Chapter 5

Phonoruns

A lack of vowels is not the only phonetic irregularity that Drsk has. Syllables too are somewhat missing, and in their place are phonoruns, which are a simpler structure. While phonoruns are also a combination of phonemes into more accessible and pronounceable groups, the difference is that the boundaries between phonoruns are determined in a fundamentally different way than syllables. Specifically, for Drsk, structures that are termed "explicit-terminator phonoruns" best describe their behaviour.

In this chapter we shall denote phonoruns by the symbol ρ, and syllables with σ. We shall discuss in section 5.1 the behaviour of phonoruns using the prototypical phonorun-only language Rattssaw. Then, in section 5.2, we discuss how the way explicit-terminator phonoruns work, and how they apply to Drsk in particular. We shall then explore the implications of phonoruns and prosody in section 5.4, and finish with a reconciliation of Drsk phonoruns and Drsk syllables in section 5.5.

It should be noted that while we speak of "Drsk syllables", the phonorun has always been the more important and fundamental part of Drsk phonology. Syllables, while playing an important role in some contexts, are highly dependent on vowel insertion (for which see chapter 6) and are therefore dialect-dependent; on the other hand, the signature (which is to a phonorun as a syllabification is to a syllable) of a word is relatively more stable cross-dialectally and is therefore the one that gets used in contexts like lexicography and education.

Where it does not introduce ambiguity, the word "phonorun" may be abbreviated to the word "run".

5.1 Phonoruns in Rattssaw

Phonoruns first appeared in Rattssaw, before being spread as an areal feature to Drsk. As such it is productive to briefly touch upon its phonorun system, and to give ourselves a little bit of an understanding of how phonoruns work in general.

In Rattssaw, every phoneme carries an enumeration feature, [Open] and [Closed]. All vowels are open, all plosives are closed, and all the other phonemes are one or the other, though sometimes this can depend on the previous phoneme. Since this is a Drsk grammar, not a Rattssaw grammar, we will work with Rattssaw orthography only, which is to a near-perfect approximation Rattssaw phonetics.

So for our purposes, the phonemes in Rattssaw are as in table 5.1, along with their open and closed status. (14 | 54)

Table 5.1: Table of phonemes in Rattssaw by their open/closed status.

English	Symbol	Rattssaw	List
Open	o	Þhaər	a e i o u á é ó ay wá yé h
Closed	c	Skkdþ	p b t d pp bb tt dd þ f ff k g kk j ss #
Inheritable	i	Fflayer	l r rr þþ z s f x ś m n ng

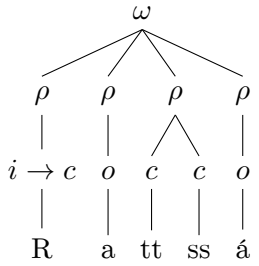
Now, the important part about phonoruns in Rattssaw is that they always consist of one or more of exactly one type of phoneme. Symbolically we say that a phonorun ρ consists of

$$\rho = \begin{cases} o + (o|i)^* \\ c + (c|i)^* \end{cases} \tag{5.1}$$

where $+$ means string concatenation and $|$ means alternation.

For instance, these Rattssaw words are "syllabified" as such:

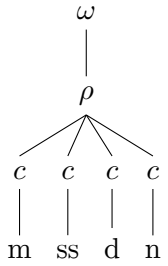
1. Rattssá



Here we have a fairly simple word with four runs. As is demonstrated here runs can be as little as one phoneme long. Furthermore, an *i* becomes a *c* if it appears word initially as the initial phonemes is always *c* unless indicated otherwise, in anticipation of the honorific prefix `#`.

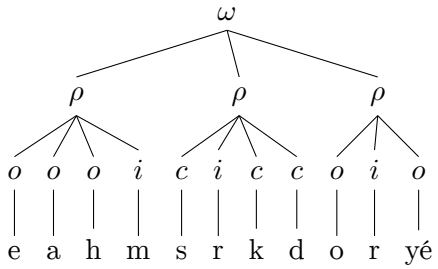
In some dialects, the initial `r` becomes an $i \rightarrow o$ instead, which means that the word now has three runs instead of four. Here the reason is that the goal is to reduce the number of runs and therefore the "inheriting" part of the inheritable phoneme can also inherit from the right if the phoneme on the left is unavailable.

2. Mssdn



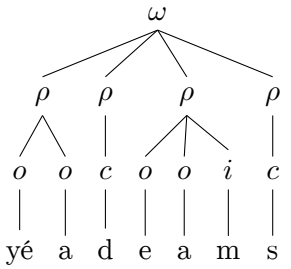
This is in some way the opposite of (1) – instead of having lots of short runs, we have one very long run. The prosody here is somewhat different as a result.

3. Eahmsrkdoryé



This is a complicated word with multiple runs that contain multiple phonemes.

4. Yéadeams



   And to round it off, here is a word that has both complex and simple runs. This should take more or less the same time as in (3), in spite of some of the objects being simpler.

In general though, apart from a difference in prosody and the simpler structure, phonoruns are more or less the same as syllables.

The important part to take away here is that a phonorun is controlled by largely regular means and therefore can be predicted in its entirety even programmatically. The trade off is that sometimes they can get very long, an that is a fact that remains true as we transition to discussing Drsk phonoruns.

5.2 Terminators and Non-Terminators

The major difference then between Rattssaw phonoruns and Drsk phonoruns is that instead of the boundary being an implicit gap caused by a switch between two types of phoneme, it is determined by a very explicit property of particular phonemes.

In other words, certain phonemes carry the property [+Terminator], which signals that a phonorun that precedes it ends with it, and a new one begins after it. (23 | 4) Hence, one can describe a phonorun using the following formula:

$$\rho = n^* + t \tag{5.2}$$

where n is a non-Terminal, t is a terminal, and * is a Kleene star.

And then, a phonological word ω is defined using the auxiliary object ω' as

$$\omega' = \rho^* + \rho \tag{5.3}$$

$$= (n^* + t)^* + n^* + t \tag{5.4}$$

in which case ω is defined to be any continuous substring of ω' which contains at least one t and ends in a t. Note that sometimes a word can end with an n, in which case an epenthetic terminator is inserted. This terminator can be implicit and is generally assimilated to the preceding consonant, or [k] otherwise.

For example, in the word tmrlħð "Rudder-TOP", the final morpheme is an n, as seen in the analysis in item 3. The entire word is pronounced /t.m.rlħð/. In order to have the last phoneme be a complete run, the epenthetic phoneme must be inserted, so for instance [tmrlhðk], or to apply an assimilation rule [tmrlhðd]. Other considerations may limit assimilation options, such as the fact that adding [t] would make things needlessly ambiguous because it would be the same as one of the suffices t /t/ [t] "TRANS" (see section 9.1). Furthermore, since we have not accounted for vowel insertion (see section 6), the final narrow transcription would be more complex than this.

As with Rattssaw phonoruns, Drsk phonoruns are also entirely regular and can be predicted programmatically, but this time run delimiters are explicitly marked out using specific phonemes. There are two main analyses that explain this particular phenomenon:

1. The **single-class hypothesis** claims that there has only ever been one class of phoneme, as most other languages that use phonoruns typically require that consonants be able to cluster together to a single class; and in the case of Drsk there are *only* consonants so naturally there's only one "extant" class. (Older versions of Drsk indeed still has vowels.) In this case there are pressures arising from overlong phonoruns that is then resolved by creating a terminator structure, explicitly terminating a run at set point.

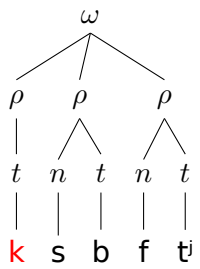
2. The **restricted second-class hypothesis** claims that the terminal phonemes have always been part of a separate class – it isn't unheard of that consonants are broken up in phonorun classes – and that the length of the run in this particular class is then restricted to 1.

It can be argued that the truth is somewhere in between these options, with a large amount of phoneme classes either melting away or combining into one, all the way up until the last few phonemes remain in a hold-out class, which promptly was restricted down to runs of 1 long. This situation is then reänalysed into the situation we have today.

Regardless, here are some examples of how "syllabification" works in Drsk, using the orthography mentioned in section 7.1. The t phoneme that is

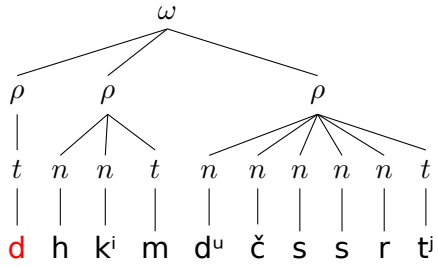
required for a word to be valid is coloured as such. Phonologically speaking, there is no particular justification as to which of the terminals can be the requirement, only that there *is* one, though overall the choice of what counts as the requirement is the first terminal in the root word.

1. ksbftʲ "Seasonal pond"



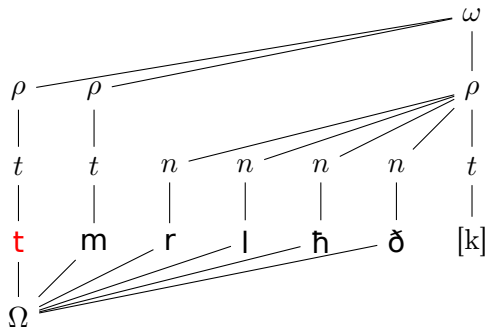
A relatively straightforward word consisting of three runs with a modest amount of phonemes each.

2. dhkⁱmdᵘčssrtʲ "*Wool$_1$"



A slightly more complex word with long runs. In principle, there is no upper bound to how long a run can be, though of course practically they are bounded by how much one can remember at any given time. That is to say, in Drsk there is no "maximal phonorun".

3. tmrlħð "Rudder-TOP"



This word is special in that the phonological word ω is not the same as the *lexical* word Ω, which means that the final [k] that was added on was merely a phonotactical consideration – it does not form part of the lexical word at all. However, both can be putative "roots" of the word, so they are drawn as trees with shared leaves here. More examples of this kind of detachment can be seen in section 5.3.

The examples above show a great difference between phonoruns and syllables; not only are phonoruns allowed to go on arbitrarily long, they also have relatively little internal structure (as previously mentioned), which implies that most if not every possible combination of phonemes up to the definition of the phonorun can be one. Other languages that use phonoruns don't have this unusual leniency, though in effect their restrictions tend to be adjacency limits (including adjacencies with the boundaries) as that is all that phonoruns can afford.

If a terminator starts a word, it is called the "incipator". Phonologically this does not change the behaviour of the phoneme in isolation, and it retains virtually all of the qualities that a terminator would have. However, incipators can sometimes become part of a "false word" (a phonorun that does not belong to any particular word) if it is preceded by a word that does not end in a terminator. All words shown here are incipators, and in fact a large amount of words have one, again due to morphological considerations on nouns.

5.3 Phonological and lexical words

As seen in example (3) in section 5.2, *phonological* words ω may not be the same as *lexical* words Ω, though the two are closely related.

As so it happens most ordinary nouns are guaranteed to have $\Omega = \omega$, as due to the suffix mechanism (see section 9.1) they all in their dictionary form have the suffix t^{j} which is class t and therefore closes the run right where it is most desirable, i.e. at the end of Ω. While its most common alternate forms replace the t^{j} with other phonemes, the replacements are also frequently closed when it terminates a word (e.g. t, p or t^{t}), which allows us to confidently have $\Omega = \omega$. When this doesn't happen, we have several ways to resolve the problem. This behaviour is fairly complex and interacts strongly with runs and syllables.

When $\Omega \neq \omega$, the situation is considered unstable and some dialects behave strangely. In particular, the phonemes that "stick out" of a lexical word can be swept up by a terminator in the next word, causing a ρ to stick

out to the next Ω, seemingly ignoring any phoneme that has been inserted stop up the phonorun at the end of Ω). This can result in a case where the phonoruns as they are pronounced not lining up with word boundaries. A concrete example of this is given in section 5.5, where we will also see common ways to resolve this instability.

5.4 Prosody

Generally speaking, prosody using phonoruns are handled by the time it takes to go through one of them, and all of them in a single word. This timing can be accurately predicted using a simple equation, which is known as the prosody equation.

First, let's see the equation itself, which is Equation 5.5.

$$T = Kn^{L+\nu n} \tag{5.5}$$

Here, K, L and ν are language- and dialect-specific constants, T is the total amount of time required to pronounce the phonorun, and n is the number of phonemes that are in that run. K in particular controls the overall speed of the speech, and since that is in turn controlled by various individual effects (as well as the choice of units used to measure T) this turns out to be basically impossible to predict. On the other hand, L and ν are much more predictable and stable from region to region, and from year to year.

L is called the "starting constant". It says that as a run goes on for longer, the successive consonants take longer and longer to come out. There is a small correction factor called ν, which is called the "corrective constant". This further penalises large runs.

Generally speaking, exponentials are bad as they quickly grow out of control, so for this reason we often have $1 < L \leq \frac{5}{4}$ and $0 \leq \nu \leq \frac{1}{40}$, and similarly it gives a soft cap as to the maximum number of phonemes in a run, namely, $n \lesssim 6$.

For example, a specific dialect of Rattssaw has the following quantities for use in the Prosody Equation:

$$L = \frac{5}{4}; \nu = \frac{1}{48} \tag{5.6}$$

For all Drsk dialects, the distribution of L and ν can be summarised as a roughly normal distribution with the following parameters:

$$L = 1.15 \pm 0.04 \tag{5.7}$$

$$\nu = 0.0139 \pm 0.0019 \tag{5.8}$$

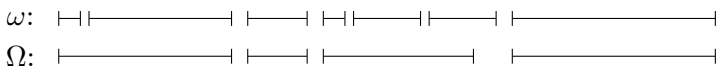
Geographically speaking, this distribution is largely continuous, with neighbouring regions having roughly equal values of L and ν. It is usually higher toward the south, and lower toward the north and near the border with the J.-Senlis.

5.5 Syllabic non-denial and reconciliation

After vowel insertion, the vowels form nuclei, and from there we have recreated syllables, σ. The ability of Drsk to somehow retain the syllable even after phonoruns have supplanted it as the main way to divide up a word is remarkable, although circumstances has never conspired again to see if another language would do the same under similar conditions. This makes Drsk a non-denial language – it does not deny that syllables exist in itself (which is definitely the case in Rattssaw, where there's no sign of a syllable anywhere.) As previously mentioned however, where and what vowels are inserted is highly dialectal and therefore cannot be made consistent throughout the language.

Generally speaking, a Drsk syllable prefers to be in the form described by section 6.3. While they can also be considered units of prosody, it is not generally done for one simple reason: their relationship with both types of words can get very counterintuitive and unusual. Compare this to phonoruns, which has a much nicer behaviour when it comes to word boundaries.

All ω have an obvious parent Ω such that no ω can simultaneously belong to two Ω. That is to say, we can construct a forest where all the roots are objects in Ω and all the leaves are objects in ω. Graphically, it would look like this:



where each line is an instance of an Ω or an ω depending on which row it is on, and their boundaries are indicated. Notice how while every Ω boundary coïncides with an ω boundary, there can be times where an ω slips out, due to the requirements of an ω needing to terminate at a t, not seen in this level of abstraction.

As for the most part we have $\Omega = \omega$, most of the lines in the two rows line up exactly. Since they're mostly plain nouns, we can therefore say that

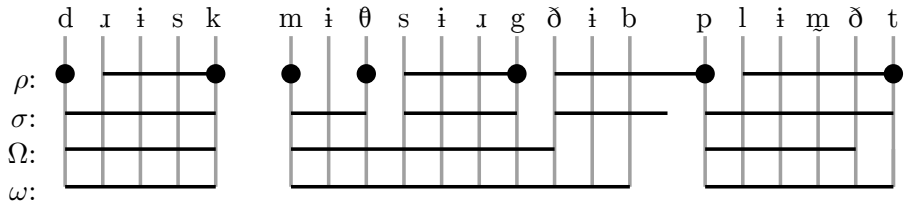


Figure 5.1: The divisions of the phrase drsk mŤSRGĐ plmᵘð in runs (ρ), syllables (σ), lexical (Ω) and phonological (ω) words. On the phonoruns row, the terminators are indicated using dots. Grey lines connect each phoneme to the rows to ease reading.

most nouns have a phonetic signature that allows them to be detected even without much external context. The caveat is that this is not a fact that is transparent to most speakers and the education system does not particularly seize upon this point.

In the cases where $\Omega \neq \omega$, things get a little bit more involved, so let's provide an example phrase. (23 | 46)

(1) drsk mŤSRGĐ plmᵘð
/dɹsk mθsɹgð plm̰ð/
[dɹi̟sk mi̟θsi̟ɹgði̟b pli̟m̰ðt]

Note that we have already inserted vowels (here, the constant [i̟]) in the narrow transcription, and opted for a certain assimilation rule to round off the phonoruns (here marked as such). Let's draw in the diagram indicating where all the syllables σ, runs ρ, phonological words ω and lexical words Ω are. This diagram is in figure 5.1.

We can see that in this phrase, the requirement that no ω can be part of two or more Ω holds, and in fact there's usually a one-to-one correspondence between them. This is the usual case, and having a many-to-one relationship is exceedingly rare.

Notice that for the most part, Drsk syllables will connect any isolated Terminal phonemes to the next run, stopping at somewhat unpredictable locations (see, for instance, how Drsk ends up being one syllable, with the incipating d being pushed in to the rest of the word which is one run long.) Indeed, most foreign speakers unaccustomed to phonoruns would use that many syllables to pronounce it.

An interesting effect comes when considering how the [p] in plmᵘð interacts with the epenthetic consonant [b]. In dialects that utilise constant

epenthesis to stop up phonoruns (using [k], for instance), it and the [p] would be combined to create a coärticulated consonant, [k͡p], or sometimes an unreleased [k̚p]. In any case, the run then ends in between words. This can result in a somewhat confusing situation where a ρ can span more than one ω and Ω, but there is no problem since ρs are determined on a phrase level, whereas the rule that they must stick within one ω or Ω is only determined on a single-Ω level.

However, when it comes to syllables, the system breaks down somewhat. Notice the awkward placement of the syllable boundary; it is because there's no consensus as to which place the boundary should be, especially if the dialect does not assimilate the epenthetic consonant. This means that it is possible that the syllable boundary can stick in-between the two consonants, causing an awkward situation where there is a time interval where both syllables are in effect, which isn't particularly favoured. Syllables are in fact only tacked on to Drsk analysis in some way, and they don't have anything hanging off of them.

As a result, syllable analysis is not particular common with Drsk, and phonorun analysis is favoured instead.