Variation in the Swedish Postalveolars: A Phonetic Investigation with a Gestural Analysis

Emanuel Karlsson

Supervisor: Pétur Helgason
Abstract

This paper presents a phonetic study of variation in postalveolar consonants in Swedish. Postalveolars appear both word-internally in the coda and at word and morpheme boundaries where a final /r/ is followed by an initial dental. Traditionally, postalveolars in all positions have been considered to be the phonetic outcome of an underlying phonological segment sequence of /r/ followed by a dental. For this investigation, the lowest spectral resonance frequency of sibilants that are eligible to be postalveolar was measured in two sets of data, one of scripted speech and one of unscripted speech. The findings reveal different behaviour in the different positions: internally there is obligatory [ʃ], whereas at boundaries there is a continuous spectrum of variation from dental to postalveolar (as opposed to a simple choice between [ɹ] and [ʃ], predicted by the traditional model). Also, the presence of a rhotic before the sibilant at boundaries may not be directly correlated to the sibilant being dental (which would have been expected in the traditional model). It is argued that the phonetic variation observed is incompatible with the traditional description. Assumptions that underlie the traditional analysis of Swedish postalveolars (e.g., biuniqueness) are criticized theoretically, as is the traditional model of speech production in general. An alternative analysis set in gestural phonology is proposed.
Table of contents

1. Introduction ...................................................................................................................... 5
2. Background .................................................................................................................... 6
   2.1 Postalveolar consonants in Swedish ........................................................................ 6
      2.1.1 Preliminaries ...................................................................................................... 6
      2.1.2 Issues in the distribution ................................................................................... 7
      2.1.3 The traditional account of underlying segment sequences ................................. 8
      2.1.4 Patterns of obligatory rhotic plus dental ............................................................ 11
      2.1.5 The traditional appeal to geminates .................................................................... 12
      2.1.6 Systemic considerations of the postalveolar sibilant ........................................... 14
   2.2 Previous treatment in the literature ................................................................. 14
      2.2.1 Eliasson (1986) .................................................................................................. 15
      2.2.2 Elert (1957, 1964, 1970) .................................................................................. 17
      2.2.3 Witting (1959) ................................................................................................ 17
      2.2.4 Malmberg (1969a, 1969b, 1968) ..................................................................... 18
      2.2.5 Other accounts .................................................................................................. 18
   2.3 Gestural phonology ................................................................................................. 19
3. Method ............................................................................................................................ 21
4. Results ............................................................................................................................ 24
5. Analysis .......................................................................................................................... 27
   5.1 Implications of the data ............................................................................................ 27
   5.2 Gestural analysis of postalveolars in the data ......................................................... 28
      5.2.1 Internal [f] .......................................................................................................... 29
      5.2.2 Boundary aligned [f], disaligned [ːs], and mixed [ːʃ] ........................................... 31
      5.2.3 [ːʃ]and vowel rhoticization as gesture weakening ............................................. 33
      5.2.4 Boundary marking of the type [ʃs] and variations .............................................. 34
6. Discussion ....................................................................................................................... 36
   6.1 Summary of the results and analysis .................................................................... 36
   6.2 Traditional assumptions of the phonology of postalveolars ................................. 37
      6.2.1 Implications of phonological status ................................................................. 37
      6.2.2 Biuniqueness .................................................................................................... 38
      6.2.3 Variation at boundaries ..................................................................................... 38
      6.2.4 Hierarchical allomorphy .................................................................................. 39
      6.2.5 Geminates ........................................................................................................ 40
      6.2.6 Segments and derivations .................................................................................. 40
   6.3 Alternative perspectives on phonology ............................................................... 41
      6.3.1 The nature of speech and implications for phonology ..................................... 41
      6.3.2 Explaining speech in the traditional model ....................................................... 42
      6.3.3 Mental concerns ............................................................................................... 43
      6.3.4 Pragmatic models of phonetic variation ........................................................... 44
6.4 Interpreting the gestural variational analysis of the postalveolars ............................ 45
      6.4.1 Variation and gestures in the postalveolars ....................................................... 45
      6.4.2 The case of obligatory rhotic plus dental ......................................................... 46
      6.4.3 Considerations in explaining the variation in r#s ............................................... 47
      6.4.4 Final remarks ................................................................................................... 48
7. Conclusions ................................................................................................................... 49
References ......................................................................................................................... 51
1. Introduction

The study identifies and describes postalveolar consonants in Standard Swedish. In the traditional analysis, these are considered phonetic outcomes of a derivational process from a sequence of underlying phonological segments /r/ plus dental. This is based on a generally implicit assumption of biuniqueness: as postalveolar phones may appear at word boundaries with a final /r/ followed by an initial dental, biuniqueness dictates that word internal postalveolar phones must be derived in the same way. It is argued here that biuniqueness cannot be assumed to be a valid concept. Also, phonetic sequence of [ɹ] plus dental does indeed occur internally, and it is argued the traditional explanation for this is inadequate.

A phonetic investigation shows how in the case of the sibilant there is a difference between how internal postalveolars and boundary postalveolars behave: internally, the postalveolar sibilant [ʃ] is obligatory, while at particular boundaries, there is a continuous span of variation basically from a sequence of [ɹ] plus a dental sibilant [s] to a unitary postalveolar [ʃ]. The investigation measures the spectral quality of all occurrences of sibilants eligible to be postalveolar in two sets of data (mainly a radio news programme). The results illustrate the difference between the categories and the continuous variational scale of [ɹʃ~ʃ] in the boundary category.

One purpose of the investigation is to show the wealth and importance of phonetic variation and it is asserted that this variation must be acknowledged and accounted for in an analysis. The paper argues that the continuous variation present in [ɹʃ] does not yield to a derivational, segmental analysis. Rather, it argues that to represent the variation adequately it can be displayed gesturally, where the special relation between [ɹ] and [ʃ] as different constellations of the same gestures is highlighted. An analysis of instances of variable [ɹʃ] in a gestural perspective also reveals that the quality of the fricative is not directly correlated to the presence of the distinct rhotic (i.e., the presence of a rhotic cue does not necessarily imply a dental quality, contrary to what is traditionally assumed).

The analysis of the study is based in gestural phonology, in which phonological treatment of phonetic form is not seen as fundamentally different from its articulatory manifestation. This means phonetic and phonological segments in the traditional sense are abandoned in favour of gestures that allow for gradual and temporal variation, and categorical derivational processes are rejected to allow for distributional variation, favouring instead associative linking between sounds and between morphological forms.

The paper starts with an introductory section that describes the occurrence and distribution of postalveolars in Swedish, accounts for the traditional analysis, presents some accounts in the literature with the traditional view, and briefly introduces gestural phonology. This is followed by two sections on the method and results of the investigation. The next section presents an analysis of the results and additional gestural analyses of some tokens in the data. The discussion evaluates traditional phonology and the meaning of phonological status, develops the gestural view of the data, and also touches on the question of explaining the variation in [ɹʃ].

The study does not treat postalveolars as they exist in different varieties of Swedish (or in Norwegian) but is concerned only with Standard Swedish, nor does it give any typological comparisons. It makes no real attempt at accounting for or explaining the phenomena diachronically. It also does not offer any quantified explanation of the variation, but makes a few suggestions on it and on how to proceed.
2. Background

2.1 Postalveolar consonants in Swedish

2.1.1 Preliminaries

Swedish is fairly average in its number of consonants (being very rich in vowels) though there are aspects of the composition of its inventory that are typologically more unusual. Two points can be considered here:

1) The possession in non-initial positions of an additional set of coronal articulation points (apico-postalveolar) slightly further back than the common dento-alveolars (laminal); this is areally very unusual, and among Germanic languages it is shared only by some varieties of the closely related Norwegian – such a contrast is otherwise virtually unknown in Europe.

2) The profusion and complicated distribution of its lingual fricatives, with at least three distinct sibilant articulations (none of which involves a voicing contrast), employed by possibly four phonemes.

Part of the complexity of #2 is a direct product of #1, as one of the sibilants is found in the new row of coronals. In what ways these two issues are linked within the context of ongoing developments in the fricative system is well outside the scope of this study (but see 2.1.6). The topic here is the synchronic treatment of the posterior set of coronals and more specifically the sibilant fricative concerned in the investigation. First, we turn to the phonetic nature and distribution of the “posterior coronals”.

In non-initial positions, Standard Swedish has two contrastive coronal places of articulation in the alveolar region. These articulations differ in their distance from the upper teeth for the point of tongue tip contact with the alveolar ridge. The anterior coronal set is the more common by distribution and diachronically the original; it is usually laminal (but apical is common at least in dialects) and largely dental rather than alveolar. The posterior coronal articulation is apical and impacts further back so that by and large the tongue tip strikes the back of the alveolar ridge or somewhat behind it (i.e., a postalveolar articulation). This contrasting distance is partly variable (as evidenced by the investigation) and often reduced.

The posterior set is variously called “supradental” and “retroflex” in traditional terminology. In the standard variety, these sounds are normally distinctly apical and not retroflex in the stricter sense of subapical. More posterior apicals and even subapicals do occur in certain dialects, though. The term “supradental” is vague and seems to imply anything alveolar and possibly beyond, useful perhaps a cover term for enveloping dialectal variation. While alveolar articulations are common, the canonical place of articulation is rather towards the back of the alveolum: the term postalveolar will be used here to refer to the articulation and the set of sounds (especially relevant in contrast to the dento-alveolars, or just “dentals”).

Thus there are two sets of coronal consonants – laminal dento-alveolar and apical postalveolar – comprising four or five elements each: the plosive voicing pair, the nasal, the sibilant fricative, and less frequently (as a postalveolar) the lateral approximant. The less predictable behaviour of the lateral of the postalveolar set will be described below. The rhotic is coronal but does not take part in the contrast in the standard variety concerned here it is

---

1 According to Livijn & Engstrand (2001) and Livijn (2010), /t/, /l/ and /n/ are consistently dental, whereas /d/ is frequently alveolar. However, /l/ can certainly also be alveolar. /l/ has extended contact area as compared to /d/ (Engstrand 1989, Lindblad & Lundqvist 1996). /t/, /d/, /l/ are called apico-laminal by Lindblad & Lundqvist (1996). For simplicity, all the anterior coronals will collectively be referred to as dentals. /l/ is probably more clearly laminal in the standard variety, whereas the others may be somewhat more prone to being apical.

2 A retroflex is taken here to mean not when the tongue simply is drawn back to reach behind the alveolus, but when the point of contact is the underside of the blade (cf. Simonsen et al 2000, contrasting Ladefoged & Maddieson 1996:25). The term may be useful typologically however for where there are two contrasting sets.
usually an apical alveolar or postalveolar with approximant, tapping or trilling manner, with some contextual and much inter- and intra-speaker variation.

Table 2.1 gives a systematized view of the Swedish coronals. As there are no dedicated IPA letters for postalveolars other than the fricatives, the alveolar letters modified with the diacritic for retraction are used here in transcription. Elsewhere, the IPA retroflex letters have commonly been used to represent these sounds.

Table 2.1. The Swedish coronals (except /ʃ̟/)

<table>
<thead>
<tr>
<th>Dental</th>
<th>Postalveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t̠</td>
</tr>
<tr>
<td>d</td>
<td>d̠</td>
</tr>
<tr>
<td>n</td>
<td>n̠</td>
</tr>
<tr>
<td>s</td>
<td>ʃ</td>
</tr>
<tr>
<td>l</td>
<td>l̠</td>
</tr>
</tbody>
</table>

Table 2.1. The Swedish coronals (except /ʃ̟/)

<table>
<thead>
<tr>
<th>Dental</th>
<th>Postalveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t̠</td>
</tr>
<tr>
<td>d</td>
<td>d̠</td>
</tr>
<tr>
<td>n</td>
<td>n̠</td>
</tr>
<tr>
<td>s</td>
<td>ʃ</td>
</tr>
<tr>
<td>l</td>
<td>l̠</td>
</tr>
</tbody>
</table>

Though the term “postalveolar” is used, the question of whether the defining trait of this set is being postalveolar or apical or both is not to be directly addressed.

It can be positively established that there are two contrasting sets of dentals vs. postalveolars in non-initial position. This is evidenced by minimal pairs such as [faːt] “plate” vs. [faːt] “speed”, [buːd] “shed” vs. [buːd] “table”, [tʰuːn] “tone” vs. [tʰuːn] “tower”, [his] “elevator” vs. [hif] “millet”, and (less consistently) [ˈpʰɔːlɐ̞] “to pile (ground)” vs. [ˈpʰɔːlɐ̞] “to purl”.

2.1.2 Issues in the distribution

With the above functional relevance in mind, it is nonetheless evident that there are distributional irregularities that single out the behaviour of these sounds:

1) As said, postalveolars appear contrastively only in final and medial positions. They do appear in initial position, but never semantically contrast the dentals; rather, they are contextually triggered (cf. #4 below). They are also much less common than dentals.

2) They seem to have a special relationship with the rhotic consonant on account of a number of factors. To start with, /r/ is the only coronal (apart from /ʃ̟/) that does not take part at all in the contrast and does not come in a pair of corresponding dental and postalveolar.

3) Postalveolars do not follow [ʃ] (except possibly in free variation with dentals, cf. below). Sequences in the coda of [ʃ] plus dental are also fairly rare but do exist, whereas coda sequences of [ʃ] plus other consonants are common (these too do not occur initially within a word as the rhotic is never followed by any consonant in the onset). [ʃ] plus dental occurs internally also across a syllable boundary when the rhotic is the only element in the coda of a rime whose nucleus is a stressed short vowel (cf. interpretations of this below).

4) The contextual trigger for a word-initial postalveolar is a lexical /r/ that occurs in the coda of the preceding word either finally or preceding applicable coronals, in which case such subsequent coronals in the coda and in the following onset will be postalveolar (i.e., maintain the place of articulation of the rhotic, assuming it is actually postalveolar), with absence of a discrete [ʃ]. So, a sequence of lexical /r/ plus dentals here can be interchanged phonetically with postalveolars. However, as shown by the investigation, the outcome is a continuous scale from dental to postalveolar. In semantically contrasting non-initial position too, the form in

---

3 Though [ʃ] is usually defined as palato-alveolar; it is uncertain if there is a palatal element here.

4 Articulated slightly front of the postalveolar, it does not align with either dentals or postalveolars; spelt <tj>.

5 [ʃ] will be used to denote the rhotic phone (as a distinct cue). It is frequently approximant, but it may also be a tap, [ɾ], or rarely (perhaps emphatically) a trill, [r].

6 Not coronals following in turn the lateral (Eliasson 1986, Elert 1997), nor /ʃ̟/ (which only occurs initially).
which they appear can frequently be linked to a root ending in /t/ plus a suffix or root beginning with a dental. The non-initial instances however will (in all but a subset of cases) not vary but be consistently postalveolar (the investigation shows they will only exhibit such variation when the two entities are elements of transparent lexical compounds).

5) Importantly, /t/ is usually thought of as basically sharing the same place of articulation as the postalveolars or at least align with them (its articulation may in fact vary, but the effect could be diachronic). This would explain what in #4 above is traditionally termed assimilation of subsequent dentals to the postalveolar articulation of /t/ (a descriptive device which on account of articulatory concerns and not least the variation displayed is rather simplistic). Put another way, postalveolars basically share some articulatory characteristics (their place) with the rhotic and some (their manner) with the corresponding dental.

6) There are contextual constraints on postalveolars: e.g., voiced postalveolars hardly occur with stressed short vowels (correspondingly, the sequence [ɻ] plus dental is rare except when the dental is [d] and the preceding vowel is stressed and short). Postalveolars normally do not contrast with a sequence of [ɻ] plus dental: both do appear – so the assimilation of dentals to the articulation of [ɻ] and subsequent “loss” of [ɻ] does not always occur – but usually in different contexts. A major group of exceptions are those that would occur from words or elements of compounds whose final-syllable rime consists of a stressed short vowel and [ɻ] (cf. #3 above). Other exceptions are usually found among proper names and loanwords.

7) There seems to be more or less free variation for most speakers between the sequence [ɻl] and the postalveolar lateral [ɭ] in internal position too. Some speakers may perhaps more consistently (but not necessarily exclusively) use [ɭ].

8) Postalveolars are cognate to sequences of rhotic (whether coronal or dorsal) plus dental in other varieties (and in some varieties to plain dental with no rhotic). This of course has only indirect synchronic relevance to the variety under consideration, but hints at the diachronic development, additional evidence for which is orthographical and historical.

9) The vowels whose usual manifestations are [ɛ] and [œ] are lower – [æ] and [œ̞] – both before postalveolars and before [ɻ].

All of this seems to indicate some connection between the postalveolars and the sequence of [ɻ] plus dental. Evidently, at some historical stage, what are now postalveolars in all contexts were [ɻ] followed by dentals. Postalveolars can be described diachronically as the results of the articulatory coalescence of these sequences – in a way, the place of articulation of [ɻ] has been applied to all directly following applicable coronal articulations that were otherwise dental (a “spreading” of place as it were). The diachronic change has progressed gradually across the different dentals (cf. e.g. Eliasson 2005), with the lateral being the last to partake (one reason for this could be its phonetic similarity to the rhotic), showing variation that arguably could be symptomatic of ongoing change. But there is apparently also a synchronous process involving postalveolars: at any rate, there is something which at first sight has the same effect (and is basically alike in origin), giving postalveolar results across word boundaries where a lexical coda /t/ (or other postalveolar) and initial dental come together.

2.1.3 The traditional account of underlying segment sequences

In the traditional account of the Swedish phonological system (cf. the overview of literature in 2.2 below), allpostalveolar articulatory patterns are considered phonetic outcomes of a derivational process of rules working upon underlying sequences of phonematic /t/ plus dental. In this view, postalveolar articulations are, as such, not phonemes in their own right, but

---

7 According to Elert (1997), final or post-tonic <rl> is less commonly [ɭ]; it also can be dental [l] with no rhotic.

8 In a gestural view, lowered /l/ and /ɭ/ occurs before the postalveolar gesture which is present in both [ɻ] and the other postalveolars. Alternatively, these could also reasonably have been due to a loss of conditioning environment, which certainly occurs in varieties with plain dentals only.
merely complex realizations formed by the synchronous coalescence of two underlying phonemes into one segmental *phone* in a rule-based derivation. So for example the word [fɔʃ], orthographically *fors*, is considered to consist of four underlying segmental units represented as /fors/, the last two of which undergo a “phonological process” producing as the phonetic “surface realization” the singular segment. (The sequences of phonetic [ɹ] plus dental that do occur must then be assumed to be in complementary distribution to postalveolars in one way or another.) It follows from this argument that both cases – across boundaries and internally – yield identical results, and so – by theoretical assumption – that all instances of postalveolars must be outcomes of the same synchronous process with identical segmental input everywhere (and so the synchronic and diachronic processes are one and the same). There is reason to question both the treatment of data and the theoretical suppositions.

Firstly, as shown in the present study (but rarely considered in the literature), the outcomes of the two categories do *not* give the *same* results. Put simply, postalveolar articulation is obligatory (and [ɹ] plus dental impossible) in “internal” cases, while those with intervening lexical boundary not only show “free” variation of a postalveolar and a sequence of [ɹ] plus dental, but in fact do so on a continuous spectrum from dental to postalveolar place (not entirely correlated to the presence of a perceptually independent rhotic cue [ɹ]). So, regardless of what conclusions to draw from this, there is undeniably some *difference* to start with: in one case, no sequence [ɹ] plus dental (except for the lateral) ever “surfaces” as it were.

Secondly, the traditional argument builds upon – and is relevant because of – preconceptions in structuralist phonology that cannot simply be assumed to be valid. The following are some interdependent notions that come into play here (cf. e.g. also the discussion in Ohala 1990):

1) Reliance on “form” as the basis for phonological categorization – a belief that phonology is solely or primarily a matter of distributions of discrete segments, a structural property that is best gauged simply from “appropriate” backtracking of the phonetic material (implicitly assuming it is bidirectional). Functional contrastiveness, then, is secondary. To make the analysis more intuitive, an arbitrarily defined proviso of phonetic similarity is introduced.\(^9\)

2) Segmentality, as necessary for crystallizing patterns of distribution according to the above: it is maintained that speech (despite phonetic appearance to the contrary) is to be understood as being a linear string of discrete and fundamentally invariant alphabetical units with strict internal structure (where the switching of some binary feature results in a different unit).

3) Computationality in derivation: in order to “properly” generalize distributions (and by extension to account for the deviance of the signal from its assumed segmental nature), surface form is taken to be derived from segmental underliers (mental or wholly abstract) that are converted through an abstract regular mechanism, whereby computational rules invoke a phonological input and an equally abstract “phonetic” output to transform it into, with no place for variation. Presumably in some more realistic sense this is the transfer of mental units composing lexical material into articulatory mechanisms – and from perceptual cues back – by regarding production and perception as linear translation or a coding and decoding process. It is in this sense, then, that it would be meaningful to speak of “two underlying units”.

4) “Biuniqueness”, meaning the supposed need – in order for perception as linear decoding to work and for invariance to be real – for a bidirectional one-to-one mapping of phonetic to phonological units by way of context rules, wherein no two different underlying phonematic specifications could compete for the same contextual allophone, and equivalent outcomes can only map to a unique phoneme. So, as regards the postalveolars, this means that because the same phonetic outcome (or what goes for being the same) also results from combinations of a final /t/ of one word followed by the initial dental of another, each and every instance of this phonetic outcome (i.e., also internal postalveolars) must be derived from this same source.\(^10\)

\(^9\) So that complementary [h] and [ŋ] in English or Swedish are not to be taken as being of the same phoneme.

\(^10\) Though a difference between the initial and the final has been established here.
5) As a consequence of computationalism, there is – corresponding to the desire in morphology to derive all related forms by simple concatenative affixing rules from invariant roots – a dislike of alternations. The underlying form must then consist of an invariable string of “radical” segments (i.e., equal to the root). Alternating forms are to be explained by (phonetic) derivational rules as far as possible – e.g., in an inflection with an internal vowel change, or a form where a segment of the root has disappeared through historical apocope, a phonological rule that in fact represents a diachronic event must be posited to explain this as a “synchronic” effect on the “mental” root. Thus alternates such as /s-ʃ/ and /-r/:/-Ø/, which then require all relevant lexical items to have additional defined allomorphs diverging from the root, are out of favour in face of a simple phonological rule of basically /rs/→[ʃ]. Arguably, alternations as being strictly derived allomorphs of fundamentals do not seem mentally relevant, as opposed to treating them under a less hierarchical and more associative model (also, the alternation here will become more transparent in a gestural treatment).

6) There is an a priori desire to reduce the number of distinct underlying units in the phonological system. This favours rules that delegate as much of the phonetic diversity as intuitively possible (by some unspecified criterion) to combinatorial effects of units “already present” over adding new contrastive units to the system. So, if there can be fewer phonological units by positing extra rules, this is seen as an intrinsic gain. (Likewise, the number of phonetic features used to specify phonemes are also to be kept to a minimum.)

Thirdly, as suggested above, the data is more complicated than is accountable for by a simple ubiquitous assimilation and deletion process as in a set of rules, even if such processes were taken to be mentally relevant and real. Continuous variation between [ʃ] and [ɹ] is not something that yields to description within this mode of representation.

The traditional rule formulation of the process is something like [+cor]→[+ret]/[+-ret] and r→Ø/[-ret], where the first rule is applied recursively (and where /r/ is [+ret] and boundaries are ignored). This means that a segment with the feature “coronal” activates the feature “retroflex” when it follows a segment with the “retroflex” feature (which includes /r/). Then the /r/ segment is removed by a separate rule. It is then necessary to apply the first rule recursively – that is, to run it again and again until all consecutive coronals following a retroflex (including those produced by the rule the last time around) are made retroflex, so that a stretch of dental segments will all become retroflex: recursion under this system must be a separate mechanism, the application of which is some sort of property of the rule.

The derivational and sequential notion implied in this arguably does not give a very faithful portrayal of the actual production mechanism. Firstly, there is clearly one cohesive process at work, retaining the tongue in its posterior position throughout a stretch of time and other articulatory specifications, and not an iterative change (even on a mental level) which has to be reapplied separately to each “segment”. In addition, it requires the postulation of a separate recursion mechanism of some kind, which is left undefined and without clear mental equivalence – it introduces intermediate stages in the process, where deficient forms in-between recursive treatments of the rule “exist” somewhere between input and output.

A derivational system of categorical, dichotomous rules fails to cover such variational patterns as seen in the present study, where the internal case has an obligatory postalveolar whereas across boundaries there is continuous articulatory variation. It also fails to account

---

11 Note that (because of the segmental approach) in the alternation of finals it would not be enough to have the suffix alternate, seeing as one of the integral units of the stem ostensibly “disappears”.

12 A more elaborate version is given in Eliasson (1986:286); in an endnote (295) he gives an alternative which seeks to incorporate recursion into the rule, “as often done by Norwegian linguists”.

13 Per biuniqueness, every phonetic segment must be an allophone of one phoneme, and so [ʃ] is said to be an allophone of /s/, with /h/ deleted; however, it ought to be equivalent saying it is an allophone of /h/ with /s/ deleted, but for the fact that this is counter-intuitive.
convincingly for the deviating cases of root-internal, lexical-specific sequence of obligatory [ɾ] plus dental (those these have admittedly been acknowledged to exist).

2.1.4 Patterns of obligatory rhotic plus dental

The following defective distributions or exceptions to the supposed regular derivation of /t/ + dental to postalveolar can be drawn (some additional examples are presented later on):

1) Voiced postalveolars almost exclusively do not occur following stressed short vowels. A sequence of stressed short vowel plus [i] and voiced dental does appear however (albeit infrequently), and is obligatory in the sense that it cannot be interchanged for a postalveolar, e.g. [tʰᵊʃdæ] “absurd”. It is seemingly more common with certain vowels, notably /ɵl/, and it occurs exclusively in non-native names and loans (but incl. acclimatized ones). The pattern is apparently productive: a novel foreign word or name with orthographical vowel plus rhotic plus dental encountered in writing will typically get a pronunciation with a short vowel and rhotic-dental sequence. In native words these short vowels rather seem to have lengthened, and some loanwords exhibit free variation of long vowel plus postalveolar and short vowel plus [i] + dental, such as [ˈɡəðɑː] alongside [ˈɡədʒ] “guard (collective)” (cf. the description below of Swedish quantity). Short vowel plus voiced postalveolar in unstressed position is obligatory in words such as [ʊd̠jˈnɔː] “ordinary” but is variable in others (see below).

2) [i] plus dental occurs when morpheme-final /t/ is the single final in a syllable with a stressed short vowel immediately across a boundary from the dental. Unlike other contexts where final /t/ contacts initial dental, there is an obligatorily distinct [ɾ], and any following dentals usually do not assimilate, though according to Elert (1997) this may occur though [ɾ] is still present. The usual outcome can be exemplified by the compound [ˈbʊʃkʊg] “coniferous forest” and the derivation [ˈbʊʃˌnʊt̠] “drilling”.

3) There are lexical cases where [ɾ] plus dental is obligatory and where it could apparently contrast with postalveolars (though minimal pairs are rare), mostly non-native names and fairly recent loans. Evidently, upon loaning, new loan words (presumably with a distinct rhotic in the source language) will be rendered with a sequence of [ɾ] plus dental and not a postalveolar (arguably because speakers do not recognise the rhotic as an actual component of postalveolars); this is productive, and is stable also in acclimatized words, while at the same time these loans are readily adapted in the pronunciation of the vowels according to Swedish spelling (suggesting perhaps that rather than approximating the rhotic in the source, instead an ⟨r⟩ in the spelling of an unfamiliar word will prompt a distinct [ɾ] in pronunciation).

4) Some words seem to have postalveolar or sequence of [ɾ] + dental in free variation (possibly patterning continuously like in /t#s/), such as unstressed [pʰæsɛpˈxʊn] or [pʰæʃɛʃˈxʊn] “perception”, stressed [pʰæsɛʃɪvʊl] or [pʰæʃɛʃɪvʊl] “Percival”, [ˈʃɔʃɪ] or [ʃɔːʃɪ] “yurt”; there are a few words with this variation, whereas root-internally most words can never take a dental sequence but have obligatory postalveolar, such as [pʰæʃɛʃɹˈnɔl] “personnel” and [pʰæʃɛʃʰɹ] “Persian (noun)”. The difference seems to be lexical; again, variable pattern words are names or, ultimately, borrowings.

---

14 Notable exceptions being the honorific herrn, a contraction of herren, and abbreviation imorn for i morgen. It may also appear with vowel shortening in an unstressed syllable, such as the place-name Hornstull.
15 For the inexistence of /t/ plus /t/ + dental, note that in borrowings ⟨æ⟩ is read as /ɵl/ and ⟨œ⟩ as /ʊl/.
16 Cf. e.g. the variations [nɔː]-[nud], both “north”, and the historical development of e.g. Sw. svärd with long vowel vs. Norw. sverd which has short vowel and rhotic plus dental sequence.
17 According to Danell (1937), this word could once also have a short vowel plus postalveolar, [ˈɡədʒɑː].
18 There is an exception in obligatory [ˈʃæʃɪkap] herrskap, lexicalized from herr and suffix -skap.
19 Garlén (1988), unlike Elert (1997), states no supradentalization occurs, but here presumably he primarily refers to the absence of “r-deletion”. Note also that [ʃ] hardly occurs in the infrequent genitive indefinite form such as dörr-s (whereas with a following relative pronoun dörr som there may be a more posterior articulation).
20 For example, cf. near-minimal [kʰʊʃ] “course” vs. [kʰʊʃk] “Kursk (Russian city)” or [ˈʊsʊl] “Ursula”.
Thus, all examples so far except those under #2 above seem to be loans or “foreign” names. This need not be a problem, but it references the issue of what makes these words different (and matters of rule synchronicity and productivity), a partial explanation of which was suggested in #3. Also, it will be seen that this is not exclusively the case but that this pattern may also occur in native but unfamiliar names (cf. 6.4.2). At any rate, most of these loans are old, common and fully assimilated.

2.1.5 The traditional appeal to geminates

The traditional solution to the cases of internal rhotic plus dental (or rather, to some of them) is that a phonologically geminate /r/ – that is, an underlying lengthened or doubled phoneme, /rː/ or /rr/ – does not trigger the postalveolarization process, i.e. it does not cause assimilation, or at least does not undergo ‘r-deletion’. To understand this use of geminates it is necessary to look into the traditional account of stress and phonological length/quantity in Swedish.

In addition to mostly intonational cues, stress in Swedish implies an increase in the duration of the stressed syllable. In traditional parlance, this surfaces complementarily either as a long consonant or a cluster (which is taken to correspond to the former), or as a long vowel. Long vowel plus multiple consonants hardly occurs root-internally but may come about from suffixation (unless this has been lexicalized, in which case the vowel has often shortened historically; many such suffixed forms also fluctuate between long and short vowel).

Phonetically there are two durational patterns (cf. Elert 1964), one in which a single consonant in the coda (at least for sonorants in isolated words) takes up about three fourths of the duration of the rime (short vowel and long consonant), and one where it lasts about half the rime (this is often called “long vowel”). This can be exemplified by the minimal pair [siːl] “sieve” (with “long” vowel, or rather an offglide) vs. [si:ˌl] “herring” (with short vowel). The quantity contrast may tend to be reduced and is perhaps not as important in continuous speech as quality differences between the “long” and “short” vowels. Arguably, length is a property of the stressed syllable, and the phonetic contrast between vowels here is a combination of traits including both relative length and vowel quality.

Phonologically, this has been treated in reductionistic ways. Reflecting the structuralist desire to reduce the number of units in the system (cf. above), analyses have been dominant that do not to give the two sets of vowels phonological status of their own, despite their quantitative and qualitative differences and concomitant semantic distinction. Rather, the vowels are seen as long and short allophone pairs conditioned by an overarching length contrast. This contrast is usually taken as segmentally triggered by way of geminate consonant segments – in representation, the consonant segment is simply doubled (as it happens, much like in Swedish orthography), thus giving the contrast as /siːl/ vs. /sil/. The fundamental conceptual assumption of segments is axiomatic here, and production then presumably articulates two abstract segments in one phone.

Another structuralist solution is positing a kind of “length segment”, which perhaps would be considered phonemic (perhaps a “chroneme” or “prosodeme”) or perhaps would work somehow on the same level as a syllable boundary. Its existence or position before or after the consonant segment (V:C vs. VC:) would determine which preceding segment gets the length, thus the contrast would be represented as /siː(ː)l/ vs. /sil/. The segmental status of these length

---

21 A transcription [sɪi̯l] does not mean there are four segments equivalent to a proposed *[sɪll]; in fact, the whole stretch [iː] in the former has more or less the same duration as the one [i] of the same word.

22 It can be noted that there is one short vowel less than there are long ones, with the short counterpart of both /eː/ and /eːl/ being /ɛ/. While long vowels are only present in stressed syllables since length pertains to stress, the quality exhibited by the vowels in long positions can also appear in unstressed position (though without the glide), such as e.g. [mʉˈsɪk] “music” (perhaps in free variation with [ʊ]). On the other hand, in e.g. Finland Swedish, the contrast is more consistently quantitative (Kuronen & Leinonen 2000).
units is unclear, much like the status of such other more or less abstract concepts as boundaries, which it seems that the derivational system can call upon basically at will.

The concept of geminates is analogous to the fact that short vowels are otherwise followed by two or more different segments while root-internal long vowels are not. Correspondingly, where the short vowel is followed by only one distinct consonant cue, this is taken to represent an equivalent string of two identical segments. This might well be equivalent to simply saying that the codas has a certain (minimal) duration for the sake of marking stress (perhaps rhythmically). Geminates may seem diachronically relevant, such as in [vɪtː] “white” vs. [vɪtː] “white (neuter)”, orthographically vit-vitt, where the latter has the neuter suffix -t, which is then argued to be a synchronic gemination. By extension, roots with short vowels followed by a single consonant would then have geminates lexically.

This means that e.g. [ˈbɹiˌskʊːɡ] is to be composed of the elements /barr/ with a geminate plus /skʊːɡ/, and the postalveolarization process is then taken to affect only /rs/ and not /frs/. Whether there is any difference between having a geminate consonant and a short vowel as the triggering condition is largely an open question, but this device is used to explain the cases like [bɹiˈsaʊd] (a geminate here would otherwise be unwarranted as the cluster performs the same function) where an extra phonological segment /t/ is added in the underlying form, giving a phonological representation /abˈsʊrd/. This solution preserves the preconceived derivation of phonetic singular postalveolars from two underlying segments, as the phonetic two-segment [t] plus dental is derived instead from three underlying segments.

What the [i] plus dental pattern in the case of [ˈbɹiˌskʊːɡ] may in fact amount to is a phonotactic constraint against open syllables with short vowels. In an outcome without [i], the sibilant would still be counted to the following syllable because of the morpheme boundary (of which this is also a cue) and so would have turned out a stressed, short, open syllable. This would favour speaker retention of [i] diachronically. One could ask why [ʃ] will count to the next syllable in cases like [ˈbɹiˌʃʊl] “bar stool”, being by some token as much [i] as [s]. Arguably, the boundary is more likely placed to retain a lexical initial in its root syllable than a final, though possibly [ʃ] is also closer to [s] perceptually (cf. also 6.4.2 on syllabification).

Clearly there is some phonological difference between the sequence /seːrd/ “can you see?” and /abˈsʊrd/ that makes the former come out as [ɪd~d] while the latter always as [ɪd], even though it might seem paradoxical that it is the lack of a boundary in the latter case which prevents a postalveolar from occurring. At any rate, the relevance and status of geminates, especially as used in the context of the postalveolars, is only as a hypothetically possible account of an abstracted distribution – and there are distributional issues as well.

Postalveolars following long vowels root-internally (which is the norm for voiced postalveolars) seem to be simple segments as clusters would produce short vowels (as noted e.g. by Braunmüller 1998:30-31). Attempting to accommodate for this through rule ordering of postalveolarization and lengthening is complicated by the fact that voiceless postalveolars can appear after both long and short vowels, as in near-minimal [fʌt] “speed” vs. [fɔtː] “fort”. A postalveolar following a short vowel would then apparently need a phonological form with a geminate dental (e.g. /fɔrtt/), if vowel lengthening is not to apply after /t/ is lost.

It is unconvincing that the variation in a word such as perception is best described by sporadically adding a third underlying segment, i.e. /per(r)sepˈxʊn/. Likewise, alternating the

---

23 Old Swedish had contrasting vowel and consonant length, with four contrastive duration patterns, so this was something like [xwiːt] vs. [xwiːtː], whereupon long vowels regularly shortened before long/multiple consonants.

24 This can be compared to the structuralist argument of /h/ being the sequence /ng/ because it is always preceded by a short vowel – this incidentally is similar to the treatment of postalveolars, as two-segment [ŋɡ] must then be derived from three-segment /ŋɡɡ/ – and again, a diachronic process where historical prevalence determines appropriateness is taken to be a synchronic constraint.

25 Some words exhibit free variation of long or short vowel, e.g. [pʰʊt] or [pʰɔtː] “gate”.

---

13
position of the syllable boundary could be utilized for describing free variation in quantity such as [ˈøglɛ]-[ˈœglɛ] “loop (figure)” as /øg, gla/–/øeg.la/, but that would still leave it unexplained, as would adding some kind of boundary, i.e. /øg(#)gl/ (which would have made it comparable to the boundary effects in /t̥rs/; cf. also the discussion of junctures in 6.2.3). Though the claim here is that /persepˈxn/ might indeed exhibit similar [ɪsʃʃ] variation as in compounds, inserting semantic boundaries inside root morphemes is not an ideal way to represent mental processes in the speaker. The traditional analysis becomes one of adding hypothetical segments at will and manipulating where syllable boundaries (or other increasingly specious boundaries) are to be drawn, rather than what consonants follow what vowels. The alternation in [tʰær] “dry” vs. [tʰær] “dry (neuter)” (with the suffix -t) is also interesting – here apparently there is some unexplained reduction of the geminate.

Finally, there are nativized loans with “long” vowel and sequence of [ɨ] plus dental root-internally, such as purser [ˈpʰɛsə] “flight attendant”, pierca [ˈpʰɛsə] “to pierce (body ornamentation)” and its derivation piercing [ˈpʰɛsɪŋ] “piercing” (there is variation in these, perhaps such as seen in perception). Here, then, at any rate, /t/ cannot be a geminate, as that would necessarily give a short vowel. (The paucity of examples is not surprising considering how diachronically recent this pattern is.) Speakers might also tend to pronounce unfamiliar words (including native place names) with <rs> as [as] even after a long vowel (cf. 6.4.2).

### 2.1.6 Systemic considerations of the postalveolar sibilant

Another issue (mentioned in 2.1.1), which will only be broached here, is how the postalveolar sibilant interacts with other fricatives: the phone [ʃ] occurs in non-initial positions for /ʃ/ <sj> in the standard variety (initially too in some varieties); also, there are difficulties delineating this sound from that of /ʃʃ/ <ʃʃ> which is slightly more front at the least. These concerns are particularly relevant as the sibilant is much more salient and more consistently distinguished from the dental than the other postalveolars, perhaps suggesting phonological implications.

/ʃ/ presents difficulties in itself, as the sounds encompassed are numerous and wide-ranging (for a phonetic survey, see Lindblad 1980); it has often been notated as a sibilant phone [ʃ]. The problem of assigning fundamental articulatory features to phonological units is evident here, and because /ʃ/ is difficult to fit nicely into an inventory, there have even been attempts at reducing all instances – of [ʃ] and [x] etc., including initially – to the underlying sequence /rs/. Again, such analysis and reasoning is conditioned by biuniqueness.

### 2.2 Previous treatment in the literature

This section presents selected accounts in the Swedish tradition that have defined and perpetuated the traditional model. The analysis of postalveolars in all positions as being the source of the syllable boundary could be utilized for describing free variation in quantity such as [ˈøglɛ]-[ˈœglɛ] “loop (figure)” as /øg, gla/–/øeg.la/, but that would still leave it unexplained, as would adding some kind of boundary, i.e. /øg(#)gl/ (which would have made it comparable to the boundary effects in /t̥rs/; cf. also the discussion of junctures in 6.2.3). Though the claim here is that /persepˈxn/ might indeed exhibit similar [ɪsʃʃ] variation as in compounds, inserting semantic boundaries inside root morphemes is not an ideal way to represent mental processes in the speaker. The traditional analysis becomes one of adding hypothetical segments at will and manipulating where syllable boundaries (or other increasingly specious boundaries) are to be drawn, rather than what consonants follow what vowels. The alternation in [tʰær] “dry” vs. [tʰær] “dry (neuter)” (with the suffix -t) is also interesting – here apparently there is some unexplained reduction of the geminate.

Finally, there are nativized loans with “long” vowel and sequence of [ɨ] plus dental root-internally, such as purser [ˈpʰɛsə] “flight attendant”, pierca [ˈpʰɛsə] “to pierce (body ornamentation)” and its derivation piercing [ˈpʰɛsɪŋ] “piercing” (there is variation in these, perhaps such as seen in perception). Here, then, at any rate, /t/ cannot be a geminate, as that would necessarily give a short vowel. (The paucity of examples is not surprising considering how diachronically recent this pattern is.) Speakers might also tend to pronounce unfamiliar words (including native place names) with <rs> as [as] even after a long vowel (cf. 6.4.2).

---

26 Some speakers might tend to exchange non-posttonic [ʃ] for [x], pronouncing e.g. personal as [pʰɛʃnəl] rather than [pʰɛʃna] nil]. The reverse is true for some speakers in some positions, with consistent [ˈkʰɛʃna] for /ˈkær/ “maybe” (evident by the misspelling karne for kanske). [ʃ] may perhaps also be more common for instances of /ʃ/ following /l/ (e.g. version). Earlier hirsch is now hirs; hutsch now has alternative spelling harts; “spontaneous” postalveolarization is evidenced in misspelling rd for d in e.g. utfordra, ordla, invardera (perhaps triggered by adjacent [ɾ] and [l] in the first two, and cf. noun invasion with /x/ in the latter).

27 Notably Elert (1957:46, 1964:15 in passing, 1970:75-77). Opposition to this was based only in phonotactic concerns of the unexpressed underlying initial /rsl/, e.g. Witting (1959:105, 117, 130-131) and Sigurd (1965:61-62), while Fant (1973:145) wavers. The generativist account of Linell et al. (1971:99-103) takes initial [x] to be derived from at least /skl/, /sj/ and /skj/; based exclusively on the orthography.
products of assimilation”. The fact that different national traditions of phonology promote different analyses for what is really the “same” data is remarkable. There is no reason for making the phonological analysis of Swedish any different from that of Norwegian.

In the Swedish literature, the argumentation usually revolves around the fact that postalveolarization can occur across word boundaries (disregarding the variation) where two phonological segments “must” be assumed — this is usually taken rather intuitively to mean that this must be the same identical process that occurs in word-internal postalveolars and so two underlying segments “should” also be the proper analysis there, thus implicitly appealing to biuniqueness. Sometimes biuniqueness is brought up more explicitly, as well as the issue of alternations in allomorphs, and the geminate argument for explaining internal [ɾ] plus dental.

If mention is made of the variability at boundaries, this is always treated as the rule deriving postalveolars from sequences being “non-obligatory” and thus simply either active or not — a dichotomy of postalveolar and dental sequence. This is usually held to occur “between words”, as opposed to internally, and thus based completely in morphology, without reference to phonetic data, which puts the demarcation between the two categories somewhat differently.

2.2.1 Eliasson (1986)
An essential traditional account is Eliasson (1986), who treats postalveolars as “secondary” and “phonetic” — a ‘sandhi’ effect and a phonetic “modification” of the dentals (277). He assumes them to be in all positions the result of one and the same process of derivation, which he calls “a typical contact assimilation” to /ɾ/, and “extremely regular and systematic” (279).

In a structural account such as his, it is assumed that an analysis should explain an articulatory process in terms of rules deriving forms from abstracts units. He defines the process as a recursive rule for assimilation and a rule for removing the /ɾ/ (279) and argues at length over whether ‘r-deletion’ should be a separate rule (283-284). He further presupposes geminates saying they will not trigger postalveolarization (280), though he acknowledges exceptions to postalveolarization in “a small number of words of usually foreign origin” (281) and does not explicitly accredit geminates for preventing postalveolars is these cases.

He does go on to note there is variation in applying or not applying the process across boundaries as opposed to word-externally, and suggests a few factors that may influence this (quoted in 6.4.3). He does not, however, seem to place any significance in the differing behaviour across boundaries, and still insists it is one and the same process everywhere. There seems to be an implicit supposition that processes may behave “differently” at boundaries (cf. 6.2.3) — perhaps a motivation for adopting a specialized term ‘sandhi’ in the first place — since effects between words in an utterance seem to fall outside the traditional scope of description (which is primarily words in isolation), as these effects are highly variable and seemingly unpredictable. This, perhaps, underlies the claim that some parameters controlling the variation “lie outside the linguistic system in a narrow sense” (282).

The reason for wanting the processes to be identical is evident in his explicit embracing of biuniqueness, or “bidirectionality”, saying (283-284) that the postalveolarization process “characteristically operates in a two-way or bidirectional manner”:

[N]ot only does it generate the surface supradentals in a unique fashion, but conversely their underlying representations can be unambiguously recovered or retrieved by means of the rule on the basis of phonetic information alone (cf. Eliasson 1981). (284)

Thus a phonetic postalveolar would always be linearly decoded into its ‘correct’ underliers without unnecessary alternatives. He obviously considers unambiguous phonetic analysability

---

28 Noting that the Swedish phonological tradition maintains two rules, as opposed to a Norwegian tradition favouring a single rule; he gives a composite rule that includes recursion in an endnote (295).

29 He apparently treats [ɾɔt̠ːtɔr] “dry (neuter)” as containing a geminate, saying “[o]nly marginally does the assimilation come into force after /ɾt/”, without further explanation. Cf. 2.1.5.
(i.e., biuniqueness) to be a considerable advantage (to perception, presumably) and as such an asset to phonological analysis.

In support of underlying sequences internally he seems to call upon some phonetic evidence that they might actually surface internally or across morpheme boundaries as phonetically “separated” sequences of the presumed component underliers:

- [P]ostalveolarization is in ordinary speech virtually compulsory within native and nativized morphemes [...] By and large the same holds true across inflectional and often also derivational boundaries [...] Tone and stress patterns may [,] favor a separation [of r plus dental], especially in true and formal compounds. (282)

But it should be noted that a sequence of [ɹ] plus dental word-internally (except in non-lexicalized compounds, as seen in the present study) would be a conspicuous archaisms or extreme spelling pronunciation and cannot be considered normal speech.

He notes that “Norwegians have been much more inclined than Swedes to grant them partial phonemic status” (289) and gives a disapproving summary of that position:

Borgstrøm (1938: 254-255, 258-259) analyzed them as (a) unitary phonemes in isolated morphemes, (b) unitary phonemes or, alternatively, phonemic clusters across inflectional boundaries, and (c) phonemic clusters across derivational and word boundaries.[.] Such mixed interpretations run into serious problems as regards free variation, distribution, and morphophonemics and miss the almost stunning regularity and productivity of postalveolarization as a living assimilatory process. (289)

He does not specify what these “serious problems” are, nor does the characterization of the process as showing “stunning regularity” in its application or his assumption of the postalveolar phenomena everywhere being identical fit very well with the data.31 Rather, it is based on the prior assumption (which seems to arise when formulating the process as a rule) that there is one and the same productive process throughout and that such an analysis is independently advantageous – indeed, having all instances being the result of a “living assimilatory” process comes across as being a goal in itself.

He dwells extensively on rule interaction (287-288) and draws up a curious argument for positing two separate rules for deriving vowel and consonant length, respectively before and after the postalveolarization and ‘r-deletion’ rules.32 He only seems to acknowledge a long vowel before a postalveolar as a result of a morpheme boundary between /r/ and the dental, where vowel lengthening would have to occur before the boundary is ‘lost’.33

Concerning himself especially with the separateness of ‘r-deletion’, there seems to be some confusion of synchrony and diachrony: he draws upon previous descriptions (283) to say that the rhotic historically disappears some time after having triggered postalveolarization and that this is an argument for having two rules, implying that diachronic facts somehow model the synchronous process (this is in line with the structuralist tendency noted above to see diachronic developments as synchronous derivations from underlying forms that equal earlier historical stages). In a similar fashion, the existence of a sequence [ɹ] plus postalveolar in one variety is taken by him to imply a synchronous process of ‘r-deletion’ in another variety.

---

30 He thinks dialects with initial [ʃ] present a problem, as he does not (284) accept initial [ʃ] to be /rs/ (as in Elert 1970, cf. 2.1.6) on phonotactic grounds (290); his fix is saying ‘r-deletion’ only occurs postvocically (284).

31 In an endnote (295-296), he quotes Sapir (1938) at length regarding forming phonemic categories from “absorption products”, but concludes that “[q]uite obviously, the Scandinavian supradentals have not yet reached the stage in their historical development described by Sapir”.

32 He apparently assumes that a stressed syllable must be given length either in the vowel or in the consonant through derivation. Consonant lengthening could not apply before ‘r-deletion’ presumably because it would produce the equivalent of /rl/, and he claims vowel lengthening would have to occur before ‘r-deletion’ to cover cases of a long vowel followed by a postalveolar when /rl/ is followed by a dental across a morpheme boundary.

33 He gives an analysis by Hellberg (1974:166) where the /rl/ segment is first assimilated to the following dental in manner, then the dental assimilates to the postalveolar place, and then the resulting geminate is simplified. In some way this arguably comes closer to describing what actually happens in articulatory terms. Eliasson argues against this as it would be unable to account for historical stages and varieties with persisting rhotic before dentals, curiously calling for “a generalized synchronic description of all supradentalizing dialects” (288). His comment that “[o]bviously, this proposal is formally very elegant” hints at the motivation for the analysis.
In addressing the diachronic origin of the process he tries to pinpoint an exact phonetic trigger, though without reference to actual articulatory mechanisms, rather exclusively using distributional criteria, with the concern being the basic quality of the rhotic. He apparently conceives of a diachronic explanation as the act of narrowing down the unique (preferably synchronic) phonetic cue that triggers a phenomenon universally. Here he embraces a simplistic and rather deterministic view of the problem of actuation, which is not necessarily as clear-cut as having the same minimal phonetic triggers apply everywhere (even if such could be isolated) to produce a consistent result universally across languages (perhaps influenced by the idea of a limited number of universal features). Diachrony is outside the scope of the present study, however, and his suggestions will not be discussed further here.

2.2.2 Elert (1957, 1964, 1970)

In a series of classic structuralist accounts of Swedish phonology, Elert (1957:45-46; 1964:13, 15; 1970) advocates the sequence analysis. He all but disregards the variation in /r#s/ (most pronounced in Elert 1957), but notes the sequence pronunciation occurring across boundaries in careful speech as an argument for underlying sequence structure (implicit biuniqueness), and seems also to imply that this can happen internally too (1970:72-74). He makes a curious note on his “phonematic” analysis as opposed to what he calls the “traditional” analysis, stressing that what in fact amounts to the same argument is actually original:

The argument here lies, though unexpressed, behind the traditional description of Swedish too. [...] But the phonematic analysis here is pursued without reference to other criteria than such that exist in the language state that is to be described. No reference to the phonetic realization in older stages of the language (that manifest themselves in the orthography) or other standard language variants has been required. (74)

Here he criticizes his predecessors of the methodological error of being based in orthography or diachrony, but arguably these continue to be the very reasons for making the sequence analysis in the first place. Also, he does himself adhere to the idea of an underlying phonological structure common to all varieties in saying that “despite the differences in phonetic realization, there is structural unity [among the variants]” (75). This point was also seen in Eliasson (1986) above, and is explicit in Sigurd (1965:19, following Elert 1964, 1957) as “minimizing the (phonemic) differences between varieties”.

Elert (1970) is notable for presenting the (otherwise implicit) argument that having underlying unitary postalveolars – since per implicit biuniqueness this implies it is necessarily the case externally as well as internally – produces an unwarranted number of allomorphs:

In morphophonematic terms this means that [...] morphemes ending in [-r] have alternating forms without this final [-r] and that this is the case when the following morpheme starts with a retroflex. One will also find that all morphemes beginning with a retroflex consonant other than [r] also have an alternating form that begins in a retroflex corresponding to the dental. [...]

[Sequence analysis] simplifies to a very considerable degree the morphological description. One would otherwise have to include a number of morpheme variants with and without final /r/ [...] and additional others with an alternation between dental and retroflex initial consonant (72-74)

As described above, this either-or argumentation arises from assumptions of biuniqueness and notions of strict hierarchical derivation of allomorphs.

He notes (75) that the geminate /r/ is not lost though he does not explicitly make this an argument. He claims, however, that dentals following geminates are still replaced by postalveolars – citing Witting (1959:116), and in contrast to most descriptions – and that in rapid speech, [ɹ] may be dropped here too (but it is unclear if herrskap is his only example).

2.2.3 Witting (1959)

Witting (1959:115-118) also embraces the sequence analysis, saying “sandhi products are functionally to be dissolved into their corresponding unassimilated entities”. He interestingly characterizes postalveolarization as “assimilation working for greater ease of articulation in speakers but conflicting with communicative needs of listeners”. The purpose of analysis is
clear from his remark that the “phonemic dissolution” of the postalveolars “reduces the preliminary phoneme inventory [sic] by no less than five units”. He touches on allomorphs, wherein “one assimilated item may be matched as a mere variant to another not assimilated”. He obliquely references biuniqueness in saying that “a general reduction is a scientifically natural corollary of the reduction brought about in an outer sandhi”.

In contrast to most accounts however, Witting highlights there is variation in /r#s/ and touches on the variability at boundaries as a product of intelligibility concerns. He says postalveolarization is not a rule (here he references Elert 1957:45 to the contrary) but “[a]t most, we are justified in speaking of a strong tendency”.34

2.2.4 Malmberg (1969a, 1969b, 1968)
Malmberg (1969a:166-168) concludes that sequence analysis is “the only reasonable one, as the sounds in question appear at word or morpheme junctions [...] and therefore should be the correct one in other cases too”, thus assuming implicit biuniqueness. Particular to his account is bringing up the linearity notion that postalveolars can be broken down into a specific order:

A prerequisite for being able to interpret a sound complex as a sequence of two phonemes is that the elements concerned occur in order. [...] It is completely clear that the rs sound in fors consists of /tʃ/ + /s/ and not of /s/ + /tʃ/. In this relationship lies the difference between a phoneme and distinctive feature. (167)

He fails to explain how that order of underliers is evident from simple [ʃ] alone (not resorting to diachrony or biuniqueness-informed conclusions of the already presumed distribution), especially as he goes on to explicitly contrast this with [y] and [b] which he claims cannot be analysed as an ordered sequence of respectively [i] followed by lip rounding and [p] followed by voicing – unlike [ʃ], apparently, these segments are for some reason to be regarded as integral.35 (Malmberg (1969b:47-48) makes the same claim that a palatalized consonant can somehow be broken down into an ordered sequence, again as opposed to a rounded vowel.)

Malmberg (1969b, 1968) indirectly touches on the allomorph argument that necessarily results from assuming biuniqueness, saying phonemic postalveolars “would in an unnecessary way strongly complicate the phonematic and especially morphophonematic description”.

2.2.5 Other accounts
Linell et al. (1971) is set in Generative Phonology, which they take care (e.g. 95-96, 103) to distinguish from the “taxonomic-structuralist” approach – much in the same way as Elert (1970) in turn distinguishes structuralist from traditional. This generative analysis is basically identical though, with some additional abstract reduction mimicking diachronic processes – seemingly, even further resolved in accounting for alternations as rule-based without having underlying forms changing (they seem to want to dispose of homophones, perhaps also a form of biuniqueness) – resulting in phonological forms that appear strikingly similar to the orthography. They arrive at the same conclusion for postalveolars, for the same reasons:

In generative phonology, the derivation [of supradentals] from /r/ plus /t, d, n, s, l/ is fairly obvious; supradentalization indeed does not occur only within a morpheme [...] but also over morpheme and word boundary [...] (103-104)

Textbooks usually treat the analysis as rather self-evident, e.g. Garlén (1988:17-18, 85-86, 94) who describes postalveolars as “assimilations whereby the dental consonants adjust their places of articulation to the place of articulation for /r/ [...], whereupon this /r/ disappears” and concludes merely that “the supradentalization rule is not applied as strictly between the words

34 He may be making a distinction of lexicalized and non-lexicalized compounds in saying of the compound bärsort that “[bærɔʈ]” is more common than “[bærɔʂt]”, “esp. perhaps if attention is paid to the intelligibility of the intervening [sic] morphemes”.

35 Swedish /y/ is even diphthongized with a falling glide, and /u/ is [iβ] in some speakers. And Witting (1959:130) on the contrary duly considers having /s/ for /rs/ to circumvent the “phonotactic” constraint on an initial /rs/ he favours (and Malmberg 1969b:48-49 opposes) as the replacement of /x/.
of an utterance” (86). He says a postalveolar outcome “seems to be obligatory” also when “only a morpheme boundary separates them” (as opposed to a word boundary). The process is taken not to be triggered when “/t/ follows directly on a short stressed vowel – and thus is phonetically long” though it is unclear what this last statement means. He gives, however, the words in which postalveolarization does not occur as due to being loanwords rather than explicitly positing geminate underliers for these.

Presumably, most “international” descriptions – being based in the work of Swedish phonologists – adhere to the sequence analysis. For example, Ladefoged & Maddieson (1996:171) say Swedish has three phonological sibilants (counting /x/ among them) and note as a matter of course that the retroflex fricative phone is phonologically the sequence /rs/.

2.3 Gestural phonology

The segmental approach to phonological analysis was criticized briefly above: it is argued there is no support for segments in the signal (cf. e.g. Tatham & Morton 2006:11ff), and in extension it cannot represent the patterns in the data. This will be treated further in the discussion. An alternative approach, which will be used in the analysis of material from the investigation, is that which works in terms of ‘gestures’, being multiple simultaneous and independent articulatory units, as opposed to linearly one-dimensional segments. One form of this view is presented in Articulatory Phonology (cf. Browman & Goldstein 1992).

A gestural perspective, and Articulatory Phonology in particular, represents phonology in terms that are production-relevant, seeing as articulators exhibit such continuousness and partial independence as characteristic of gestures, but it commonly also makes the claim that this is how phonological form is fundamentally stored. This ties in with the controversial subject of defining the phoneme (as treated further in the discussion). A gestural approach basically precludes derivational phonological rules in segmental terms, including assimilation as the spread of properties from one segment to another, mentioned above. Interestingly, it corresponds to an extent – but for all articulatory specifications – to how intonational features are traditionally treated as “suprasegmentals”.

Phonetic information in articulation and the signal can be represented graphically as parallel tiers of different articulatory activity over time. As can be glanced from figures 2.1a-b, units are not tied to segments but overlap them, which corresponds to production specifications.

```
Nasality
Velar closure
```

Figure 2.1a. Schematic gestural score for nasal place in the English word rank.

```
Nasality
Bilabial closure
Coronal closure
```

Figure 2.1b. Schematic gestural score for the English word man with nasalized vowel.

What would be traditionally seen as an assimilation of a nasal segment to the place of the following plosive is rather one and the same place gesture spanning the gestures of nasal and

---

36 He then goes on to make the curious argument that assuming underlying sequences of phonemes would agree with the pronunciations of some varieties and with the orthography.

37 This seems to contradict ‘vowel shortening’, as in [ hæʃ] or [ hæʃː] “is heard”, next to [ˈhɛːr] “to hear”.

38 Gestures also basically correspond to the ‘prosodies’ of Prosodic Phonology.
plosive. Assimilation is not an abstract process by which a feature extends to a neighbouring segment, but a product of the articulatory dynamic of speech and simply the overlap of different gestures. Some variational processes can favourably be explained as effects of overlap, e.g. epentheses: in the gradual transition from ['biln] to ['biln] bil-en “the car” in Swedish, or ['batn] to ['btn] button in English, being vocalic or nasal-syllabic is really the relative onset of the nasal gesture rather than “insertion” or “deletion” of a schwa segment.

This is very relevant in the representation of postalveolars, especially of that phenomenon which in traditional treatments is referred to as “recursion” of the rule, but which is really just one articulatory specification maintained over a sequence of other specifications, without having to be redefined for every new simultaneous articulation. This is illustrated in figure 2.2.

![Figure 2.2. Schematic gestural score for postalveolars in the Swedish word bärnsten “amber”](image)

Articulatory Phonology as presented by Browman & Goldstein (1992) make some particular claims about the nature of gestures (not necessarily present in the representations given here): Manners of articulation are largely seen as points on a scale of degree of constriction, so an approximant, a fricative and a plosive differ from one another only in degree of the same parameter. So, gestures are seen as composed of a position and a degree. They are supposed to work in absolute time, and phonological material is presumably stored as such. Conspicuously, articulatory features that are present in production can be physically “hidden” by other gestures – e.g. a labial closure will mask any tongue activity – and these production specifications, while phonologically relevant, can not be represented in segmental transcriptions as only the masking gesture is perceived at the surface.

While it seems obvious that a fricative is more forceful than an approximant and overrules it, this may well become problematic (e.g., flapped and trilled articulation). Also, while representing places of consonants in gestures is straightforward, it is less clear how to do the same for more complex articulations such as vowels and secondary articulations (and even how nasality or laterality is to fit with the definition of position and degree). While gestures may appear in absolute time they may also act in relation to one another. Gestural onsets and offsets can be phased, and what appears as segments represents particular overlaps of gestures. Also, Browman & Goldstein seem to imply that masking gestures explains reduction, but it is unfeasible to account for massive reductions, such as syllable deletions, in this way.

Because a gesture is defined not as binary but as a level of activation, and because they are treated independently in time, it is possible to represent gradience in a way not possible in segments. What to give the status of a separate gesture and what to consider a gradient within one gesture is not obvious however. Seeing phonological units as identical to temporally real articulatory events could also neglect the fact that there would still have to be mental representations linking them to areas outside production – it is not obvious that aspects of perception and storage should stem from articulatory primes.

---

39 In gradually decreasing the force of the airstream, fricatives do become approximants; this is common in the reduction of voiced fricatives as an effect of the articulatory effort in producing both friction and voicing.
3. Method

The investigation was carried out on two sets of spoken material of Standard Swedish. The larger set consists of broadcasts from the most popular Swedish daily news radio programme Ekot (also known as Dagens Eko), mostly scripted speech. This is complemented by unscripted spontaneous speech material from the Swedish Map Task Corpus (Helgason 2006).

The Ekot data comprise three midday broadcasts of 20-25 minutes each: one from November 2010 and two from November 2011. This is the main part of the investigation and includes several speakers. The broadcasts consist of a news anchor giving brief presentations of news usually followed by clips with different reporters, including interviews. Most of this data is thus scripted – that is, written text read aloud, usually in real time but not spontaneous as opposed to conversational speech – it also includes some more spontaneous speech in interviews. Only speakers subjectively judged to speak Standard Swedish were included, so speakers of regional varieties (including relatively slight accents) were excluded. Also, to ensure quality, none of the material that was provided in low-quality phone line recordings was included. The duration of the sound material used in the investigation is thus less than the total duration of the full recordings.

The SMTC data were recorded in 1997 and is a set of elicited conversations between two pairs of two people (totalling four subjects) performing first a “map task” where one subject is set to give directions to the other, each having a slightly different map, then followed by a set of exercises describing different objects in pictures. This material is in the form of ordinary conversational speech and subjects principally speak the standard language. The SMTC data set is smaller (by duration and word count) and complements the results of the Ekot data.

All speakers in the investigation were adults and spoke Standard Swedish with minor variations. Speakers were given an index number and were tagged as male or female. Speaker age was not measured but was judged to range from roughly 25 to 60 years.

Ekot was chosen in part because of its normative impact on Standard Swedish due to its spread and influence, and because of the wide availability of speech material in high-quality recordings (from the website of Sveriges Radio). The more limited SMTC data supplemented this material and was used as a check that the results were the same in spontaneous speech.

The target of the investigation was the phonetic outcomes of what in writing is represented as <rs>, or <t> + <s> with intervening dental consonants <t, d, n> (excluding <l>, cf. 2.1.2) and punctuation – this includes across word and sentence boundaries. What makes up an investigation token is a stretch of sibilant noise with variable spectral composition – i.e., a sibilant fricative which perceptually varies in pitch – with the aim to present a value of the perceptual quality of the fricative over the morphosyntactic position of the token (i.e., what degree of boundary there is between <t> and <s>). The hypothesis was that the variation ranging from a perceptually more high-pitched (brighter) noise to a more low-pitched (darker) would be correlated to morphosyntactic position, such that the greater the morphosyntactic closeness of <t> and <s> (i.e., their occurrence over a less distinctive boundary) the lower the perceptual pitch would be (and thus being closer to [ʃ]).

Tokens had to be measured according to some criterion corresponding to perceived pitch. The acoustic nature of the fricative is characterized by a high energy spectrum, with noise starting at about 2 kHz, usually with a formant-like intensity region close to this lower end of the spectrum, but spectral composition is quite variable, making objective measurement difficult. In general terms, a [ʃ] is at least perceived as having lower pitch than laminal [s], with apical [ʂ] falling somewhere in between on what appears to be a continuous scale between the two. How /ʃ/, which is perceived as having higher pitch than the postalveolar, relates to this scale is unclear.
intensity concentration (which may appear as a formant) is in some way lower in the spectrum for [j] than in [s]. The resonant frequency at the lower end of the spectrum was taken to be indicative of perceived pitch, and any other qualities of the spectrum that could affect the overall percept were ignored.

First, a manual measurement in one part of the data (the 2010 recording) was made of the lower spectral termination of intensive fricative noise impressionistically using WaveSurfer (Sjölander & Beskow 2000), taking the average over the duration of the fricative which was set approximately, and frequency values for the points in the spectrum where the noise “plateau” was judged to terminate were recorded. This measuring served as a comparison to the following automated analysis.

The analysis was automated using Praat (Boersma & Weenink 2001). The frequency of the lowest resonance in the spectrum was found by specifying the number of resonances the program will attempt to find below a specified frequency. The sampling frequency of the audio files was at 31 kHz, thus a Nyquist frequency at 15.5 kHz. Setting the program to find 5 resonances below 15500 Hz yielded the most satisfactory results. Increasing the number of formants had the effect in some cases of introducing a ghost formant below the actual lower spectral boundary of the noise.

The relevant sections of fricative noise where tagged using labels in WaveSurfer. To find the occurrences, the content was put together in orthographic transcription (partial transcriptions of the radio material were already available and were supplemented manually) in a text document which was then searched using regular expressions to find all instances of `<r+s>` in the text including intervening spaces, punctuation and other dentals. The words and their contexts were listed in the labels. The context labels were then applied to each occurrence in the file and then each consecutive instance of `<r+s>` was tagged at the appropriate text label. Determining time spans of each occurrence to be analysed was done trying to include the whole of the noise section, but end points had to be set somewhat impressionistically. All tokens were given an identification number.

The tagged sections were then read by Praat and a script was used to automatically get the frequency of the lowest resonance of each token, by taking an average of the formant points from Praat that were 25 ms into the defined section from either end point. Thus the script provided a list of tokens with values of the frequency in kHz; this was compiled into the list of labels and speaker index in a spreadsheet file. Each token can, by reference to the text labels appended in the data in the spreadsheet file, be sorted according to speaker and orthographic context, thus enabling analysis for individual speaker, gender and (simplistically) phonetic environment.

Comparison of the frequency data from the automated results with the initial manual measurements that had been made of one subset of the data showed that the manual measurements, apart from a few anomalies, yielded slightly higher results fairly consistently; this is in line with the fact that the manual measurements measured what was judged to be the lower spectral termination point of the fricative noise while the automated analysis estimated the lowest resonance.

The SMTC spontaneous speech data had already been orthographically transcribed and labelled. The SMTC stereo data files were split into the separate mono channels (one for each speaker) and resampled to 31 kHz, so as to conform to the Ekot data. Here again, instances of `<r+s>` were labelled and the end points of the sections were marked, and then the files were processed (separately for each channel) using the same script in Praat to yield another list of frequency data tokens. Tokens were further indexed for speaker, conforming to the Ekot data.

---

41 This is a source of error especially in a few shorter duration tokens, as they come quite close to being 25 ms themselves, but this might also possibly be a desired effect as it excludes transitions – though some tokens exhibit continuous change not easily accounted for in a singular measurement value (cf. examples in 5.2.4).
Each token was categorized for morphosyntactic position and different types of boundaries. The categories are, in order of increasing morphosyntactic distance: root-internal (tautomorphematic); across derivational boundary, which may be either lexicalized or transparent; across inflectional boundary; across lexical boundary (i.e., in a compound), lexicalized or transparent; across word boundary but phrase-internal, which may be semi-lexicalized (for periphrastic constructions); clause internal word boundary; utterance internal word boundary; and, non-utterance internal word boundary (e.g. end of quote).

Additional levels of distinction between transparent and lexicalized were judged to be necessary for derivational and lexical boundaries for uncertain or intermediary cases, totalling thirteen categories. There were no formalized criteria for assignment to a particular category and individual instances required subjective evaluation (as the thirteen categories are basically collapsed into two in the analysis, the effect of subjectively judged tokens is minimized). Table 3.1 gives a list of the categories with orthographic example tokens from the data.

Table 3.1. The thirteen morphosyntactic categories of boundary distinctiveness used in token categorization, arranged in order of increasing morphosyntactic distance, with examples.

<table>
<thead>
<tr>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tautomorphematic (root-internal)</td>
<td>korsar, person, annars, orsakade</td>
</tr>
<tr>
<td>2. Non-transparent derivational (lexicalized)</td>
<td>förslag, först</td>
</tr>
<tr>
<td>3. Semi-transparent derivational</td>
<td>undersökning, förskräcklig</td>
</tr>
<tr>
<td>4. Transparent derivational</td>
<td>understiger, luftsatsmyndighet</td>
</tr>
<tr>
<td>5. Inflectional</td>
<td>görs, uteckningsministerns</td>
</tr>
<tr>
<td>6. Non-transparent lexical (lexicalized)</td>
<td>Larsson</td>
</tr>
<tr>
<td>7. Semi-transparent lexical</td>
<td>ordspråket, kvarstå</td>
</tr>
<tr>
<td>8. Transparent lexical</td>
<td>ledarsidor, klätterställe</td>
</tr>
<tr>
<td>9. Semi-lexicalized phrase (e.g. periphrastic verb forms)</td>
<td>har satt, tar slut, rör sig</td>
</tr>
<tr>
<td>10. Phrase internal</td>
<td>för stor, har stöd, eller snö, mer svenska</td>
</tr>
<tr>
<td>11. Clause internal</td>
<td>Pétur sa, låter som</td>
</tr>
<tr>
<td>12. Utterance internal</td>
<td>fåglar som</td>
</tr>
<tr>
<td>13. Non-utterance internal (e.g. end of quote)</td>
<td>...när man säger nåt sånt här, sa Tysklands förbundskansler...</td>
</tr>
</tbody>
</table>

The definition of what is a transparent derivational and lexical boundary is partly semantic and partly morphological (and could also be based in phonetic data) – factors in judging compounds and derivations as transparent or lexicalized are to what extent they are reducible into lexical components or affixes that can stand as independent words and whether their meaning can be connected to these free forms or whether they are idiomatic (e.g. if a verb with a prefix can be split into its component lexical verb and a particle). In many cases it is not clear to what extent a diachronic explanation is synchronically valid, e.g. whether a certain ending is derivational or inflectional (as in en sorts ...).

Some tokens were discarded, mainly because /t/ (i.e., a postalveolar gesture), though expected and present orthographically as <r>, did not occur. Lack of /t/ is ubiquitous in the copula är, which is simply /eː/, though orthographically with the rhotic. In other cases, lack of postalveolar is a more deviant pattern. Examples of this in the data include [visəˈsifər] visär siffror, where the second sibilant is laminal despite the fact that the preceding present tense suffix -ar lexically has a final /t/, and [fəˈsɔːrəɹə] försvarare, the same thing for the prefix för-. This is a cause of a few anomalous tokens in the data.
4. Results

The radio data totalled 65 minutes. The time actually used is shorter, due to the material discarded for speakers with regional accent or technical limitations to recordings. The length of the SMTC recordings is in total 48:30 minutes. The scripted speech is faster and pauses are much less frequent than in the unscripted speech and so the number of words (and thus potential tokens) per minute is much higher in the Ekot data than the SMTC data.

In total, 436 tokens were identified in the orthographic transcriptions. 321 were from the Ekot data and 115 were from the SMTC data. Of the total number of tokens, 25 were removed having no postalveolar gesture, most notably the copula (cf. above). The final data amounts to 411 tokens of fricative frequency over category of morphological distance.

The spread of the frequencies according to each category can be seen in figure 4.1.

![Figure 4.1](image-url)

*Figure 4.1. Box plot of spread of all token occurrences over fricative noise low resonance frequency arranged in the 13 morphosyntactic categories of context (in alphabetical order)*

All tokens were labelled in the sound files so that spectrograms of the context of any token could easily be retrieved and drawn. The length of each fricative token could also be found as a result of end points being marked.

The data have been represented in a diagram in figure 4.2a-c where morphosyntactic categories have been ordered for increasing distance between elements and plotted against measured frequency of the fricatives. One dot represents one token. Figure 4.2a-c first presents separately the radio (Ekot) data (4.2a) and the spontaneous speech (SMTC) data (4.2b), and then both are taken together in total (4.2c).
Figure 4.2a. Graph of token occurrences in the scripted (radio) data by morphosyntactic categories arranged in relative order of increasing morphosyntactic distance.

Figure 4.2b. Same as 4.2a but for the unscripted (spontaneous speech) data.

Figure 4.2c. The scripted (4.2a) and unscripted material (4.2b) taken together.
While the order of the categories is fairly straightforward, their relative values on the y axis require explanation. The intermediate categories (semi-transparent derivational and lexical, and semi-lexicalized) have been given only a 0.5 point difference. Also, the difference between utterance internal and utterance external word boundary is slight and was taken not to warrant more than 0.5 point increase. The values of the categories are listed in Table 4.1.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root-internal</td>
<td>1</td>
</tr>
<tr>
<td>Non-transparent derivational</td>
<td>2</td>
</tr>
<tr>
<td>Semi-transparent derivational</td>
<td>2.5</td>
</tr>
<tr>
<td>Transparent derivational</td>
<td>3</td>
</tr>
<tr>
<td>Inflectional</td>
<td>4</td>
</tr>
<tr>
<td>Non-transparent lexical</td>
<td>5</td>
</tr>
<tr>
<td>Semi-transparent lexical</td>
<td>5.5</td>
</tr>
<tr>
<td>Transparent lexical</td>
<td>6</td>
</tr>
<tr>
<td>Semi-lexicalized</td>
<td>7</td>
</tr>
<tr>
<td>Phrase-internal</td>
<td>7.5</td>
</tr>
<tr>
<td>Clause-internal</td>
<td>8.5</td>
</tr>
<tr>
<td>Utterance-internal</td>
<td>9</td>
</tr>
<tr>
<td>Utterance-external</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Various factors can be addressed as possible sources of variability in the data, such as speaker sex and absence or presence of rounding in the phonetic context of the sibilant. In Figure 4.3a-b the data have been separated into tokens produced by male and female speakers respectively (the total number of tokens is 148 for the former and 287 for the latter). Figure 4.4a-b shows the data by phonetic context of rounded vs. unrounded tokens (mostly according to spelling with some complementary judgment of vowel reductions).

![Figure 4.3a. Total tokens (of 4.2c) for female speakers.](image)

![Figure 4.3b. Total tokens (of 4.2c) for male speakers.](image)
5. Analysis

5.1 Implications of the data
The spread in the plots above (cf. figure 4.2c) illustrates that the data fall into two broader categories, where one has a span of ca 2.25-4 kHz and one ca 2.5-5.5 kHz, and where the divider falls between transparent lexical boundary and semi-transparent lexical boundary. Collapsing the 13 categories into these two, with word internal boundaries except transparent lexical boundaries on the one hand, and transparent lexical boundaries and word boundaries (i.e., a significant boundary) on the other, has been done as a box plot in figure 5.1.

The presentation of the data in figure 5.1 clearly illustrates a difference between two groups, one in which there is variation spanning over a larger frequency range than in the other, correlated to the presence of a significant boundary. This basically supports the hypothesis in that significant articulatory variation exists in the boundary category but does not exist in the internal category and that there is not a categorical choice of two isolated articulatory positions but that the range of articulatory variation between these two extremes is continuous, as evident from figure 4.2c.

Thus, the two groups distinguish consistent [ʃ] vs. variational [s~ʃ]. Orthographically, both [ʃ] and [s~ʃ] correspond to <r>+/<s> (potentially intersected by other dentals), but the latter
may occur across a word boundary. The second group exhibits variation on a scale from [s] to
[ʃ] and the fricative occurs morphologically immediately following a significant boundary,
that is between two distinguished morphological units (either non-lexicalized compound
elements or words). A fricative in this group may – though it is not visible from this data,
which only measures fricative quality – be preceded by a more or less distinct rhotic [ɹ], only
roughly more likely to be distinct the higher the fricative frequency. This rhotic might have
been supposed to precede the fricative [s], ostensibly across the morphological boundary
(while [ʃ] can be said to be concurrent to the boundary), but as seen below there is no
requirement that [ʃ] cannot be preceded by [ɹ], as might have been expected, and there are
effects (5.2.4) that dispute the positional attribute of the boundary.

The group of tokens that manifest only as [ʃ] shows that within a morpheme or across less
significant boundaries (between elements of lexicalized compounds, and derivational or
inflectional affixes) there is no such variation between [ʃ] and some supposed *[ɹ(s)]. Instead,
this group of tokens (matching written <rs> internally) consistently manifests as [ʃ], and are in
fact contrastive of [ɹs] (2.1.4, 6.4.2).

Importantly, lexical r#s is not simply undergoing a non-obligatory rule process to be
rendered as either [ɹs] or [ʃ] as usually assumed in the literature (to the extent variation is
considered at all, cf. 2.2), but there is continuous (gradual) variation in [ɹs~ʃ] (the r#s group)
both for quality of fricative frequency and presence of [ɹ]. There is not just either [ɹs] or [ʃ]
but apparently variation in between these extremes, and there is a continuous spread of
occurrences within the range, as apparent in figure 4.2c, where tokens exhibit a fairly even
spread on the scale. How this continuousness might work will be treated in a gestural analysis
below, where the position of each occurrence within the continuous variation can be
represented faithfully. It will also be illustrated how the presence of [ɹ] is gradual and not
directly correlated to [s] vs. [ʃ].

The same conclusions seem to hold true both for read and spontaneous speech. The patterns
show similar characteristics (compare figures 4.2a and 4.2b). In the unscripted data, however,
the lower extreme of the [ɹs~ʃ] category (r#s) – that is, the lowest frequency extent of [ʃ] – is
at 3 kHz while in the [ʃ] category this is at 2.5 kHz, in contrast with the scripted data where
these values largely coincide. Also, there is no substantial difference in the patterns for
speaker sex (figures 4.3a-b) or rounded context (figures 4.4a-b).

The internal category also exhibits some limited gradual variational range in its [ʃ] fricative,
with an upper extent of 4.75 kHz (cf. figure 5.1). When excluding a small number of tokens
that can be accounted for as having reduced or missing postalveolar gesture (cf. 3. above),
that extent is rather at 3.75 kHz. This range represents negligible physical and individual
variation present within any articulation target, well within the scope of the description [ʃ].

5.2 Gestural analysis of postalveolars in the data
As already noted, postalveolars can favourably be treated under a gestural analysis. The
temporal flexibility and continuous articulatory variation can be faithfully represented within
a gestural approach in a way that is not possible with segments.

Clearly demonstrating the explanatory power of treating phonetic data as gestural in a tier-
based representation is the need to represent that the tongue position gesture of postalveolar is
continued over a stretch of other complementary articulations, seen in the abstraction in figure
2.2 above and applied to data in figure 5.3 below. Arguably, tongue position is not redefined
with each consecutive segment as it were – and does not “spread” its activation into
neighbouring segments (as presumed by “assimilation”) – but it is defined once for
postalveolariness and this is simply retained until that target is no longer defined and the
gesture ends. This is one articulatory unit set over other simultaneous units, not a reiterated
activation of the postalveolar definition for each consecutive other articulation.
Moreover, it becomes evident gesturally that it is in fact just the merger (simultaneousness) of the two gestures for postalveolar position and sibilance that result in [ʃ], not by way of “deletion” of the rhotic segment as it is usually phrased in the literature. Simply, in [ʂ] the gestures are separate, whereas in [ʃ] they are simultaneous. The postalveolar component of the rhotic is very much still present but coincides with the fricative (the manner component has been lost, or this has been superseded by frication) – only the temporal separateness of the place component of the rhotic from the manner component of the fricative is absent in [ʃ] (conceivably, some division between them has been lost diachronically). Assimilation is separate from deletion in derivational accounts, but here these are rather effects of the same process, as the postalveolar place coincides with the fricative manner.

Though some accounts make explicit that [ʃ] has qualities of both the rhotic and the dental, they still adhere to the notion that an /r/ is deleted and that [ʃ] is a modified /s/. In the gestural analysis presented here, [ʃ] is not seen simply as some modification of /s/ but instead as a product of the unification of gestures of both [ɹ] and [s].

The continuous variation between [ɹs] and [ʃ] could perhaps be explained, to some extent at least, as the gradual covering of the postalveolar gesture of the rhotic with the sibilant gesture of [s] – that is, the closer the postalveolar onset is to coincide with the fricative onset rather than precede it (offsets almost always align), the more [ʃ]-like the fricative would be. No test of this hypothesis has been made here, however.

In assuming segments and categorical rules that do not allow for gradualness, the outcome of a derivation is either a segment [ʃ] or two segments [ɹ]+[s] – in articulatory terms the two gestures should thus either coincide completely or be entirely separate. It is clear from the continuous variation seen in the phonetic data that this is not the case (as illustrated in the gestural representations below). Further, the gradual presence of a distinct rhotic does not seem to completely correlate to [ʃ]-quality in the fricative. Only in a gestural representation can this more complicated picture be adequately shown.

In the following, spectrograms and rough gestural representations of different categories of tokens from the data will be presented. Above, a categorical difference was made between internal [ʃ] and boundary [ɹ~ʃ]. In the latter category, as shown below, there can be completely aligned gestures [ʃ], separated gestures [ɹs], and a number of gradually varying patterns in between (some of which will be displayed below). The gestures given are voicing, glottal airflow (spread glottis with non-occluded airstream), and postalveolar position; they are meant to be illustrative of the issue at hand, and are thus ad hoc rather than explicitly theoretically defined.42

5.2.1 Internal [ʃ]

Figure 5.2 shows a morpheme-internal token from the data, illustrating the internal [ʃ] category. Here the frication manner gesture and the postalveolar place gesture coincide neatly (as seen in the gestural representational score) in a unified and rather delineable fricative section with little change over time. It is most naturally thought of as a postalveolar sibilant fricative, represented as [ʃ], bordered by two vowels. There is negligible if any trace of rhoticity in the vowel preceding the frication.

42 One can note briefly that the gestures presented here contrast with the gestures of Articulatory Phonology, where e.g. frication is thought of as a step on a gradual scale of constriction or gestural degree. Here, apart from illustrative gradient end points, gesture gradualness is represented merely as three schematic degrees of activation (full or partial or none).
Here, the postalveolar place and fricative manner gestures coincide with each other, as both their ends align, and can be said to be phased. In this sense, they constitute a unit, though the postalveolar gesture could extend outside the span of the fricative if a relevant coronal articulation were to directly follow.

Figure 5.3 shows postalveolar spread over multiple segments as it were, illustrating the postalveolar gesture being one single unit continued over a stretch of other articulations, with a defined start and end point at the ends of the entire sequence (cf. figure 2.2). Here, this token has been taken to belong to the internal singular \[\text{ʃ}\] category, being set in a reasonably lexicalized compound (though this might be a somewhat subjective judgement).

Nominaly, the same pattern should occur for any subsequent articulation within the stretch, but longer activation duration of the gesture might be subject to additional constraints on top of those presented below for the boundary category of [ɹ~ʃ], when a stretch of articulations is coupled with an intervening boundary (and this is seen below in figure 5.9a) – that is, extension of the gesture might be more likely to be halted at length where there is a boundary – but no examples of this have been found in the present data. Also, an internal [ʃ] followed by what would constitute a stretch but across a boundary, where the postalveolar gesture would be expected to continue, is prone to be subject to change more in line with the boundary category, thus less postalveolar activation (as in figure 5.9b below).
In figure 5.3, the postalveolar gesture of tongue tip place of articulation is initiated at the first stop closure, persists through the stop release and frication and the second stop closure and is disengaged with the second stop release (tiers for closure gestures flanking the frication are not included in the figure). Here, again, the postalveolar gesture appears to align clearly with coronal articulation gestures – i.e., there is little trace of any rhotic element in the preceding vowel resembling [ɪ], and the place of the second plosive is indeed postalveolar.

What this shows is that the tongue assumes the postalveolar position, and retains it throughout consecutive different articulatory manners. There is no redefinition of its position for each other independent articulation trait stop-fricative-stop. This could be production planning a row of coronals ahead, with a postalveolar definition that supersedes the less explicit alveolar-dental and follows through for the coronal position unless there is some definition making marked use of non-postalveolar (such as a condition that elicits marking a juncture, cf. 5.2.4) – this could be viewed as a marked cue by nature of being more of an articulatory effort. This would then be seeing postalveolars (like “assimilation” overall) ultimately as an economic matter where a gesture extends over a longer period of compatible articulations that allow for it, as it is not defined against.

Compared to figure 5.2 above, where it was said the gestures were phased and constituted a unit of sorts, there is obviously some complication arising here, where the postalveolar gesture is phased not to one but across three consecutive manners. What is one unit seen from one angle (in one tier) corresponds to three units in another. But the nature of gestures is such that this is not really a problem in a representation where gestures are the basis of analysis and appear on independent tiers (though tiers are interrelated and gestures and their end points may align to those in other tiers). The problem only surfaces if we were to try to force a segmental view on the data.

5.2.2 Boundary aligned [ʃ], disaligned [ʌʃ], and mixed [ɻʃ]

In figure 5.4 there is an instance of the [ʌʃ–ʃ] category where gestures align, giving a unitary [ʃ]. (It may be necessary to point out that what could appear to be a small break or dip at the end of the vowel before [ʃ] is not a tap but an effect of breathy voice.)

![Figure 5.4. ...länder soʃm... Male speaker.](image)

This illustrates one end of the variation [ʌʃ–ʃ] and corresponds to one of the possible outcomes in the traditional description. Here, the postalveolar and fricative gestures are aligned at the end points and coincide completely. The boundary between the words here (with a final /ɻ/ and a following initial /s/ in the lexical forms) is one between a main clause and a secondary clause.
In figure 5.5, the construction environment is mostly the same (though a different speaker) but the outcome is a clearly separated approximant [ɹ] and laminal [s] (which is the norm of initial /s/ elsewhere, with a markedly high-spectrum fricative, as opposed to the apical).

![Figure 5.5: Postalveolar gesture with onset preceding the fricative.](image)

Here, instead, the gestures are “disaligned” – that is, the end point of the postalveolar gesture is synchronized with the start of the fricative gesture. So the fricative is entirely distinct from the postalveolar – in terms of articulatory effort this pattern may be considered marked (cf. 6.4.1). A relevant question is what makes the manner of the rhotic appear here when the postalveolar gesture is separate, but not when it coincides with frication – this could be because frication overrules approximation one way or another (as in degree of constriction).

One may also consider the fact that what a postalveolar gesture means is moving the tongue towards the relevant place – approximant manner is then merely bringing the tongue at a certain distance relevant to its specified position. Thus manner could be more of a direct consequence of position, and increase in postalveolar gesture activation may well lead to what appears to be a more prominent manner (though this would not mean the rhotic is less postalveolar than the sibilant, only less strongly constricted). The rhotic exists on a gradual scale from rhoticized vowel to approximant and further perhaps to voiced frication or tapping or even trilling, where the latter are basically greater activation of the gesture, showing the gradual increase in the constriction level resulting in apparently different manners.

Manner differences as a scale of increased activation of one gesture is an idea of Articulatory Phonology which could certainly be relevant here. However, how such activation is used must also be described, and how it contrasts with cases of /s/ elsewhere where little variation is possible. Another question is whether the perceptually more high-pitched laminal [s] (in marked use to contrast postalveolarness at junctures) can also be classified as increased gesture activation compared to an apical – this would mean gestures can also have perceptual targets and not simply be articulatory (in maximizing a perceptual cue and not an articulation).

At any rate, apart from the two extremes [ɹs] and [ʃ], there are numerous kinds of gradual variation in between, which is not predicted (nor can it be) in a segmental derivational account. We can note that there seems to be semi-independent variables for rhotic presence and fricative quality, which again supports their gestural nature in existing on independent (but correlated) tiers.

In figure 5.6, there is a postalveolar gesture whose onset precedes the fricative, but continues throughout the duration of the latter – the result being a distinctive [ɹ] before the fricative while the fricative itself is also postalveolar, that is a sequence [ʃ].
5.2.3 Apical sibilant and varying rhotic as degree of gesture activation

In figure 5.7, there is a postalveolar gesture preceding the fricative and thus a distinct rhotic. But this rhotic is not followed by a truly postalveolar [ʃ], nor is the sibilant laminal as in figure 5.5. Instead, it is an apical-alveolar [ʂ], an articulation and tongue tip place of contact that falls within the gradual scope between [ʃ] and [ʂ]. Thus it has some aspects of being postalveolar, but this is not as pronounced as in [ʃ], nor is it laminal-dental, typical of /s/. This illustrates a point in the middle of the variational scale of fricative perceived height and postalveolarness. Not continuing postalveolarness but not defined in contrast as dental, this could perhaps be defined as partial postalveolar and partial dental. It is given as a partially active postalveolar gesture in the representation.

In figure 5.7, there is a tap preceding the fricative, but this is in usage largely indiscriminate from an approximant. Possibly, a tap could be considered a higher (or lower) degree of gesture activation, perhaps speaker-related.

The question here is whether the termination of the postalveolar gesture is not so much tied to the start of the fricative as it is a deactivation, meaning a gradual fading of its effect. Perhaps, tongue position is no longer actively defined as postalveolar, but its apical position remains (perhaps as an economic-inertial effect), as it has not been actively contrasted by defining postalveolarness (whether that is laminal articulation or pitch height perceptually).
This is arguably the postalveolar gesture gradually weakening, as seen perhaps in the variation of tap and approximant, and duration and intensity of approximants. This is seen in figure 5.8, where the rhotic is present merely as a faint rhotic colouring on the final part of the vowel. Here, too, the fricative is apical [ʂ], suggesting that the weaker postalveolar activation continues through the fricative. This is seen in the representation as a partially active postalveolar gesture stretching from inside vowel through the fricative. The transition between rhotic vowel colouring and approximant [ɻ] is clearly gradual.

![Figure 5.8...](image)

**Figure 5.8.** ...[uppfa]ttar som den s[tora]... Female speaker.

### 5.2.4 Boundary marking of the type [fʂ] and variations

We now turn to cases where effects show up due to boundaries, both between the rhotic and the sibilant, and when following a postalveolar and sibilant across a boundary. Arguably, in more formal speech, there may be various kinds of distinctive juncture marking at such boundaries (no estimates of register were made for the present data, but juncture marking is hypothesized to be less frequent in the spontaneous speech data).

In figure 5.9a, there is a final internal postalveolar followed by a boundary and a lexical initial sibilant /s/ in the following word. What happens here is a fully postalveolar final fricative that instantaneously changes into a laminal, marking the juncture (in the traditional derivational account, this apparently aborted application of the rule would be difficult to explain). This highlights the obligatory postalveolaranness in the internal category and the possibility of using non-postalveolarness in the variational boundary category to signal a juncture. A more extreme version of this juncture marking was a pronunciation [jeʃɛ:ʃo] görs så, with an epenthetic schwa between the final internal postalveolar and the laminal sibilant following the boundary.

In figure 5.9b, the boundary juncture follows the internal final sibilant before an initial plosive. Here, the fricative quality changes from being postalveolar to apico-alveolar as it approaches the plosive [t]. This is possibly an inertial or preparatory transition from [ʃ] to dental [t], considering the short duration, and so it is less consequential.

---

43 But cf. what was said above for figure 5.7 of postalveolarness being undefined in the fricative.
44 Gårding (1967) provides some data on variation of postalveolars over boundaries when signalling junctures. Postalveolars are less prone to be split by a juncture and are most frequently perceived to be followed by the juncture (38f), even when the postalveolar stops are aspirated, though aspiration otherwise is a clear indication of juncture position.
Voicing
Glottal flow
Postalveolar

Figure 5.9a. "[d]agars tå[mm]a"... Female speaker.

Figure 5.9b. "[en] sorts t[ack]... Female speaker.

Figures 5.9a-b showed juncture marking in the case of an initial sibilant or plosive following an internal postalveolar fricative across a boundary. In figure 5.10, however, it is shown that juncture marking of this type of deactivating postalveolarness inside a fricative, as seen above where there are final and initial fricatives across the boundary, can also be present when there is no internal postalveolar or fricative preceding the boundary, but only a final /r/ – even so, the fricative is at first postalveolar.

That is, there is a postalveolar gesture initiating in the preceding vowel as a rhotic approximant and continuing through the start of the fricative producing a postalveolar sibilant. Then, inside the fricative, the juncture is signalled as the sibilant transforms into non-postalveolar (though still apical here), despite the fact that the sibilant is supposedly only on one side of the boundary. Thus, the postalveolar gesture starts inside the vowel and continues into almost half the fricative, and then it is deactivated to mark the juncture within the duration of the fricative. Out of a phonological /r#/s/, we thus get [Ⓐf]$\theta$. This illustrates how it is the deactivation of the postalveolar gesture that is used to mark the juncture. The boundary does not fall “between segments” as expected lexically, but inside the fricative segment as it were. This pattern contradicts the positional nature of boundaries as being segment-like.
The marked junctures in figures 5.9a-b and 5.10 can be contrasted with figure 5.11 where no juncture marking is made in a sequence of internal postalveolar followed by a boundary and an initial sibilant and plosive. The postalveolar gesture starts with the fricative onset and is continued throughout the duration of the frication (including the subsequent lexical initial /s/, the whole being durationally not longer than a normal single [s]) and the stop.

Figure 5.11. ... för si[naj]... Female speaker.

6. Discussion

6.1 Summary of the results and analysis
The distribution of postalveolars and differing perspectives on their phonological treatment were brought up in section 2, with reference to the abstract nature of traditional explanations and giving some reasons for considering phonological status of internal postalveolars.\(^{45}\) As shown in the investigation (cf. 5.1), the two patterns demonstrated in variable [is-ʃ] at a boundary and obligatory [ʃ] internally are indeed phonetically different. In section 2, they were also described as functionally different, and the latter contrasting with [is]. It was also

\(^{45}\)Except in the lateral where there is variation [dl-ʃ] internally and no contrast. This variation is perhaps comparable to the [is-ʃ] variation, and perhaps illustrative of diachronic change still in progress.
argued in section 2 that traditional concepts, especially biuniqueness – the necessary mapping of a phone (in a context) to one phoneme only (if one assumes it is at all applicable to [ɹʃ]) because of [ʃ] – cannot be taken for granted.

The gestural analysis in section 5 illustrated the variation in [ɹʃ], as opposed to showing the dichotomous results of a rule, and demonstrated that description is more feasible in a gestural than in a segmental analysis. The section displayed some of what can be gained in understanding – of postalveolars and the different kinds of variation present, but also boundaries and juncture marking – from analysis of actual instances in terms of gestures.

It was argued that the different effects can be described as different temporal constellations of the same gestures. In the internal postalveolar category the two gestures are obligatorily aligned, whereas in the boundary category they are not, meaning there is variation in the onset and the internal relation of gestures. There are a number of different effects within the variation, which does not fit with traditional derivational accounts. Thus we do get e.g. [ʃ] in the gradual gestural variation, and not just either [ʃ] or [ɹʃ] as predicted by the traditional model. In addition to those two extreme points, the main types of variation found are rhotic followed by fully postalveolar fricative [ʃ], a rhotic followed by an apical fricative [ɹʃ], and a weak postalveolar (rhotic) colouring of the vowel followed by an apical fricative [V-ʃ]. It is not known whether a weak rhotic colouring followed a fully postalveolar fricative will occur.

The traditional analysis was noted in section 2.1.3 to depend on a number of factors that will be discussed in the following. It is argued a segmental and derivational nature of phonology should not be presumed, and a clear alternative has been introduced in gestures. The derivational rule for postalveolars apparently does not adequately represent the continuous variation, or all instances of <rs>, including internally, where it was argued that positing geminates is merely an ad hoc fix that is not satisfactory either. Variation as is shown to exist here cannot be accounted for in the traditional model of speech. Not only is postalveolar articulation non-obligatory at boundaries, but it is rather articulatorily variable on a continuous scale (or more than one, as regards the rhotic) that is also relying temporally on gestural coordination. This is a problem to the concept of the segmental phonological unit and the derivational model connecting phonology to the output – such variation cannot be described in terms of segments and derivations.

6.2 Traditional assumptions of the phonology of postalveolars

6.2.1 Implications of phonological status
Assigning phonological status to internal postalveolars (other than the lateral) means introducing a new set of consonants in the Swedish phonological system. Seen from a perspective which emphasizes functional contrast as the goal of phonology, positing a set of postalveolars really only means one group of articulations is used functionally in some way fundamentally distinct from others.

In the traditional analysis however – where there is an ambition of having as few underlying units “as possible” because the purpose of analysis is to “find generalizations” (cf. e.g. Lass 1984) – introducing a new postalveolar feature is then in itself a negative thing. It is often assumed there is to be synchronic symmetry in phonological systems, meaning that intersections of features (conceived as rows and columns in a table) are to be maximally filled when constructing phonological systems in order to reduce the number of features – adding an extra feature makes it more difficult to fill all junctions and so is to be avoided.

---

46 There are certainly diachronic tendencies for symmetry, seemingly pre-empted by a synchronic absolute.
‘Generalization’ is thus synonymous with reducing the number of fundamentals by deriving surface forms through rules. Giving “phonological status” to units is restrictive. But there seems to be little reflection on the phonological status of rules, on the other hand – the legitimacy of appropriating diachronic rules that actually account for change over time to pass as synchronous mental processes never seems to be called into question. Neither is the claim that phonological status is identical for production and perception.

6.2.2 Biuniqueness
There is no convincing evidence that the same phonetic outcomes cannot result from different underlying sources as it were. Biuniqueness is an untested assumption that is only motivated by the rule-based framework as contextual rules are desired to work both ways. But perception is not just production in reverse and it works without assuming biuniqueness, partly because it is not necessarily a serial process. Though not commonly explicitly adopted, biuniqueness is often (as in this case) assumed implicitly when this provides an argument (or opportunity) for explanation through rules rather than granting phonological status to units.

As shown by the investigation, what is more in this case is that there is in fact a difference here between two phenomena, assumed to be identical as this is a difference not properly recognized in traditional terms – it has been largely ignored, or it is disregarded just by nature of being at a boundary. So for biuniqueness even to apply here one would only have to consider only the [ʃ] case of [ɹʃ] which is “identical” to internal [ʃ]. But at the same time, the existence of [ɹʃ] is seen as evidence of the underlying form surfacing.

Looking at the appeal to biuniqueness in the traditional literature on the postalveolars, only Eliasson (1986) seems to bring it up explicitly. All earlier accounts only take on a more or less implicit assumption. From Eliasson’s argument (see quote in section 2.2.1) it would seem that it is in some way advantageous or even necessary for phonological processes to run unambiguously both ways – i.e., the phonological form is to be recoverable from phonetic form by backtracking the same derivational rule that produced it – but it is not clear in what way this is so. Conceivably, to state this is to be making some claim of how perception works, though this is not made explicit. Arguably, this concept is an effect of seeing phonology as primarily a matter of rules in the traditional sense – and of wanting to utilize the same set of rules both ways, as perception is seen as necessarily being the mirror image of production.

6.2.3 Variation at boundaries
Although it is usually recognized that the coalescence of rhotic and dental across word boundaries not always occurs, this seems to have had little effect on the presupposition that all postalveolarization is one and the same process, and thus identical, nor has it therefore prompted any consideration as to why postalveolars behave differently internally and across boundaries. The derivational rule cannot show how in one position it will apply only “sometimes”; accounting for the continuous variation is completely impossible. Even when accounts say something of the variation, there is little attempt to explain it, and for the sake of the argument it is simply disregarded, asserting that the two pertain to the same process.

A reason (mentioned in 2.2.1) that variation can be so easily overlooked is that traditional phonology is primarily occupied with isolated words and not continuous utterances, thus has little need for explaining why things happen as they do in the “contact” of words in real-time production. In this perspective, the fact that effects between words do not turn out as expected is not given much attention, either because the subject is never dwelled upon, or because one tacitly assumes in a self-explanatory way that whatever happens does so simply because it is between words and not in an “ordinary” setting (that is, word-internally). One posits

---

47 There is an outspoken objective to “minimize storage [and] maximize computation” (Lass 1984:25-26).
“boundaries” or “junctures” where some diffuse operation occurs, presumably on utterance
level, and as such it is not applicable to traditional phonology – so something non-
phonological is responsible for spaces between words in an uncontrollable and phonologically
irrelevant manner, evidenced in pauses and other “non-linguistic” elements. Rather,
production is very much a continuous but nonetheless phonological matter, and where
separating words is a marked feature (not necessarily segmental, cf. junctures in 5.2.4). And it
is rather a point of interest what factors really control the continuous variation in /r#s/.

The traditional analysis uses junctures as a “solution” to the different patterns. The input
definition of the derivational rule for the phonological sequence could then be formulated not
to apply (or simply apply “differently”) where there is an intervening juncture, thus appealing
to a critical difference of phonetic vs. phonological level: postalveolar for /r#s/ would then
merely be a (more shallow) phonetic effect, as opposed to internally. But then the principle
argument from implicit biuniqueness, as presented in the literature, for positing an underlying
sequence internally in the first place would be lost, as they are recognized as different.

6.2.4 Hierarchical allomorphy
Another argument encountered in the literature is that excessive derivational allomorphy
would result from phonological postalveolars, as this is commonly pictured to imply there
must be gratuitous allomorphs in \( \text{-} \) for \( \text{-s} \) and \( \text{-Ø} \) for \( \text{-r} \). Firstly, when not resorting to
biuniqueness, and as they are demonstrably not identical anyway, positing internal
phonological postalveolars does not necessitate positing them at boundaries.

Secondly, this type of reasoning is based in the assumption that affixed forms are necessarily
seen as sequential strings in hierarchical relation to the root. Phonological derivational rules
are tied to serial morphological concerns. Morphological elements are considered to be
underlyingly just the “chain” of roots and affixes in isolation, even though this is often not
evident from phonetic data. For example, the combination of Swedish root /hør/, “to hear”,
and passive suffix /s/ then must be the completely transparent /hør+s/.

The combinatorial idea is that a phonological and morphological form must be analysed as
the strict sum of discrete constituent parts consisting of phonological segments. The
underlying form should then always be constant – an invariable root with concatenative
affixing forming a linear string neatly divided into its morphological components. Any
material contradicting this is taken to be derived from it by phonological rules at the surface,
making alternation (a change conditioned only morphologically) hard to explain and
undesirable. Perception is believed to work in a similar way whereby it cannot take in
something that cannot be properly and uniquely reduced – hence biuniqueness.

In the approach pursued here, it is not assumed allomorphs work in the traditional sense of
being stored as phonemic absolutes, hierarchically on a level beneath the base form. It is
unclear what empirical support there is for seeing forms and parts of forms mentally as
inflexible compositional discretes derived from non-alternating stems and affixes – especially,
for why derivation rather than association is the more plausible mental mechanism. Forms
might favourably be regarded rather as constituting a network of connections between them
that are gradual and associative and where phonetic and semantic cues each play part: there is
similarity, but it does not have to be serially derivable as in the traditional rules.

Incidentally, too, as seen in the data, there would be more support for an alternation /rl/ to Ø
before a suffix /s/ in intra-word than in inter-word position or than an initial alternation /s/ to
/ʃ/ as it were. Again, the phonological consequences of the difference between e.g. [ˈbʊəˌʃtʊl] vs. [ˈbʊəˌʃʊl] can be radically reappraised in a gestural perspective, where the same two

---

48 As mentioned in 2.1.5, it is also not unheard of to insert spurious morphological boundaries inside roots, in
order to provide a “solution” to explain some aberrant pattern.
gestural units are present in both [ɹs] and [ʃ] (and this should be relevant for gestural treatment of fusional inflection overall).

6.2.5 Geminates

Resorting to geminates (cf. 2.1.5) did not constructively explain the data and introduced alternations of its own cases of variation. It also presumes without argument that speakers proceed from an abstract solution, in introducing the gratuitous extra /r/ in e.g. /abˈsurrɪl/, and that this is advantageous to another explanation. And geminates are contradicted by the occurrence of [ɹs] also after a long vowel. One can also ask whether there is any meaningful difference representing the distinction as /ʃl-/rsl/ or as /rsl/-/rrs/ from a gestural perspective.

The geminate argument makes the claim that speakers would “utilize” gemination in such a way as to better render loans such as *absurd*: saying in such explicit terms that speakers analyze words abstractly as containing otherwise unwarranted geminates (the justification for which is itself doubtful), just for the sake of being able to pronounce [ɹ], seems teleological and to be ascribing speakers abstract metalinguistic creativity.

Obviously there is a difference between cases of [ɹs~ʃ] variation, of obligatory alignment, and of obligatory separation of gestures. But speakers do not proceed from abstract analyses in utilizing existing patterns – it may, however, be more appropriate to see would-be gemination as analogizing. That is, geminates might possibly be held to be responsible for the disparate patterns of [ɹ] plus dental to the extent that speakers when producing this pattern could be argued to analogize to syllables that have a rhotic before a dental when the absence of this rhotic would result in a stressed short open syllable (and as that rhotic would be considered a geminate). But it is then a contributing analogy effect, not the adoption of an abstract solution.

Regarding the scarcity of some of the patterns here, it can be noted in general that patterns being absent in some distribution does not necessarily mean a synchronic constraint against them, but this may rather be the diachronic product of this being the case historically. New formations and loans can show if a pattern is productive, though speakers will always analogize to existing patterns, whether this is in terms of rules or just by association. That something does not occur need not imply a synchronic process, though absence of a pattern will be a reason in itself for seeing it as aberrant and making it disfavoured and unproductive.

6.2.6 Segments and derivations

The traditional analysis is heavily dependent on viewing phonological processes as derivational rules producing phonetic segments from phonemes – the latter being abstract segmental units of contrast that are phonetically defined. The “phonetic” material used in analysis, being both the outcome of the underlying structure and the basis for it, is usually rather far removed from the properties of actual speech signals and is in itself quite abstract. And, as previously mentioned, segments and derivations largely preclude describing variation as it was seen under a gestural treatment. Segmental derivational models can only generate strict homogenous outcomes, not patterns of systematic variation, especially temporal detail. Variation encountered is then by default treated as some visibly haphazard and irrelevant aberration from the “phonetic form” predicted by the segmental rules.

But even the outcomes that are indeed possible from the rules are necessarily considered aberrations from underlying phonological units, though these also have intrinsic phonetic specification, phonemes being abstract or mental “sounds” (cf. 6.3.3). This means that phonetic phenomena that do not appear to “contrast” with one another (they cannot be functionally exchanged but are predictably conditioned) are definitionally equal and their distinction is fundamentally seen as irrelevant. An example of such a phenomenon that can be “undone” through rules is when an articulation type is extended over concurrent articulatory specifications. In traditional terms, segments are said to be “assimilated” to one another in
some respect, or a trait having “spread” into surrounding segments – a segment takes on a feature from another segment, without altering their status as fully independent segments.

Phonemes are then taken to embody the “proper” content of the signal, in that they capture generalizations while still being “sounds” in some sense, even while being endowed with (implicit or explicit) mental reality. As such, phonemes are supposedly relevant for both storage, production and perception as units of basically the same kind as phonetic segments, only abstract and at a generalized (“contrastive”) level. Rather than phonemes being a mental collection of phonetic phenomena corresponding to the same semantic contrast, phonetic segments are derivational products (“allophones”) of mental sound segments of the same type.

Positing segments is necessary for traditional derivational mechanisms because these, along with any statements about distributional patterns, which is the goal of traditional phonology, are rendered in fundamentally segmental terms. The prior assumption of segments is then perhaps one reason for traditional phonological distinctions to be based primarily on formal distributions of segments within a segmented material. Illustrative of the traditional view is Jones’ (1950) motivation for segments, though they physically do not exist, being that “the conception of the chain of speech sounds is indispensable in all linguistic investigations”.

6.3 Alternative perspectives on phonology

6.3.1 The nature of speech and implications for phonology

The presumed invariant phonetic segments (or phonological units) cannot be matched to any consistent parts of the actual signal. The direct source of the signal is speech production working in terms of articulations. The speech apparatus cannot move from one specification to another seamlessly because articulatory movement is continuous. A fundamental property of the phonological system therefore is the fact that the signal is a continuous dynamic progression characterized by gradual transitions in the way that articulators move and interact through inherent articulatory coordination and inertia.

The signal cannot be sliced conveniently into significant portions (even if perhaps specifications could be individual targets) and temporal phonetic cues cannot be delineated linearly but overlap one another. Independent articulators imply that some phonetic specifications can be independent of others, and these specifications can then apparently overarch “segments” or terminate inside them. Production specifications, then, are arguably best represented as an array of tiers allowing for overlap of independent articulations, as in a gestural treatment. Overlap is the coordination of gestures causing effects based on concomitant articulations.

The traditional concept of “assimilation” of segments to other segments can be explained as gestures expanding over multiple other gestures. Part of the reason for this happening is that targets are not processed linearly or distinctly one-by-one. Rather, production at any one point is always shaped by previous specifications and physically tuned towards what comes next; future specifications are planned over some stretch and may exist in parallel as independent articulators continuously prepare for upcoming targets.

Assuming underlying segments necessitates placing these in relation to the articulators – this raises the question how to translate abstract segments into physical dynamics and whether this is at all necessary (cf. Tatham & Morton 2006:11ff). Segmentation cannot proceed from the signal without prior analysis of it (as in perception), and the need is not obvious of a process in production wherein segmental units have to be converted into articulatory specifications rather than being stored as those specifications (or reasonable abstractions of them) in the first place.
In terms of production, mental specifications should rather be assumed to be articulatorily relevant, as what it comes down to is instructions to the articulators – and articulatory independence, continuousness and gradualness of activation are indeed a definitional property of that system. Phonetics is traditionally seen as merely the end product of language, set in an imperfect medium – the first objective in serially processing the signal is to convert it “back” into its “fundamentals” (i.e., phonological segments) before any further linguistic analysis is done. Rather, phonetic information should be present at a much deeper level of analysis, where variable phonetic cues that are not by necessity contrastive in a strict sense are processed along with any “segmental” information.49

6.3.2 Explaining speech in the traditional model
Variation found in actual speech cannot be properly explained or even notated in the context of segments and rules. In the traditional model, it is largely ignored in favour of analysing the hyperarticulated citation form of the solitary word. This is because when phonological analysis is taken to be the generalization of distributional patterns of segments, it becomes an end in itself to do away with surface diversity as well as to reduce the number of primitive units by means of derivational rules. A homogenous segmental output simply cannot describe variation and makes it irrelevant. In the words of Lass (1984:29), “variation [...] is outside the ambit of any phoneme theory in the strict sense”. The disregard for variation also follows from the bias that variation is deviation from a prescriptivistically “correct” fixed form. But invariance is not the norm from which variation deviates.

Reductionism is also seemingly necessary for derivational processes to be universal, as they are taken to operate on the “same” segments and with the same set of mechanisms in any language. As put by Lodge (2009:11), “the search for a universal set of phonetic descriptors that can also be used for phonological analysis has clouded the issue of phonetic variability.”

It has been argued the traditional view results from alphabetical bias and the unreflective preservation of historical methods of data collection (orthographical transcription based on impressionistic listening or mere supposition), for example in Gafos & Goldstein (2012):

Work has largely proceeded from the assumption (occasionally made explicit as a hypothesis) that the observables used to build the theory are transcriptions of speech as a sequence of segmental units. Not surprisingly, therefore, the internal representation of speech that theories of phonology have traditionally postulated is likewise a sequence of segmental units [...] [...] The acoustics revealed by spectrographic analysis appeared so unlike the transcriptions of speech, and so apparently incompatible with it (Harris 1953, Hockett 1955), that it was rejected as a primary observable. It was assumed that somehow a listener must re-construct a segmental transcription when listening to speech.

The segmental view sees speech as the linear progression of strings of integral sound units resembling an alphabetic script, wherein multiple articulatory and perceptual traits are merged into preset units, and the continuous, non-localized nature of phonetic properties and articulation is ignored. In a traditional segmental view, a particular articulation will be redefined for every segment (with “assimilation” setting binary values of multiple segments, as in stretches of postalveolars), even though this entails no additional articulatory activity, whether or not there is some perceptual relevance to redefining the cue for each other articulation that occurs (as in perceiving a segment as a change).

Trying to account for the non-segmental nature of the signal has been the focus of accounts assuming invariance, holding the signal to exhibit “coarticulation”, transition effects joining segments in the signal, attempting to find mechanisms for “restoring” the continuous signal into its presumed segments (cf. Tatham & Morton 2006:22ff). But because phonology is directly related to the properties of the system, these effects cannot be considered secondary but integral: cues are inherently spread out unevenly in the signal and transitions are not a

49 Indeed, it is the properties of the speech apparatus that shape phonological systems, wherein variation is meaningful (cf. Lindblom 1990) and a prerequisite of sound change (cf. Ohala 1989).
negative side-effect but in fact a feature.\(^{50}\) As an integral property of the signal, coarticulation is actually necessary for perception, not a drawback, and allows for redundancy by leaving traces over longer stretches of the signal, which is an aid to perception.

Lack of invariance is traditionally a “problem” for phonological theory, because of the presupposition that speakers start with a string of discrete invariant objects and that the objective is for the listener to convert serially the “imperfect” speech signal back to those same mental objects. The segments in the signal have “acquired” coarticulatory effects and become blurred, and this needs to be undone. Appealing to coarticulation is seeing speech phenomena as deviant from underlying segmentality.

6.3.3 Mental concerns
Segmentation into a finite set of discrete units could be argued as necessary for rendering a large lexicon through combinatorial means – basically, it would save space in memory not having each word as a separate entity but a string of the combination of a few basic units – this would also differentiate it from “ordinary” memory. Storage would seemingly be less tasking storing abstract units rather than sound data.

But apparently much sound data (metadata) is also stored, and it has already been suggested perception is not a serial process but that phonetic cues matter at a more profound level. There is evidence to suggest that words are not treated in memory as linear strings of absolute segments first hand (cf. “groping” in Johnson 2001), meaning that storage of lexical entries is not fundamentally segmental but in some sense rather organized in terms of phonetic wholes. A more radical view is exemplar-based phonology (e.g. Johnson 2005), which sees linguistic storage as being fundamentally phonetic and equates it more with other forms of memory. Phonological units have also been shown experimentally (Houde & Jordan 2002) to be neither integral nor absolute with respect to phonetic specifications.

Abandoning segments does not necessarily mean abandoning phonological units, however, though it means there is less separation of mental units from the nature of phonetic elements. Also, gestures need perhaps not be less discrete than segments (only multi-dimensional), although they are usually taken to be temporally more absolute.

There is clearly some more or less conscious level of psychological categorization where segments have relevance (cf. Tatham & Morton 2003:13). But while segmental categorization is psychologically real to some extent, a gestural dynamic might well be how things work at a fundamental level of mental representation of production, with segmentation being a superimposed categorization of patterns. And perception is then best considered in terms of dynamic cues or perhaps diffuse checkpoints with faint resemblance to segments. Conscious browsing of the lexicon segmentally is far removed from fluent speech production.

Thus segmentation should be a secondary categorization procedure: the chunking up of the signal is then an after-the-fact mental superstructure taking place late in the process, not until after analysis. Reasonably, the signal must be analysed before it can be segmented, and this analysis is what is important; segments are not something inherent in the signal, and cannot be something which must be “retrieved” before it is analysed.

Perception uses an integrated approach of phonetic and semantic-contextual cues to receive and interpret the signal. Here, the concern is how to discern the different phonological patterns out of the hugely variable signal and how to find the cues to rely on, whereas the

\(^{50}\) A non-standard alternative having “segments” instead be momentane targets, with stretches as transitions between such points, may have some relevance as perceptual “milestones”, but also is not fully adequate: here too, as information given from articulation and used in perception is rather spread out all over the signal, it is not simply the question of linearly transiting from one state to another and where only the states matter – the phonetic effects are in the transitions. It also retains the one-dimensional view of not accommodating for specifications stretching across other specifications as in gestural units.
traditional approach for the analysis would be to try to “restore” the underlying segments. While it remains difficult to come to any conclusions of how perceptual mechanisms work to arrive at categorization (cf. Tatham & Morton 2006:13ff, Johnson 2003:70), it should have to be in terms of dynamic cues rather than segments first hand.

Phonological units should be mentally meaningful, where “the identification of sounds in a linguistic system as the same must be based on the function of the sounds” (Lodge 2009:11). Phonology cannot be simply distributional patterns of presumed segments in an abstract material but must be conceived of as the mental linking of production, perception and contrastive function. Also, though semantic function is linked to production instructions and perception patterns, phonological units are not “sounds”, because they are mental abstracts. The traditional phoneme, however, is a “sound unit” (defined in terms of a limited set of binary values) that distinguishes meaning – it is an abstract that is still a sound and is fundamentally phonetically specified. Lodge (2009:17) further suggests that “an assumption of phonetically based phonological features forces a derivational account of the relationship between the lexical forms and their realizations”.

Mental phonological specifications encompass production, perception and memory storage and retrieval (and semantic contrastivity). In this sense, it becomes meaningless to speak of phonemes as “mental sounds”. Some phonetic patterns are capable of upholding distinctions and some are not, where different contrasting patterns mark the limits of phonemes, and non-contrasting patterns pertain in some sense to these phonemes – by functional association and perceptual pattern recognition rather than the traditional abstract derivation into allophones. As seen above, this is supported by evidence that phonological units are indeed not phonetically absolute and that storage of lexical entries is not fundamentally segmental.

6.3.4 Pragmatic models of phonetic variation
Articulatory transition effects can vary in degree of overlap and are partly language-specific. In gestural terms, there can be more or less temporal alignment. A related articulatory process is “reduction” of articulations. Features of the vocal apparatus can be manipulated to achieve marked results (as in segmentally enunciated speech) where phonetic cues are used variationally to signal pragmatic aspects.

This is the focus of pragmatics-based theories of phonetic variation such as the H&H theory of Lindblom (1990) and Cognitive Phonetics of Tatham & Morton (2006). The latter stresses variation as in cognitively influence on physical aspects of production – though not cognitively initiated – in a phonologically relevant way.

H&H theory sets out with the claim that speakers habitually provide as much information as is assumed to be needed in each particular case (automatically adjusting for various factors such as environment interference and background knowledge) – providing too much information is then marked. This is possible as participants in conversations continuously construct scenarios of what the other can be expected to say. Reduction is “possible” because of the presuppositional nature of language and predictability of content based on context and construction patterns. Reduction is made when a sufficient amount of cues and contrasts for identification is judged to be met to allow simplification of articulatory targets.

Arguably, there is no obvious reason for regarding clearly articulated speech, being “more” segmental, as in being closer to some mental template – rather, clearly distinguished gestures are less spontaneous and more marked – they reflect an additional effort that is sometimes necessary (e.g., it might add redundancy in a noisy environment) but otherwise out of place. Ordinary speech falls somewhere between two extremes within a variable continuum – in

---

51 E.g., a difference can be seen between greater reduction in expected functional markers, whereas non-predictable content words etc. must be articulated more thoroughly.
terms of the H&H theory, normal speech should be an optimal choice somewhere between hypo and hyper, and not descended from hyperspeech.

6.4 Interpreting the gestural variational analysis of the postalveolars

6.4.1 Variation and gestures in the postalveolars

A pragmatics-based view of phonetic variation can be applied to the boundary postalveolar gestural variation. The postalveolar and sibilant gestures show variation in alignment as well as intensity when at a boundary, which might suggest this variation plays some pragmatic role as a device for marking a juncture (a possible feature of hyperspeech).

The articulatory and acoustic variation span allowed in the internal postalveolar is much more constrained than in the boundary postalveolar, evident in the greater span of tongue positions – it is unknown exactly how alignment of gestures relates to backness of articulation. As seen, in the internal category, the postalveolar gesture is rather coherently aligned or phased with the fricative gesture. In the boundary category, concomitant or separated gestures are ends on a variational spectrum, and complete separation is more marked.

It remains to be explained how the discreteness of the gestures is discontinued diachronically – that is, how the gestural specification changes from alignment of postalveolar end point to sibilant starting point and into either alignment of gestures to one another (in the case of internal postalveolar), or to variation of postalveolar onset and to some extent offset more or less freely relative to the sibilant onset (in the case of boundary postalveolar). What mechanism governs alignment of gestures also remains to be explained – how different constellations of gestural alignment and alignment of terminal points are differentiated – and how this is to be represented in a production specification. For this to work, phonological specifications should indeed be organized in terms of gestures at a fundamental level.

There may still be some connection of the full alignment end of the variational span to the internal unit [ʃ] where gestures obligatorily align. Likewise, the separation of gestures as [ʃs] can be seen as resulting from superimposed influence of analogical nature caused by the separate [ɹ] and [s] of the lexical constituents in other positions. There is indeed a problem to what extent to treat the phased internal category as an integral unit, and if so whether cases of alignment in the boundary category are to be taken as identical to this unit. This can be compared to the traditional question of whether [ʃ] is a phoneme, but it is increasingly clear this is an oversimplification of the issue – the difficulty of how to relate the case where variation occurs and the one where it does not cannot even be phrased traditionally because the mechanisms of variation are ignored. In gestural terms, both cases of [ʃs] and [ʃ] are consistently composed of the same two phonologically relevant units, exhibiting contrastive differences in their alignment.

At any rate, this means that there is more leeway in the boundary postalveolar for marking pragmatic cues (in this case junctures) by influencing the gestural constellation. Arguably, more precise alignment of gestures is associated with hyperspeech (enunciation), whereas increased reduction and gesture overlapping and spreading is a feature of hypospeech. In this case, however, there are two modes of alignment, either of the whole gestures to each other or of end point to starting point. The detailed variation of precise gestural alignment should be a factor here too in both alignment modes, the separation of gestures into [ʃs] is arguably itself also an independent marker of hyperspeech. There is, at any rate, evidently less spread allowed in the internal postalveolar. It is therefore relevant to speak of two different scales of significant variation in the two. Exactly what synchronic conditions favour gestural separation and which favour gradual overlap or merger remains to be explored however (see below for some suggestions).
This arguably also reflects the diachronic drift from a state where disalignment was less marked. Upholding gestural distinction is an upper bound of a synchronic variational scale that can be described as a product of a diachronic change of normative patterns with a movement of the span of H&H variation – the most frequent range (i.e., normal speech) “descends” into what was previously hypospeech and the former normative range becomes more and more hyperspeech. The internal category differs in that it has a constraint on the variation where gestures can only align, and so the drift of the variational span has played out its course: there has been a development over a period of variation from gestures locked as separate to where the span has finally ended up in always aligning the gestures.

The distribution and difference of [ɪl] and [ɪ], where gestures in the former are separate and in the latter merged (but irregardless of position relative to a boundary), may perhaps be similar in character and behaviour to the boundary category – there may well be a corresponding continuous variation of gestural alignment. So H&H variation still would apply also to internal rl, but not to the other historical cases of r plus dental. This could mean this change is still in progress, representing a stage in development where the other postalveolars were previously. There may be phonetic reasons for special treatment of the lateral.

6.4.2 The case of obligatory rhotic plus dental
In addition to being one end of the variation in r#s, there are importantly internal cases where the gestures are obligatorily separate. This seems to occur lexically, as in absurd, and at an internal boundary with the rhotic as a final in a short stressed syllable, orthographically <rrs>.

It is not really possible to have <rrs> belong to the same variational pattern as r#s, as the presence of [ɪ] in the former is obligatory (even though variation in the sibilant might be found). It was argued in 2.1.5 that there is a phonotactic reason for retaining [ɪ] – that is, to prevent a short stressed open syllable. A phonotactic-based constraint on the variation makes sense where there is such a need, but not in the likewise obligatory absurd which then certainly must be regarded as different, and arguably purely lexical.

This phonotactic reasoning can also be questioned in view of what has been said of the gestural view, as it apparently presupposes a segment or, in a way, seems to define a segment as a metrical unit. But this can be resolved if taken not as an intrinsic quality but something which emerges out of syllable structure: that is, the need for a metrical unit compels a separated postalveolar gesture on which to place this metrical weight (the notion of absolute time in Articulatory Phonology may be relevant here).

This also touches on the more general problem of syllabification. Why is it that the following sibilant in <rrs> cannot rather be the coda of the preceding syllable (or be ambisyllabic)? It seems that syllabification is very much governed by morphology – a boundary implies a syllable break, perhaps as a kind of juncture marking. Treatment of the boundary has so far been rather segmental, as being some sort of positional entity between two segments (as evidenced by the notation r#/s) – without reflecting over what a boundary really is. Obviously in one sense it is where one lexical item ends and another begins, but how this manifests in actual production and perception is much less clear-cut. Apparently, it is relevant as seen in the case of juncture marking – but in general, production does not need to make a distinction of articulations at boundaries and works instead with specifications of
longer utterance blocks than words. In junctures, specifications and cues will be run against lexical material in both perception and production – so the boundary emerges as a product of interference from the respective words.

The traditional explanation of the <rrs> case was seen above in postulating gemination of the /tr/, which, by virtue of this, is not removed by the derivational process – or if gemination is to be segment doubling, perhaps one of the rhotic segments is removed. While geminates in the sense of being double segments make no sense, as they are clearly one continuous articulatory unit, geminates as long segments can be represented gesturally as longer duration gestures. It is not entirely clear that this is actually the case for Swedish consonants, or at least that it is a fundamental trait – it could be a direct consequence of vowel length contrast, or perhaps rather that a stressed syllable carries length, which is patterned so that some vowels take it and some cannot (and whether this is best coded in gesture length is an open question).

What happens, anyway, is that the gestures are locked as separate. Geminates could perhaps be taken to mean simply that – as a descriptive device, even if in more abstract terms. Relating this issue to vowel length, however, is incongruent with the data. Firstly, in cases of obligatory gestural separation where no such open syllable would occur and no boundary exists, as in absurd, geminates seem unwarranted if they are at all related to vowel length. Arguably then, this is a lexical feature: that is, it is part of the stored lexical specification that the gestures are to be aligned as separate. In origin, this may be considered reinstating a rhotic cue no longer felt to be present in unitary postalveolars and needed to properly approximate the source pronunciation of the loanword. In yet other cases, such as perception, there is variation [as-ʃ], which may be akin to that of the boundary category.

Secondly, there are indeed instances of a conspicuous [ɾ] followed by a laminal denti-alveolar sibilant after a long vowel, such as piercing [ˈpʰɪrɪŋ] (cf. 2.1.5). An interesting example of this is the spelling pronunciation of unfamiliar words, such as proper names – e.g., the place-name Irsta is pronounced by speakers familiar with the name (i.e., the inhabitants of the locality) always as [ˈɪʃtə], whereas speakers introduced to it in writing can apparently pronounce it [ˈiːstə]. If it is a boundary effect, not only do [ˈɪʃtə] speakers place a boundary between two supposed elements, they also make the separation of gestures obligatory. At any rate, these speakers here feel the need of a distinct rhotic cue – comparable to the treatment of loanwords (or perhaps as a separation of components, thus juncture marking) – precipitated by the spelling. This could be seen as an indication of an historical development gradually disentangling the patterns of postalveolars and the rhotic – this is perhaps also seen in the variational drift in 6.4.1, as well as speaker reanalyses (e.g., how it relates to /x/).

6.4.3 Considerations in explaining the variation in r#s
It has already been suggested that explaining how any particular instance of r#s manifests – that is, its position in the variational span – can be done in an H&H perspective, where pragmatic factors are exercised in gestural coordination, and where separation of the gestures is a marked hyperspeech feature. Some statements partly to this effect, though not in these terms, can also be found in the literature (e.g. in Witting 1959, cf. 2.2.3, Gårding 1967 on the inconsistency of juncture marking, and Eliasson 1986). For example, Eliasson (1986) says:

The variability [of postalveolarization at a boundary] is contingent on factors of three kinds: (a) the phonetic nature of the input sentence, (b) the immediate structural environment, notable grammatical boundaries, and (c) wider, partly structure-external parameters. [...]  

56 Such as in saying e.g. that in loaning absurd, to retain a distinct [ɾ], the rhotic would have had to be analysed as a geminate to represent what is just different entities.

57 A structuralist “solution” could doubtlessly be conceived, such as /per#s#ning/ or /per.rsning/ with appropriate rule ordering (unless it is for phonotactic considerations of the underlying form, as was the case with /r#s/ for [x]).

58 This in not the case e.g. with the more generally familiar locality Mårsta, which is uniformly [ˈmeɾʃta].
In general, it can be said that the closer the transition is between two words, the more easily the assimilation will take place. Among the relevant parameters are speech rate including pause and hesitations patterns, speech style, frequency of word collocations, individual speech habits, etc. Allegro speech and ordinary conversation [...] and causal speech [...] regularly favor sandhi. (281-282)

Gårding (1967:46), in investigating how speakers identify junctures, notes that they “waver” over postalveolar or sequence across boundaries, giving some suggestions as to factors that may govern variation in r#s. Malmberg (1971:78) says the boundary process “depends on speech tempo” or “individual habits” and that postalveolar may be favoured in “slow, emphatic speech”. Kuronen (2003) says it is “somewhat dependent on the combination [of words], the phonetic context, the idiolect, and – possibly – speech tempo”. According to Stausland-Johnsen (2011), whether postalveolarization in Norwegian occurs across word boundaries is context-dependent.

Exactly how the variation manifests may be difficult to quantify, and showing how or to what extent different concerns influence separation or simultaneousness of gestures is beyond the scope of this study. A study of the [ʌs~ʃ] category with a view to finding correlates for spectral characteristics would need to control for both syntax and phonotactics, as well as factors of prosody and pragmatics.

Arguably though, a gestural analysis is what is needed to describe it. Gestures enable measurement of variation that a segmental and derivational view cannot, as there are obviously one or more gradual variables, and not just application of a rule or not. As for explaining how variation is organized, it is suggested some pragmatic-based model (akin to H&H or Cognitive Phonetics) can provide a framework.

Finding the correlates of the variation in r#s has been identified here as one area for further study. To complement this, examining corresponding variation in different varieties (including Norwegian and Finland Swedish) may also be addressed, as could the variation in [ʃl~l] (and the possible variation in <rrs>). Setting up a model for diachronic explanation in a gestural and variationist perspective is another topic of interest.

Another survey that can be done, partly on the present data, is to measure the duration of the postalveolar and sibilant gestures relative to one another. The data already include the duration of the sibilant section, but no analysis has been made in the present study. There can be several different approaches to measuring this, including what percentage is the sibilant of the total duration. This could then also be compared to the results of the frequency of the sibilant presented here.

There is some difficulty how to represent some of the possible gestural outcomes that appeared in the data. At any rate, to fully determine how the variation may manifest, there is need to make detailed gestural analysis of more tokens to establish the hypothesis that the gradual variation is correlated to gesture alignment (and what other effects occur).

6.4.4 Final remarks

In conclusion then, there are two distinct postalveolar phenomena at work: internal [ʃ] and external [ʌs~ʃ]. The traditional analysis that equates them describes the latter as /r#s/ – that is, the variation is somehow caused by the presence of a boundary. There is no attempt at an actual mechanism for explaining how the boundary affects the pronunciation – this is arguably an effect of the word being the ubiquitous unit of analysis, so that explaining effects occurring at the intersection of words in an utterance is neglected. This is not to say that boundaries do not mean anything, but merely calling on a boundary cannot be a justifiable explanation in itself. Even with the presence of a boundary or juncture, an explanation is still required as to how that brings about the patterns encountered.

Positing a form /r#s/ is also in fact saying there is indeed a difference between the patterns, only that in a traditionalist perspective it is taken to say something of the contents of a phonological inventory of units, because it vacates as it were the sequence /rs/ for use as [ʃ].
This also conflicts with production, as the boundary only seems present when there is a marked juncture, whereas word boundaries in themselves are not normally produced as production is seamless. The only point then seems to be a desire not to have production-relevant units in the phonological inventory. Again this is linked to the question of what phonemes really are about.

There are, at any rate, two different production specifications – one with variation and one without, and where the latter contrasts with a pattern that exists within the variation of the former – and already in this there is a difference of phonological concern. The phonological nature of sounds is to what degree patterns are associated with one another into manageable units. As mentioned above, the traditional model contrary to this originally did not purport to describe a mental process, but only a “generalization” of abstract data.

Both categories (and all outcomes of the boundary category) are comprised of the same gestures, but with different timing as to how the gestures relate. In the internal category, gestures are locked as simultaneous to each other, in the boundary category they can be simultaneous or entirely separated or possibly anything in between, arguably as exposed to pragmatic influence as in H&H variation. So in some sense they are the same underlying units, and merely their timing shifts (with some effect at the overlap as a result). There is nonetheless a significant difference between the patterns, but one which is unstateable (as is the variation) in a representation based in the assumption of segments.

On the phonological status of postalveolars then (as opposed to two “underlying segments”), one could ask whether it is not an open issue depending on one’s point of view of what is the relevant unit. In a gestural perspective, there are indeed two identical “units” at work here in both cases, as nominally also in the traditional sequence analysis, but the same gestures are also present in the different allomorphs. Gestures really take on a different dimension compared to segments, with the difference between the categories being distinctive in temporal relation of gestures to each other – indeed, the singular [ʃ], even “phonetically”, is also characterized by the same two gestural units. A gestural perspective thus represents a radically different and rather incompatible way of thinking of phonetic and phonological form compared with the segmental view.

In a larger context, the answer depends on whether one is willing to accept segments as mentally ubiquitous units, and the gestural properties of the signal as merely surface effects, and if variation is to be treated as irrelevant, or if it is to be acknowledged as systematic (in some form) and meaningful, and as a feature that merits description. It depends on what one wants to gain out of phonology and what is the ontological nature of phonological analyses and processes – and what is really meant by “underlying”. And with this whether phonology is taken to be mental specifications relevant to production and perception, or only abstract generalization of data. It has been argued here at any rate that phonology is more complicated and exerts itself in other ways than can be displayed in derivational segmental models.

7. Conclusions

The study concludes that there is a phonetic difference between internal [ʃ] and a variable [ʃ~ɹ] occurring at boundaries. The former is always a postalveolar, while the latter is a variational continuum from a postalveolar [ʃ] to a [ɹ] sequence. There is further a lexical contrast of internal [ʃ] vs. a largely non-variable [ɹ]. It is possible that cases of [ʃ~ɹ] variation might occur internally too.

59 The traditional description of “two sounds becoming one sound”, or the “deletion of one sound”, or the “spread” of something “from one sound into another”, becomes increasingly strange under a gestural treatment.
These conclusions are based on data experimentally obtained from an investigation that measured the fricative quality of both categories, and of gestural phonetic analysis of individual cases that illustrate the aspects of continuous variation. What is said here of the postalveolar sibilant should also hold for the other postalveolars (apart from the lateral).

The variation in [ɹʃ~ʃ] is continuous, not a dichotomous choice between [ɹ] and [ʃ]. There are in fact two semi-independent parameters: the quality of the fricative and the distinct presence of [ʃ] before the fricative. These are not directly linked, as evidenced by the presence of e.g. [ɹʃ]. Both are continuous: the fricative can be anywhere on a gradual scale from laminal alveolar to apical postalveolar, and this is also suggested for the rhotic as a conspicuous approximant, vowel colouring or not present at all.

Fixed [ʃ] is internal and [ɹʃ~ʃ] occurs at a boundary; the separation of these two groups seems strict. The delineation relevant here was seen to be one between words or between elements of a transparent (less lexicalized) compound, while other morpheme boundaries seem to fall into the internal [ʃ] category.

It is claimed that a description of this kind of variation cannot be accomplished under the segmental and derivational model of traditional phonology. Instead, some form of gestural phonology is advocated. Segments in the traditional alphabetical sense are not articulatorily meaningful, and it is argued there is no obvious reason for giving them mental status. The difference between [ɹʃ] and [ʃʃ] can be seen as simply one of alignment of the same component gestures, rather than – as in the traditional view – exchanges of three altogether different segments. Also, derivational rules (in their usual form at least) cannot account for the continuousness of the variation observed, as their output can only be segments and because of the “either-or” nature of their application.

The sequence claim cannot explain cases of internal obligatory [ɹʃ], except when introducing otherwise unwarranted underlying geminates; it is argued this concept has little phonological value and is nothing more than an ad hoc fix.

The traditional argument saying postalveolars are phonologically sequences is based in axiomatic conceptions of a segmental and derivational nature of phonology; it is informed by a concept of implicit or explicit biuniqueness, meaning identical allophones in context could not be mapped to different phonemes; it has an idealistic view of the phoneme as a mental “unit of sound”, a phonetic segment as an exalted abstract; it is diachronic in appropriating historical sound change claimed as mental processes; it is driven by a preconceived desire to reduce the number of phonological units.

Though there is some sort of relationship between postalveolars and /rl/, it need not be hierarchical, nor synchronous, nor biunique. The statement that this is not the case is in line with a view of phonology (or indeed language processing overall) as mentally being analogical or associative (or ‘connectionist’) rather than derivational and serial, and that it is not mentally relevant to consider phonemes as “sounds”.

Gestural phonology models mental storage of form as being closer to articulation, rejecting phonological segments, thus precluding derivation as a meaningful device. When acknowledging that [ɹʃ] and [ʃʃ] are merely different temporal constellations of the same gestures anyway, the question of whether it represents one or two phonemes (in the traditional sense of phonological segments) largely becomes moot.

Factors controlling the variation of [ɹʃ~ʃ] have only been hinted at and require additional research. It is suggested that some form of a pragmatic model of phonetic variation may be used as an explanatory framework.
References


