Phonology **18** (2001) 387–420. © 2001 Cambridge University Press
DOI: 10.1017/S0952675701004213 Printed in the United Kingdom

*The phonological status of Dutch epenthetic schwa**

Natasha Warner
University of Arizona and Max Planck Institute for Psycholinguistics, Nijmegen

Allard Jongman
University of Kansas

Anne Cutler
Max Planck Institute for Psycholinguistics, Nijmegen

Doris Mücke
University of Cologne

In this paper, we use articulatory measures to determine whether Dutch schwa epenthesis is an abstract phonological process or a concrete phonetic process depending on articulatory timing. We examine tongue position during /l/ before underlying schwa and epenthetic schwa and in coda position. We find greater tip raising before both types of schwa, indicating light /l/ before schwa and dark /l/ in coda position. We argue that the ability of epenthetic schwa to condition the /l/ alternation shows that Dutch schwa epenthesis is an abstract phonological process involving insertion of some unit, and cannot be accounted for within Articulatory Phonology.

1 Introduction

In this paper, we use articulatory measures to determine whether a particular alternation, schwa epenthesis in Dutch, is a phonological or a phonetic process. Although it is often very difficult to tell whether a given

* We would like to thank Bryan Gick and Taehong Cho for advice on articulograph methodology and for discussion of this material, as well as three anonymous reviewers for their very helpful comments. We are also grateful to audiences at UCLA, the University of Arizona, LabPhon VII and the Max Planck Institute for Psycholinguistics for their comments on earlier versions of this work. We would like to thank Theo Klinker, Niels Janssen, Keren Shatzman, Rachèl Kemps, Aoju Chen, Tau van Dijck and Anne Pier Salverda for their help with various aspects of this project, and Inge Doehring and Herbert Baumann for technical assistance. Finally, we are particularly indebted to our speakers. Any errors are, of course, our own responsibility. This work was carried out while the second author was a visitor at the Max Planck Institute.

alternation is the result of a phonological rule or constraint, or simply reflects variation in the phonetics of speech production, facts about Dutch phonology make this particular case amenable to experimental investigation. In addition to testing the phonological status of the alternation, our results have implications for the theory of Articulatory Phonology, particularly for the claim that all gestures which are present in any surface form of a word must also be present in the underlying representation. The first section of the paper provides background on several topics which are important for the study, and also presents two possible hypotheses for how Dutch schwa epenthesis takes place.

1.1 Dutch schwa epenthesis

In Dutch, a schwa is variably inserted into a consonant cluster, between /l/ or /r/ and a following non-coronal consonant (1a), as well as in the cluster /rn/. Epenthesis is possible whether the consonant cluster is entirely within the coda of one syllable, or crosses a syllable boundary (1b), although it is more common in the former environment and may be non-standard in the latter (Collins & Mees 1999, Swerts *et al.* 2001, Kloots *et al.*, in press).

(1)	a.	*melk*	/mɛlk/	[mɛlk] ~ [mɛlək][1]	'milk'
		wilg	/ʋɪlx/	[ʋɪlx] ~ [ʋɪləx]	'willow'
		hulp	/hʏlp/	[hʏlp] ~ [hʏləp]	'help'
		berg	/bɛrx/	[bɛʁx] ~ [bɛʁəx]	'mountain'
		korf	/kɔrf/	[kɔʁf] ~ [kɔʁəf]	'basket'
	b.	*filmer*	/fɪlmər/	[fɪlməʁ] ~ [fɪləməʁ]	'cameraman'
		ergens	/ɛrxəns/	[ɛʁɣəns] ~ [ɛʁəɣəns]	'somewhere'

Schwa epenthesis is an optional and highly variable process. Dialect is clearly a factor (Swerts *et al.* 2001, Kloots *et al.*, in press), but even within a dialect and within a speaker there is considerable variation. Some speakers epenthesise very often, and some rarely, but many epenthesise variably (Kuijpers & van Donselaar 1998). Epenthesis is apparently not conditioned by register or speech rate, and at least in word-final position in dialects within the Netherlands, it is not sociolinguistically marked (Collins & Mees 1999). It can frequently be heard in newscasts as well as in casual conversation. Van Reenen & Jongkind (2000), Swerts *et al.* (2001) and Kloots *et al.* (in press) show that epenthesis is common among Dutch speakers of all ages, and among both men and women.

Kuijpers & van Donselaar (1998) show that surrounding prosody influences the likelihood of epenthesis, at least word-finally. They placed a word with the environment for epenthesis, e.g. *tulp* /tʏlp/ 'tulip', after either several strong syllables or several strong–weak feet.

[1] In phonetic transcriptions, dark [ɫ] will not be marked except where articulatorily documented, since the darkness and lightness of /l/ in particular environments is what our experiment tests.

(2) a. Naast al dat onkruid zag ik nog maar één tulp.
/nast ɑl dɑt ɔnkrœyd zɑx ɪk nɔx mar en tʏlp/ (*strong syllables*)
'Alongside the weeds I saw just one tulip.'

b. Tussen deze bloemen stond een hele mooie tulp.
/tʏsən dezə blumən stɔnd ən helə mojə tʏlp/ (*strong–weak feet*)
'Amidst these flowers there was a very nice tulip.'

Speakers were more likely to produce epenthesis after strong–weak feet than after exclusively strong syllables. Thus speakers are more likely to epenthesise if this will produce a form which matches the prosodic structure of surrounding words. The preference for epenthesis where it will match surrounding foot structure held true in Kuijpers & van Donselaar's (1998) study separately for speakers who epenthesise quite frequently, those who rarely epenthesise, and those who epenthesise most variably. However, for words with a non-final epenthesis environment (e.g. *tulpen* /tʏlpən/ 'tulips'), placing the word after strong–weak feet (which would match the non-epenthesis form [tʏlpən]) *vs*. after strong–weak–weak feet (which would match the epenthesis form [tʏləpən]) had no significant effect. Epenthesis was, overall, far less likely in such words.

Psycholinguistic investigation of Dutch schwa epenthesis has indicated that the form without epenthesis is the more basic of the two alternants (van Donselaar *et al*. 1999). This conclusion is based on a study in which subjects were trained to produce words backwards by reversing the string of syllables for polysyllabic words (e.g. *hotel* /hotɛl/ 'hotel' becomes /tɛlho/), but the string of segments within the syllable for monosyllabic words (e.g. *tap* /tɑp/ 'beer tap' becomes /pɑt/). When presented with forms such as either [tʏlp] or [tʏləp], subjects produced /plʏt/ significantly more often than /ləptʏ/, suggesting that the mental representation of the word is /tʏlp/ rather than /tʏləp/ (although this result could be influenced by orthographic knowledge). Goetry *et al*. (in press) provide evidence from preliterate children's judgements of syllable counts which suggests that even before learning the orthography, the form without epenthesis is more basic, at least for most children. Children usually judge non-epenthesis forms such as *melk* [mɛlk] 'milk' as containing one syllable, and also judge epenthesis productions such as [mɛlək] to contain just one syllable approximately half the time (although conditions with longer words show that they can perform the task). Traditional linguistic criteria would also lead to the conclusion that the non-epenthesis form is basic: the form [tʏləp] is predictable from the underlying representation /tʏlp/, but the form [tʏlp] would not be predictable from an underlying representation /tʏləp/, because there are minimal pairs which differ in whether schwa is variable or not (3):

(3)	*wilg*	/ʋɪlx/	[ʋɪlx] ~ [ʋɪləx]	'willow'
	willig	/ʋɪləx/	[ʋɪləx] (*[ʋɪlx])	'willing'

Van Donselaar *et al*. (1999) further find that even though the non-epenthesis form is more basic, listeners are faster to recognise the word when it is produced with epenthesis than when it is produced without. They attribute this result to the /l/ having stronger acoustic cues when it is prevocalic, in the epenthesis form, than when it is preconsonantal. In support of this claim, they show that listeners are also faster to spot the /l/ (in a phoneme-monitoring task) in forms with epenthesis.

1.2 Articulatory Phonology

The theory of Articulatory Phonology (henceforth AP) (Browman & Goldstein 1990a, b, 1992a, b, 1995, Gick 1999a, b) unifies many aspects of phonetics and phonology by proposing that underlying representations consist of articulatory gestures rather than phonemes, features, CV slots, etc. The assumption that gestures are the basic units of speech allows the theory to model many alternations in speech elegantly, and to account for otherwise surprising experimental data. For example, articulatory measures show movement of the tongue tip toward the alveolar target of the /t/ in a phrase like *perfect memory*, even if the cluster is simplified such that there is no acoustic trace of the /t/ (i.e. [pʰɚfəkmɛmɚi]) (Browman & Goldstein 1990b). Browman & Goldstein show that this is because the tongue tip gesture is completely overlapped by the preceding velar gesture and the following labial gesture.

To constrain the alternations which manipulation of gestures would allow, researchers have proposed that speech variation, at least allophonic and speech-rate variation, involves only adjustments to the size and timing of underlying gestures. That is, gestures which are present in the underlying representation can be extended in time to overlap other gestures, and their magnitude can be increased or reduced, but no gestures can be added to or deleted from those in the underlying representation. Any gesture which occurs in the production of any form of a word must also be present in the word's underlying form. (Deletion of gestures might be possible as an extreme case of reduction of size of the gesture, however.) Gick (1999a: 3) makes this position explicit: 'According to the principles of AP, all gestures present in the phonetic output of a word are specified in its lexical representation; thus, while the magnitude or timing of a gesture may vary, gestures may not be inserted. It is in this way that the traditional domains of phonetics and phonology are inextricably linked in this framework'. Because this point is important for the hypotheses tested in the current paper, we will discuss past research on gestural insertion or deletion in some detail.

1.2.1 *Modelling apparent insertions and deletions in AP*. Considering the range of processes which seem to involve insertion or deletion of segments in speech, the claim that gestures cannot be inserted or deleted might seem extreme. However, AP has proven capable of modelling many such processes. For example, Gick (1999b) discusses the apparent insertion of /r/ in the non-standard American English form *warsh* for *wash*: if the pharyngeal gesture for /ɑ/ overlaps the palatal and labial gestures for /ʃ/, there will be labial, palatal and pharyngeal constrictions in the vocal tract at the same time, which are characteristic of the American /r/. Although listeners and learners may reinterpret the acoustic result as including gestures for an /r/ and thus arrive at a new underlying representation, it is not necessary to insert any gestures in order to account for the apparent epenthesis of a segment. Epenthetic stops between a nasal and an obstruent (e.g. [t] in *prince*) are amenable to a similar analysis.

Gick (1999b) also discusses coda /r/-deletion and prevocalic /r/-insertion in a variety of dialects, and argues that in dialects where /r/ seems to be inserted between vowels even though it is not historically expected (e.g. *idea*[ɹ] *is* but *idea*[ə] *which*), all words which are subject to the alternation have the gestures for /r/ in their underlying representation. These gestures are reduced in magnitude and moved toward the edge of the syllable when in coda position, making them inaudible. In syllable onset position, the /r/ gestures are not reduced. Thus, no insertion of gestures is necessary to produce the alternation. Gick *et al*. (in press) present further evidence for the gestural similarity of [ə] and [ɹ], strengthening the case for why this particular alternation should occur.

Jannedy (1994) discusses a case somewhat similar to Dutch schwa epenthesis, slow speech insertion of schwa in German, by which *braten* 'to roast' can become homophonous with *beraten* 'to advise'. She presents evidence based on gradience of misperceptions that this speech-rate alternation involves not the insertion of a schwa segment, but rather a re-timing of the existing gestures. In slow speech, the labial gesture for /b/ pulls apart in time from the gestures for /r/, leaving an intervening time without gestural specifications. A short voiced sound without gestural specifications sounds like a schwa ('targetless schwa'; Browman & Goldstein 1992a), leading to the percept of schwa insertion. This is not entirely parallel to Dutch schwa epenthesis, since the German epenthesis only occurs in slow speech, but it represents another case in which apparent insertion of segments can plausibly be achieved without insertion of gestures. Gick & Wilson (2001) consider a different type of schwa insertion, the type which makes words such as *heel*, *hail* and *hire* bisyllabic in some dialects of English. They argue that this involves not the insertion of schwa gestures, but rather the tongue moving through the position it would occupy for a schwa during the transition from the tongue-root position for the preceding high front vowel to that of the following liquid.

Browman & Goldstein (1990a, 1992b) discuss the inverse of Jannedy's case, schwa deletion in fast speech (e.g. *beret* becoming homophonous with *bray*), as an example of apparent deletion of a segment without deletion of gestures. They present evidence that such reduction is gradient, and that it involves the labial gesture of the /b/ and the gestures of the /r/ coming to overlap, rather than a categorical deletion of a segment or syllable nucleus position (Browman & Goldstein 1990a). The *perfect memory* cluster-simplification example mentioned above, as well as cases of place assimilation across word boundaries (e.g. *seven plus*, in which the cluster sounds like [mp] rather than [np] because the labial gesture overlaps the alveolar gesture; Browman & Goldstein 1990b, 1992b), provide further examples of how the apparent deletion of a segment or a feature can be modelled through gestural overlap without deletion of any gestures.

1.2.2 *Challenges to the claim of no gestural insertion or deletion.* Kohler (1992) mentions the intrusive /r/ of some English dialects as an alternation which might require gestural insertion, and McMahon and colleagues (McMahon *et al.* 1994, McMahon & Foulkes 1995) take this topic up in more detail. Both argue that the underlying representation of words with no historical /r/, such as *idea*, must not include an /r/ even in dialects where an /r/ is inserted if the word is followed by a vowel (i.e. *idea*[ɹ] *is*). If the gestures for /r/ are not underlyingly present in words such as *idea* in these dialects, then insertion of gestures would clearly be necessary to produce the prevocalic forms. However, Gick (1999b) (see above) not only argues that it is reasonable to assume the /r/ is underlying, but also presents articulatory data showing that the particular pattern of gestural timing in the /r/ alternation parallels the behaviour of other sonorant consonants in these positions.

Kingston & Cohen (1992) argue that glottal adduction during final stops in English requires gestural insertion. Browman & Goldstein (1992b) treat the allophonic variation between aspirated and voiceless unaspirated stops in English as adjustments to the timing and magnitude of the glottal abduction gesture relative to the oral gesture of the stop, but glottalised syllable-final stops require active closure of the glottis, which cannot be viewed as a reduction of the underlying abduction gesture. Kohler (1992) also mentions morphophonological variation as a class of phenomena which would often require deletion, insertion or replacement of gestures. There are certainly morphophonological alternations which involve changes to the speech string too radical to treat either by adjustments to timing and size of gestures or by including in the underlying representation all the gestures present in any form of the morpheme. To avoid gestural reorganisation, one would either have to posit multiple lexically listed allomorphs, or allow a more abstract stage of (morpho-)phonology in addition to those aspects of phonology and phonetics which are treated by AP (Kohler 1992). However, research in AP has mostly addressed cases of allophonic or speech rate variation, or such relatively transparent phenomena as /r/ intrusion. It is not surprising if more abstract alternations cannot be treated as adjustments to gestures.

1.3 Articulation of /l/

Sproat & Fujimura (1993), Browman & Goldstein (1995) and Gick (1999a, b) provide a wealth of information on how English /l/ is articulated. The most important finding for the current study is that the articulation of /l/ involves two component gestures, a raising and/or fronting of the tongue tip and a lowering and/or backing of the tongue dorsum or body. Sproat & Fujimura (1993) term the former the 'consonantal gesture' of the /l/ and the latter the 'vocalic gesture' of the /l/. They find that both light and dark realisations of /l/ involve both of these component gestures, but that in light /l/, the consonantal gesture is stronger, and the vocalic gesture is reduced in magnitude. The reverse is true of dark /l/. Furthermore, they find a difference in gestural timing between light and dark /l/: in a light /l/, the two component gestures are timed close together, while in a dark /l/ the consonantal (tip) gesture is delayed relative to the vocalic gesture.

Krakow (1989) found, similarly, that velum lowering is earlier relative to the oral closure gesture in word-final nasals than in word-initial nasals. Gick (1999a, b) shows that such syllable position-dependent patterns of gestural timing are typical for sonorant segments involving more than one gesture. The pattern also holds for /w/, and depends on syllable structure: consonantal gestures are attracted toward the edge of the syllable (away from the gestures of the preceding vowel), while vocalic gestures are attracted toward the nucleus (toward the gestures of the vowel). Thus, consonantal gestures, when in coda position, move away from the nucleus toward the end of the syllable, and if they are also reduced in magnitude, this can lead, in extreme cases, to vocalisation of the /l/ in coda position. Krakow (1999) gives a useful review of the evidence regarding syllable position effects on gestures. Sproat & Fujimura (1993) also show that variation between light and dark /l/ in English is gradient rather than categorical, and is not conditioned exclusively by syllable position. For example, word-final /l/ is darker before stronger prosodic boundaries.

Work on /l/ in Dutch has been more sociolinguistic than phonetic, and we do not know of any studies which establish whether the results on English /l/ also hold for Dutch. Studies in which speakers make judgements about their articulations do show that Dutch speakers often lack tongue-tip contact in /l/ in coda position (van Reenen 1986, van Reenen & Jongkind 2000). This does not preclude the possibility of a tongue-tip gesture of reduced magnitude, but it does indicate that the alternation between Dutch light and dark /l/ is in general articulatorily similar to English.

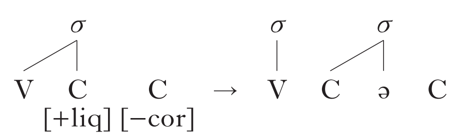
The exact environment for the light/dark /l/ alternation in Dutch is not clear. Booij states that /l/ is 'rather clear' and has alveolar contact word-initially before vowels, and is dark both in coda position and intervocalically (1995: 8). Van Reenen (1986) and van Reenen & Jongkind (2000) do not investigate intervocalic /l/, but they describe the environment for light /l/ as prevocalic, implying that intervocalic /l/ would be light. Collins & Mees (1999) state that /l/ is dark in coda position and generally light prevocalically, but that intervocalic /l/ varies, tending to be dark after non-high back vowels but otherwise light. They also mention that some dialects have dark /l/ in all environments, while others have light /l/ in all environments. Perhaps the disagreement about the darkness of intervocalic /l/ reflects a gradient alternation, as in English. It is worth noting that there was also lack of agreement about the exact environment for this alternation in English before Sproat & Fujimura (1993) clarified the situation. Krakow (1989) also documented substantial variation in whether intervocalic nasals in English pattern articulatorily with syllable-initial or syllable-final nasals. In any case, it seems that for many Dutch speakers intervocalic /l/ may be, like word-initial /l/, lighter than coda /l/.

1.4 Two hypotheses for Dutch schwa epenthesis

In this paper we consider two possible accounts, one within the theory of AP and one not, of how schwa epenthesis should be incorporated into the grammar of Dutch. An AP account would be parallel to the account Jannedy (1994) provides for German slow speech schwa epenthesis, discussed above. In a word such as *film*, when produced as [fɪləm], the tongue gestures for the /l/ pull apart in time from the lip gesture for the following /m/, leaving a period of time with voicing but without any oral gestural specifications. This part of the signal would sound like a schwa, although it would have no gestural targets. (Although Browman & Goldstein (1992a) find that underlyingly present schwa does have gestural specifications, in that its tongue position cannot be entirely predicted from tongue position of surrounding sounds, they also find that a brief time period with no gestural specifications is perceived as a schwa. Gick (in press) also finds that schwa does involve an active gesture, namely a pharyngeal constriction.) This account, represented schematically in Fig. 1, assumes that gestures for the epenthetic schwa are not underlyingly present, and requires no insertion of gestures to produce epenthesis.

The alternative hypothesis is that schwa epenthesis does involve insertion of some linguistic unit. This alternative is represented in (4).[2]

(4)



Under this hypothesis, some linguistic unit with schwa as its realisation is inserted into clusters of a liquid followed by a non-coronal consonant,

[2] With the schematic representation in (4), we do not mean to make any claims about how syllables are structured. (The medial consonant in epenthesis forms might be ambisyllabic, for example, and this is not represented here.) We mean only to show that a schwa is inserted, and that this adds a syllable. We also do not mean to address rule-based *vs.* constraint-based approaches to phonology; this insertion is depicted as a rule here, but could be done in Optimality Theory as well.

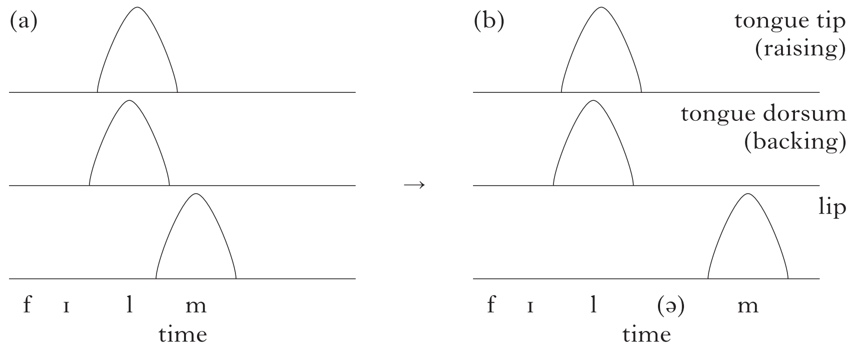


Figure 1
Schematic representation of schwa epenthesis using only adjustments to the timing of gestures. In (a), a non-epenthesis production, the onset of the labial gesture overlaps the end of the lingual gestures for the /l/. In (b), where epenthesis occurs, the labial gesture does not begin until after the lingual gestures have ended, leaving a gesturally unspecified time period.

and the result is a form with an additional syllable, with schwa as the new syllable's nucleus. This latter hypothesis would not be possible within a standard view of AP. There are several possibilities for what might be inserted, such as a collection of gestures for the schwa, the phoneme schwa, a V slot (an underspecified vowel) or a syllable nucleus. (In some phonological theories, these may not all be distinct from each other.) The first of these, a collection of gestures, would be consistent with AP's view of the gesture as the basic unit of phonology, but would present a problem for the theory's claim that gestures cannot be inserted between the underlying and surface representations. Insertion of a schwa phoneme or an underspecified vowel would be problematic because AP does not recognise higher-level units such as segments. Rather, in AP, several gestures may stand in a particular timing relationship, but they do not form a segmental unit as such. Inserting a syllable nucleus which would be filled in by default material would be equally problematic, because AP does not recognise the syllable as a unit of phonological representation. Furthermore, gestures can be timed with respect to other gestures, but not with respect to syllables.

The light/dark /l/ alternation in Dutch makes it possible to test these two hypotheses (retiming *vs.* insertion of some phonological unit) experimentally. If the /l/ in forms with epenthesis, e.g. /fɪlm/ [fɪləm], is realised as a relatively light /l/, while /l/ in non-epenthesis forms (e.g. [fɪlm]) is dark, this would indicate that the variation in realisation of /l/ in the two pronunciations of the word is conditioned by the presence of the epenthetic schwa. In order for the schwa to condition allophonic variation, it must be present as a phonological unit, because the allophonic variation involves timing of the /l/ gestures relative to the vowel. If the schwa were simply a period of time without gestural specifications, which happens to be interpreted perceptually as a schwa-like sound, this targetless schwa would be merely a perceptual epiphenomenon, not a linguistic unit, and could not possibly condition allophonic variation. In AP, syllables are defined by the phasing relations of the gestures of consonants and vowels. Consonant gestures are phased relative to the gestures of the nuclear vowel (Browman & Goldstein 1995). (Gick 1999a discusses phasing of onset consonant gestures to the gestures of the following vowel, of coda consonant gestures to the preceding vowel and of ambisyllabic consonant gestures to both surrounding vowels, for example.) Since schwa epenthesis by separation of the two consonant gestures in time would create a schwa-like sound without any vocalic gestures, the gestures of the preceding /l/ could not be phased to the epenthetic schwa. Therefore, simply separating the gestures of the /l/ and the following consonant in time could not cause allophonic variation in the /l/.

On the other hand, if schwa epenthesis involves the insertion of some linguistic unit, such as the schwa phoneme, one would expect allophonic variation in the /l/. As shown in (4), insertion of the schwa implies the addition of a syllable, with the schwa as its nucleus. If the epenthetic schwa is an inserted vowel, it should condition variation in the /l/ just as an underlying vowel would. Van Donselaar *et al*. (1999) imply that they support the view of schwa epenthesis as insertion of a linguistic unit, since they conclude that increased acoustic perceptibility of /l/ prevocalically is the reason for listeners' faster detection of /l/ and recognition of words in the epenthesis pronunciation. However, they do not test directly what makes the /l/ more perceptible before epenthetic schwa than in coda position. Furthermore, considering past work in AP which shows that speakers often continue to produce gestures even in environments where those gestures can have no acoustic effect (e.g. the *perfect memory* case; Browman & Goldstein 1990b), it is very difficult to draw conclusions about whether the articulatory gestures reflect a light or a dark /l/ solely on the basis of acoustic or perceptual data.

Therefore, we designed an experiment to determine what articulations are involved in the production of /l/ in forms with and without epenthesis in Dutch. To allow for the possibility that /l/ before epenthetic schwa might be like neither coda /l/ nor intervocalic /l/ between underlying vowels, we included a control condition, phonetically similar to an epenthesis production, but with the schwa underlyingly present in the word. Thus, we compared triples consisting of words such as *film* /fɪlm/ 'film' produced with and without epenthesis ([fɪlm, fɪləm]) and a matched word *Willem* /ʋɪləm/ [ʋɪləm] 'William'. If schwa epenthesis involves insertion of a phonological unit, /l/ in epenthesis forms and /l/ before underlying schwa should both be light, while /l/ in non-epenthesis forms (in coda position) should be dark. However, if epenthesis involves only retiming of gestures and not insertion of a schwa, /l/ in epenthesis forms should be articulatorily identical to the dark /l/ in non-epenthesis forms (because gestures of the /l/ would still be phased to the preceding vowel, just as they are in non-epenthesis productions). The /l/ before underlying schwa would be light and different from the other two conditions, however.

2 Methods

2.1 Materials

We chose 24 Dutch words with an environment for epenthesis after /l/ and 24 phonologically matched words with an underlying schwa in place of epenthesis (5). In order to minimise influence of following sounds on the articulation of /l/, only words with labials or labiodentals as the next consonant after the /l/ were used. Although epenthesis is more common after /r/ than /l/ (Swerts *et al*. 2001, Kloots *et al*., in press), we restricted the investigation to epenthesis after /l/ rather than /r/ because of limitations of the articulograph methodology for obtaining data about articulations at the back of the mouth, and because of the importance of uvular or pharyngeal articulation for several allophones of /r/ in Dutch. Although the environment for schwa epenthesis is very common in the Dutch lexicon, the need for a following labial and a matched underlying schwa word made appropriate pairs rare. These 24 pairs may represent all relevant items in the Dutch lexicon.[3]

(5) a.

Epenthesis words			*Underlying schwa words*		
film	/ˈfɪlm/	‘film’	*Willem*	/ˈʋɪləm/	‘William’
olm	/ˈɔlm/	‘elm’	*column*	/ˈkɔləm/	‘column’
holm	/ˈhɔlm/	‘spiral’	*Hollum*	/ˈhɔləm/	(place name)
molm	/ˈmɔlm/	‘mulch’	*Gollem*	/ˈxɔləm/	‘Gollum’
helm	/ˈhɛlm/	‘helmet’	*cerebellum*	/ˌserəˈbɛləm/	‘cerebellum’
schelm	/ˈsxɛlm/	‘rogue’	*Zelhem*	/ˈzɛləm/	(place name)
kalm	/ˈkɑlm/	‘calm’	*stal’m*	/ˈstɑləm/	‘put it up’
galm	/ˈxɑlm/	‘echo’	*val’m*	/ˈvɑləm/	‘fell him’
zalm	/ˈzɑlm/	‘salmon’	*zal’m*	/ˈzɑləm/	‘shall him’

[3] Some factors are not matched across the test words. For example, a few items have additional syllables before the matched /VləC/ test sequence, and the vowels in the syllable after the /VləC/ sequence are usually full vowels with secondary stress in the underlying schwa condition, but are unstressed schwa in the epenthesis words. However, the degree of matching across conditions we have achieved (identical /VləC/ sequence with primary stress on the vowel preceding /l/) should remove most irrelevant variation. Any effects of, for example, the presence or absence of a syllable with secondary stress elsewhere in the word should be far smaller than the effects of whether /l/ is followed by a consonant or by an epenthetic or underlying schwa. With the additional restrictions discussed in the text (e.g. the requirement that the following consonant be labial), these pairs are the best matches the Dutch lexicon allows.

b.

filmer 'cameraman'	/ˈfɪlmər/	*brillenmaker* 'optician'	/ˈbrɪləˌmakər/
Wilmer (name)	/ˈʋɪlmər/	*schillenmand* 'peel basket'	/ˈsxɪləˌmɑnd/
olmen 'elms'	/ˈɔlmən/[4]	*dolleman* 'madman'	/ˈdɔləˌmɑn/
holmen 'spirals'	/ˈhɔlmən/	*jolleman* 'ferryman'	/ˈjɔləˌmɑn/
bedwelmen 'drug'	/bəˈdʋɛlmən/	*bellenman* 'bell-man'	/ˈbɛləˌmɑn/
Wilhelmus (name)	/ʋɪlˈhɛlmʏs/	*ellenmaat* 'measure'	/ˈɛləˌmat/
kalme 'calm' (M/F)	/ˈkalmə/	*mallemolen* 'merry-go-round'	/ˈmɑləˌmolə/
halve 'half' (M/F)	/ˈhɑlvə/	*ballenvanger* 'catcher'	/ˈbɑləˌvɑŋər/
delven 'dig'	/ˈdɛlvən/	*hellevorst* 'prince of darkness'	/ˈhɛləˌvɔrst/
elven 'elves'	/ˈɛlvən/	*helleveeg* 'shrew'	/ˈhɛləˌvex/
dolven 'dug'	/ˈdɔlvən/	*drollenvanger* 'baggy trousers'	/ˈdrɔləˌvaŋər/
golven 'waves'	/ˈxɔlvən/	*molleval* 'mole-trap'	/ˈmɔləˌvɑl/
kolven 'cobs'	/ˈkɔlvən/	*bollenveld* 'bulb-field'	/ˈbɔləˌvɛld/
wolven 'wolves'	/ˈʋɔlvən/	*mollevel* 'moleskin'	/ˈmɔləˌvɛl/
kolvers 'milkers'	/ˈkɔlvərs/	*schollevaar* 'large cormorant'	/ˈsxɔləˌvar/

Each pair has the same vowel preceding /l/, and the same consonant following the schwa or the /l/. The vowel preceding /l/ is always the one with primary stress. It was only rarely possible to also match the initial consonant of the word, or even its place of articulation. For the nine pairs in (5a), the environment for epenthesis and the matched underlying /CəC/ occur at the end of the word. We will refer to these as 'short' words. For the fifteen pairs in (5b), at least one additional syllable follows the relevant environment, and we will refer to these as 'long' words.

[4] Morpheme-final /n/ after schwa is usually deleted in Dutch, although in this careful speech situation, our speakers produced word-final /n/ relatively often. Many of the underlying schwa words orthographically contain such an /n/ between their morphemes, e.g. *brillenmaker* 'optician', which would alter the /Cə/-labial environment if produced. Speakers were reminded that the *n* is not normally pronounced in such words. However, it is of little consequence if speakers produce word-final /n/ in the epenthesis words.

Although epenthesis is more common in shorter words (Kuijpers & van Donselaar 1998), very few matched pairs exist in this category, so the longer words were used as well. A few of the items (/stɑləm, vɑləm, zɑləm/ in the underlying condition) actually consist of a verb followed by a clitic pronoun, rather than of a single word.

Because comparison across speakers is often difficult for articulatory measures, it was important to elicit the epenthesis environment words from each speaker both with and without epenthesis. We placed each word containing an epenthesis environment into two short phrases (6), with the prosody of one favouring epenthesis, and the prosody of the other disfavouring it (Kuijpers & van Donselaar 1998). Each phrase consisted of two frame words with the target word between them, and was semantically possible. The frames were varied to avoid special prosody that might be induced by a repetitive frame sentence. A full list of items and frames appears in the Appendix. (A few additional items, also listed in the Appendix, were included in case speakers did not epenthesise spontaneously often enough in the longer words.)

(6) a. geen film zien /xen fɪlm zin/
'see no movie' (*epenthesis disfavoured*)

b. samen film huren /samən fɪlm hyrən/
'rent a movie together' (*epenthesis favoured*)

c. nieuwe olmen planten /niʋə ɔlmən plɑntən/
'plant new elms' (*epenthesis disfavoured*)

d. droevige olmen jammeren /druvəxə ɔlmən jɑmərən/
'sad elms moan' (*epenthesis favoured*)

The underlying schwa words have only one possible pronunciation, so each was placed in a single frame.[5]

Short words (5a) produced without epenthesis were subsequently excluded from the data, because the /l/ is only separated from the first consonant of the following frame word by the final labial (e.g. [ɔlm] in *geen olm daar* /xen ɔlm dar/ 'no elm there'), so the gestures of the /l/ were strongly influenced by the gestures of the consonant in the following word. Thus, for the coda /l/ condition, only the longer words (5b) could be used. The epenthesis productions, the underlying schwa words and the longer words (5b) are all exempt from this problem, since they have a schwa separating the /l/ from the gestures of the following consonants (e.g.

[5] Since the syllable count of the frame words was not being used to manipulate pronunciation of the underlying schwa words, as it was for the epenthesis words, the exact prosodic characteristics of the frames for underlying schwa words were allowed to vary more than those for the epenthesis words. The frames for underlying schwa words were chosen to be similar to other frames in the experiment. The number of syllables in the frames for underlying schwa words is unlikely to have had any substantial effect on articulation of the /l/, as the /l/ in these words is not variable.

[ɔləm] in an epenthesis production of *grote olm planten* /xrotə ɔlm plɑntən/ 'plant a big elm').

2.2 Speakers

Eight native Dutch speakers participated in the experiment. Because we wished to elicit both the epenthesis and the non-epenthesis pronunciation of each epenthesis word from each speaker, we screened a larger pool of native Dutch speakers for variability of epenthesis by having them read a list containing words with the epenthesis environment (in coda position) and filler words. The first author judged auditorily whether epenthesis had been produced in each form, and we chose the speakers who epenthesised most variably, i.e. those who came closest to epenthesising in half of the words with an epenthesis environment. These were six women (Speakers A, E, M, L, K and I) and two men (Speakers D and G), all between 18 and 35 years old, none with training in phonetics. The speakers represented various dialects, since we considered it more important to control for variability of epenthesis than for dialect. All were raised in a monolingual Dutch environment within the Netherlands, and had acquired foreign languages only after childhood.

2.3 Procedures

The speakers were recorded individually at the Institute for Phonetics of the University of Cologne, using a Carstens Articulograph. An articulograph uses magnets held in place by a helmet on the speaker's head, and records the distance from each magnet of small receivers glued to the speaker's tongue. From this information, the horizontal and vertical position of the receivers over time can be calculated, allowing for observation of the position of various parts of the tongue during connected speech. For each speaker, three receivers were placed on the midline of the tongue, one just behind the tip of the tongue, one as far back as possible and one approximately in the middle of the distance between those two. A reference receiver was also placed on the upper teeth. For two speakers (Speakers K and I), we were able to place an additional reference receiver on the bridge of the nose, but for the remaining speakers this reference receiver could not be placed within the recording range of the helmet. For the two speakers who had both reference receivers, the bite plane of the mouth relative to the helmet was recorded before the receivers were positioned in the mouth. For these two speakers, the sampling rate was 500 Hz, while for the other speakers it was 40 Hz.[6] Acoustic data was recorded simultaneously.

We wished to obtain articulatory data about epenthesis in relatively natural speech, within the limits of articulograph methodology. Therefore,

[6] The difference in sampling rate was due to an error in the recording set-up, but even the lower sampling rate seems to have been sufficient for locating the peak of tip raising which is measured in this study.

we at first did not point out to the speakers that the epenthesis words can be produced in two ways. Since some of the words in the materials are of quite low frequency, we did allow speakers to familiarise themselves with the entire set of materials in advance. Either the second author or an assistant, both native Dutch-speaking phoneticians, also pointed out a few items which speakers sometimes mispronounced in ways unrelated to epenthesis (e.g. orthographic *h* in the place name *Zelhem* /zɛləm/ and orthographic *n* in words like *brillenmaker* 'optician' are not pronounced). Otherwise, speakers were simply told to speak naturally and at a comfortable speed.

Materials were presented in three blocks, the epenthesis words in the frame that favours epenthesis, the same words in the frame that disfavours epenthesis and the underlying schwa words. Within each of the former two blocks, the short words (5a) were placed before the long words (5b), as pilot work showed that it was easier for speakers to produce the phrases this way. Six different orders for the materials were created, with the order of the three blocks as well as the pseudo-random order of items within them varied. Each speaker was recorded using one of these orders. Some additional materials were included in the recording but will not be reported here.

After this first recording, either the second author or a native Dutch-speaking phonetically trained assistant pointed out to each speaker that the epenthesis words could be produced two ways, with or without an 'extra [ə] sound', and demonstrated the two pronunciations of a few words. The speaker was then instructed to produce all of the words in a block either with or without epenthesis. The block in which speakers were told to epenthesise was always the block with frame phrases favouring epenthesis. During both recordings, the native Dutch-speaking experimenter judged auditorily whether each item was produced with or without epenthesis, and recorded this information. During the instructed recording only, if the speaker failed to epenthesise in an item where epenthesis was desired, or did epenthesise in an item where epenthesis was not desired, he or she was reminded about the 'extra sound' and asked to repeat that item at the end of the recording session. In a few cases, it was necessary for the Dutch-speaking experimenter to model the epenthesis or non-epenthesis pronunciation of a particular word. Most speakers had little or no difficulty producing words both with and without epenthesis during the instructed recording. Speaker E, however, could not learn to produce most items in more than one way, so her data was excluded from the analysis.

The instructed recording allowed us to obtain epenthesis and non-epenthesis productions of a large number of items for each speaker, and the first, uninstructed recording allowed us to obtain data on epenthesis when it is produced spontaneously. Since the prosodic effect found by Kuijpers & van Donselaar (1998) is not an extremely large effect, and we collected only one uninstructed recording of each item in each environment, we rarely obtained large numbers of words both with and without epenthesis in the uninstructed recording. We will discuss below what comparisons we made with which subsets of the data. As much as possible, the receivers were left attached to the tongue during the entire recording session, including both instructed and uninstructed recordings. However, for Speaker L (all receivers) and Speaker A (dorsal receiver), receivers loosened from the tongue and had to be replaced between the uninstructed and instructed recordings, so that data from the two recordings cannot be compared.

2.4 Measurements

For the two speakers for whom two reference receivers and bite-plane data were available, the data was first downsampled to 250 Hz and then smoothed (using low pass filters with cut-off frequencies of 10 Hz for the reference receivers and 25 Hz for the tongue receivers). Dynamic correction was then applied based on the two reference receivers to correct for any small shifts of the helmet relative to the head during recording. The data was then rotated by an angle calculated from the bite-plane measurement in order to bring the bite plane to horizontal (instead of horizontal relative to the helmet). The Carstens Emalyse and Tailor software were used for all data processing and measurements.

For the remaining five speakers these adjustments to the data were not possible. As an alternative correction for any small shifts of the helmet relative to the head, for these five speakers, the position of the upper teeth reference receiver (in the horizontal and vertical dimensions) was subtracted from the positions of each tongue receiver, rather than analysing raw tongue-receiver positions. This should remove most variability introduced by any helmet shifts. Since bite-plane correction is not possible for these speakers, one cannot be sure in their data whether the 'horizontal' and 'vertical' dimensions are actually horizontal and vertical, since they are measured relative to the helmet. However, comparison of the palatal contours for the two speakers for whom rotation is possible and the five for whom it is not suggests that the helmet was not severely rotated relative to the head in these five speakers.[7] Furthermore, while the terms 'horizontal' and 'vertical' may not be entirely accurate for these five speakers, what is of most importance is whether /l/ before epenthetic schwa patterns with /l/ before underlying schwa or with coda /l/, and this comparison within each speaker does not rely on whether movement is horizontal, vertical or both.

After these corrections, position of the three tongue receivers was measured for each token which was used in any of the comparisons below

[7] Palatal contours show the position of the palate and alveolar ridge for each speaker, and are obtained by having the speaker move the tongue tip, with a receiver attached to it, along the palate to the teeth. The approximate angle of the speaker's head relative to the helmet can be judged, and this allows us to conclude that the 'horizontal' and 'vertical' dimensions for the speakers for whom rotation correction is not possible are indeed close to horizontal and vertical despite the lack of correction.

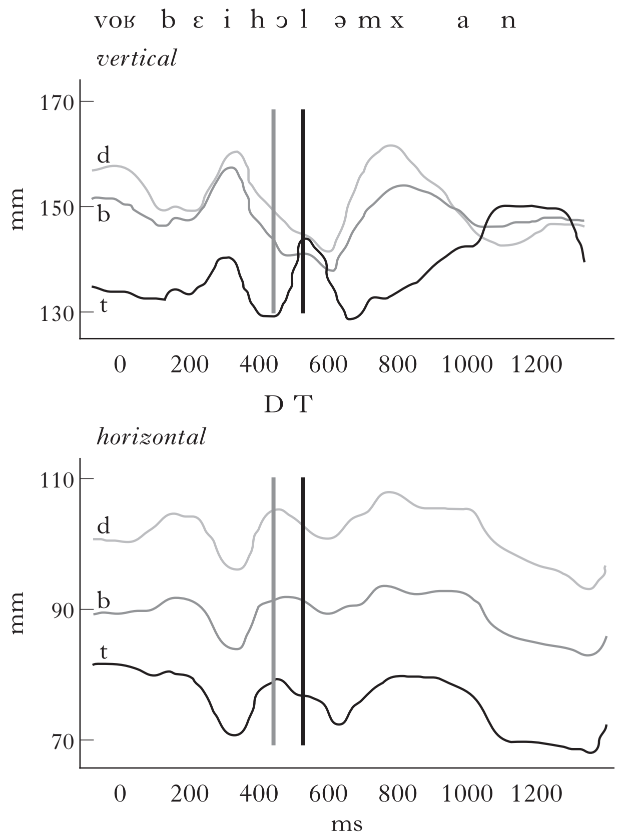


Figure 2

Vertical (upper panel) and horizontal (lower panel) motion of the tongue tip, body and dorsum during the phrase *voorbij Hollum gaan*. The vertical line 'T' marks the peak of tip raising, and 'D' marks the peak of dorsal backing. The curve 'd' shows dorsum position, 'b' shows tongue-body position and 't' shows tongue-tip position. In the vertical panel, higher position on the y axis represents higher position of the tongue. In the horizontal panel, higher position on the y axis represents a more back position of the tongue (i.e. up is back). This token exemplifies the frequent difficulty with locating a peak of dorsal backing associated with the /l/ rather than the preceding vowel.

at two manually located time points, the peak of the tongue-tip raising gesture and the peak of the dorsal backing gesture (Fig. 2). The choice of these measurement points was based on previous literature on /l/ articulation (Sproat & Fujimura 1993, Browman & Goldstein 1995, Gick 1999a). Browman & Goldstein report tip raising and dorsal backing. Gick reports dorsal backing and either tip fronting or tip raising and fronting for /l/, depending on the direction of the gesture for the particular speaker. Sproat & Fujimura focus on tongue-body lowering and the

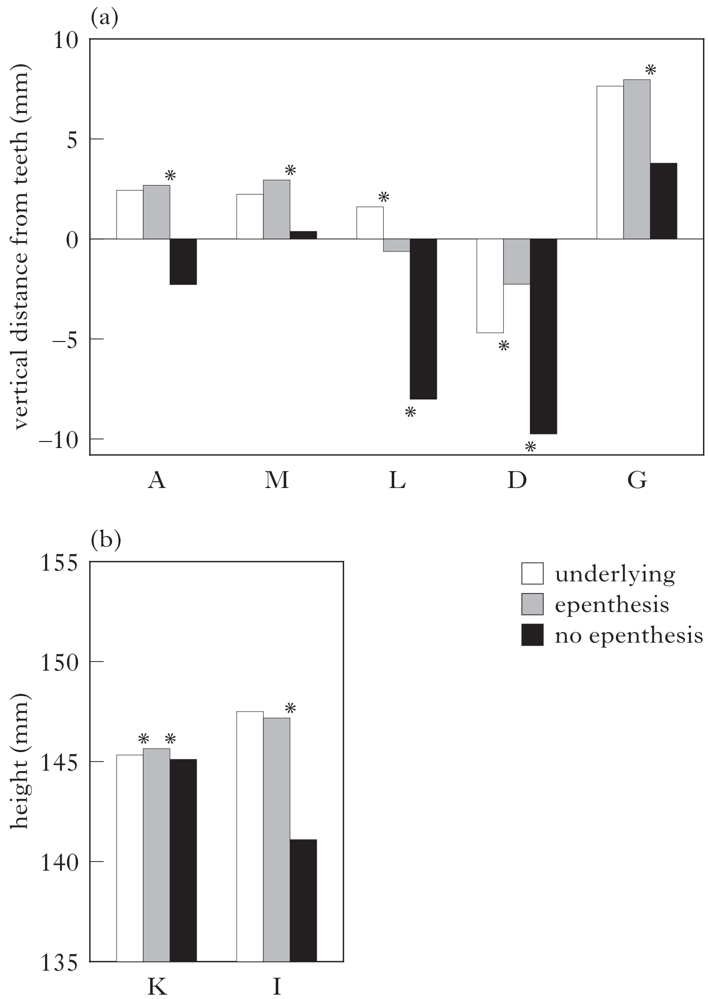


Figure 3

Vertical position of tongue tip in the underlying schwa, epenthesis and coda /l/ (no epenthesis) conditions, in instructed speech. Asterisks mark differences which are significant in the pairwise comparisons of epenthesis *vs*. coda and epenthesis *vs*. underlying schwa conditions (Table II). (a) The five speakers for whom the upper tooth receiver position was used as a correction, in millimetres from the upper tooth receiver. (b) The two speakers for whom dynamic correction was used, raw position in millimetres after correction.

extremum of forward and upward motion for the tongue tip. For our data, however, it was often difficult to locate a peak of dorsal backing associated with the /l/ rather than with the preceding vowel, so data measured at this time point is not reliable and will not be reported here. Therefore, separation of the tip and dorsal gestures in time, a common measure in the previous literature, also cannot be considered.

environment	Speaker A	Speaker M	Speaker L	Speaker D	Speaker G	Speaker K	Speaker I
vertical: underlying	243·13 (123·34)	223·33 (142·94)	160·27 (229·07)	−469·53 (170·10)	764·07 (139·99)	14533·00 (43·18)	14750·00 (159·58)
vertical: epenthesis	268·33 (61·30)	294·53 (146·93)	−62·60 (156·29)	−226·13 (172·58)	796·57 (164·28)	14565·00 (46·75)	14718·00 (67·89)
vertical: no epenthesis	−228·13 (192·66)	37·60 (128·17)	−801·40 (265·26)	−975·20 (186·71)	378·86 (132·12)	14511·00 (54·65)	14110·00 (312·82)
horiz: underlying	1216·13 (119·07)	1407·93 (308·37)	2130·80 (196·68)	1337·80 (338·40)	2051·07 (119·98)	4819·00 (121·87)	7435·00 (123·59)
horiz: epenthesis	1189·00 (83·87)	1513·87 (322·60)	1889·27 (80·54)	1274·27 (219·86)	2051·71 (135·83)	4702·00 (183·80)	7381·00 (129·86)
horiz: no epenthesis	1451·07 (223·85)	974·93 (198·86)	1919·20 (185·84)	754·00 (256·90)	2063·71 (238·71)	4579·00 (104·39)	7020·00 (224·36)

Table I

Means and standard deviations for the comparison within instructed speech, for vertical and horizontal position of the tongue tip in each condition, by speaker. Standard deviations appear in parentheses. Means for Speakers K and I reflect raw receiver position in hundredths of millimetres, while those for all other speakers reflect distance (vertical or horizontal) from teeth in hundredths of millimetres.

3 Results

3.1 Comparison of /l/ before epenthetic and underlying schwa and in coda position

In this section, we present a comparison of only the longer words from the instructed recording (5b), and briefly discuss the uninstructed speech data. Since we asked speakers to repeat misproduced items in the instructed recording, we obtained usable tokens of all materials from that recording, allowing the use of fifteen triples (epenthesis, non-epenthesis and underlying schwa productions) for each of the seven speakers.[8] We will analyse this data for each speaker separately, with items as the repeated measure, because this provides the closest comparison to previous work. However, in order to evaluate the similarity of results across speakers, we will also provide statistical tests with speaker as the repeated measure. The number of speakers for this study (seven) is larger than is common for articulograph research (Krakow 1999), and facilitates comparison across speakers.

[8] In Speaker G's production of one non-epenthesis item, however, no tip-raising gesture at all could be found, so the triple with this item was excluded from his data.

dependent variable	Speak-er A	Speak-er M	Speak-er L	Speak-er D	Speak-er G	Speak-er K	Speak-er I
vertical epenthesis vs. *no epenthesis*	‡	‡	‡	‡	‡	*	‡
vertical epenthesis vs. *underlying*	F<1	n.s.	‡	†	F<1	*	F<1
vertical underlying vs. *no epenthesis*	‡	†	‡	‡	‡	n.s.	‡
horizontal epenthesis vs. *no epenthesis*	‡	‡	F<1	‡	n/a	*	‡
horizontal epenthesis vs. *underlying*	F<1	n.s.	‡	F<1	n/a	*	n.s.
horizontal underlying vs. *no epenthesis*	†	‡	*	‡	n/a	‡	‡

Table II

Significance of pairwise comparisons (results of ANOVAs) for comparison of the data presented in Table I, for each speaker separately. Pairwise comparisons are only performed if the overall ANOVA shows a significant effect of environment, but this is the case for all except the horizontal data of Speaker G. Degrees of freedom are (1,13) for Speaker G and (1,14) for all others.

* Significant at $p<\cdot05$, † significant at $p<\cdot01$, ‡ significant at $p<\cdot001$.

Figure 3 shows the vertical position of the tongue tip, measured at the peak of tip raising, for each speaker in /l/ before underlying schwa, before epenthetic schwa and in the non-epenthesis productions (coda position). For each of the seven speakers, the overall effect of environment on height of the tongue tip is significant. (Means and standard deviations appear in Table I, and results of statistical tests in Table II.) Furthermore, for each speaker, the tip is significantly higher during /l/ before epenthetic schwa than during coda /l/, although the difference is small for Speaker K. For three of the speakers (L, D and K), the difference between the epenthesis condition and the underlying schwa condition is also significant, with the tip higher in the epenthesis condition for Speakers D and K and lower for Speaker L.

For all speakers except D and K, it is clear that the epenthesis condition is more similar to the underlying schwa condition than to the coda /l/ condition. For those two speakers, tip height of /l/ before epenthetic schwa is more extreme than in /l/ before underlying schwa, and clearly is not more similar to /l/ in coda position. For all speakers except K, the tip is also significantly higher before underlying schwa than in coda position, confirming that there is indeed a light/dark /l/ alternation in Dutch.

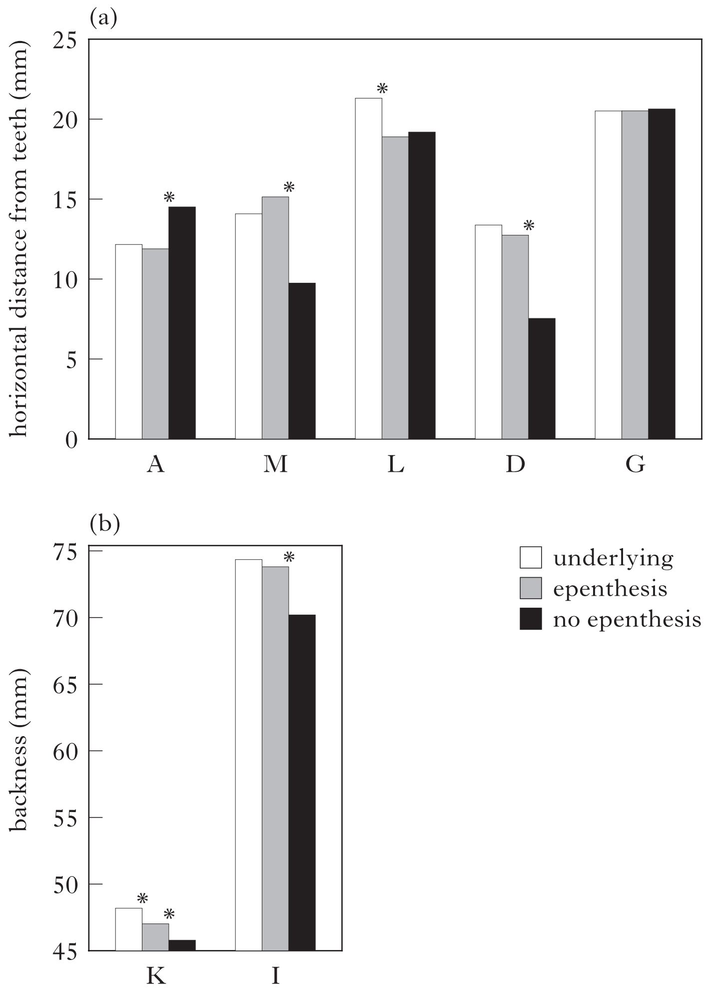
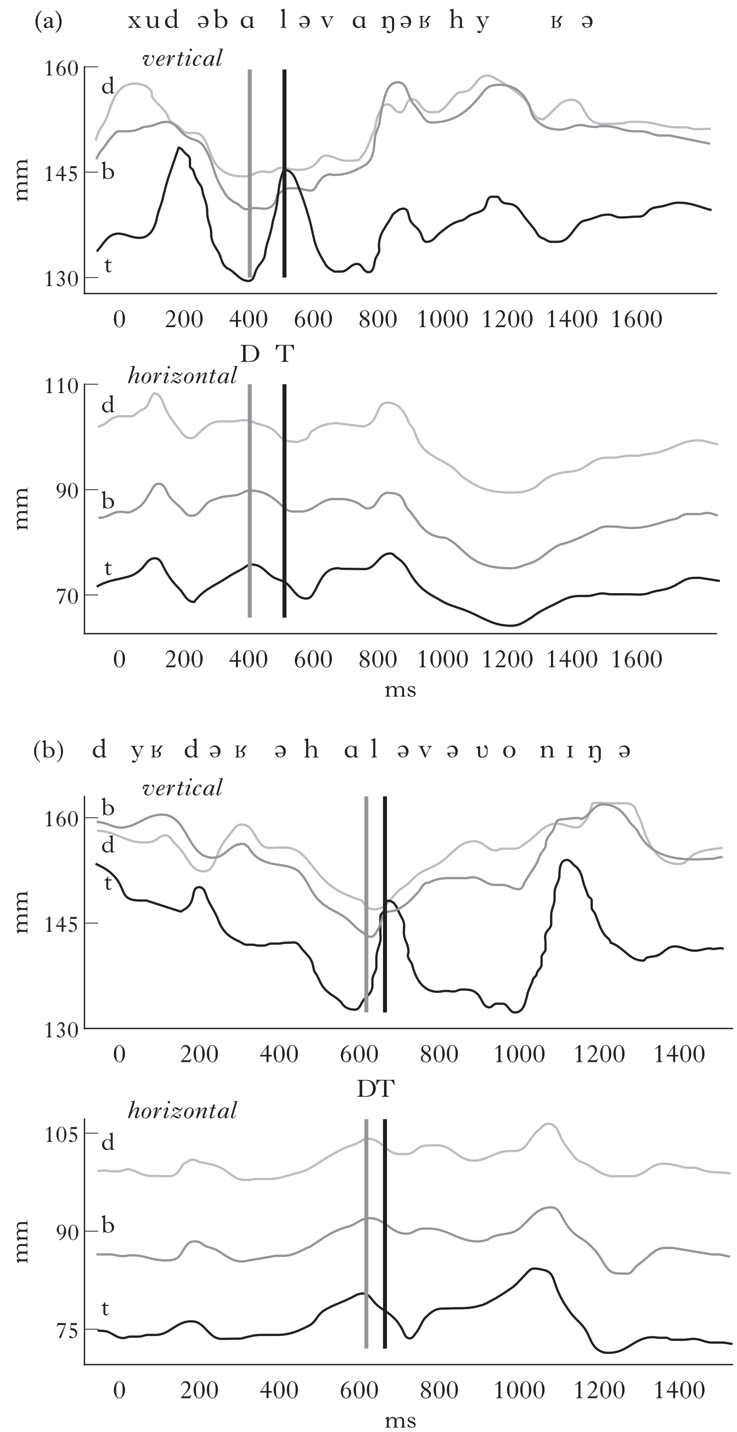


Figure 4

Horizontal position of the tongue tip in instructed speech. See caption to Fig. 3 for further details.

Figure 4 shows the horizontal position of the tongue tip. The direction of effects for the three conditions varies considerably across speakers, but it does appear that for most speakers, the epenthesis and underlying schwa conditions are the most similar. For six of the speakers (all except Speaker G), the overall effect of environment is significant. For five of those speakers (A, M, D, K and I), the epenthesis condition differs significantly from the coda /l/ condition (four – A, D, K and I – with the tongue tip further back in the epenthesis condition, Speaker M with it further forward). For two speakers, L and K, the tip is significantly further


(a)
xud əb ɑ l ə v ɑ ŋ ə ʁ h y ʁ ə
vertical
160
145
130
mm
d
b
t
0 200 400 600 800 1000 1200 1400 1600
D T
horizontal
110
90
70
mm
d
b
t
0 200 400 600 800 1000 1200 1400 1600
ms
(b)
d y ʁ d ə ʁ ə h ɑ l ə v ə ʊ o n ɪ ŋ ə
vertical
160
145
130
mm
b
d
t
0 200 400 600 800 1000 1200 1400
DT
horizontal
105
90
75
mm
d
b
t
0 200 400 600 800 1000 1200 1400
ms

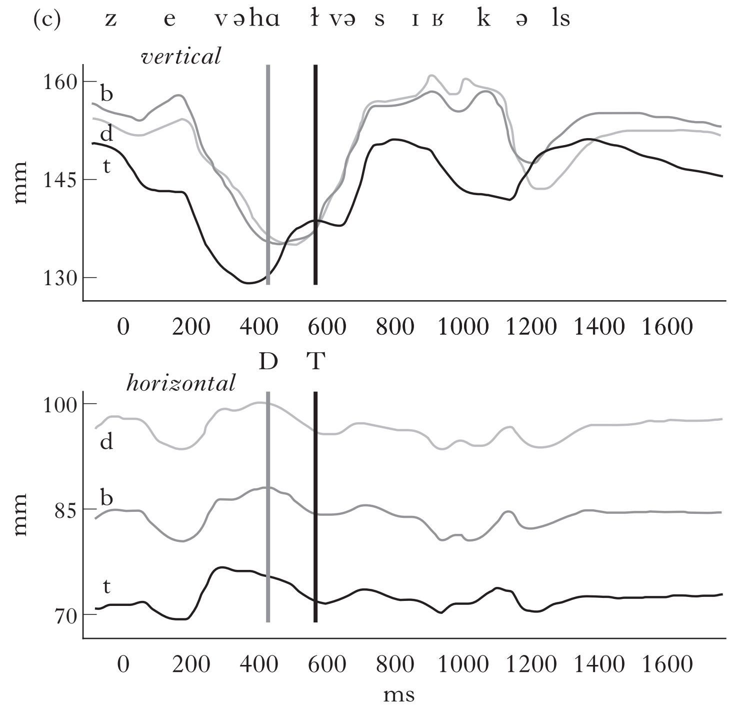


Figure 5

Vertical (upper panel) and horizontal (lower panel) motion of the tongue tip, body, and dorsum during a phrase of each condition. (a) /l/ before underlying schwa in the phrase *goede ballenvanger huren*. (b) /l/ before epenthetic schwa in the phrase *duurdere hal*[ə]*ve woningen*. (c) /l/ in a non-epenthesis production, in the phrase *zeven halve cirkels*. See caption to Fig. 2 for further details.

forward in the epenthesis condition than in the underlying schwa condition. For four speakers (A, M, D and I), the epenthesis condition clearly patterns with the underlying schwa condition and not with the coda /l/ condition. For all six speakers who have a significant overall difference in tip-horizontal position, the underlying schwa condition also differs significantly from the coda /l/ condition, although the direction of the effect differs across subjects.

Tip raising is greater before both types of schwa than in coda /l/, and differences between /l/ before epenthetic and underlying schwa are minimal and not consistent. Figure 5 shows the motion of the tongue receivers during articulation of each of the conditions. The quick, large vertical movement of the tip upward is characteristic of /l/ before both types of schwa, and the extreme lowering of the dorsum during the preceding vowel and /l/ is typical of the coda /l/ (no epenthesis) condition.

There are no clear effects on horizontal position across speakers. The lack of clear effects on horizontal position could reflect the lack of a rotation correction for five of the speakers.[9] However, since it is not the case that speakers with a large effect in the vertical dimension have a small effect in the horizontal dimension and *vice versa*, this is unlikely. It seems that the effect of following schwa (either type) *vs.* coda position on articulation of /l/ in Dutch is primarily in the vertical dimension. Statistical tests across subjects (averaged across items) confirm this: the overall effect of environment on tip-vertical position is significant ($F(2,12) = 17{\cdot}00$, $p < {\cdot}001$), but the overall effect on tip-horizontal position is not ($F(2,12) = 3{\cdot}40$, $p > {\cdot}05$). The tip is significantly higher during /l/ before epenthetic schwa than in coda /l/ ($F(1,6) = 24{\cdot}18$, $p < {\cdot}01$) and during /l/ before underlying schwa than in coda /l/ ($F(1,6) = 15{\cdot}43$, $p < {\cdot}01$). Vertical tip position does not differ significantly between the epenthesis and underlying schwa conditions ($F < 1$).

We wished to further support these results with data from more natural productions of epenthesis, such as is available from our uninstructed recordings. However, most speakers did not epenthesise spontaneously in more than a few of the longer words. Therefore, we compared short words which were spontaneously produced with epenthesis and phonologically matched longer words in which the speaker did not epenthesise spontaneously, as well as underlying schwa words. For example, in this comparison *olm* [ɔləm] /ɔlm/ 'elm' (epenthesis), *olmen* [ɔlmən] /ɔlmən/ 'elms' (no epenthesis) and *column* [kɔləm] /kɔləm/ 'column' (underlying schwa), all from the uninstructed recording, might form a triple. A smaller number of such triples was available for each speaker, and these triples are not as well matched as those from the instructed recording, so the data from this comparison is less clear. However, the overall pattern of results is similar to the results for the instructed speech, showing greater tip raising before both types of schwa than in coda /l/, and similarity of the epenthesis and underlying schwa conditions.

3.2 Comparison of spontaneous and instructed epenthesis

To assess whether speakers produced epenthesis in a natural way in both recordings, we compared the articulation of /l/ in short words where epenthesis was produced spontaneously *vs.* in long words where it was only produced under instruction (e.g. [fɪləm] /fɪlm/ from the uninstructed recording and [fɪləməʁ] /fɪlmər/ from the instructed recording). Since this requires comparing across the two recordings, data from Speaker L was excluded because receivers had to be reattached between recordings.[10]

[9] The lack of a consistent direction for the small differences between the epenthesis and underlying schwa conditions in the vertical dimension, however, could not be a result of the lack of a rotation correction, since the direction of the larger difference between the epenthesis and coda conditions is consistent across speakers.

[10] Although Speaker A's dorsal receiver also had to be reattached, as mentioned above, the tip receiver did not have to be reattached for her, so her data can be used.

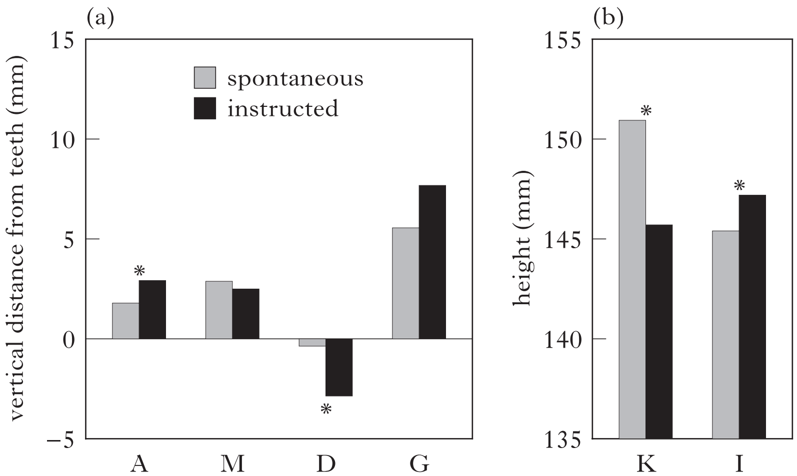


Figure 6
Vertical position of the tip for /l/ before epenthetic schwa in uninstructed and instructed speech. See caption to Fig. 3 for further details.

There were six to nine pairs of such matched words available for the remaining six speakers.

The results appear in Fig. 6. Two of the six speakers show no significant difference in tip-vertical position between /l/ before spontaneous and instructed epenthetic schwa; the other four (A, D, K and I) do show a significant difference (Table III). However, the direction of this effect is not consistent across these four speakers. (Since the same speakers have differences in consistent directions across conditions in the instructed speech comparison above (§3.1), the inconsistency in the direction of this effect cannot result from the lack of a correction for rotation.) The lack of consistent effects in this comparison suggests that the articulation of /l/ before epenthetic schwa is the same, whether the speaker epenthesises spontaneously or under instruction.

4 Conclusions

The data above shows that for most speakers, Dutch /l/ is articulated as a light /l/ before a schwa, whether that schwa is epenthetic or underlyingly present, and that /l/ is darker in productions without epenthesis (coda position). The data from the instructed speech recording indicates, with the exception of Speaker K, that the tip-raising gesture is larger in /l/ before either type of schwa than in coda /l/. Data from uninstructed speech shows a similar pattern. The comparison of /l/ before

tip vertical	Speaker A	Speaker M	Speaker D	Speaker G	Speaker K	Speaker I
spontaneous	178·89	288·11	−36·86	555·67	15093·83	14540·25
(*sd*)	(117·74)	(116·00)	(159·40)	(225·46)	(57·36)	(166·88)
instructed	291·89	249·78	−286·29	768·00	14570·50	14719·25
(*sd*)	(69·71)	(115·71)	(234·26)	(150·49)	(41·50)	(71·77)
significance	*	F<1	*	n.s.	‡	*

Table III

Means, standard deviations and significance of comparison of tongue-tip vertical position of /l/ before spontaneous *vs.* instructed epenthetic schwa. Means for Speakers K and I reflect raw receiver position in hundredths of millimetres, while those for all other speakers reflect vertical distance from teeth in hundredths of millimetres. Standard deviations appear in parentheses. Degrees of freedom are (1,8) for Speakers A, M and G, (1,7) for Speaker I, (1,6) for Speaker D and (1,5) for Speaker K.

* Significant at p<·05, † significant at p<·01, ‡ significant at p<·001.

spontaneous and instructed epenthesis confirms that the finding in the instructed speech is not an artefact of having told the speakers to produce epenthesis. This difference in tip raising is consistent with past findings about articulation of light and dark /l/. The difference between Speaker K's behaviour and that of the other speakers can probably be attributed to a dialectal difference in realisation of light and dark /l/ (cf. Collins & Mees' 1999 claim that the light/dark /l/ alternation varies across dialects), although the realisation of /l/ in Speaker K's dialect is not yet understood.

In addition to the implications of these results for the mechanism of schwa epenthesis, the finding that /l/ before underlying schwa (the control condition) is indeed lighter than /l/ in coda position clarifies one aspect of the light/dark /l/ alternation in Dutch. As Collins & Mees (1999) and van Reenen (1986) state, but unlike Booij's (1995) description, intervocalic /l/ is light, or at least lighter than coda /l/, in Dutch. Since we did not test utterance-initial /l/ or manipulate strength of prosodic boundary as Sproat & Fujimura (1993) did for English, we cannot be sure, but it is certainly possible that the /l/ alternation in Dutch is actually a continuum as in English, rather than a binary light/dark change. This would explain the inconsistencies in the literature on whether intervocalic /l/ is light or dark.

The finding that /l/ before both epenthetic and underlying schwa is light and coda /l/ is dark demonstrates that epenthetic schwa can condition the light/dark /l/ alternation. The ability of epenthetic schwa to condition such an alternation indicates that schwa epenthesis involves the insertion of a phonological unit, such as the schwa phoneme or a syllable nucleus, rather than just the retiming of existing gestures.[11] Retiming of the gestures (of /l/ and the following labial consonant in this case), although it has provided plausible explanations for several other phonological and phonetic alternations as discussed in §1.2.1 above, cannot account for the change from dark /l/ in non-epenthesis productions to light /l/ in epenthesis productions. Simply drawing the gestures of the /l/ and those of the following labial consonant apart in time might well create a sound perceptually similar to an epenthetic schwa, but there is no reason for it to change the magnitude of the tip-raising gesture of the /l/, as we see in this data. Since overlap with a labial gesture does not usually prevent the realisation of lingual gestures, separating the preceding tongue-tip gesture from the labial gesture in time would not influence the magnitude of the tip gesture.[12]

Furthermore, separating the gestures of the two consonants in time could not, in and of itself, add a syllable to the word and thus condition the alternation that way, even though the light/dark /l/ alternation has the typical characteristics of a syllable-structure dependent alternation (Gick 1999b). Syllables in AP are a matter of phasing relations between gestures, particularly of consonant gestures being phased with respect to vowel gestures (Browman & Goldstein 1995, Gick 1999a). A targetless schwa consisting only of a gesturally unspecified gap in the signal could not have the gestures of the /l/ phased to it. Therefore, Dutch schwa epenthesis cannot be a matter of separating the gestures of the two consonants in time.

One of the most persistent past challenges to the claim that alternations do not require inserting gestures involves English intrusive /r/ (Kohler 1992, McMahon *et al*. 1994, McMahon & Foulkes 1995). Gick (1999b) provides an elegant analysis of this phenomenon within AP, requiring no insertion of gestures, by arguing that the gestures of the apparently inserted consonant are actually underlyingly present, but that because of normal articulatory patterns caused by syllable structure, they have no audible consequence in the environment where the consonant seems not to be inserted. That is, he claims that the apparently intrusive /r/ is part of the underlying representation, and that rather than being deleted in coda position, it is vocalised through adjustments to the timing and magnitude of the existing gestures. Further research on the gestural similarity of [ə] and [ɹ] supports this analysis (Gick *et al*., in press), as well

[11] Thus forms with epenthesis have, articulatorily, an extra syllable, even though the findings of van Donselaar *et al*. (1999) and Goetry *et al*. (in press) indicate that the underlying representations have neither epenthesis nor the additional syllable. This supports the idea that underlying representations are more abstract than phonetic forms.

[12] Furthermore, data for the position of the tongue dorsum, although not presented here, suggests that the dorsal lowering gesture is larger in coda /l/ than in /l/ before either type of schwa. Even if separation in time from the following labial could account for the increase in magnitude of the tongue-tip gesture when schwa is inserted, it could not possibly account for a simultaneous reduction of the magnitude of the dorsal gesture.

as pointing out the similarity of final /l/-vocalisation. This approach cannot be applied to Dutch schwa epenthesis or its accompanying light/dark /l/ alternation, though. First, there is phonological and psycholinguistic evidence that Dutch epenthetic schwa is not underlying (van Donselaar *et al*. 1999, Goetry *et al*., in press; see §1.1 above). Second, even if the apparently epenthetic schwa were present underlyingly, retiming of gestures could not produce the alternation we have observed. A retiming analysis with epenthetic schwa as underlying would involve moving the gestures of the /l/ and the post-schwa consonant closer in time to overlap the gestures of the schwa. Since the gestures of the schwa would be overlapped but not deleted, the gestures of the /l/ would still be prevocalic, and they would not shift to the gestures of a dark /l/.

Since retiming of gestures and adjusting their magnitude cannot produce the alternation between light and dark /l/ depending on epenthetic schwa, and there is evidence that the epenthetic schwa is not present in the underlying representation, it must be possible to insert the schwa into the form. The results we report here have no bearing on what sort of unit is inserted. It could be that what is inserted is the gesture(s) of a schwa. This would require a modification to AP to allow gestural insertion, but would maintain the gesture as the basic unit of phonological representations. This would be similar to the approach McMahon and colleagues suggest (McMahon *et al*. 1994, McMahon & Foulkes 1995) of blending AP and Lexical Phonology by allowing gestural insertion, substitution or deletion during a more abstract stage of the phonology, perhaps corresponding to the lexical level. However, it could also be that the phoneme schwa is inserted. A third possibility is that an underspecified vowel is inserted. Yet another is the insertion of a syllable nucleus position.

Our data does not directly show that gestures for a schwa are inserted in forms with epenthesis. Although Browman & Goldstein (1992a), Gick (in press) and Gick *et al*. (in press) show that schwa does generally have gestures, those gestures are quite difficult to measure, since they either are not extreme relative to gestures of other speech sounds or take place in the pharynx (Gick, in press, Gick *et al*., in press). We argue that the alternation between dark /l/ in non-epenthesis forms and light /l/ in epenthesis forms indicates that epenthesis inserts a schwa, with whatever gestures it involves. Since we do not measure the gestures of the schwa itself, one might argue that in our data, no gestures have been inserted, rather the underlying gestures have been rearranged to exactly the pattern of gestures we find in the epenthesis forms. However, this would mean that the gestures of a dark preconsonantal /l/ would not only be separated in time from the gestures of the following consonant, but would also coincidentally be rearranged to exactly the gestures of a light /l/, which one would otherwise only expect to occur prevocalically. Although some differences between /l/ before epenthetic and underlying schwa do reach significance for some speakers, the overall pattern is of similarity of /l/ before the two types of schwa, indicating that the schwa is phonologically present in both conditions.

Furthermore, examination of the standard deviations (Table I) shows in general no greater variability for /l/ before epenthetic schwa than before underlying schwa. This also suggests that /l/ before the two types of schwa is the same thing, and the insertion of the schwa is categorical and phonological, not gradient.

Thus, some phonological unit has been inserted in epenthesis forms, and if schwa has gestures, then its gestures have been inserted. This would require only one modification of AP, allowing gestural insertion (although it might seem suspicious that the insertion happens to correspond to the size of a phoneme). If what is inserted in epenthesis is not a collection of gestures but rather the phoneme schwa, an underspecified vowel, or an empty syllable nucleus, more fundamental revision of AP would be necessary, since that theory does not recognise either phonemes or syllables as phonological units.

Our results indicate that Dutch schwa epenthesis is primarily a phonological rather than a phonetic process. By treating many phonological alternations as adjustments to articulatory gestures, AP integrates phonetics and phonology in many cases. However, since Dutch schwa epenthesis must involve the insertion of some phonological unit into the string, it is clearly a relatively abstract change in the form of a word, not a phonetic one. Although Dutch does allow a wide variety of CC clusters, in words where epenthesis applies, it reduces syllable-structure complexity. One could, for example, analyse Dutch epenthesis in optimality-theoretic terms, probably through a constraint against certain complex syllable structures which would be more highly ranked than faithfulness constraints, resulting in the insertion of a default vowel, schwa.

Zsiga (1997) also shows experimentally that some alternations are abstract, phonological, and cannot be readily treated as adjustments to gestures. She examines vowel harmony and assimilation of contiguous vowels in Igbo. She finds that assimilation of contiguous vowels is a gradient process which can be clearly described through adjustments to gestures, but that vowel harmony is a categorical, phonological process which does not take place through adjustments to underlyingly present gestures. It should be noted that not all work which analyses speech in terms of gestures, or which proposes gestures as the basic units of phonological representations, maintains the strong position of AP that alternations can be modelled exclusively through adjustments to timing and magnitude of gestures (Zsiga 1997). Ladd & Scobbie (in press) also test whether a particular pattern, an across-word sandhi phenomenon in Sardinian, is phonetic and able to be modelled through adjustments to gestures, or categorical and phonological. They find that it is in fact categorical and therefore phonological rather than phonetic. Both Zsiga (1997) and Ladd & Scobbie (in press) argue against gestural analyses for particular cases by showing that the alternation is categorical rather than gradient. We use a different approach, arguing that Dutch schwa epenthesis is phonological and not a matter of adjustments to gestures not because it is categorical (which it may well be – we do not test this), but because it conditions a syllable-structure based alternation in a neighbouring segment.

We suspect that in general, epenthesis which brings syllable structure closer to a CV pattern and which is not limited to slow speech is more likely to be phonological and to involve insertion of the epenthetic segment, while other types of epenthesis may indeed involve only adjustments to the timing of existing gestures. For example, epenthetic stops in nasal–obstruent clusters ([t] in *prince*) (Warner & Weber 2001, Warner, in press), as well as German slow-speech schwa insertion (Jannedy 1994), probably do result from overlap or separation of gestures rather than insertion of any unit. Schwa insertion in *hail* and *hire* is a related case, resulting from the tongue moving through the position for schwa during the transition between surrounding sounds (Gick & Wilson 2001).

English intrusive /r/ (Gick 1999b) does give syllables CV structure, and is not conditioned by speech rate. This seems to be a case of a rather abstract phonological alternation which can be analysed within the mechanisms of AP (adjustments to timing and magnitude of gestures). Because AP draws the distinction between phonological and phonetic alternations differently than other theories, many abstract phonological alternations do fall within it. However, the data above makes it clear that Dutch schwa insertion is a different kind of phenomenon. Because of the fortuitous co-occurrence of schwa epenthesis and the /l/ alternation in Dutch, this is a rare case in which it is possible to conclude that a particular alternation is abstract and phonological rather than articulatorily based and phonetic.

Appendix

Materials, in Dutch orthography and IPA transcription, with glosses. Transcriptions show epenthesis words with epenthesis in the environment which favours epenthesis and without in the environment which disfavours it. The (i) forms are those with epenthesis, (ii) without epenthesis and (iii) with underlying schwa.

Shorter words

a.	i.	samen film huren	[ˈsamən ˈfɪləm ˈhyʁən]	‘rent a movie together’
	ii.	geen film zien	[ˈxen ˈfɪlm ˈzin]	‘not see a movie’
	iii.	vandaag Willem zien	[vɑnˈdax ˈʋɪləm ˈzin]	‘see Willem today’
b.	i.	grote olm planten	[ˈxʁotə ˈɔləm ˈplɑntən]	‘plant a large elm’
	ii.	geen olm daar	[ˈxen ˈɔlm ˈdaʁ]	‘no elm there’
	iii.	Renée’s column laatst	[ʁəˈnes ˈkɔləm ˈlatst]	‘Renée’s column at the end’
c.	i.	mooie holm maken	[ˈmojə ˈhɔləm ˈmakən]	‘make a beautiful spiral ornament’
	ii.	geen holm daar	[ˈxen ˈhɔlm ˈdaʁ]	‘no spiral ornament there’
	iii.	voorbij Hollum gaan	[voʁˈbɛi ˈhɔləm ˈxan]	‘go past Hollum’
d.	i.	onder molm leggen	[ˈɔndəʁ ˈmɔləm ˈlɛɣən]	‘lay under mulch’
	ii.	veel molm hier	[ˈvel ˈmɔlm ˈhiʁ]	‘a lot of mulch here’
	iii.	helaas Gollem zien	[heˈlas ˈxɔləm ˈzin]	‘unfortunately see Gollum’

e.	i.	altijd helm brengen	[ˈɑltɛit ˈhɛləm ˈbʁɛŋən]	'always bring a helmet'
	ii.	geen helm meer	[ˈxen ˈhɛlm ˈmeʁ]	'no more helmet'
	iii.	oud cerebellum-beeld	[ˈaut ˌseʁəˈbɛləm ˈbelt]	'old cerebellum image'
f.	i.	grote schelm worden	[ˈxʁotə ˈsxɛləm ˈʋɔʁdən]	'become a big rogue'
	ii.	geen schelm daar	[ˈxen ˈsxɛlm ˈdaʁ]	'no rogue there'
	iii.	voorbij Zelhem gaan	[voʁˈbɛi ˈzɛləm ˈxan]	'go past Zelhem'
g.	i.	altijd kalm blijven	[ˈɑltɛit ˈkɑləm ˈblɛivən]	'always stay calm'
	ii.	steeds kalm zijn	[ˈstets ˈkɑlm ˈzɛin]	'always be calm'
	iii.	Renée zal'm zijn	[ʁəˈne ˈzɑləm ˈzɛin]	'Renée's gonna be it'
h.	i.	nare galm horen	[ˈnaʁə ˈxaləm ˈhoʁən]	'hear a nasty echo'
	ii.	geen galm klonk	[ˈxen ˈxɑlm ˈklɔŋk]	'no echo sounded'
	iii.	helaas val'm aan	[heˈlas ˈvɑləm ˈan]	'unfortunately attack him'
i.	i.	verse zalm eten	[ˈvɛʁsə ˈzɑləm ˈetən]	'eat fresh salmon'
	ii.	die zalm zwemt	[ˈdi ˈzɑlm ˈzʋɛmt]	'that salmon swims'
	iii.	pas op stal'm hier	[pɑs ˈɔp ˈstɑləm ˈhiʁ]	'watch out put it up here'

Longer words

a.	i.	lastige filmers mopperen	[ˈlɑstəɣə ˈfɪləməʁs ˈmɔpəʁən]	'demanding cameramen grumble'
	ii.	goede filmer worden	[ˈxudə ˈfɪlməʁ ˈʋɔʁdən]	'become a good cameraman'
	iii.	morgen brillenmaker bellen	[ˈmɔʁɣən ˈbʁɪləˌmakəʁ ˈbɛlən]	'call the optician tomorrow'
b.	i.	eventjes Wilmer knuffelen	[ˈevəntjəs ˈʋɪləməʁ ˈknʏfələn]	'cuddle Wilmer for a bit'
	ii.	morgen Wilmer bellen	[ˈmɔʁɣən ˈʋɪlməʁ ˈbɛlən]	'call Wilmer tomorrow'
	iii.	Hennies schillenmand legen	[ˈhɛnis ˈsxɪləˌmɑnt ˈleɣən]	'empty Hennie's peel basket'
c.	i.	droevige olmen jammeren	[ˈdʁuvəɣə ˈɔləmən ˈjɑməʁən]	'sad elms moan'
	ii.	nieuwe olmen planten	[ˈniʋə ˈɔlmən ˈplɑntən]	'plant new elms'
	iii.	tegen jolleman roeien	[ˈteɣən ˈjɔləˌman ˈʁujən]	'row against the ferryman'
d.	i.	grotere holmen tekenen	[ˈxʁotəʁə ˈhɔləmən ˈtekənən]	'draw larger spiral ornaments'
	ii.	dure holmen maken	[ˈdyʁə ˈhɔlmən ˈmakən]	'make expensive spiral ornaments'
	iii.	gekke dolleman spelen	[ˈxɛkə ˈdɔləˌmɑn ˈspelən]	'play a crazy madman'
e.	i.	pillen bedwelmen veiliger	[ˈpɪlən bəˈdʋɛləmən ˈvɛiləɣəʁ]	'pills drug more safely'
	ii.	doktoren bedwelmen patienten	[dɔkˈtoʁən bəˈdʋɛlmən paˈʃjɛntən]	'doctors drug patients'
	iii.	zachtjes bellenman horen	[ˈzɑxjəs ˈbɛləˌmɑn ˈhoʁən]	'quietly hear the bell-man'
f.	i.	Mauritius Wilhelmus Verhoogendyck	[ˌmauʁˈitsiʏs ʋɪlˈhɛləmʏs vəʁˈhoɣənˌdɛik]	(proper name)
	ii.	gevoelig Wilhelmus vertolker	[xəˈvuləx ʋɪlˈhɛlmʏs vəʁˈtɔlkəʁ]	'a sensitive performer of the 'Wilhelmus''
	iii.	Karel's ellenmaat nemen	[ˈkaʁəls ˈɛləˌmat ˈnemən]	'take Karel's measure'
g.	i.	rustige kalme bladeren	[ˈʁʏstəɣə ˈkɑləmə ˈbladəʁən]	'peaceful calm leaves'
	ii.	koele kalme meren	[ˈkulə ˈkɑlmə ˈmeʁən]	'cool calm lakes'
	iii.	even mallemolen zeggen	[ˈevən ˈmɑləˌmolən ˈzɛɣən]	'just say 'merry-go-round''
h.	i.	duurdere halve woningen	[ˈdyʁdəʁə ˈhɑləvə ˈʋonɪŋən]	'more expensive half houses'
	ii.	zeven halve cirkels	[ˈzevən ˈhɑlvə ˈsɪʁkəls]	'seven semicircles'
	iii.	goede ballenvanger huren	[ˈxudə ˈbɑləˌvɑŋəʁ ˈhyʁən]	'hire a good catcher'
i.	i.	jongeren delven vrolijker	[ˈjɔŋəʁən ˈdɛləvən ˈvʁoləkəʁ]	'young people dig more merrily'
	ii.	vrouwen delven sneller	[ˈvʁauʋən ˈdɛlvən ˈsnɛləʁ]	'women dig faster'
	iii.	nimmer hellevorst tarten	[ˈnɪməʁ ˈhɛləˌvɔʁst ˈtɑʁtən]	'never defy the prince of darkness'

j.	i.	oudere elven bibberen	[ˈaudəʁə ˈɛləvən ˈbɪbəʁən]	‘older elves tremble’
	ii.	tegen elven komen	[ˈteɣən ˈɛlvən ˈkomən]	‘come towards eleven o’clock’
	iii.	nimmer helleveeg trouwen	[ˈnɪməʁ ˈhɛləˌvex ˈtʁauʊən]	‘never marry a shrew’
k.	i.	anderen dolven gisteren	[ˈɑndəʁən ˈdɔləvən ˈxɪstəʁən]	‘others dug yesterday’
	ii.	spaden dolven dieper	[ˈspadən ˈdɔlvən ˈdipəʁ]	‘spades dug deeper’
	iii.	wollen drollenvanger kopen	[ˈʊɔlən ˈdʁɔləˌvɑŋəʁ ˈkopən]	‘buy woollen baggy trousers’
l.	i.	groenige golven kabbelen	[ˈxʁunəɣə ˈxɔləvən ˈkɑbələn]	‘greenish waves ripple’
	ii.	wilde golven maken	[ˈʊɪldə ˈxɔlvən ˈmakən]	‘make wild waves’
	iii.	grote molleval zetten	[ˈxʁotə ˈmɔləˌvɑl ˈzɛtən]	‘set a big mole-trap’
m.	i.	honderden kolven rekenen	[ˈhɔndəʁdən ˈkɔləvən ˈʁekənən]	‘charge hundreds of cobs’
	ii.	morgen kolven oogsten	[ˈmɔʁɣən ˈkɔlvən ˈoxstən]	‘harvest cobs tomorrow’
	iii.	kleurig bollenveld planten	[ˈkløʁəx ˈbɔləˌvɛlt ˈplɑntən]	‘plant a colourful bulb-field’
n.	i.	rennende wolven naderen	[ˈʁɛnəndə ˈʊɔləvən ˈnadəʁən]	‘running wolves approach’
	ii.	boze wolven bijten	[ˈbozə ˈʊɔlvən ˈbɛitən]	‘angry wolves bite’
	iii.	zwarte mollenvel dragen	[ˈzwɑʁtə ˈmɔləˌvɛl ˈdʁaɣən]	‘wear black moleskin’
o.	i.	winnende kolvers jubelen	[ˈʊɪnəndə kɔləvəʁs ˈjybələn]	‘winning milkers rejoice’
	ii.	moeie kolvers rusten	[ˈmujə ˈkɔlvəʁs ˈʁʏstən]	‘tired milkers rest’
	iii.	kleine schollevaar vangen	[ˈklɛinə ˈsxɔləˌvaʁ ˈvɑŋən]	‘catch a small cormorant’

Additional longer words used only for comparison of spontaneous and instructed epenthesis

ruisende olmpjes fluisteren	[ˈʁœysəndə ˈɔləmpjəs ˈflœystəʁən]	‘rustling little elms whisper’
heerlijke zalmpjes roosteren	[ˈheʁləkə ˈzɑləmpjəs ˈʁostəʁən]	‘grill delicious little salmon’
prachtige palmpjes schilderen	[ˈpʁɑxtəɣə ˈpɑləmpjəs ˈsxɪldəʁən]	‘paint beautiful little palm trees’

REFERENCES

Booij, Geert (1995). *The phonology of Dutch*. Oxford: Oxford University Press.

Browman, Catherine & Louis Goldstein (1990a). Gestural specification using dynamically-defined articulatory structures. *JPh* **18**. 299–320.

Browman, Catherine & Louis Goldstein (1990b). Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 341–376.

Browman, Catherine & Louis Goldstein (1992a). ‘Targetless’ schwa: an articulatory analysis. In Gerard Docherty & D. Robert Ladd (eds.) *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 26–56.

Browman, Catherine & Louis Goldstein (1992b). Articulatory phonology: an overview. *Phonetica* **49**. 155–180.

Browman, Catherine & Louis Goldstein (1995). Gestural syllable position effects in American English. In Fredericka Bell-Berti & Lawrence Raphael (eds.) *Producing speech: contemporary issues. For Katherine Safford Harris*. Woodbury, NY: American Institute of Physics Press. 19–33.

Collins, Beverley & Inger M. Mees (1999). *The phonetics of English and Dutch*. Leiden: Brill.

Donselaar, Wilma van, Cecile Kuijpers & Anne Cutler (1999). Facilitatory effects of vowel epenthesis on word processing in Dutch. *Journal of Memory and Language* **41**. 59–77.

Gick, Bryan (1999a). *The articulatory basis of syllable structure: a study of English glides and liquids*. PhD dissertation, Yale University.

Gick, Bryan (1999b). A gesture-based account of intrusive consonants in English. *Phonology* **16**. 29–54.

Gick, Bryan (in press). An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* **59**.

Gick, Bryan, A. Min Kang & D. H. Whalen (in press). MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *JPh* **30**.

Gick, Bryan & Ian Wilson (2001). Pre-liquid excrescent schwa: what happens when vocalic targets conflict. In *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*. Aalborg, Denmark.

Goetry, Vincent, Philippe Mousty & Régine Kolinsky (in press). The influence of the language-specific orthographic representations on the conceptualization of the phonological structures of words: a cross-linguistic comparison between pre-readers and readers. Paper presented at the 7th Annual Meeting of the Society for the Scientific Study of Reading, Boulder, Colorado.

Jannedy, Stefanie (1994). Rate effects on German unstressed syllables. *OSU Working Papers in Linguistics* **44**. 105–124.

Kingston, John & Avis Cohen (1992). Extending articulatory phonology. *Phonetica* **49**. 194–204.

Kloots, Hanne, Georges De Schutter, Steven Gillis & Marc Swerts (in press). Sjwa-insertie in eindclusters: variatiepatronen in het Standaardnederlands. *Nederlandse Taalkunde*.

Kohler, K. J. (1992). Gestural reorganization in connected speech: a functional viewpoint on 'articulatory phonology'. *Phonetica* **49**. 205–211.

Krakow, Rena A. (1989). *The articulatory organization of syllables: a kinematic analysis of labial and velar gestures*. PhD dissertation, Yale University.

Krakow, Rena A. (1999). Physiological organization of syllables: a review. *JPh* **27**. 23–54.

Kuijpers, Cecile & Wilma van Donselaar (1998). The influence of rhythmic context on schwa epenthesis and schwa deletion. *Language and Speech* **41**. 87–108.

Ladd, D. Robert & James M. Scobbie (in press). External sandhi as gestural overlap? Counter-evidence from Sardinian. In J. Local (ed.) *Papers in laboratory phonology* VI. Cambridge: Cambridge University Press.

McMahon, April & Paul Foulkes (1995). Sound change, phonological rules, and articulatory phonology. *Belgian Journal of Linguistics* **9**. 1–20.

McMahon, April, Paul Foulkes & Laura Tollfree (1994). Gestural representation and Lexical Phonology. *Phonology* **11**. 277–316.

Reenen, Pieter van (1986). The vocalization of /l/ in standard Dutch, a pilot study of an ongoing change. In Frits Beukema & Aafke Hulk (eds.) *Linguistics in the Netherlands 1986*. Dordrecht: Foris. 189–198.

Reenen, Pieter van & Anke Jongkind (2000). De vocalisering van de /l/ in het Standaard Nederlands. *Taal en Tongval* **52**. 189–199.

Sproat, Richard & Osamu Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *JPh* **21**. 291–311.

Swerts, Marc, Hanne Kloots, Steven Gillis & Georges De Schutter (2001). Factors affecting sjwa-insertion in final consonant clusters in Standard Dutch. In *Proceedings of the 7th European Conference on Speech Communication and Technology* (*Eurospeech*). Vol. 1. Aalborg, Denmark. 75–78.

Warner, Natasha (in press). The phonology of epenthetic stops: implications for the phonetics–phonology interface in Optimality Theory. *Linguistics* **40**. 1–27.

Warner, Natasha & Andrea Weber (2001). Perception of epenthetic stops. *JPh* **29**. 53–87.

Zsiga, Elizabeth (1997). Features, gestures, and Igbo vowel assimilation: an approach to the phonology/phonetics mapping. *Lg* **73**. 227–274.