Articulatory dynamics of (de)gemination in Dutch

DOI:
10.1016/j.wocn.2018.03.005

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Journal of Phonetics

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Abstract

So-called ‘fake’ or derived geminates differ in the extent to which they behave like lexical geminates, both across and within languages. Data that enable us to study the gestural as well as the durational properties of fake geminates can shed light on their status as long consonants or clusters, and on the effects of any degemination rules present in the language. Our focus is on Standard Dutch, which has been said to exhibit categorical degemination of derived identical consonant clusters. The specific interest is in the post-lexical gemination of /r/, which involves gesturally strongly distinct allophones in the language, such that the members of /r#r/ sequences are identical phonologically, but not phonetically. Ultrasound data from 8 speakers of Standard Dutch show that fake geminate /r#r/ sequences present combined articulatory characteristics of coda (bunched approximant) and onset (uvular fricative or alveolar tap) allophones, which can be captured in terms of gestural overlap or ambisyllabicity. At the same time, however, concatenated /r#r/ does not show an increased duration relative to singleton /r/. These results suggest that Dutch degemination does not involve categorical segmental deletion, as previously suggested, but instead it has constraints on phonetic consonant duration that limit temporal gemination. Such specific duration constraints have a more general role to play in any account of gemination, as accumulating evidence suggests that fake geminates at word boundaries tend to behave as consonant clusters, subject to constraints on degree of overlap.

Keywords: Dutch; fake geminates; rhotics; ultrasound; gestural dynamics

*The results of a pilot study on which the present work builds were published as Strycharczuk, P. and Sebregts, K. (2014). Erring on the side of phonology. In A. Auer and B. Köhnlein (Eds.), Linguistics in the Netherlands 2014, 145–164. The present article supercedes this earlier work. All the data and the analyses presented here are new.
1 Introduction

Fake geminates can be defined as sequences of identical consonants derived by morphological concatenation (Hayes, 1986). They are typologically less restricted than lexical geminates, and may occur in languages where lexical geminates are not present. However, languages vary in the extent to which fake geminates behave phonetically like lexical geminates. At one end of the spectrum, English and German do not have lexical gemination, but fake geminates are phonetically long (Oh & Redford, 2012; Kotzor et al., 2016). In contrast, Dutch is reported to prohibit long consonants, including those derived by external sandhi. In generative accounts, Dutch degemination is described as a categorical, phonological process, consisting in the deletion of one of the two consonants (Booij, 1995, 69). This analysis, however, is challenged by Martens & Quené (1994), who report a subtle increase in duration for fake geminates (e.g. raaf fraai ‘raven beautiful’ compared to ra fraai ‘yard beautiful’, mean difference of 18ms in normal speech). Based on this result, Martens & Quené argue against phonological degemination rules in Dutch, and propose treating degemination as a gradual process instead.

In the present study, we provide new evidence on the nature of fake geminates in Dutch, based on their articulatory characteristics. We focus on the case of /r/, where an on-going sound change has created an allophony pattern that differentiates strongly between the two members of a geminate sequence. In recent years, Standard Dutch in the Netherlands has developed an approximant /r/ realization in the coda (Sebregts, 2015), which is articulatorily either bunched or retroflex (Scobbie & Sebregts, 2010). This variant is increasingly prevalent in the linguistic repertoire of Standard Dutch speakers, but it is restricted to coda position. Onset /r/ is typically uvular (normally a weak fricative or an approximant), or alveolar (typically a tap). The resulting patterns of onset-coda allophony involve robust articulatory distinctions, which allows us to investigate to what extent fake geminates resemble coda and onset variants, and it provides a window on the phonetics-phonology interface: the r#r sequence contains identical consonants at the lexical-phonological level, but highly dissimilar ones at the phonetic level. We also consider to what extent the phonetic characteristics of fake geminates are consistent with phonological approaches to geminates (and potential degemination) in different frameworks.

Lexical geminates are traditionally analysed as a single melodic unit linked to multiple timing slots (Kenstowicz & Pyle, 1973; Schein & Steriade, 1986). This is largely because geminates are argued to behave as a single unit phonologically. For instance, geminate sequences cannot be broken apart by epenthesis, and the two parts of a geminate sequence cannot undergo phonological rules independently of each other. It is less clear, however, to what extent the same type of structure may be appropriate for
fake geminates. According to McCarthy (1986), derived geminates can be analysed as identical to lexical ones. However, experimental research suggests there may be some differences between the two types of geminates. Ridouane (2010) compares lexical and fake geminates in Tashlhiyt Berber, and finds that fake geminates pattern phonetically with lexical geminates, but only when they are derived through assimilation. Geminates that arise through concatenation of identical consonants (like the Dutch fake geminate case) show phonetic differences compared to lexical or assimilated geminates. Specifically, the differences concerned RMS amplitude (lower in concatenated geminates) and the duration of the preceding vowel (ca. 15ms increase in concatenated geminates). Hence, Ridouane (2010) concludes that true geminates and fake geminates can be distinguished at the melodic level: both are associated with two timing slots, but with only one melodic position in the case of true geminates, and two melodic positions for fake geminates.

Similarly to Tashlhiyt, different types of derived geminates behave differently in American English. Oh & Redford (2012) study the acoustic duration of nasals in mono-morphemic words such as *immensely*, poly-morphemic words such as *immemorial*, where a geminate is created by suffixation, and in sequences such as *prim memorial*, where a geminate is created across word boundaries. Whereas the absolute duration of the nasal was on average ca. 40 ms greater in both types of geminates than in the baseline singleton consonant, a different temporal pattern emerges once relative duration is considered. The C:V1 duration ratio was considerably increased in word-medial geminates (with a mean of ca. 1.8 in normal speech) compared to word-boundary geminates (mean CV ration of ca. 1.4 in normal speech), and no significant difference was detected in C:V1 ratio between word-boundary geminates and singletons. Oh & Redford propose that the observed difference between word-internal and word-boundary geminates calls for differentiating between them at the analytical level. They argue that word-internal geminates are best analysed as long consonants which have intrinsic timing properties, whereas word-boundary geminates are underlyingly clusters realized with overlapping gestures. Further support for this analysis comes from the observation that word-boundary geminates, but not word-internal geminates, may be pulled apart in careful speech.

Both Ridouane (2010) and Oh & Redford (2012) argue for an analysis of fake geminates as consonant sequences, but they propose different representational accounts for this phenomenon. Whereas Ridouane (2010) argues that his findings are consistent with an autosegmental model, Oh & Redford propose an Articulatory Phonology style representation, where gemination is subsumed under a more general process of gestural overlap (Browman & Goldstein, 1986, 1989, 1990, 1992). Gemination effects across word boundaries can be captured in this model as partial overlap between two inde-
pendent consonantal gestures which are blended, but which do not overlap completely, hence the increased duration.

Scobbie & Pouplier (2010) argue for a gestural interpretation of fake geminates in Southern Standard British English and Scottish English, as part of a larger analysis of how multiple segmental and prosodic factors affect the articulation of /l/. Scobbie & Pouplier present EPG data on the production of /l/ across a range of segmental and prosodic conditions, including word-initial (pee leeward), word-final pre-consonantal (peel#beavers), word-final pre-vocalic (peel#Eve) and the fake-geminate context (peel#lemurs). They report that the fake geminate context shows considerable dorsal retraction, typical of /l/ in canonical coda positions (word-final pre-consonantal). At the same time, however, the fake geminate contexts always show tongue tip contact, which is typical in onset (word-initial) /l/, but which may be reduced in word-final pre-consonantal context. The presence of both advanced dorsal retraction and tongue tip raising makes the fake geminate context unique, and it is consistent with analysing it as a gestural blend of word-initial and word-final /l/. Although Scobbie & Pouplier (2010) argue for a gestural overlap analysis mainly on the basis of comparison between word-final pre-consonantal and word-final pre-vocalic /l/, their argument is reinforced by the presence of articulatory blending in /l#l/ sequences.

The Dutch /r/-allophony case is similar to the English /l/-allophony one, in the sense that it involves two distinct articulatory configurations depending on prosodic position. Cases like this provide a unique window into testing the gestural dynamics of gemination, and potential degemination. In the extreme case of a categorical, phonological process of degemination (i.e., deletion of one of the two segments or timing slots), it stands to reason that only the gestures associated with one of the two allophones would be recoverable. If the process is gradient, the interest lies in the interplay between gestural overlap and temporal overlap. Again, there is an extreme case imaginable in which the coda and onset gestures are no longer sequential, but truly blended, and the fake geminate r#r forms its own allophonic category. In all but these extreme cases, however, we would see a more sequential presence of coda and onset gestures. The resulting question would then be to what extent the fake geminate differs from other consonant sequences, i.e. those where the two consonants are phonologically non-identical.

2 Materials and method

2.1 Test items

The test items we used included /r/ in the following contexts. The key context was fake geminate /r/ (Vr#rV). The baseline contexts were canonical onsets (word-initial
Table 1: Test items used in the study

<table>
<thead>
<tr>
<th>Category</th>
<th>Word Form</th>
<th>Syllabic Representation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-initial onset</td>
<td>V#r</td>
<td>pa reizen /pa'r.ei.zən/</td>
<td>‘father travels’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pa raden /pa'ra.dən/</td>
<td>‘father guess’</td>
</tr>
<tr>
<td>Word-medial onset</td>
<td>VrV</td>
<td>Parijzenaar /pa.ri.zə.nər/</td>
<td>‘Parisian’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>parade /pa'.ra.də/</td>
<td>‘parade’</td>
</tr>
<tr>
<td>Canonical coda</td>
<td>Vr#C</td>
<td>paar meisjes /par.mei.səs/</td>
<td>‘couple of girls’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>paar baden /par'ba.dən/</td>
<td>‘couple bathe’</td>
</tr>
<tr>
<td>Word-final prevocalic</td>
<td>Vr#V</td>
<td>paar eisen /par'e.i.sən/</td>
<td>‘couple of demands’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>paar ademen /par'a.do.mən/</td>
<td>‘couple breathe’</td>
</tr>
<tr>
<td>Fake geminate</td>
<td>Vr#r</td>
<td>paar reizen /par'rei.zən/</td>
<td>‘couple of trips’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>paar raden /par'ra.dən/</td>
<td>‘couple guess’</td>
</tr>
</tbody>
</table>

V#rV, and word-medial VrV), canonical coda (Vr#C), and word-final pre-vocalic /r/ (Vr#V, which is not typically resyllabified in Dutch). The preceding and the following vowels were controlled for. The preceding vowel was always /a/, and the following vowels were /a/ or /ei/ (balanced across the set). We were limited in our choice of the preceding vowel by /r/-colouring that affects most long vowels in Dutch, e.g. the vowel /o/ followed by a coda /r/ becomes [oa] (Gussenhoven, 1992). Such effects would make Vr# and V#r relatively less comparable for most vowels. Consistent use of low vowels as segmental context also had the advantage of increasing the articulatory distance between the vowels and the /r/ making the /r/ gesture easier to identify (see Section 2.4.2). The full list of test items is in Table 1.

The test items were embedded in meaningful carrier sentences, as exemplified in (1). All the carrier sentence were questions to ensure comparable prosody.

(1) Example prompt sentence: Kon dat paar raden wat er in de doos zat?
   ‘Could that couple guess what was in the box?’

In addition to the 10 test items, listed in Table 1, we included 14 distractor sentences to obscure the purpose of the experiment to the participants. The distractor sentences had a similar structure to the target sentences. An example is in (2).

(2) Example distractor sentence:
   Kon dat feest niet doorgaan zonder de gasten?
   ‘Could that party not go ahead without the guests?’
2.2 Participants

The participants were eight native speakers of Dutch (5 females, aged 19–39, $M=26$). Since the experiment took place in Scotland, where we had access to a high-speed ultrasound system (see 2.3 below), we were restricted to recruiting Dutch expats in the UK. The time spent abroad varied from a few months to 11 years, with a mean of 4 years. They came from a variety of locations in the Netherlands. They were all speakers of Standard Dutch: They identified as such, and their self-evaluation was corroborated by judgement of the second author, who is a native speaker of Dutch familiar with the sociolinguistic situation in The Netherlands. The sociolinguistic features we relied on as hallmarks of the standard accent included the diphthongal realization of /e/ and /o/ and non-fronted place of articulation of /x/ (Smakman, 2006). Note that accents currently considered as standard in the Netherlands may involve fairly diverse /r/-allophony patterns. We did not impose any participant restrictions in this respect, since we were interested in the potential interactions between /r/-allophony types and the realization of fake geminates. We recorded three additional participants, but discarded their data for reasons including poor ultrasound image, detectable foreign accent in one participant, and non-standard accent features.

2.3 Procedure

The recordings took place in the Centre for Audiology, Speech and Language (CASL) research centre at Queen Margaret University, Edinburgh. We obtained simultaneous ultrasound and audio signal. The synchronization of the two types of signal was controlled by the Articulate Assistant Advanced\textsuperscript{TM} (AAA) software, version 2.16 (Articulate Instruments Ltd, 2014). Ultrasound images of the participants’ tongues in the midsagittal plane were captured using a high-speed Sonix RP ultrasound system (frame rate = 121.5 fps, scanlines = 63, pixels per scanline = 412, field of vision = 134.9°, pixel offset = 51, depth = 80 mm). A stabilization headset was used in order to minimize the movement of the ultrasound probe throughout the recording (Articulate Instruments Ltd, 2008). The audio data were recorded with a lavalier Audio-Technica AT803 condenser microphone connected to a synchronization unit. The sampling rate for the audio data was 22kHz. For further details on this particular recording set-up, the reader is referred to Wrench & Scobbie (2016).

Each recording session began with fitting the stabilization headset. Following that, we recorded the participant swallowing water in order to obtain an image of their hard palate (Epstein & Stone, 2005). Next, we recorded the participant biting on a plastic bite plate, pushing their tongue up against the plate, so that we could image the occlusal plane (Scobbie et al., 2011). The hard palate and occlusal plane tracings are illustrated
for an example participant, DF2, in Figure 1. Following those steps, we commenced the recording session. Four participants read five repetitions of the experimental material, whereas for the remaining five, we were able to record six repetitions. In sum, this yielded 500 tokens. We acknowledge that this overall count of stimuli is relatively low. This was necessitated by restrictions on the duration of recording sessions (ca. 20 mins), due to discomfort from prolonged wear of the stabilization headset. We could have increased the number of tokens at the cost of fillers, but the use of fillers was motivated by considerations of task effects. The prompts were pseudo-randomized in blocks, excluding the immediate succession of fake geminates by the minimally different baselines, avoiding, for instance, *paar reizen* neighbouring *pa reizen*. The prompts were presented on a computer screen, one at a time. Participants were instructed to speak as naturally as possible. If a participant made a mistake, we repeated the prompt, and discarded the error token.

Figure 1: Raw ultrasound frames showing a frame representative of /r/-constriction for speaker DF2. The frame one the left shows a uvular onset /r/ (V#rV), whereas the frame on the rights shows a bunched approximant coda /r/ (Vr#C). The tracings show the position of the hard palate and the occlusal plane. Tongue tip is on the right.

2.4 Analysis

2.4.1 Acoustic analysis

The acoustic data were exported from AAA into Praat (Boersma & Weenink, 2009) and forced-aligned using the Munich Automatic Segmentation System MAUS (Kisler et al., 2012). The automatic segmentation was then hand-corrected by the second author, paying particular attention to the onset of /a/ (the vowel preceding /r/) and the offset of /r/. Example segmentation is illustrated in Figure 2. In all cases, /a/ was preceded by /p/. The vowel onset was placed at the offset of the stop burst (Dutch is a non-aspirating language). The offset of /r/ was determined based on intensity criteria, as /r/ could be followed by either a consonant (/b/ or /m/), or a vowel. For the Vr#C condition, we also marked the offset of the consonant following /r/. We made
no attempt to identify a boundary between /a/ and /r/, as the acoustic transition was entirely gradient in cases where /r/ was realized as an approximant, which is a substantial portion of our data. This limits us in our analysis of duration, considering that earlier studies distinguish between the absolute effect that gemination has on duration and the relative effect on the ratio of consonant to vowel duration. However, due to the continuous nature of transitions in our data (see Figure 2), the segmentation was simply not possible.

Figure 2: Example segmentation of tokens including onset tap /r/ (left) and coda approximant /r/ (right) for speaker DF1

We used the segmentation to analyse durational differences based on two linear mixed effects regression models predicting the acoustic duration of /ar/, by which we refer here to the entire sequence of the coda vowel plus either singleton-/r/ (itself either in coda or onset), or the fake geminate cluster. The models were fitted using the lme4 package (Bates et al., 2015) in R version 3.3.2 (R Development Core Team, 2005). The model specifications are detailed in Section 3.1. We then imported the acoustic annotations into AAA, and used the time points to select the ultrasonic frames that formed the basis of our articulatory analysis.

2.4.2 Articulatory contour tracking

For the articulatory analysis, we tracked the tongue contour for all the frames corresponding to the acoustic /ar/ interval, using the automatic tracker in AAA version 2.16. The tracking was supervised by the first author in real time, and manual corrections were applied if required. We then exported the Cartesian coordinates for all the tongue contours, rotating and offsetting the image relative to the occlusal plane. This ensures that the upward movement of the tongue corresponds to the vertical dimension in the diagrams, and in the input to our analysis.
For each token, we extracted three ultrasonic frames at three landmarks: (i) acoustic onset of the vowel; (ii) maximum /r/ constriction; (iii) acoustic offset of /r/. We defined the maximum /r/ constriction as the point of maximum vertical tongue displacement. This was appropriate given that tongue rotation was standardized across the speakers using the occlusal plane, and given the nature of our stimuli, which invariably involved tongue body raising in the transition from the low vowel /a/ to /r/, followed by lowering in anticipation of the following low vowel /a/ or /ei/ (note that consonantal contexts always involved labials /b/ or /m/, which are not expected to trigger any lingual coarticulation on the preceding /r/). We analysed the tongue shapes using Smoothing Splines Analysis of Variance (SS-ANOVA, Davidson, 2006; Gu, 2013, 2014), applied per speaker and per each of three chosen time points. The SS-ANOVA was calculated based on a polar transformation of Cartesian Coordinates exported from AAA (Mielke, 2015).

2.4.3 PCA articulatory analysis

We followed up the static analysis of averaged tongue contours with a dynamic analysis of the articulatory transition throughout /ar/. We used Principle Component Analysis (PCA) of pixel intensity data to reduce the information in the ultrasonic frames (Hueber et al., 2007; Carignan et al., 2016; Hoole & Pouplier, 2017; Mielke et al., 2017). We used all ultrasonic frames (in JPG format) corresponding to a speech signal in our recordings as input to the PCA. The images in these frames are filtered and reduced to manually-specified regions of interest (polygons outlining the space for lingual movement) and filtered. The PCA was run on data from each speaker separately, across all frames, using the TRACTUS software (Carignan, 2014; Carignan et al., 2016). For a detailed description of the algorithm, the reader is referred to Mielke et al. (2017). The images were not rotated in this case.

Although the PCA provides a considerable reduction of the dimensions encoded in ultrasonic image, we find that it still returns a rather large number of PCs if we are to capture 80% of the data. A representative example is speaker DF2, for whom this corresponded to 51 PCs. In order to reduce this information further, and to make the PCA output interpretable, we performed a Linear Discriminant Analysis (LDA) on the PCA output (see Strycharczuk & Scobbie 2017 for a similar approach). LDA provides a linear transformation of numerical predictors that best expresses a given category. In our case, the predictors were PCs. For each speaker, we extracted the PCs that capture 80% of the variance and used them as LDA input. We trained the LDA to distinguish the word-final coda context (Vr#C) from the word-initial onset (V#rV). This was done on the frames corresponding to the acoustic /ar/ region in each token. In each case, the analysis was able to reliably distinguish coda and onset (the LD1 distributions for the
two categories are non-overlapping), which is not surprising given the large articulatory differences between these two contexts.

For all speakers, low LD scores were assigned to the onset conditions, whereas high LD scores were assigned to the onset condition. Having trained the LDA to distinguish onset and coda, we used the PCA-LDA loadings for each speaker to predict the LD values for new data: word-medial onsets (VrV), word-final pre-vocalic rhotics (Vr#V), and fake geminates (Vr#r). We then compared the LD values in this context against the two baselines (canonical onset and canonical coda) using Generalized Additive Mixed Modelling (GAMM, Wood 2006; van Rij et al. 2015). We chose to use GAMM, since exploration of the raw data revealed non-linearities in the progression of LD values over time, and GAMM is particularly well suited to analysing non-linear time-series data.

3 Results

3.1 Acoustic duration

The first model compared the acoustic duration of the /ar/ sequence across all five experimental conditions. The model was testing the hypothesis that the duration of /ar/ in the fake geminate condition (Vr#rV, e.g. paar reizen) is longer compared to the /ar/ sequence of the canonical onset condition (V#rV, e.g. pa reizen). The two remaining contexts (VrV and Vr#V) were also included. There were two fixed effects: condition and the vowel following /r/ (/a/ or /eɪ/). The random structure included random intercepts for speaker, item and block, as well as a random coefficient for condition within speaker.

The effect of condition is illustrated in the left panel of Figure 3. According to the model, there was no increase in duration in the fake geminate context (Vr#rV) compared to word-initial onset (V#rV). The average difference (18.43 ms) was in fact in the opposite direction, but it was not significant (SE =17.01, t = −1.08). This result provides no evidence in favour of temporal gemination, as we might expect that temporal gemination would manifest itself as increase in /ar/ duration in the Vr#rV context. The left panel also shows that both the canonical onset and the fake geminate context exhibited a longer duration than the remaining three contexts, word-medial intervocalic, canonical pre-consonantal coda, and pre-vocalic word-final coda.

The model above suggests that fake geminate /r/ is not lengthened compared to singleton word-initial onset /r/. In order to fully understand the duration facts, however, a further comparison is needed that also takes into account cluster duration. We built another model, this time comparing the combined duration of coda-onset clusters. We measured the duration of such clusters in the word-final coda context (Vr#C), and compared it with the duration of fake geminate /r/ and word-initial /r/. The vocalic
portion (preceding /a/) was once again included in the measurements, due to segmentation issues discussed in Section 2.4.1.

As shown in the right panel of Figure 3, the combined duration in contexts including /r#C clusters does not exceed the duration of vowel + /r/ in the fake geminate context (Vr#rV), or the word-initial onset context (V#rV). Taken together, the results of the two models suggest that acoustic duration is ambiguous as a source of evidence concerning degemination rules in the case of Dutch /r/. While there is no clear lengthening effect in fake geminates compared to singleton onset, we cannot take this as support of temporal degemination, because we do not observe lengthening either in the duration of /r#C clusters. For this reason, it is crucial to also consider the spatial properties of fake geminates before we can conclude whether they behave as geminate or singleton consonants.

### 3.2 Tongue contour comparison

Figures 4 and 5 provide an articulatory comparison between the fake geminate context Vr#rV and two baselines: word-initial onset V#rV and word-final coda Vr#C. Specifically, the plots show averaged tongue contours for individual speakers by context at three different time points: the onset of the preceding vowel, maximum /r/-constriction and the offset of /r/. Although the SS-ANOVA was calculated based on data from all five contexts, only three are plotted to improve legibility. Note that information from
the endpoints of the tongue contour should be interpreted with caution, since the automatic spline tracker may return data in regions with low confidence (endpoints of the tongue contour are typically obscured by jaw and hyoid shadows).

Focusing on the tongue shapes from the frame showing maximum /r/ constriction (the middle panel in all plots), we can see a variety of allophonic patterns. As far as the onset /r/ goes, uvular is the dominant pattern, seen in six out of eight speakers: DF1, DF2, DF3, DF5, DM2 and DM4. DM5 consistently used an alveolar trill in the onset, whereas DF4 alternated between a uvular and an alveolar variant. These two onset variants may occur with different coda /r/-allophones. For speakers DF5 and DM4, the coda /r/ seems to be a reduced articulatory variant of the onset. Arguably, we can see a similar reduction pattern in DF1, although in this case, the tongue shape for coda /r/ can also be interpreted as a case of weak bunching. For the five remaining speakers, the coda /r/ is a distinct, and not clearly reduced articulation compared to onset. DF2, DF4, DM2 and DM5 have clear bunching of coda /r/ with the characteristic saddle shape. DF3 has a more forward constriction in coda /r/ compared to uvular onset, but there is no clear bunching.

Strikingly, the speakers who display articulatory reduction of coda /r/, DF1, DF5 and DM4, also share another articulatory pattern. For all those speakers, the fake geminate /r/ patterns closely with the onset /r/ at maximum constriction. All the remaining speakers, those with a more distinct coda variant, produce a more intermediate /r/ shape in the fake geminate context, combining features of their onset and coda variants. For instance, DF2 has a fake geminate /r/ that shows raising of the tongue dorsum (typical of a uvular sound), but also raising of the pre-dorsum, as seen in her coda bunched /r/. DM2 also shows this general pattern. DM5 combines raising of the pre-dorsum with tongue tip raising, which is a characteristic of his alveolar onset /r/. DF3 shows some raising of the pre-dorsum in the fake geminate condition compared to her uvular onset variant, but it is less raising than we see in the coda.

Looking at the first frame we analysed for each speaker, we find that vowels preceding onset, coda and fake-geminate /r/ tend to be very similar, except for speakers DF3 and DF4. For these two speakers, the fake geminate context patterns closely with that of the coda, whereas onset remains distinct. At the offset of /r/, the fake geminate /r/ has an average shape that is very similar to the average shape in the onset condition, whereas coda is different. DF2, DF3 and DM5 maintain a three-way distinction until the offset of /r/.

In summary, although we observe different allophony patterns in our speaker sample, fake geminate /r/ tends to follow an articulatory progression from either a neutral, or a coda-like variant early in the vowel, to an onset-like variant later on. For speakers with

\[1\] Trilling is evident in the acoustic signal, but it cannot of course be seen in a single ultrasonic frame.
Figure 4: SS-ANOVA averaged time contours for speakers DF1, DF2, DF3 and DF4 by context and by time point. Tongue tip is on the right.
Figure 5: SS-ANOVA averaged time contours for speakers DF5, DM2, DM4 and DM5 by context and by time point. Tongue tip is on the right.
a relatively weak coda \( /r/ \), the fake geminate becomes fairly onset-like already at the \( /r/ \) maximum. The speakers who have a stronger, bunched, coda \( /r/ \) in their repertoire, tend to produce a fake geminate variant that blends the articulatory characteristics of their onset and coda \( /r/ \). Towards the end of \( /r/ \), the fake geminate may become more like the onset, or it may remain distinct.

### 3.3 PCA-LDA results

A generalized additive mixed model was built to analyse the LD1 scores. Because the LDA had been applied separately for each speaker, and the speakers differ in their allophony patterns, LD1 does not have a consistent articulatory interpretation. Instead, it should be seen as a more abstract similarity scale, where relatively higher LD1 values correspond to more coda-like realizations, whereas relatively lower LD1 values correspond to more onset-like realizations. The predictors included a parametric term of prosodic context (five levels: V#r, VrV, Vr#C, Vr#V, Vr#r), a smooth for normalized time, as well as a smooth for normalized time by context. In addition, we included by-item random smooths for normalized time, as well as random smooths for each speaker at each level of context. We had tested for the effect of a following vowel (\( /a/ \) or \( /i/ \)), and found that it was not significant, based on model comparison. We corrected for autocorrelation by including a first-order autoregressive model (Baayen et al., 2016; Sóskuthy, 2017; Wieling et al., 2016).

The effect of prosodic context in time is plotted in Figure 6. Word-final coda (Vr#C) and word-initial onset (V#rV) are the two most distinct contexts characterized by low and high LD1 scores respectively. Word-medial onset (VrV) generally patterns closely with word-initial onset. Word-final pre-vocalic \( /r/ \) (Vr#V) is similar to canonical word-final coda (Vr#C). Note that word-medial onsets and word-final pre-vocalic codas were new data classified by the LDA-based predictor, and that they are correctly identified by the model as onset-like and coda-like respectively. The fake geminate context (Vr#rV) changes over time: it starts off as an intermediate between onset and coda categories, to become more onset-like over time. This gradual progression is consistent with the results of the SS-ANOVA reported in Section 3.2.

Figure 7 illustrates the individual LD1 trajectories in the production of the fake geminate context. In general, all speakers followed a similar progression in the articulation of the fake geminate, starting off with a more coda-like variant, and then moving towards an onset-like realization. There were, however, individual differences with respect to when the peak coda-like articulation (relatively highest LD1 value) was achieved. For DF1, DF2 and DF5, for instance, this was at the onset of \( /a/ \). In contrast, for DM4 and DM5, there was a peak between 30 and 40% into the \( /ar/ \) sequence. The shape of individual trajectories also varied.
Figure 6: Averaged trajectories of change in LD1 over normalized time depending on context

The dynamic approach presented in this section adds more detailed timing information to the static results of the SS-ANOVA, and it provides a way of obtaining across-speaker means (as illustrated in Figure 6), while still maintaining a way to examine individual differences (Figure 7). While the primary data are considerably transformed in the process, the analysis receives a partial validation from the SS-ANOVA, which, as already noted, replicates the generalization concerning the dynamic progression of fake geminates over the /ar/ sequence (from more coda-like to more onset-like). We may see the two techniques as complementary in the sense that SS-ANOVA provides more detailed spatial information, while the time-series analysis of PCA-LDA results allows us to form more principled dynamic generalizations.

4 Discussion

4.1 Main findings

The articulatory data on /r#r/ sequences suggest that fake geminates show a blending of coda and onset /r/ allophones. The blending is clearly present for the speakers who have a bunched allophone in the coda. Such speakers produce fake geminate /r/ that shows some residual bunching combined with articulatory characteristics of the onset /r/ allophone: dorsum raising for speakers with a uvular onset, or tongue tip raising for the speakers with an alveolar onset /r/. Speakers who do not have a bunched /r/ in the coda, but more of a reduced uvular approximant, show relatively more onset-like
Figure 7: By-speaker trajectories of LD1 over normalized time for the fake geminate context, including a random intercept for item (paar raden vs. paar reizen).
/r/ variants in the fake geminate context at /r/ maximum. Nevertheless, regardless of the specific allophonic pattern, all speakers produce a fake geminate /r/ through articulatory progression from a more coda-like realization to a more onset-like one. Thus, we can generalize that fake geminate /r/ combines articulatory characteristics of both coda and onset /r/. This kind of gestural blending, however, is not clearly reflected in segmental duration. Fake geminate /r/ does not have increased duration compared to singular onset /r/. This observation comes with two caveats. Firstly, we are limited in fully investigating the duration effects by segmentation issues. Since we cannot reliably determine a boundary between the coda /r/ and the preceding vowel, we decided to evaluate the duration based on combined vowel + /r/ sequences. Consequently, if there are durational trade-offs between the vowel and the consonant, we cannot report on them. Secondly, while there is no lengthening in Vr#rV compared to V#rV, neither do we find lengthening comparing the duration of an entire Vr#C sequence to a combined Vr#r. This would suggest that heterosyllabic clusters at word boundaries undergo considerable shortening, which makes the durational reduction of fake geminates less idiosyncratic.

4.2 An overlap-based account

The gestural blending we observe between coda and onset /r/ makes our data highly compatible with an Articulatory Phonology gestural account. In this account, fake geminates behave essentially as a cluster of coda and onset /r/ with a high degree of temporal overlap. This timing is language-specific, although possibly not specific to fake geminates within Dutch. Considering that r#C clusters were equal in duration to fake geminates (r#r), we do not need to posit specific timing rules for fake geminates, although ultimately, this would require support from the duration of other types of clusters. Crucially, coda and onset /r/ are treated as any cluster of distinct consonants would be, without regard for their phonemic status; i.e. the fact that coda and onset /r/ share an identity at a higher level of abstraction is irrelevant to both the gestural and temporal phonetic implementation.

The timing restrictions on the duration of fake geminates (or more generally clusters) result in complete temporal overlap between coda /r/ and onset /r/. A consequence of this is gestural blending, especially affecting the coda component which is not fully realized. Hence, we find residual bunching, as seen for instance in speaker DF2 or DM5, but there is less bunching for fake geminates than there is for canonical coda. It is important to note, however, that coda /r/ is not deleted. That is evident from residual bunching for some speakers, as already discussed, and from coarticulatory influence on the preceding vowel. For instance, speaker DF3 shows a pattern where the pre-/r/ vowel in the fake geminate context patterns distinctly from the onset context,
but together with the coda context. DF1 shows a similar pattern, although the scale is smaller. Note that DF1 has a fake geminate /r/ that looks very much onset-like, but nevertheless, the preceding vowel in the fake geminate context patterns with the coda.

In comparison to the coda component, the onset component is relatively more prominent. This is expected, since anticipatory coarticulation is typically stronger than perseverative coarticulation (Farnetani, 1997). In our case, this effect is even more prominent due to prosodic factors: word-initial /r/ is strengthened, whereas word-final /r/ may be reduced. What is more, the onset /r/ gesture only partially overlaps with the coda one, so towards the end of the /r#r/ sequence, the fake geminate increasingly resembles a canonical onset.

Another pattern we observe in the data is that speakers with bunched /r/ show more clear gestural blending, whereas the speakers who show constriction reduction in the coda, but not bunching as such, have a fake geminate /r/ that is canonical onset-like. A potential explanation for this phenomenon is that bunched /r/ is a relatively stronger articulation, and as such, it is more resistant to coarticulation to the following gesture. This prediction is consistent with the DAC model (degree of articulatory constraint; Recasens et al. 1997) which proposes that greater lingual involvement in a consonantal gesture will make it less susceptible to coarticulation. Extending these predictions to the cases where there is apparently no trace of the coda /r/ in the fake geminate context, we can generalize that coda /r/ is not deleted in these cases, but rather, it is obscured by coarticulation with the onset /r/-gesture. This interpretation is also supported by the results of the dynamic PCA-LDA analysis, which suggest that the dynamic trajectory of fake geminates is similar for all speakers, regardless of their individual allophony patterns.

### 4.3 Autosegmental account

Having seen how the data can be analysed in an overlap-based framework, let us now consider to what extent they can also be analysed within a more abstractionist autosegmental framework. Evidently, a representation that involves a deletion of one of the underlying segments in a fake geminate is not appropriate, as elements of both are present in the articulatory output. An ambisyllabic representation, as traditionally proposed for lexical geminates, does not seem appropriate either. In an ambisyllabic representation, a geminate is a single unit, linked to both coda and onset position. This captures the fact that a geminate may combine onset and coda properties, but it also predicts that a geminate may not be pulled apart. This is true of some types of geminates, e.g. geminates across suffix boundaries in English, as noted by Oh & Redford (2012). Geminates created across word boundaries, on the other hand, may be broken apart, and in that sense, they behave like a cluster.
We do not have systematic evidence from speech style manipulations in our study, but there is little doubt that fake geminate /r/ in Dutch can be broken apart in slow speech. This is also supported by the effect of speech rate on the production of fake geminate fricatives in the study by Martens & Quené (1994). Martens & Quené find that the ratio of consonant to vowel duration for geminates varied between speech rates: it was close to 1 in normal speech, but almost double that in slow speech. This is in stark contrast to singleton onset fricatives, in which the consonant and vowel duration changed proportionally across different speech rates. This asymmetry can be accounted for if fake geminates are analysed as clusters: in slow speech, the individual consonant gestures are longer, and the degree of overlap between them decreases, and hence the duration increase for the fake geminate is greater than that for the preceding vowel.

The considerations of fake geminate separability prompt us to treat fake geminate /r/ in Dutch as a cluster, similarly to the proposal by Ridouane (2010) for Tashlhiyt Berber and by Oh & Redford (2012) for American English. We must, however, note that while we arrive at essentially the same symbolic representation, Dutch still stands in contrast to Berber or English, in the sense that the temporal effect of gemination is very limited. There is no lengthening in fake geminate /r/ compared to singleton onset in our study, whereas for fricatives, Martens & Quené (1994) find 11 ms difference in normal speech. In comparison, fake geminates in Berber, American English, Anglo English and German are approximately double the length of a singleton consonant (Ridouane, 2010; Oh & Redford, 2012; Kotzor et al., 2016).

We propose that the observed differences in the duration of fake geminates in different languages are due to different phonetic implementation rules that determine cluster duration, and in consequence also the degree of overlap between two neighbouring consonant gestures. Where the duration is greater, there is less overlap, whereas when the duration is constrained, neighbouring gestures overlap more. In that sense, the two accounts we have discussed are not so much competing, as complementary. The abstract generalization is that fake geminates are coda-onset clusters, whereas the gestural representation allows us to derive the observed degree of spatial overlap, subject to language-specific timing constraints on gestures.

4.4 Residual issues

What remains to be determined is whether the constraints on degree of overlap are specific to fake geminates. We do not have sufficient evidence to propose this based on our current study, since the duration of fake geminates did not differ from r#C clusters. However, evidence from different types of clusters is needed before we decide whether fake geminates are in any sense unique. This is true of Dutch as well as other languages. Earlier studies have proposed to treat fake geminates as clusters, based on
comparison with lexical geminates. Further comparisons with different types of clusters are necessary to fully understand the implications of this analysis.

Furthermore, we must consider to what extent our discussion of fake geminates in Dutch can be generalized beyond the case of /r/. It is certainly the case that /r/ is special, in that it has a very distinct allophony pattern. A reviewer suggests that another special feature of /r/ is that the consonantal status of the coda allophone is not obvious. The ongoing increase of approximant coda /r/-variants in Dutch may signal ongoing coda vocalization, similarly to English /r/. While it is certainly true that the phonetic identity between onset and coda /r/ in Dutch is being obliterated, there are no reasons to doubt that coda /r/ is there at the phonological level. One argument supporting this view is alternations: coda /r/ regularly alternates with an onset in plural forms such as *paar*∼*paar-en* ('pair'∼'pairs'), in which case the clearly consonantal onset allophone is substituted for the coda one. Furthermore, even if we can speculate that /r/ in Dutch is headed for a vocalic secondary articulation type variant, it has not reached that stage yet. Note that, in our data, we find a variety of coda variants, which can be interpreted as representing different stages of change: some speakers have a coda /r/ that seems to be (DF5, DM4) a slightly reduced variant of their onset uvular fricative, while others have clearly adopted the new bunched variant (DF2). For the former, more conservative, types of speaker, the articulatory coda reduction is arguably undershoot of the fricative constriction. Importantly, coda and onset are distinguished in this case not by the type of gesture (consonantal vs. vocalic), but rather, by gestural magnitude. Hence, we cannot conclude that for such speakers, the consonantal gesture is erased and substituted by a vocalic one.

Note also that the two types of speakers are much alike in their treatment of fake geminates, which is consistent with treating them as an /r#r/ cluster in all cases. The main difference between the two types of speakers is that the ones with uvular coda reduction produce an /r/ in the Vr#r context that is completely onset-like. In contrast, the speakers with a clearer bunched variant, maintain a three-way distinction between coda, onset and fake-geminate context at the point of maximum /r/-constriction. As argued previously in Section 4.2, this can be independently explained by the coarticulatory resistance of the bunched variant. This argument could be strengthened if we also find a similar pattern in speakers with an alveolar variant in both onset and coda. This is a possible pattern in Standard Dutch, albeit a conservative one, and hence it is not surprising that we did not find it among the speakers in our study. Alveolar /r/ is clearly consonantal (since it involves tongue tip raising). According to our interpretation of the degemination patterns in Standard Dutch, we predict that such speakers will have no duration distinction between Vr#r and V#r, that their fake geminate /r/ will be alveolar, and that the vowels preceding fake geminate /r/ may pattern with
vowel preceding coda /r/, to the exclusion of vowels in the V#rV context.

5 Conclusion

The data presented in this paper show that the r#r fake geminate in Dutch does not undergo degemination in the sense of deletion of one of the two elements in the cluster. That is, there is no categorical degemination at the highest (lexical, segmental) phonological level, which is the level at which coda /r/ and onset /r/ are identical. Instead, gestures of both coda and onset /r/ are preserved. However, we do find durational reduction of the cluster to the effect that the fake geminate is not longer compared to a singleton onset /r/. This is manifested as gestural overlap and reduction. Dutch differs here from other languages in which fake geminates have also been analysed as clusters and these show increased duration vis-à-vis singleton consonants. It is unclear whether the durational reduction itself is a form of degemination, operating at the temporal but not the gestural level. The fact that there is also no durational difference between r#r clusters and other r#C clusters suggests that the temporal reduction is not limited to the fake geminate situation. Based on this, we argue that the r#r fake geminate is best analysed as a cluster.

The Dutch fake geminate case also shows the importance of dynamic articulatory data in the analysis of sandhi phenomena. Apart from revealing the gestural overlap in the first place, it elucidates the complexities involved: the different strategies speakers employ to effectuate gestural overlap, as well as the similarities between speakers with different allophonic patterns.

Acknowledgments

We wish to thank our participants, and colleagues who helped at various stages of this project: Chris Carignan, Stefano Coretta, Steve Cowen, Jim Scobbie and Alan Wrench. We are grateful to the editor, Marija Tabain, and three anonymous reviewers for their helpful comments. We also thank audiences at Ultrafest VII, Edinburgh University P-workshop and LabPhon15 for comments and discussion. The first author acknowledges the support of British Academy grant PDF/pf130029.

References


