locus)).
```

Note that it is possible (and desirable) not to confine the temporal structure of the coda to within the limits of Start and End, although we do not allow parameters to extend beyond End in nuclei and codas or occur before Start in nuclei and onsets. Note also that there are no 'hard numbers' for the Values of the liquid; they are all relative and all depend on the value which is initially picked up. (We call this the 'pick-up' value; 'pick-up' can refer to a time or a value).

The equations we have presented above are a form of Klatt's (1980) modified locus equation. For reasons of contractual confidentiality we are not allowed to publish the values of Locus and Const. We can say however that it is possible to model the formant values and transitions for English laterals extremely well using this method.

The results of this exponency statement are shown in dotted lines on the diagram below.

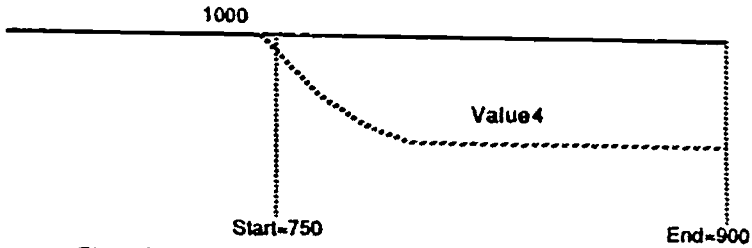

Fig 4: 'Overlaying' coda parameters on to nucleus parameters

In what sense is this 'overlaying'? The values of the exponents of the overlaid coda are all determined in relation to the value of the parameter on to which they have been overlaid (which was calculated when the nucleus was p_interpreted); and the timing does not put the new parameter values *next to* the ones from the nucleus, but on top of them. This is significantly different from the conventional method of synthesis by rule, which adjoins segments in linear sequence and smooths over the join. The YorkTalk method is to build up a database of equations containing <Time, Value> pairs calculated from the exponency statements. Once the whole word has been p_interpreted, the equations are consulted and a synthesis file is generated. In no sense then is anything deleted, because all the exponency information is always present in the database.

The above form of interpretation, where *all* the Values for exponents of a constituent are worked out from Values for another constituent's exponents, is not very common in YorkTalk. More usual is that *some* values are worked out this way, usually the ones nearest the 'pick-up' value, while other values are 'hard numbers'. This is the case for stops, for instance, where the parameter values nearest the 'pick-up' depend on the value picked up, but the values of the parameters at the point of eg the burst are 'hard' numbers. Even this is not as rigid as it sounds; each coda and onset constituent has a voc field which determines the resonance of that constituent. The voc field is inherited from the nucleus, so that the overlaying is treated as phonological with a phonetic interpretation. (Feature structures which are identical in every respect apart from the voc field are as logically distinct from each other as, say, fricatives from nasals; their feature structures do not necessarily match, and the program only considers something as unifiable or not.)

Parametric Interpretation of Onsets

The parametric interpretation of onsets is carried out in just the same way as that of codas, except of course that the exponents are overlaid from the Start of the syllable rather than the End, and the pick-up Times and Values occur latest in the temporal structure of the exponency statement, while in codas they occur earliest. Strictly speaking, the onset is overlaid on to the *rime*, therefore the onset exponents are overlaid on to the exponents of the rime. This has important consequences for vowel quality, as the schematic diagrams below illustrate. Note that while part of the exponency of a coda is to 'know' how to get *into* the coda, and not out of it, part of the exponency of an onset is to 'know' how to get *out of* the onset, but not into it (at least, not in any sophisticated way).[6]

Schematic diagrams showing the parametric interpretation of a syllable

Below are diagrams showing in stages how parametric interpretation for two parameters might progress. The parameters aren't named because all parameters are instantiated in the same way. The parameters in the diagrams can be taken as anonymous typical representatives.

In Fig 5a only the nucleus has been interpreted. In 5b, the coda exponents are overlaid on to the nucleus exponents. In 5c, the onset exponents have been overlaid on to the rime exponents. The dotted lines in Figs 5a-c represent parameters that are overlaid. 5a is only possible when only the nucleus is present with empty onset and coda; 5b only when nucleus and coda (ie rime), but empty onset; and 5c only with a syllable with an onset and rime. Understanding interpretation as happening in stages is not quite right; theoretically it happens all at once.

---

[6]This is somewhat overstating the case. Of course, a coda that is utterance-final has to join into silence, and this is as much a part of the exponency of a coda as are the sophisticated transitions which lead into the coda. The point is that onsets necessarily look 'rightwards' while codas necessarily look 'leftwards'. In a segmental synthesiser the joins to the right and left would be equally important regardless of phonological structure.

Note that by the end of the interpretation, there is less steady state present than there was when only the nucleus had been interpreted; and also that by overlaying the coda and onset exponents, the vocalic quality of the syllable is not changed because the coda and onset exponents are calculated with reference to the nucleus, either because their Values use the parameter values directly, or because the 'voc' field in their phonological description ensures the correct values for the interpreted structure. On the other hand, the vocalic quality is not identical in the last diagram to that of the first diagram; it is by appropriate overlaying of parameters that we achieve small variations in vowel quality such as between eg 'tap' and 'tack'; or larger differences such as 'fees' and 'feel'.

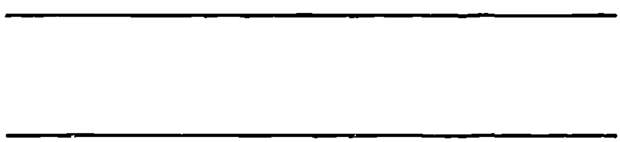

Fig 5a: Nucleus exponents

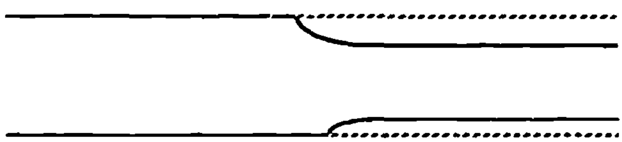

Fig 5b: Coda exponents overlaid on nucleus exponents

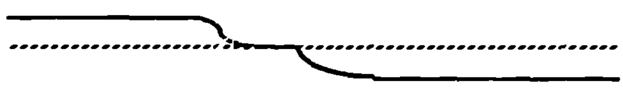

Fig 5c: Onset exponents overlaid on rime exponents

5. Parametric Interpretation of Syllable Overlay

In polysyllabic words, syllables are interpreted as being overlaid on each other. There are two kinds of syllable join; ambisyllabic and non-ambisyllabic[7]. p_interpreting *any* syllable join consists of just p_interpreting the individual syllables to be overlaid, since the real work of 'overlaying' is handled in t_interpret: the Start of the second syllable is the same as the Start of the Coda of the first syllable plus a degree of Overlap. So the parametric join is in the way that onset exponents are made to pick up from the coda exponents.

There is an essential difference between onset and coda exponency; remember that when one syllable is overlaid on another, the Start of the second syllable is the Start of the coda of the first syllable, plus a degree of overlap. In other words, the transitions out of the coda are not so important as the transitions in, whereas the onset exponents of the syllable being interpreted have to pick up from the coda of the syllable being overlaid and provide a suitable join. In other words, we have to define the transitions *into* the onset as well as the ones out *of* it; whereas for a coda, we just have to state the transitions in.

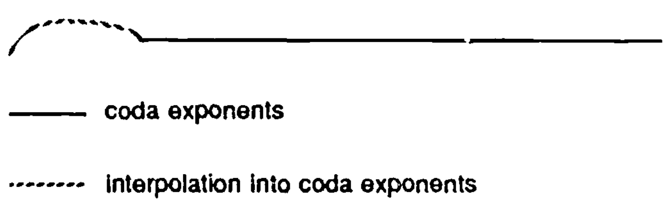

Fig 6: Transitions into the coda are more important than transitions out

It turns out that in order to join two syllables we need do nothing very sophisticated, because there are just two sorts of syllable join; ambisyllabic and non-ambisyllabic. In the ambisyllabic case, the exponents of the coda of the first syllable and the onset of the second are rather similar, and all that is needed is to ensure a smooth transition from coda to onset exponents; the sophisticated 'ways in', such as formant transitions and offset of voicing etc, are taken care of by the coda exponents, while the 'ways out' are taken care of by the onset exponents.

In the non-ambisyllabic case we do not predict any need for a sophisticated join between coda and onset exponents, and in fact the

[7]YorkTalk assumes maximal ambisyllabicity.

'pick-ups' for the onsets just look a certain distance 'back' and interpolate in a straightforward manner from 'pick up' to the first relevant value. So there is only one sort of join for overlaying onset on to coda (ie syllable on to syllable), and it is rather a simple one compared to the more sophisticated exponency statements used for overlaying onsets and codas on to nuclei and rimes respectively.

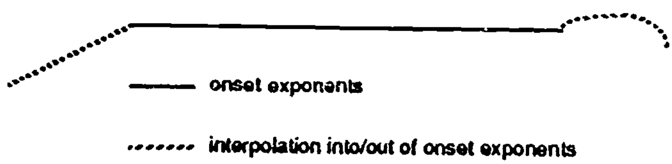

Fig 7: Interpolation out of onsets is like that into codas; but interpolation <u>into</u> onsets need not be sophisticated

Fig 8 shows the parametric interpretation of overlaying syllables; the onset exponents have been overlaid on to the coda exponents. Let us imagine in this instance that the coda and onset are ambisyllabic. Note the straightforward interpolation into the onset (ie simple interpolation between two points), but the more complex interpolation out of the onset exponents, which may require a more refined exponency statement than the interpolation into the onset. Note also the lack of interpolation out of the coda. The onset/coda join is handled only by the onset.

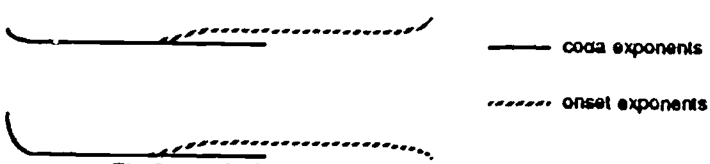

Fig 8: overlaying onset on to coda (ambisyllabic)

In Fig 9 is shown a possible non-ambisyllabic overlay; the onset exponents are the same as in the preceding diagram, but the coda exponents are different. The interpolation in the onset exponents however is the same as in the ambisyllabic case - a straightforward interpolation between two points.

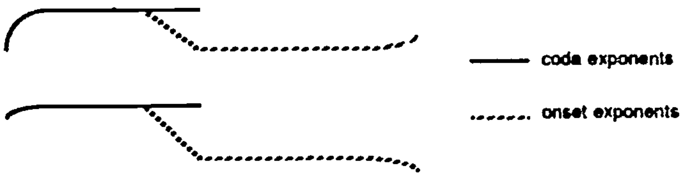

Fig 9: Overlaying onset on to coda (non-ambisyllabic)

The first case (fig 8) might be an interpretation of, say, the structure for 'ri(bb)on', while the second (fig 9) might be an interpretation of the structure for 'hus)(band', ie the first one has ambisyllabic structure, the second has non-ambisyllabic structure.

6. The Quality of Laterals in English

English, in 'classic' phonology, is said to have one lateral phoneme /l/ with two allophones [l] and [ɫ], the latter of which is found syllable-finally (see well-known descriptions of this in eg Gimson 1962 and Jones 1962). Lehiste, however, in an instrumental study (Lehiste 1964) found that the formant values of [l] in American English varied according to two things: the position in the syllable, and the vocalic environment. Syllable-initial [l]-sounds were found to be clearer on the whole than syllable-final ones; in broad acoustic terms, the difference between f2 and f3 was found to be lower on the whole at the end of the syllable. But the other strand of the analysis was that the f3-f2 difference (which can be seen as a correlate of darkness or clearness) also depended on the vowel before or after the acoustic segment identified as 'lateral' by Lehiste.

Different lateral qualities are modelled in YorkTalk in the following way: there is one set of exponency statements for all the onset laterals and another for the coda laterals.[8] The two statements are identical in form; they take the value of the formants of the nucleus on to which they are overlaid and they calculate from that the value of the formants which are the exponents of the lateral. The formulae are the same in each case; the difference is the value of the Locus and Constants which

[8]There is nothing exceptional in this. It should be clear that onset and coda exponents are logically very separate in YorkTalk, and have to be stated separately for each structural position.

are used in the formula to relate the nucleus exponent to the onset or coda exponent.

There is another difference between onset and coda laterals; in the onset, the lateral stands in a particular place in the phonological system. It is the clear member of a two-term system of liquids which commute in onset position. Whereas in the coda position it is the only member (in the variety of British English which we are modelling) of a liquid system, therefore the clearness or darkness is not phonologically relevant.

The quality which Lehiste detected in her study of medial laterals (ie in structures of the form VCV) was neither clear nor dark; it was somewhere in between. We replicate this by overlaying the onset and coda laterals in the right way to produce a period of laterality which starts off comparatively dark (from the coda) and ends up comparatively clear (from the onset), a phenomenon observed by Lehiste. So the lateral in 'silly' is not as dark as in 'sill', but not as clear as in 'lee' - it is somewhere in between because it is composed of a dark coda lateral and a clear onset lateral.

By using a different syllable overlap it is possible to produce a difference in the quality of the laterals in word pairs such as 'tieless' and 'tileless', so that in 'tieless' the laterality is shorter and clearer than in 'tileless'. This is illustrated in the figures below.

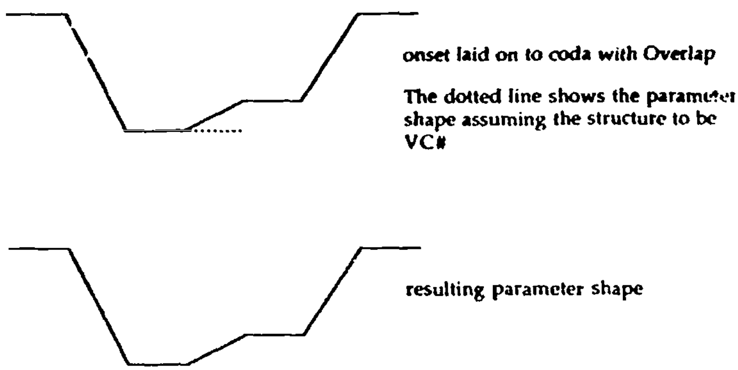

Fig 10a: lateral onset laid on to lateral coda (relatively small overlap)

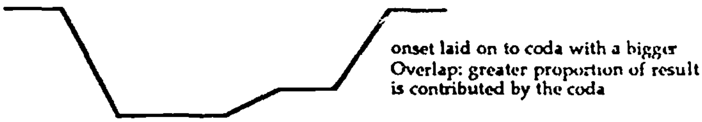

Fig 10b: lateral onset laid on to lateral coda (relatively large overlap)

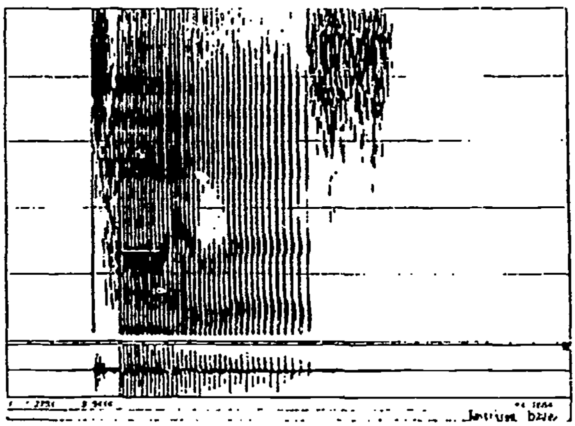

Fig 10c: Spectrogram of synthetic 'tieless' (cf Fig 10a)

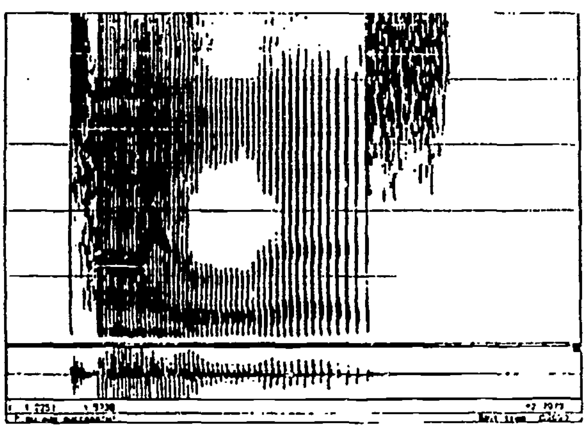

Fig 10d: Spectrogram of synthetic 'tileless' (cf Fig 10b)
Note and compare the value of f2 in the period of maximal laterality for Figs 10c & d. This is achieved solely by use of different degrees of syllable overlap (cf Figs 10a & b).

So we have only two exponency statements for English laterals, and they are determined by the phonological structure of the language; yet we can produce as many laterals as we can produce nuclei and sequences of nuclei.

Vowel Allophony

It can be observed that the final vowels of 'Henry' and 'Henley' do not have the same phonetic qualities. One is clearer and closer than the other, which is retracted and more open. Which is which will depend on the speaker's dialect, and is connected with the status of liquids in the speaker's phonological system (Kelly & Local 1986, 1989). YorkTalk models 'Henry' as dark and 'Henley' as clear. How do we achieve different phonetic qualities but have the same interpretation of the nucleus, which is not distinctive in the second syllable of these words?

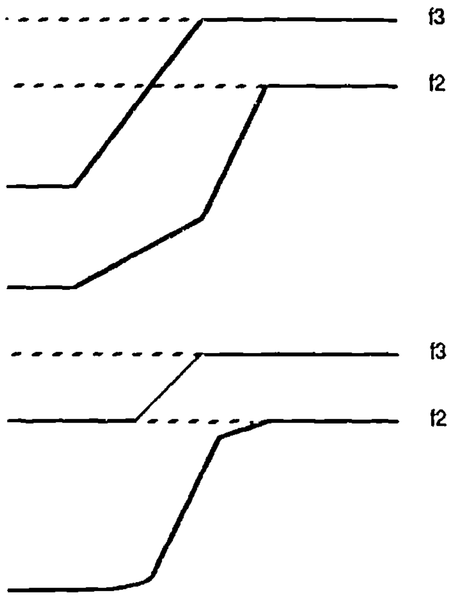

Fig 11a & b: Clear and dark liquids have different temporal structures and are overlaid in different ways, to give different qualities to the rimes on to which they are overlaid

To achieve the formant parameters which sound like the vowels described above, we use overlaying and timing. Overlaying a clear onset liquid on to a rime quite simply has a different effect from overlaying a dark onset liquid on to a rime; the off-glides are different in each case, and produce formant tracks that mimic what happens in natural speech. It is *completely unnecessary* in the YorkTalk system to handle vowel allophony by having separate exponency statements for vowels in differing phonological environments; the correct phonetic results are achieved by making sure that the components of the interpretation are as accurate as possible.

7. Summary

We have described parametric interpretation in the YorkTalk system in some detail. We have shown that parametric interpretation can be done compositionally and declaratively, and that it is possible to generate natural-sounding synthetic speech by rule.

REFERENCES

Coleman, J S & Local, J K (1987): Monostratal phonology and speech synthesis. in C C Mock, M Davies (eds): *Studies in systemic phonology*. London: Francis Pinter.

Coleman, J S (1990): Unification Phonology: another look at 'synthesis by rule'. *Proc. of CoLing*, Helsinki 1990, Vol 3, 79-84.

Firth J R (1937): *Sounds and Prosodies.* in F R Palmer (1970): *Prosodic Analysis*, Oxford: OUP.

Gazdar G, C Mellish (1989): *Natural Language Processing in Prolog*. Wokingham: Addison-Wesley.

Gimson A C (1962): *Introduction to the Pronunciation of English*. London: Arnold.

Jones, Daniel (1972): *An Outline of English Phonetics*, Cambridge: Heffer's

Kelly, J & J K Local (1986): Long-Domain Resonance Patterns in English. In *International Conference on Speech Input/Output: Techniques and Applications*, IEE Conference Publication No 258

Kelly, J & J K Local (1989): *Doing Phonology*. Manchester: Manchester University Press

Klatt D H (1980): Software for a cascade/parallel formant synthesiser. *JASA 67* (3), 971-995

Klatt, Dennis H (1987): Review of text-to-speech conversion for English. *JASA* 82 (3) pp 737-793

Lehiste, Ilse (1964): *Acoustical Characteristics of Selected English Consonants*. Indiana University, Mouton & Co., The Hague, Netherlands, pp 10-50

Local, J K (1989): Modelling assimilation in non-segmental rule-free synthesis. in D R Ladd, G Docherty (eds) *Papers in Laboratory Phonology II*, Cambridge University Press.

Robins R H (1957): Vowel Nasality in Sundanese: A Phonological and Grammatical Study. *Studies in Linguistic Analysis* pp 87-103 Oxford: Blackwell.

Shieber, Stuart M (1986): *An Introduction to Unification-Based Approaches to Grammar*. Chicago: Chicago University Press.