The ‘tight approximant’ variant of the Arabic *‘ayn*

Barry Heselwood
University of Leeds
B.C.Heselwood@leeds.ac.uk

Previous studies of the Arabic pharyngeal *‘ayn* have reported stop, fricative and approximant realisations within and across dialects. This paper presents an approximant variant that has not been specifically identified before. Described as a ‘tight approximant’, it is characterised acoustically by a pattern of filtering in which the lowest half dozen or so harmonics, including the fundamental, are markedly reduced in amplitude. Auditory responses to this kind of spectrum are characterised in terms of residue pitch. The variant can be thought of as compressed or ‘squeezed’ in the articulatory, acoustic and auditory domains in a chain of phonetic cause and effect. Laryngopharyngeal tension creates a low F0 and a compressed bandpass-filtered spectrum, which is processed auditorily by adjacent harmonics being squeezed through the same auditory filter. Acoustic and laryngographic data from selected tokens of a corpus of 436 realisations of *‘ayn*, produced by 21 speakers from 11 different North African and Middle Eastern countries, are presented in support of the description of this variant. Its frequency of occurrence in the corpus suggests it is a common variant in Arabic speech. Some suggestions concerning the diachrony of *‘ayn* are made in light of the reported observations.

1 Introduction

‘Ayn is the reflex of a guttural phoneme that was almost certainly present in Proto-Semitic (Cantineau 1960: 288). It is found in most modern varieties of Arabic (McCarthy 1991: 63, Watson 2002: 18) although in some varieties it has merged with the glottal stop, e.g. in the Tihāmah region of eastern Yemen (Greenman 1979: 54f.) and in others, e.g. western areas of Sudan, *‘ayn* and glottal stop are in free variation (Gasim 1965: 44).

This paper presents a study of one particular allophonic variant, referred to here as a ‘tight approximant’, which appears to be quite widespread cross-dialectally. Phonologically, it is in free variation with variants exhibiting other manners of articulation. Phonetically, it can be described as ‘compressed’ or ‘squeezed’ in each of the articulatory, acoustic and auditory domains. In articulatory terms, it involves compression of laryngopharyngeal structures, including the ventricles and pyriform sinuses, as a consequence of contraction of the laryngeal constrictor which brings about close approximation between the aryepiglottic folds and the epiglottis, and between the epiglottis and the rear pharyngeal wall (Esling 1996) (see figure 1). Airflow through the constriction remains laminar due to a low volume velocity. That is to say, the stricture is of the kind associated with fricatives but airflow through the stricture does not reach the turbulence threshold because transglottal airflow is restricted by laryngeal

Journal of the International Phonetic Association (2007) 37/1
doi:10.1017/S0025100306002787

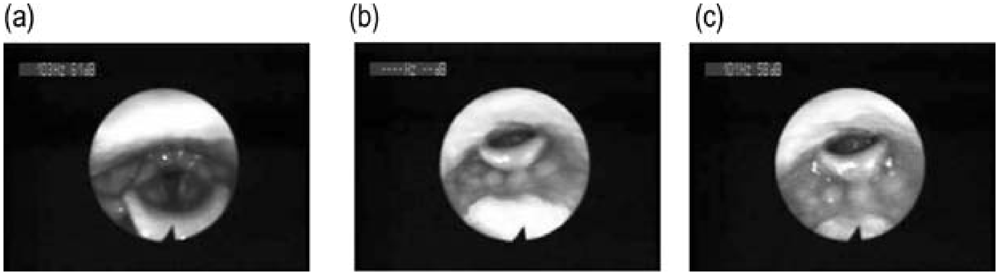


Figure 1 Nasoendoscopic images showing a) 'at rest' position of the epiglottis, b) maximum epiglotto-pharyngeal approximation during a 'tight approximant' *'ayn* in the word /waʕad/ 'to promise', c) maximum epiglotto-pharyngeal approximation during the voiceless fricative /ħ/ in the word /waħad/ 'alone'. Female speaker from Amman, Jordan. (From Heselwood & Al-Tamimi 2006.)

tension. Acoustically, a low F0 causes the harmonics to be pressed close together in frequency while a bandpass filter effect associated with laryngopharyngeal compression, particularly compression of the ventricles, leaves a narrow band of resonance centred in the vicinity of 1 kHz. In the auditory domain the particular pattern of cochlear stimulation results in adjacent harmonics being squeezed through a single auditory filter. The relationships between these phenomena will be examined and discussed within the model of laryngeal articulation developed by Esling (1996, 1999) and outlined in section 1.1 below.

The prevalence and distribution of the tight approximant allophone will be discussed in section 4, after considering its phonetic properties in section 3. A general survey of the literature concerning the phonetics of *'ayn* is presented in section 2.

1.1 Pharyngeal articulation and the laryngeal sphincter mechanism

Esling (1996, 1999) presents a model of activity in the laryngopharynx which he claims can account for the observed components of the articulation of sounds traditionally classified as 'pharyngeal'. The model views the structures comprising the larynx as a single complex mechanism extending from the glottis upwards past the ventricular bands and through the laryngeal vestibule to the aryepiglottic folds forming the rim of the laryngeal aditus and attaching anteriorly to the lateral edges of the epiglottis. At the centre of this articulator, therefore, is a tube the lower end of which is the glottis and the upper end the laryngeal aditus.

Pharyngeal articulation according to this model is brought about by a sphincteric contraction of the tube which can be controlled at three levels: the true vocal folds, the ventricular bands, and the aryepiglottic folds together with the epiglottis (Esling, Fraser & Harris 2005: 402). The uppermost of these levels is engaged in pharyngeal articulation by approximating the aryepiglottic folds while the tongue and epiglottis retract and descend towards the arytenoid cartilages and, at the same time, approximate to the rear wall of the pharynx. Different manners of articulation are achieved by different degrees of constriction of the tube in both the antero-posterior plane and the vertical plane. A complete pharyngeal stop is achieved not by occlusion between the epiglottis and the pharyngeal wall (though such an occlusion has been observed in Arabic – see section 2.2 below) but when 'the aryepiglottic folds move up and forward to meet the base of the epiglottis' (Esling 1996: 84). Degrees of constriction that do not produce full occlusion either create friction and/or trilling, or a glide approximant (Esling 1999: 353f.). Of most relevance to this paper are the states of the various parts of the laryngeal sphincter mechanism during the production of the 'tight approximant' variant of the Arabic *'ayn*. Attention will be drawn to these at appropriate points in the ensuing discussions.

2 Previous studies of *‘ayn*

Instrumental studies have concluded that the Arabic *‘ayn*, usually represented by the IPA symbol [ʕ], exhibits extensive cross-dialectal, intra-dialectal, inter-speaker and intra-speaker variation in place and manner of articulation, and in phonation type. Observers have agreed it is always a pharyngeal sound of some kind but there seem to be few other constraints on its realisation except those serving to distinguish it from the other phonemes classed as gutturals, /χ/, /ʁ/, /ħ/, /h/ and /ʔ/ (McCarthy 1994: 202, Watson 2002: 37) and particularly the fricative /ħ/, its voiceless pharyngeal counterpart. Previous findings will be briefly reviewed and re-interpreted in terms of Esling’s laryngeal articulator model. Each of the three main dimensions of phonetic classification will be taken in turn, beginning with place of articulation.

2.1 Place of articulation

Disagreement over the precise place of articulation of *‘ayn* goes back to the earliest known descriptions. Al-Khaliyl (aka Al-Farahiydi), writing around 760 AD, described it as the farthest sound in the throat (El-Saaran 1951: 204) and tradition has it that he created the letter-shape ء for *hamza* (glottal stop) by adapting the top portion of the *‘ayn* letter ع, because he regarded them as homorganic (Al-Nassir 1993: 12). According to Garbell, ‘the Arab phoneticians expressly warned against an “emphatization” [i.e. pharyngealisation, BH] of /ʔ/, which would otherwise become /ʕ/’ (Garbell 1958: 307). Sibawayh, however, writing only a decade or two later than Al-Khaliyl, classifies *‘ayn* as a sound made in the middle part of the throat while *hamza* is from the lower part (Al-Nassir 1993: 14). It is Sibawayh’s classification which has persisted into most modern accounts of the place of articulation of *‘ayn*, placing it with /ħ/ between the glottals /h, ʔ/ and the uvulars /χ, ʁ/. In terms of the laryngeal articulator model, Sibawayh’s classification takes better account of the role of the aryepiglottic folds and the narrowing of the laryngeal aditus in the production of *‘ayn*, although Al-Khaliyl was no doubt correct in identifying a lower component to the articulation as well.

Constriction at the level of the epiglottis has been reported by several modern researchers for different varieties of Arabic, e.g. Laradi (1983: 123–126) for the Tripoli variety of Libyan Arabic, Adamson (1981: 89) for Sudanese Arabic, Bukshaisha (1985: 304–306) for Qatari Arabic, and Laufer & Baer (1988: 190) for speakers of Iraqi, Lebanese and Palestinian Arabic. Fibreoptic pictures studied by El-Halees (1985: 288) also ‘show that during the production of pharyngeals, the epiglottis is so far back and so low that it covers the laryngeal vestibule and makes a very narrow stricture with the back wall of the pharynx’, a description that agrees with the images in figure 1 above. Pettorino & Giannini (1984: 203) describe *‘ayn* as ‘a laryngeal stop’ formed by the descent of the epiglottis over the laryngeal opening. Jakobson (1957: 112) and Watson (2002: 44f.) both describe it as a pharyngealised laryngeal, i.e. as the emphatic counterpart of the glottal stop phoneme. Ladefoged & Maddieson (1996: 168f.) argue, on the basis of their review of some of the experimental literature, that *‘ayn*, and its voiceless correlate /ħ/, should be described as epiglottal rather than pharyngeal.

The focus on the spatial relationship between the epiglottis and the rear pharyngeal wall in the accounts cited above is understandable because it is this relationship which is most easily observed using traditional laryngoscopy. However, according to Esling’s laryngeal articulator model, it is only part of the story of pharyngeal sounds, and not the main part at that. Although the position of the epiglottis ‘contributes critically to the cavity shaping’ (Esling 1999: 364), the most important thing to consider is the action of the aryepiglottic sphincter mechanism beneath the lowered epiglottis and the constrictions in the epilaryngeal tube that may accompany it.

2.2 Manner of articulation

Regarding manner of articulation, all degrees of constriction have been reported (McCarthy 1994: 194). Stopped realisations have been found in Iraqi Arabic in all contexts (Al-Ani

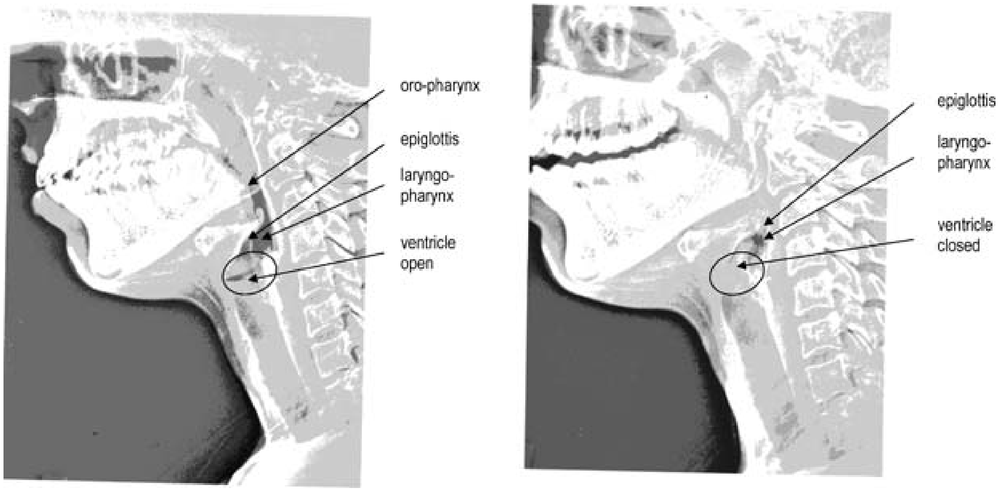


Figure 2 X-ray of vocal tract in rest position (L) and during the hold phase of a stopped geminate *'ayn* (R). Reproduced with kind permission of the author from Bukshaisha (1985: 290, 305). Labels added by present author.

1970: 63f.; see also MacCurtain 1982, cited in Laradi 1983: 120f.; also Adamson 1981: 89, for Sudanese Arabic). According to Al-Ani (1978: 95), the occlusion of the stop variant may be formed by the vocal folds or the ventricular bands, while MacCurtain's xeroradiographic data reveal evidence of a complete epiglotto-pharyngeal closure as does the X-ray photograph in El-Halees (1985: 289). Lowering of the epiglottis onto the arytenoids to close the laryngeal vestibule is reported in Pettorino & Giannini (1984: 203). While Shaheen (1979: 215) did not find any stopped realisations in Cairo Arabic, May (1981: 275f.) did observe some in initial position but did not identify the location of the occlusion.

Further evidence of stop realisations comes from X-ray photographs of geminates in Qatari Arabic, in which

> /ʕʕ/ seems to exhibit a double stricture or closure, one is an epiglotto-pharyngeal which involves the epiglottis tip and the posterior pharyngeal wall and the other is glottal which involves a sphentric [*sic* – sphincteric?] constriction of the ventricle, the vestibular and the aryepiglottic fold. This glottal constriction seems to support the delicate epiglotto-pharyngeal constriction. (Bukshaisha 1985: 304–306)

An example is provided in figure 2 which compares the mid-sagittal vocal tract configuration during a geminate *'ayn* in the word /naʕʕam/ 'to grind' with a neutral rest position. In it, we can see the compression, and even obliteration, of the resonating spaces in the larynx and pharynx while *'ayn* is being produced. A spectrogram is provided in figure 3 of a stopped realisation by a speaker of the Muslim dialect of Baghdad, Iraq, which gives the auditory impression of a 'strong' or 'massive' glottal stop of the kind described by Catford (1977: 163) and Esling (1999: 360). A clear release burst is visible at around 1 kHz.

From the literature, it appears that stopped articulations are more likely in some dialects than others, with Iraqi and Egyptian varieties representing perhaps the most and least likely respectively. Those varieties in which stops are more common tend to be the more conservative dialects, so it is little surprise if they are rarer in the Arabic of Cairo, a phonologically innovative variety (Watson 2002: 10). Phonological context is also a factor, with intervocalic singletons the least prone to full occlusion (see section 4 below).

A suggestion that *'ayn* be classed as a trill is made by Ghali (1983: 441f.) who interprets Sibawayh's term *taraddudiyyah* ('frequentative') as evidence that *'ayn* was a trill back in the

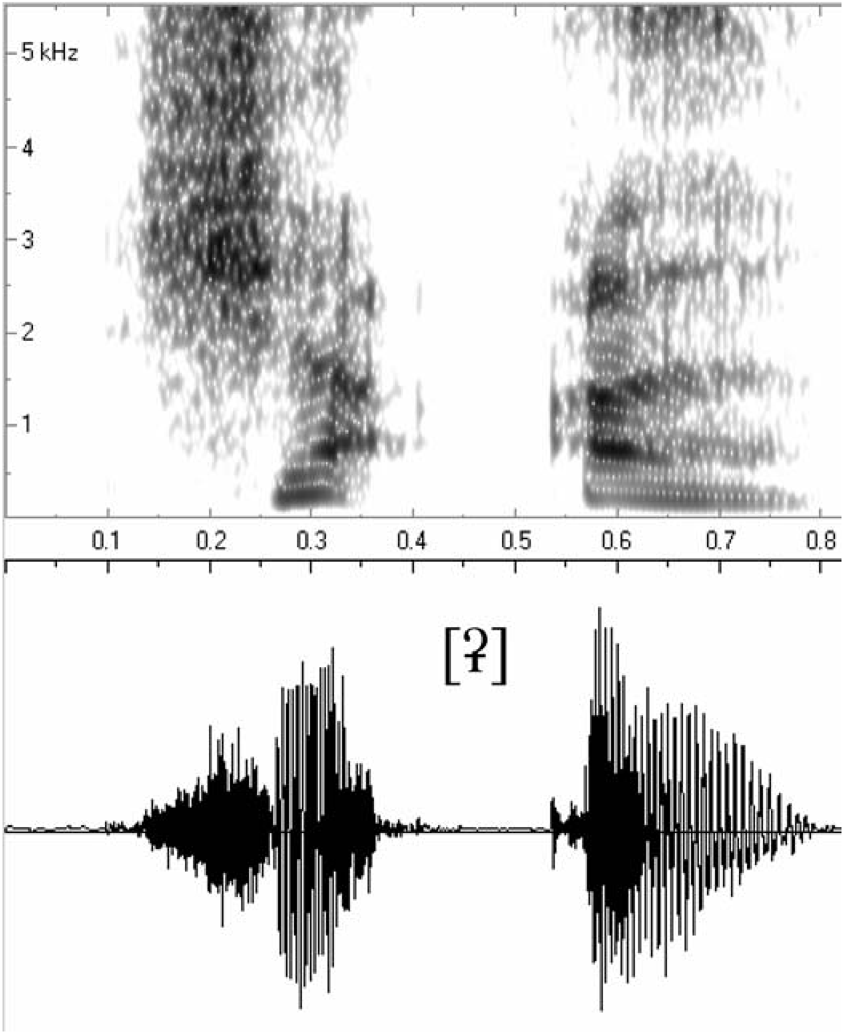


Figure 3 Spectrogram of a stopped geminate *‘ayn* in the word /ʃaʕʕa/ produced by a male speaker from Baghdad, Iraq. The auditory impression is of a ‘massive’ glottal stop involving glottal, ventricular, aryepiglottic and possibly epiglotto-pharyngeal occlusion.

eighth century. This suggestion finds little support in the literature although the voiceless /ħ/ has been observed as a trill in a speaker of Iraqi Arabic (John Esling, personal communication).

The Arabic *‘ayn* has often been classed as a fricative (Gairdner 1925: 15, Harrell 1957: 26, Garbell 1958: 307, Cantineau 1960: 73, Ghazeli 1977: 46f., May 1981: 275f., who notes however that it is realised with friction only ‘occasionally’). Laufer (1996: 114), however, claims that in Hebrew and Arabic ‘/ʕ/ is never a fricative in fluent speech’, a claim supported by the lack of any fricative realisations in the 436 realisations reported on in the current study. The motivation for classifying *‘ayn* as a fricative most likely comes from pairing it phonologically with /ħ/ which is typically realised as a voiceless fricative in all varieties.

Whether fricative realisations are ever found in ‘fluent speech’ or not, approximant or glide-like realisations do seem to be the most common. Laufer (1996) in fact suggests that the IPA symbol [ʕ] be used to denote a pharyngeal approximant rather than a voiced pharyngeal fricative, a practice which will be adopted henceforth in this paper. Reports of approximant realisations are abundant in the literature: (Al-Ani 1970: 63, Paddock 1970: 198, Gaber 1972: 271, Shaheen 1979: 215f., Adamson 1981: 94, May 1981: 275f., Bukshaisha 1985: 317, Butcher & Ahmad 1987: 170, Al-Nassir 1993: 47f.). Heselwood (1992: 235f.) found that from a sample of 40 speakers from across the Arabic-speaking region, 233 of 306 *'ayn* realisations (76%) were glides with laryngopharyngeal constriction, a substantial number of which had creaky phonation. He also found tokens that were difficult to classify, involving 'glottal stop onset followed by a constricted glide, constricted glide with glottal stop termination, sequences of creak and glottal stop' (Heselwood 1992: 235).

Similar ambiguities regarding evidence for manner (and place) of articulation attends the pharyngeal /ʕ/ phoneme in the Wakashan language Nuuchahnulth spoken in western Canada. It has been described as a pharyngealised glottal stop (Swadesh 1939: 78, Jacobsen 1969: 126) and more recently as a glottalised pharyngeal glide (Shank & Wilson 2000: 187). Shank & Wilson base their description mainly on spectrographic observation of the relative timing of the glottal and pharyngeal gestures which do seem to be temporally more separated than is generally the case in the Arabic *'ayn*, although we should be careful in generalising from their single informant. They report that the glottal gesture precedes pharyngeal narrowing in the same way that it precedes the articulatory gestures in the realisation of the 'glottalised resonant' phonemes /w'/ and /j'/ respectively (Shank & Wilson 2000: 188f.). Carlson & Esling (2003: 188f.) attribute this timing sequence to the fact that /ʕ/, like the glottalised resonants, is restricted in its phonotactic distribution to syllable onsets in Nuuchahnulth. A detailed account of the sequence of events in Nuuchahnulth pharyngeals and glottalised resonants based on laryngoscopic observation is given in Esling et al. (2005: 392–397) where they describe the Nuuchahnulth /ʕ/ as an 'epiglottal stop [ʡ] with a voiced pharyngeal offglide [ˤ]'.

Differences of manner of articulation in pharyngeal consonants are accounted for in the laryngeal articulator model by different degrees of constriction by the aryepiglottic sphincter mechanism (Esling 1999: 369). Maximum constriction brings about contact between the aryepiglottic folds and the tubercle at the base of the epiglottis (Esling 1996: 73), closing off the laryngeal vestibule. As part of the same constrictive action, the larynx may be raised up to the under surface of the retracted and lowered epiglottis (Esling 1999: 369). Fricative and approximant articulations result from less extreme action of the laryngeal sphincter mechanism.

2.3 Phonation

As well as modal voice, creaky phonation has repeatedly been reported in realisations of *'ayn* (Ghazeli 1977: 49, Shaheen 1979: 223, Adamson 1981: 89, Bukshaisha 1985: 317, Butcher & Ahmad 1987: 170, Elgendy 2001: 83f.). Agius (1992: 130) notes that creaky voice occurs as a reflex of Arabic *'ayn* in Maltese dialects.

The evidence from the literature suggests an intimate connection between articulation and phonation in productions of *'ayn* such that, typically, the more constricted the pharynx, the less modal-like the phonation. Ladefoged (2001: 146) offers the explanation that 'the necessary constriction in the pharynx also causes a constriction in the larynx' though it may be more accurate to regard laryngeal constriction as causing pharyngeal constriction. According to the laryngeal articulator model, non-modal phonation types such as creaky voice and tense voice result from approximation of the laryngeal constrictor short of full closure (Esling et al. 2005: 388) and can be interpreted as the glottal level of the mechanism exhibiting a secondary effect of the constriction at the aryepiglottal level (Esling et al. 2005: 402f.). Odisho (1973: 71) states that 'certain degrees of pharyngealisation and laryngealisation are always present' in realisations of *'ayn*, a point also made by Ghazeli (1977: 49). The most extreme situation is one in which there is closure at all levels of the laryngeal constrictor mechanism at the same time, as seems to be the case in Bukshaisha's description quoted above. As the high level of tension responsible for these occlusions is relaxed, the vocal folds become freer to vibrate. Moderate tension facilitates modal voice (Laver 1980: 111) while degrees of tension intermediate between moderate and extreme produce phonatory patterns that deviate from normal modal voice to a lesser or greater extent.

2.4 Larynx raising

Several researchers have noted that the larynx is raised during productions of *'ayn* (Catford 1977: 163, Ghazeli 1977, 36, Laradi 1983: 105, Bukshaisha 1985: 299, Butcher & Ahmad 1987: 167, Elgendy 2001: 83). Phonetically trained Arabic speakers do confirm that their larynx rises for its production. Ghazeli (1977: 36) measured the rise to reach 7 mm above its normal speech position. In a laryngoscopic investigation of pharyngeal articulations in general, Esling (1999: 357) found that larynx raising brings about approximation and even closure of the aryepiglottic folds, tongue retraction and retraction of the epiglottis. He concludes that 'raised larynx setting thus appears to be an inherent trait of the pharyngealized posture' (Esling 1999: 359). Laver (1980: 27f.) draws attention to the difficulty of maintaining a raised larynx without also constricting the pharynx and affecting the pitch of phonation, although Trigo (1991) presents evidence, mainly from Nilotic languages, that larynx elevation, tongue-root retraction and pitch are independently controllable during speech. In Arabic, however, it seems that larynx-raising does not occur without pharyngeal constriction, usually involving the root of the tongue and the epiglottis. On its own, a raised larynx should increase formant frequency values in open vowels due to shortening of the supralaryngeal vocal tract (Laver 1980: 27), but, with the possible exception of /u/, vowels adjacent to *'ayn* overwhelmingly exhibit a lowering of F2 strongly indicative of pharyngeal constriction.

2.5 Is there an articulatory 'target' for *'ayn*?

The range of variation found in the manner of articulation for *'ayn* has prompted speculation about what the target degree of approximation might be. Butcher & Ahmad (1987: 171) raise the possibility that it is a target fricative 'in an area of the vocal tract not particularly well suited to such articulations' but point out that it could also be an approximant 'involving general constriction of the pharynx . . . or a stop articulation, either at the epiglottis or, perhaps more likely, at the glottis.' Accounting for approximant realisations in terms of undershooting a fricative target is plausible enough, for example, in terms of the gestural framework developed in Articulatory Phonology (Browman & Goldstein 1991), or of Mowrey & Pagliuca's (1995) modelling of articulatory change as essentially reductive. It is problematic, however, to account for stopped realisations in these terms if a fricative target is assumed. Adamson (1981: 94) argues that 'the stop articulation seems to be the ultimate target of /ʕ/, but the gesture does not need to be completed to allow the phoneme to be perceived as /ʕ/' (see section 3.5 below). The tendency for stop articulations to occur more in careful citation style (Laufer & Condax 1979: 54) is consistent with Adamson's view, although it should not be taken as conclusive evidence of a target. We should ask whether in fact there is any need to specify a phonological manner of articulation for *'ayn* at all, given that different degrees of pharyngeal constriction are not distinctive in Arabic (McCarthy 1994: 195). Sibawayh seems to have reached a similar conclusion some twelve hundred years before when he classified *'ayn* as neither a stop (*shadiyd*) nor a continuant (*rixwah*) (Al-Nassir 1993: 47). What we are probably seeing in the variation in manner, place and phonation in Arabic *'ayn* are effects of various stages in diachronic change brought about by progressive weakening of pharyngeal and laryngeal tension over time of the kind hypothesised to be responsible for changes in the realisation of the Arabic emphatic consonants as they evolved from Proto-Semitic ejectives (Heselwood 1996: 32–34); speculation that these historical processes may have taken place does not necessarily presuppose that all modern variants came from one single historical source, however. No doubt this weakening interacts with stylistic and sociolinguistic influences to produce the quite complicated cross-dialectal and intra-dialectal picture which we find characterised by rather more gross variability than we are normally used to. (It is somewhat surprising, given such extensive variation, that *'ayn* does not seem to figure widely as a sociolinguistic stereotype in those Arabic-speaking communities that maintain an *'ayn* phoneme compared, say, to the etymological interdental fricatives and uvular stop (e.g. Daher 1998, Al-Shareef 2002, Al-Tamimi 2001)). It would perhaps be a mistake to try to identify any particular variant in this diachronic – synchronic *mélange* as the putative target which is missed most of the time by most speakers. If we are to think in terms of a target, it needs to be one which is highly tolerant and rather loosely defined in phonetic terms.

A particular problem in attempting to classify *'ayn*, and one which has implications for the notion of a target type, is distinguishing between articulatory events which are central and 'intended' by the speaker on the one hand, and those which are epiphenomenal on the other. Returning to Esling's observation that larynx raising co-occurs with aryepiglottic approximation, tongue retraction and retraction of the epiglottis, can we identify one of these events as central, and therefore more robustly represented in the speaker's intention than the others? If so, then the other events are there either to bring about the central event, or as automatic but unintended side effects, or for purposes of additional auditory enhancement. The debate over whether the epiglottis functions as an articulator tends sometimes to be couched in these kinds of terms (e.g. Laufer & Baer 1988: 198). Alternatively, we could regard each identifiable event in the total complex of events as making an equally important contribution to the whole regardless of physiological cause-and-effect relations. There is not the space to argue the issue out in this paper, but the latter view seems more reasonable when we are dealing with sounds like *'ayn* where several interconnected events and adjustments over a relatively long section of the vocal tract are implicated in the workings of the laryngeal articulator mechanism. On this point, McCarthy (1991, 1994) links the lack of place-of-articulation distinctions within the pharynx to research that suggests that proprioceptive sensation is less localisable than in the buccal chamber. Place distinctions cross-linguistically are finer at the front of the mouth, areas that are more richly innervated and associated with larger cortical projections in the brain than areas further back. Following Perkell (1980), McCarthy (1994: 201) speculates that '[W]ith respect to orosensation, the regions of the vocal tract covered by each of the features [labial], [coronal], [dorsal], and [pharyngeal] may be equal in subjective size' and that 'orosensation provides a kind of quantization of the vocal tract'. It would seem there may be more scope for different articulatory activities in the pharynx to count as 'the same' than is the case elsewhere, and more scope for several spatially distributed structures to be actively involved in the production of a single sound. Pike (1943: 121), for example, claims that faucalisation is associated with 'a type of lower pharyngeal constriction, glottal tension, and usually a raising of the larynx', all events which are observable in productions of *'ayn*.

Rather than a target expressible in terms of a single traditional place and manner of articulation, *'ayn* appears to involve multiple articulations, none of which is much constrained phonologically for degree of constriction. The multiple articulations include glottal, ventricular, aryepiglottic, epiglottal-arytenoidal, epiglotto-pharyngeal and faucal, all of them built on a raised larynx setting, and most of them capable of reaching full occlusion. The common factor is muscular tension in the whole laryngeal-pharyngeal area as a result of the engagement of the laryngeal sphincter mechanism and a consequent compression of all the resonating spaces in this part of the vocal tract: laryngeal vestibule, pyriform sinuses, vallecula and the ventricles, as can be seen in the stopped geminate in figure 2.

Like most variants of *'ayn*, the tight approximant variant displays the articulatory complexity described above but with an overall tension which is less than is found with stopped variants, although greater than is found with 'loose' approximants. By looking at the acoustic and auditory properties of the tight approximant variant we may gain some insights into why *'ayn* in general is characterised by such articulatory complexity.

3 The tight approximant variant

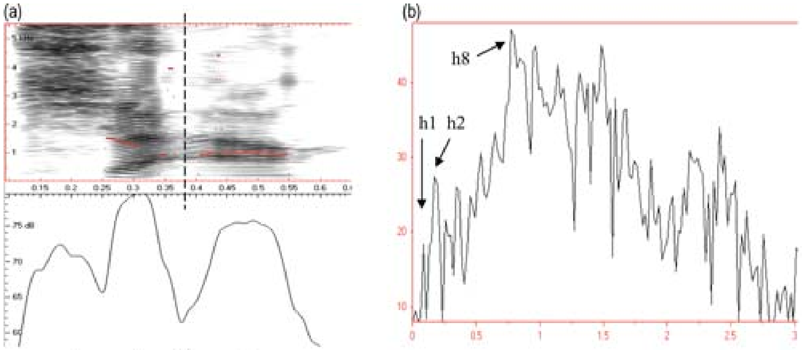
Having outlined Esling’s model of the laryngeal constrictor and how it accounts for pharyngeal articulation, and having reviewed previous descriptions of the production of Arabic *‘ayn*, we are now in a position to give an articulatory characterisation of the tight approximant variant. That it is an approximant is clear from the presence of continuous acoustic energy and the absence of any auditory or acoustic evidence of turbulent airflow. The tightness refers to the impression that a high degree of constriction is present in the articulatory mechanism, higher than is normally associated with strictures of open approximation. In terms of Esling’s model, it is likely that the laryngeal sphincter contracts to a degree just short of what is required for an epiglottal stop [ʡ], i.e. just short of the contraction required to bring the aryepiglottic folds together and to close against the laryngeal tubercle with the larynx raised up such that it brings about approximation between the inferior surface of the epiglottis and the arytenoids, and between the tip of the epiglottis and the rear pharyngeal wall. For [ʡ], complete aryepiglottic adduction is thus added to a lower glottal adduction, and possibly also ventricular adduction. In the tight approximant state, the glottis is adducted, the ventricular bands closely approximated, the laryngeal aditus reduced in area and the larynx raised towards a retracted epiglottis. This degree of constriction is not enough to prevent phonation but is sufficient to prevent the volume velocity of airflow needed to generate turbulence through the articulation. The effect on the resonating spaces is to compress them: ventricular tension compresses the ventricles, larynx raising compresses the length of the epilaryngeal tube and the retracted epiglottis compresses the pyriform sinuses. It is this total configuration rather than the approximation of any particular two components of it which is said to be ‘tight’. In short, there is tight constriction of the whole mechanism responsible for pharyngeal articulation but not tight enough to effect complete closure. Spectrographic and laryngographic data relating to the tight approximant are presented in sections 3.2 and 3.3. In section 3.4 attention is turned to the auditory processing of this sound.

To understand more about the phonetic structure of the tight approximant variant, we shall look in considerable detail at just a few selected tokens, mostly geminates in the words /ˈʃaʕʕa/ ‘to disperse’ and /ˈwaʕʕad/ ‘to make someone promise’, which show some inter-speaker variation in certain parameters, namely extent of bandpass filtering, the closed quotient of the glottal cycle, and fundamental frequency.

3.1 Data and analysis

The data for the study comprise in all 436 tokens of *‘ayn* provided by 21 speakers (17 males (M), 4 females (F)) from eleven countries spanning almost the entire Arab region: Morocco (2M), Algeria (1M), Tunisia (1M), Egypt (1M, 1F), Palestine (5M), Jordan (3M), Syria (1M, 1F), Iraq (1M, 1F), Saudi Arabia (1M), Kuwait (1M), Qatar (1F). All speakers were aged 20–50 years and were educated to at least university undergraduate level. They produced words containing *‘ayn* in the carrier sentence *‘_____, maa maʕna _____ bil luʁah al-ingliziyyah’* (‘_____, what is the meaning of _____ in English?’). The full word-list is given in the appendix. Each word was thus produced twice: once in utterance initial position followed by pause at the intonation-group boundary, and once between a vowel-final word and a consonant-initial word. In each case the target word contained the tonic syllable.

Style of speech is always an issue in collecting Arabic data, with speakers differing to some extent in how much they move towards a ‘standard’ or prestige pronunciation (Ibrahim 1986), and which prestige form they move towards – regional, supraregional, or literary. The impression is that most speakers, and all the ones represented in the figures, used what could be described as a semi-formal style of a prestige form of their own dialect, or what Daher (1998: 188) describes as that part of the stylistic continuum where the vernacular meets an informal variety of Standard Arabic. The purpose of the investigation, however, is not to compare dialects or to characterise speech styles but to gain some preliminary insight into the



	ʃ	a	ʕʕ	a
F0 (Hz)		146	94	98
Power (dB)		0	–18	–4

(a) A geminate *'ayn* realised as a tight approximant in the word /ʃaʕʕa/ by a male speaker from Fez, Morocco, showing a drop in amplitude and F0 (superimposed trace at F0 x 10), and a bandpass filter effect.

(b) Spectrum of (a) at the dashed line showing positive spectral tilt h1–h2 of +9.1dB and h1–h8 (F1 harmonic) of +28.7dB.

Figure 4 A tight approximant realisation of geminate *'ayn* by a speaker from Fez, Morocco, showing a reduction in amplitude and F0, and a bandpass filter effect. Narrowband spectrogram (512-sample window size), Hamming window spectrum.

general prevalence of the tight approximant variant and, above all, to understand its phonetic structure.

All acoustically analysed tokens were digitised at a sampling rate of 11025 Hz from analogue tape recordings, some of which were made in a recording studio and some in the field. The low sampling rate was chosen in order to get better resolution of the frequency range in which the spectral events associated with vowels and approximants, and pitch, take place, i.e. below 5 kHz (Ladefoged 2003: 26). Laboratory laryngographic recordings were obtained from six (5 males, 1 female) of the 21 informants.

3.2 Resonance properties

Figure 4(a) shows a narrowband spectrogram of the tight approximant variant in a geminate context by a male speaker from Fez, Morocco. The characteristic acoustic feature of the variant is the weakness of the lowest few harmonics compared to those between about 1–1.7 kHz, giving a steeply positive spectral tilt as can be seen in the spectrum in figure 4(b). Similar productions by a male speaker from Khan Younis, Gaza Strip, and a male speaker from Irbid, Jordan, are presented in figures 5 and 6, respectively.

Where spectral tilt is measured, two values are given. The first is the difference in amplitude between the first and second harmonics (Hayward 2000: 235, Gordon & Ladefoged 2001: 397), the second being the difference in amplitude between the first harmonic and the lowest-frequency harmonic contributing to the first formant resonance (F1), a variation on the measure given by Gordon & Ladefoged that takes the harmonic closest to F1 centre-frequency, and that given by Hayward taking the strongest F1 harmonic. The reason for this slight difference is that it better captures the sudden jump in amplitude from the lower harmonics to the higher more prominent ones which, it is claimed, is a defining acoustic characteristic of the tight approximant variant. The harmonics above the more prominent ones are similarly weak such that the signal has the appearance of having been bandpass filtered with a half-power bandwidth of about 130–150 Hz centred fairly close to 1 kHz. With a low fundamental frequency, the harmonics are brought close together such that more of them are passed through the filter than would be the case with a higher F0. In females, where F0 is typically about 80% higher, fewer harmonics are involved in each frequency region.

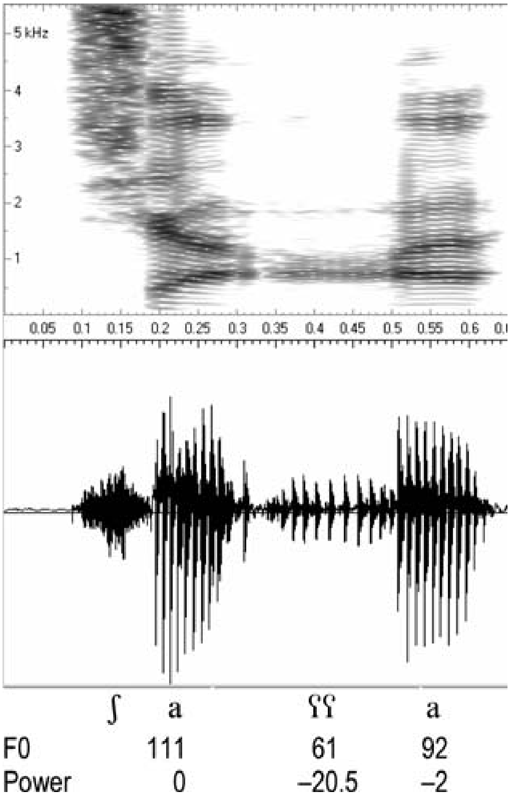


Figure 5 A geminate *'ayn* realised as a tight approximant by a male speaker from Khan Younis, Gaza Strip, also showing a drop in amplitude and F0, with creak pulses on the waveform and a bandpass filter effect.

The harmonics form a band of resonance similar to that reported for *'ayn* by Adamson (1981: 87). This is not just a consequence of F1 and F2 converging and thus concentrating acoustic energy in this part of the spectrum, as can be seen by comparing the spectrographic pattern for *'ayn* with the vowel on each side in figures 4, 5 and 6. Approximation of the formants tends to occur before, and persist after, the filtering of the harmonics, indicating that co-articulatory pharyngeal narrowing takes place independently of the adjustments responsible for the bandpass filtering effect. If Carlson & Esling (2003: 190f.) are correct in claiming that tongue root retraction only occurs in the presence of a contracted laryngeal sphincter, and if contraction of the laryngeal sphincter is part of the mechanism by which the reduction in amplitude of lower harmonics is achieved (see below), then for the F1–F2 approximation to be independent of the laryngeal sphincter contraction it must be brought about by an alternative means of pharyngeal narrowing, i.e. a general sphincteric action of the pharyngeal wall (Shaheen 1979: 216), involving the middle pharyngeal constrictor (Laver 1980: 61). Further research is needed to shed more light on this.

Interestingly, this filtering phenomenon seems to occur in other languages with pharyngeal articulations. As far as one can tell from the wideband spectrograms in Shank & Wilson (2000) it is found in Nuuchahnulth – see especially their figure 5 (p. 193) where /ʕ/ is intervocalic. It also appears in the spectrograms of the 'pharyngealised vowel' [ɑ^{ʕ}] in the southern African Khoisan language !Xóõ presented in Ladefoged & Maddieson (1996: 309), although these show a raising of F2 indicative of an even lower constriction location and shorter narrowed section of the pharynx (Stevens 1998: 274–277). A raised F2 is also evident in figure 4,

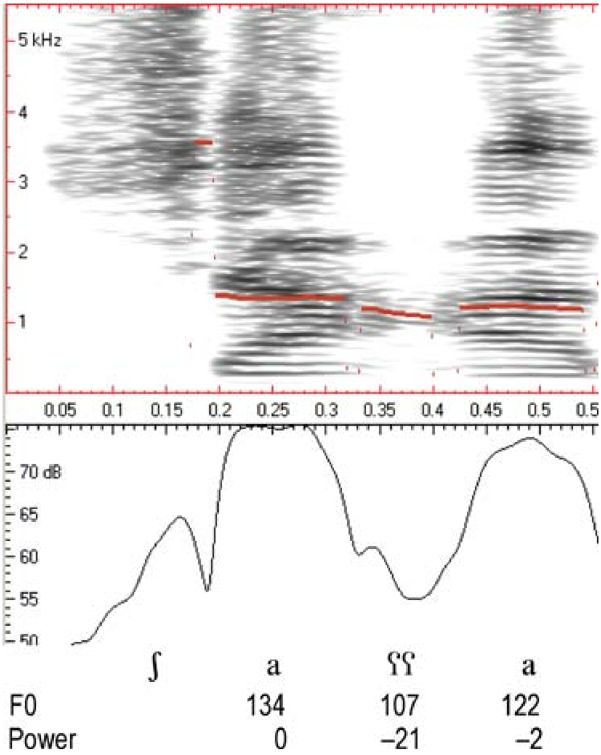


Figure 6 A geminate *ʻayn* realised as a tight approximant by a male speaker from Irbid, Jordan, again showing a drop in amplitude and F0 and a bandpass filter effect. Laryngograms of this token are shown in figure 8.

contrasting with the lowered F2 in the other figures presented here. According to perturbation theory (Chiba & Kajiyama 1941), for F2 to raise, a stricture in the pharyngeal region must be closer to the velocity antinode at the glottis than to the node of velocity in the mid-pharynx. In a vocal tract of 17 cm, this means the constriction has to be no more than 2.83 cm above the glottis, a height corresponding to around the level of the aryepiglottic folds.

What might be responsible for this bandpass filtering effect? Laver (1980: 141–156), in discussing ‘tense voice’, reviews a number of writers’ comments concerning the contribution of tense musculature in the pharynx and upper larynx to the suppression of lower harmonics. The role of the ventricles of Morgagni, located between the vocal folds and ventricular bands, in acting as a low-pass filter has been suggested by Pepinsky (1942: 32) who identifies the external thyroarytenoid and stylopharyngeus muscles as capable of adjusting the effective size of the ventricles under the speaker’s control. He describes the ventricles as Helmholtz resonators, remarking that the mucosal linings give them a broad natural resonance frequency band. Van den Berg (1955: 63) states that the descent of the ventricular bands onto the true vocal folds prevents the ventricles’ low-pass filtering, which in effect turns the ventricles into band-stop filters for the lower harmonics. This ventricular configuration is apparent in figure 2, and hypothesised as characteristic of the tight approximant variant, albeit in a less extreme form. Consequently, the lowest harmonics including the fundamental are weakened. According to Kaplan (1960, cited in Laver 1980: 143), the pharynx can have a band-stop filtering effect when it is constricted and its musculature contracted to create hard reflective surfaces. The contraction of the pharynx may in fact be reinforced by ventricular constriction (Pepinsky 1942: 32). The various components of the articulation of *ʻayn* located in the epilaryngeal tube can be seen to conspire to reduce the amplitudes of frequencies below the prominent harmonics in tokens such as those in figures 4, 5, 6 (above) and 7, 9, 11 and 13 (below). That is to say, when the laryngeal sphincter mechanism assumes the degree of tension responsible for the tight approximant variant, the ventricles are compressed.

The weak spectrum above the band of harmonics can be attributed to the narrowness of the tight approximant constriction. Because glides involve closer approximation than their corresponding vowels, the higher frequencies in the spectrum are reduced in amplitude (Stevens 1998: 532). In the case of the tight approximant variant of *‘ayn*, approximation is even closer (see e.g. figure 1) and only avoids crossing the turbulence threshold because the constricted glottis reduces the airflow.

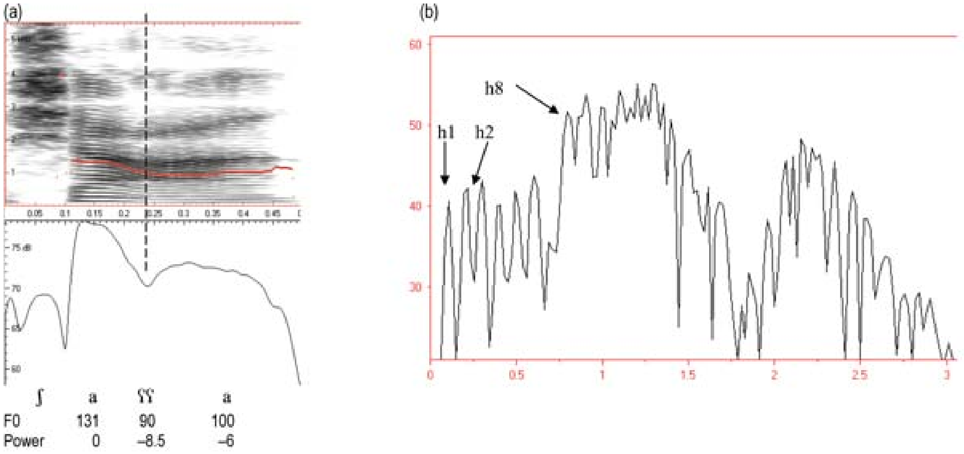
Suggestions have been made that *‘ayn* is characterised by some nasality (Laradi 1983: 164–168) which could account for reductions in higher-frequency amplitudes where pharyngeal articulation is implicated (Stevens 1998: 313–315). Elgendy (2001: 54f.) observed nasal-oral coupling during geminate /ʕʕ/, and it may well be that laryngeal constriction, pulling down on the palatoglossus muscles, could result in opening of the velopharyngeal port. Other researchers, however, have disconfirmed the presence of nasal resonance (Ghazeli 1977: 41, Bukshaisha 1985: 322, Butcher & Ahmad 1987: 167). Laver (1980: 86) identifies various possible kinds of side chamber resonance which may be perceived as nasality, including pharyngeal and laryngeal. Laryngeal can most probably be ruled out in the case of *‘ayn* because the ventricular side chambers are compressed and we find the opposite of the expected boosting of the harmonics below F1 (Laver 1980: 92). We can provisionally identify a general sphincteric narrowing of the pharyngeal tube as being responsible for the band-stop filter effect affecting those frequencies above the prominent harmonics in the spectrograms and spectra shown in the figures. This may be responsible for perceived nasality and for the historical development of nasal cognates of *‘ayn* reported by Rabin (1951: 31f.) in some West Arabian dialects, and by Hetzron (1969) for the Ethiopian Semitic language East Gurage.

The reduction in amplitude of the harmonics above and below the band-pass frequencies together reduce the overall power in the signal for the duration of the tight approximant as can be seen in the power displays in figures 4, 6, 7, 9, 11, 13, and the waveform display in figure 5.

There can be different degrees of tightness in the tight approximant, depending on the amount of muscular tension in its articulation. The spectrograms in figures 4, 5 and 6 represent particularly tight tokens, but a ‘looser’ token is shown in figure 7. Here, the harmonics outside the band of resonance formed by the approximation of F1 and F2 do exhibit a weakening of energy but not to the same extent. We can interpret this to the effect that the compression of the ventricles, vallecula, and laryngeal vestibule, and the overall tension of the pharyngeal constrictor muscles, is reduced, and therefore they are not as effective in inducing the band-stop filter effect. The auditory effect is still that of a ‘strained’ sound, but less so compared to tighter realisations. These differences are the basis for recognising weak and strong versions of the tight approximant discussed in section 4.

3.3 Phonatory properties

The important need for data relating to vocal fold behaviour in the production of pharyngeal sounds is acknowledged by Esling (1996: 74f.) and Shank & Wilson (2000: 190). As Esling points out, the vocal folds cannot be observed laryngoscopically if the epiglottis is retracted far enough to obscure them, or if the aryepiglottic folds are closely approximated. Information on vocal fold behaviour can be gleaned from spectrograms, but with less confidence than from more direct means such as electrolaryngography. For example, creaky phonation shows up on spectrograms as irregular striations, but it is not always easy to tell these from other events such as epiglottal trilling or a glottal stop release burst. Such confusions have been noted before specifically in relation to *‘ayn* (e.g. Adamson 1981: 87). With electrolaryngography we can be more certain about the contribution of the vocal folds at any given point in time.

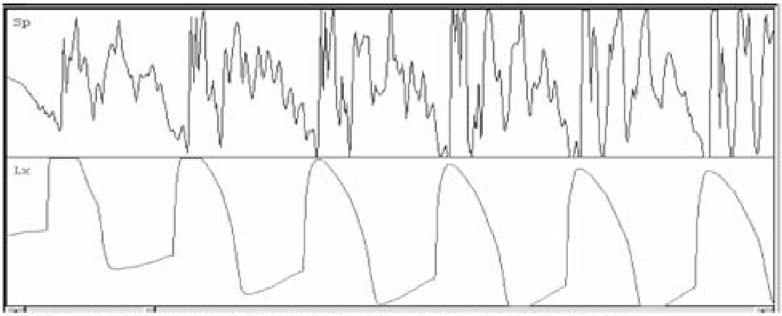


(a) The amplitude reduction outside of the F1 and F2 regions is less extreme than in figures 3 and 4, but there is a similarly extensive drop in F0.
(b) Spectrum of (a) at the dashed line showing positive spectral tilt h1–h2 of +1.7dB, and h1–h8 (first F1 harmonic) of +11.1dB.

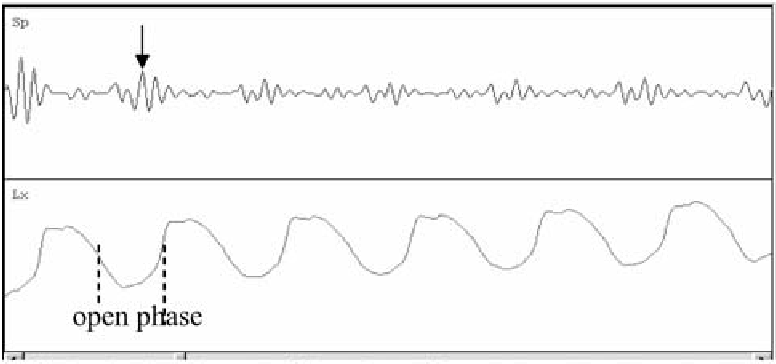
Figure 7 A 'looser' example of the tight approximant allophone. The speaker is a male from Amman, Jordan.

Figure 8 contains three six-cycle segments taken from the production of the word /ʃaʕʕa/ ('to disperse') shown spectrographically in figure 6. In each panel the upper display shows the acoustic waveform and the lower one a synchronised electrolaryngographic (Lx) waveform. The rising slope of the waveform corresponds to the closing of the vocal folds, and the descending slope to the opening of the vocal folds. Maximum vocal fold closure corresponds to the crests of the wave, maximum opening to the troughs. It can be seen clearly that both waveforms are qualitatively different during the *'ayn* (middle panel) compared to the immediately preceding and following vowels. Glottal cycles are longer, resulting in a lower Fx value (the laryngographic correlate of acoustic Fo), and exhibit an increased closed quotient (Qx – the ratio of the closed phase to the open phase, taking the closed phase to correspond to the upper 70% of the wave height). The amplitude of the acoustic pulses is considerably reduced, as can be seen in all the waveforms of *'ayn* shown in this paper, which is typical of glides compared to vowels (Stevens 1998: 519f.). The shallower slopes of the Lx wave indicate slower movement of the vocal folds during both adduction and abduction, and it can be seen that the alignment between glottal closure and acoustic excitation is shifted. Slower adduction is not what is predicted in the presence of an increased Qx value (Gordon & Ladefoged 2001: 399). However, there are peculiarities that may explain this discrepancy. During the vowels, maximum excitation occurs just after closure is attained as is expected in normal modal voice. By contrast, it occurs in this token during the open phase in the *'ayn* (marked by an arrow in figure 8). The slower adduction suggests there is not the same acceleration of airflow through the glottis to create the 'snapping shut' of the folds which is normally linked to the acoustic excitation of the supraglottal chambers. If the ventricular bands are closely approximated, a result of the laryngeal sphincter action, the Bernoulli effect is nullified (van den Berg 1955: 62), which would explain the slow closure of the true folds. The notches in the closed part of the Lx waveform in the *'ayn* indicate instability of contact between the vocal folds during the closed phase. Sluggish adduction and unstable contact are consistent with the presence of constraints on normal vocal fold movement due to increased laryngeal tension. Precisely what is responsible for the asynchronous Lx peaks and acoustic peaks in this waveform remains a matter for speculation.

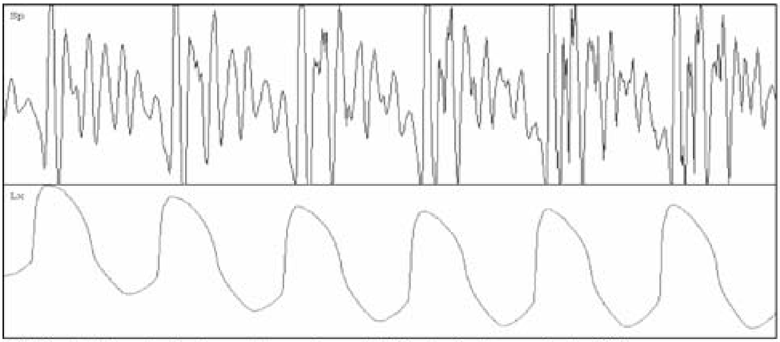
Another narrowband spectrogram of a tight approximant token can be seen in figure 9. Figure 10 shows the synchronised Lx and acoustic waveforms of the same production where



(a) Six glottal cycles in the stressed vowel [a] just prior to the geminate [ʕʕ].
Average Fx = 134 Hz. Average Qx = 52%. Lx waveform typical of modal voice.

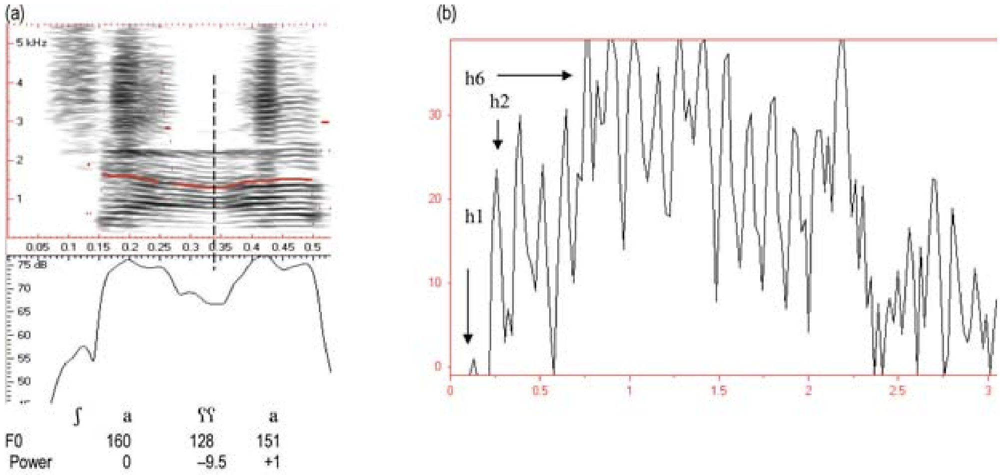


(b) Six glottal cycles in the centre of the geminate [ʕʕ].
Average Fx = 110 Hz. Average Qx = 55%. Lx waveform shows slower closing and opening movements, and unstable closure. Acoustic waveform shows very low amplitude pulses shifted to the open phase of the Lx cycle (e.g. at arrow).



(c) Six glottal cycles in the unstressed vowel [a] just after the geminate [ʕʕ].
Average Fx = 122 Hz. Average Qx = 50%. Waveform typical of modal voice.

Figure 8 Panels showing acoustic (top) and Lx (bottom) waveforms comparing geminate [ʕʕ] with preceding and following vowels. Speaker from Irbid, Jordan. Same token as figure 6.

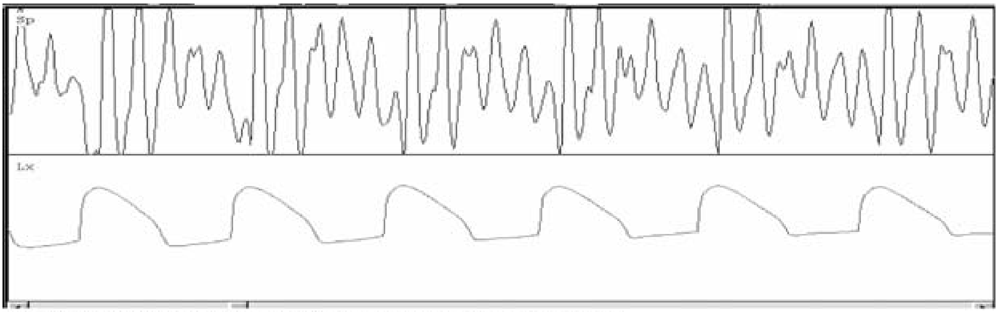


(a) Reduction in F0 and amplitude can be seen as well as the bandpass filter effect.
(b) Spectrum of (a) at dashed line showing positive spectral tilt h1–h2 of +22.6dB, and h1–h6 (first F1 harmonic – off the scale) of +46.6dB.

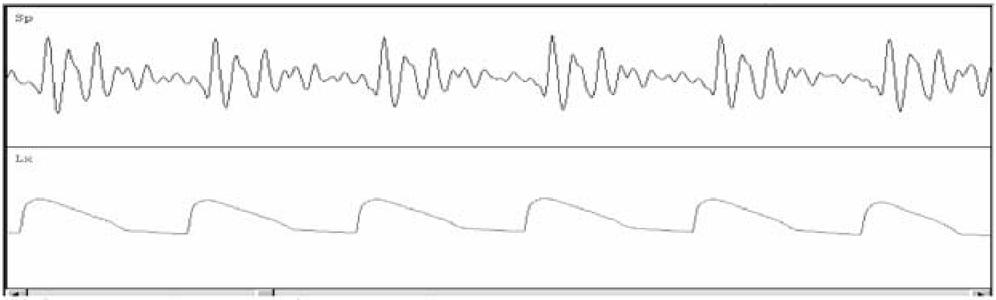
Figure 9 A tight approximant realisation of geminate *ʿayn* by a male speaker from Saudi Arabia. Laryngograms of this token are shown in figure 10.

a similar situation to that in figure 8 can be seen, but with some differences. While there is a drop in Fx and an increase in Qx, the Lx waveform is noticeably different in shape and in its alignment with peaks of acoustic excitation which is not shifted from that found in the adjacent vowels. The Lx waveform during the *ʿayn* in figure 10 has the typical shape of a pulse of creak but without the irregularity and with none of the ‘double-pulsing’ often found in creak (Laver 1980: 124f., Gerratt & Kreiman 2001: 376). Note that although the Qx value only drops 1% in the first six cycles of the following vowel, the waveshape is very different from that in the *ʿayn*, being instead like that in the first vowel.

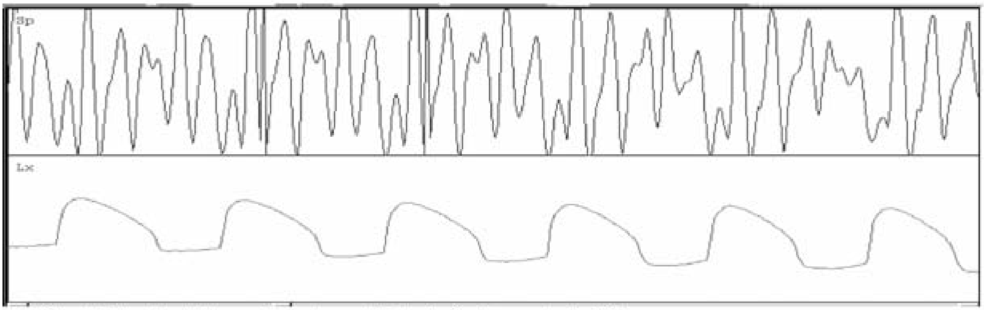
In figures 8 and 10, there is clear indication of a deviation from normal modal voice phonation due to greater glottal tension during the production of *ʿayn* than during the production of the adjacent vowels. Lower rate of vibration and longer closed phase are features of creak (Gerratt & Kreiman 2001: 376) but the Lx waveforms are not typical of creak in that there is no obvious irregularity, F0 is not down in the typical creak range below 100 Hz, there is no double-pulsing, and the Qx value has only increased relatively slightly. The evidence suggests, therefore, that the greater glottal tension is not of the kind to produce creak, but that it does produce phonation which is markedly different from that employed for the vowels. A mild form of the harsh voice described by Laver (1980: 126–132) seems a possible candidate. Harsh voice is brought about by compression of the ventricles and general laryngeal tension, and is characterised by an F0 ‘consistently (but not markedly) above 100 Hz’ (Laver 1980: 130). Esling & Harris (2005: 372–377) subdivide harsh voice into low-pitch, mid-pitch and high-pitch types, giving an example of mid-pitch with F0 of 100 Hz. However, the F0 irregularity associated with harshness is not appreciably present in the tight approximant variant of *ʿayn*, which may indicate we are dealing instead with the ‘pressed voice’ as described by Hayward. She characterises pressed voice as involving greater vocal fold adduction which ‘typically results in a rather tense quality’ but does not exhibit marked irregularity (Hayward 2000: 223f.). Her spectrum of pressed phonation (Hayward 2000: 236) shows a similar positive H1–H2 spectral tilt (+ 11.4 dB) as the spectra in figures 4(b), 7(b), 9(b) and 13(b). Stevens (1998: 85) explains that in pressed voice the time period of glottal cycles is increased and low frequency components have reduced amplitude, features which



(a) Six glottal cycles in the stressed vowel [a] just prior to the geminate [ʕʕ].
Average Fx = 150 Hz. Average Qx = 49%. Lx waveform typical of modal voice.



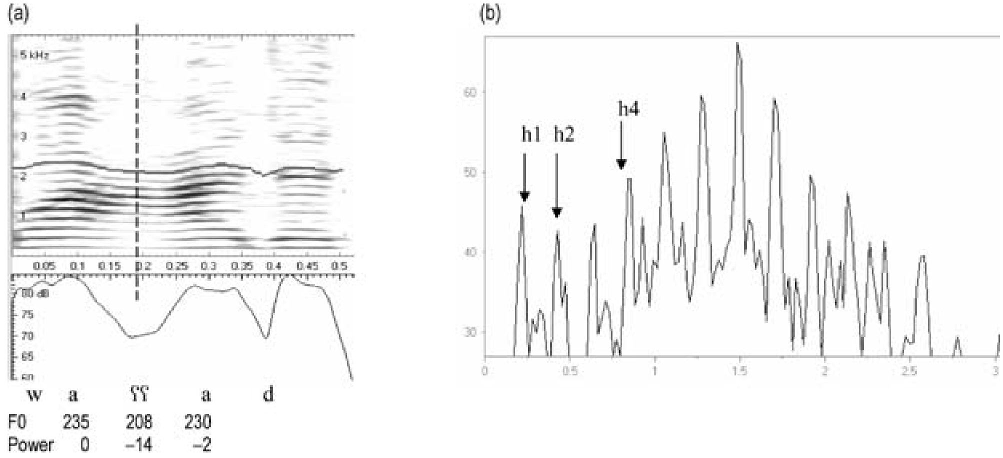
(b) Six glottal cycles in the centre of the geminate [ʕʕ].
Average Fx = 129 Hz. Average Qx = 53%. Lx waveform shows shorter excursions and slower opening movements. Acoustic waveform shows lower amplitude pulses.



(c) Six glottal cycles in the unstressed vowel [a] just after the geminate [ʕʕ].
Average Fx = 143 Hz. Average Qx = 52%. Waveform typical of modal voice.

Figure 10 Panels showing acoustic (top) and Lx (bottom) waveforms comparing geminate [ʕʕ] with preceding and following vowels. Male speaker, Saudi Arabia. Same token as figure 9.

are certainly characteristic of the tight approximant allophone of *‘ayn*. Esling (1999: 368) notes that ‘it is likely that phonation affected by the sphincter will be creaky or ventricular – “tight” or “pressed” in some terminologies’. (In a later publication, Esling & Harris (2005: 376) claim that the traditional term ‘ventricular voice’ refers to a phonation in which the ventricular bands do not vibrate but are compressed. They identify its auditory quality with their high-pitch harsh voice category.)



(a) Another 'looser' token showing the characteristic drop in F0 and amplitude.
(b) Spectrum of (a) at dashed line showing negative spectral tilt h1–h2 of −3.1dB, and positive tilt h1–h4 (first F1 harmonic) of +5.8dB.

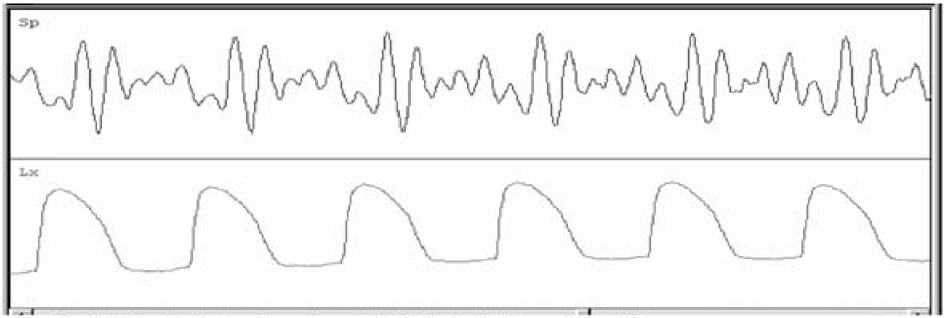
Figure 11 A tight approximant realisation of geminate *'ayn* in the word /waʕʕad/ by a female speaker from Qatar. Laryngograms of this token are shown in figure 12.

It may be useful to recognise categories of progressively increasing tension and irregularity from modal voice through whispery voice (see below), pressed voice, harshness and creak, ending with the complete occlusion in [ʡ]. Tight approximant realisations of *'ayn* are found to employ one or other of the intermediate categories, depending on the degree of constriction of the laryngeal sphincter.

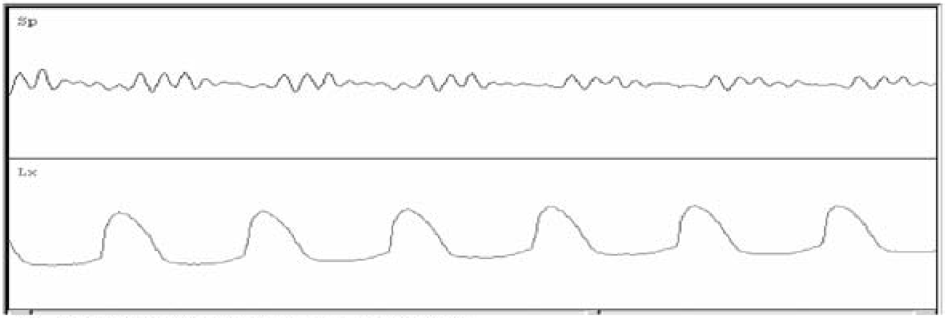
In some tokens of the tight approximant variant, F0 does descend further and can become creak-like as in figure 5 (no laryngographic recordings were obtained from this speaker). Creak also involves compression of the ventricles and laryngeal tension according to studies cited in Laver (1980: 123). Very low F0 has long been seen as a typical property of *'ayn*. Gairdner, for instance, describes the auditory quality of *'ayn* as 'resembling a sort of growl rather than a musical note' and advises learners of Arabic to '[S]ing down to your bottom note – and then one lower' (Gairdner 1925: 28). He pitches the *'ayn* at D2 on the musical scale (i.e. the D two octaves below middle C), about 73 Hz. It is possible that Gairdner's 'growl' involved aryepiglottic trilling which is set in train at about 50–60 Hz if airflow reaches a critical threshold (Esling & Harris 2005: 374). 55 Hz is A3, a musical fourth below D2. In the data examined for this study no tokens of aryepiglottic trilling were identified.

With low frequency and low amplitude, tokens such as those in figures 4 and 5 can sound like stops, a perceptual effect which is not unexpected in the light of results reported by Hillenbrand & Houde (1996: 1187–1189) who found that rapid decreases of F0 of about 10 Hz or more and sudden amplitude drops of at least 8 dB tend to give rise to the perception of a stop. This phenomenon raises an important question regarding how to categorise such tokens. If we class them according to instrumental data, they fall into the tight approximant category, but from an auditory-perceptual point of view they should perhaps be classed as stops if that is how people hear them. We will return to this issue in section 3.5 below.

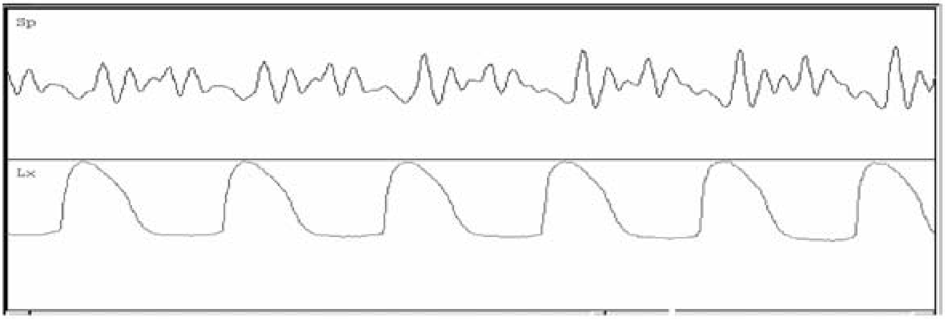
Instead of a slight increase in Qx values, some speakers exhibit a reduced closed quotient during *'ayn* compared to adjacent vowels. Figure 11 shows a spectrogram and spectrum of the word /ˈwaʕʕad/ produced by a female speaker from Qatar. Synchronised Lx and acoustic waveforms of the same token can be seen in figure 12. The Qx value has dropped considerably and the relatively long open phase can be clearly seen on the Lx waveform which has the typical appearance shared by breathy and whispery voice. The presence of the filtering



(a) Six glottal cycles in the stressed vowel [a] just prior to the geminate [ʕʕ].
Average Fx = 234 Hz. Average Qx = 46%. Lx waveform typical of modal voice.



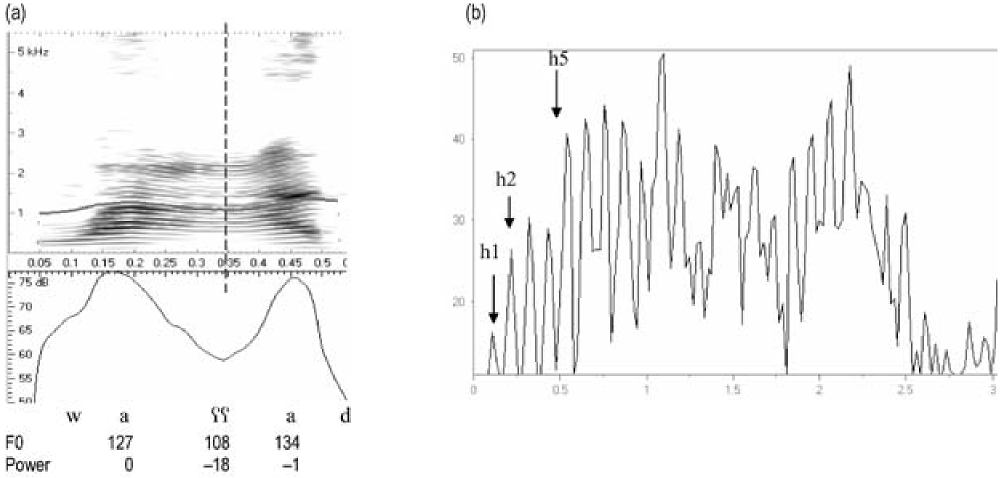
(b) Six glottal cycles in the centre of the geminate [ʕʕ].
Average Fx = 211 Hz. Average Qx = 36%. Lx waveform shows shorter excursions and longer open phase. Acoustic waveform shows very low amplitude pulses.



(c) Six glottal cycles in the unstressed vowel [a] just after the geminate [ʕʕ].
Average Fx = 220 Hz. Average Qx = 44%. Waveform typical of modal voice.

Figure 12 Panels showing acoustic (top) and Lx (bottom) waveforms comparing geminate [ʕʕ] with preceding and following vowels. Female speaker, Qatar. Same token as in figure 11.

characteristic of the tight approximant, albeit in a weaker form, means that it is much more likely to be whispery voice than breathy voice. Laver (1980: 134) describes the former as involving ‘a greater degree of laryngeal effort’ while the latter has ‘a low degree of laryngeal effort’. Esling & Harris (2005: 367) describe how for whispery voice there is a constricted

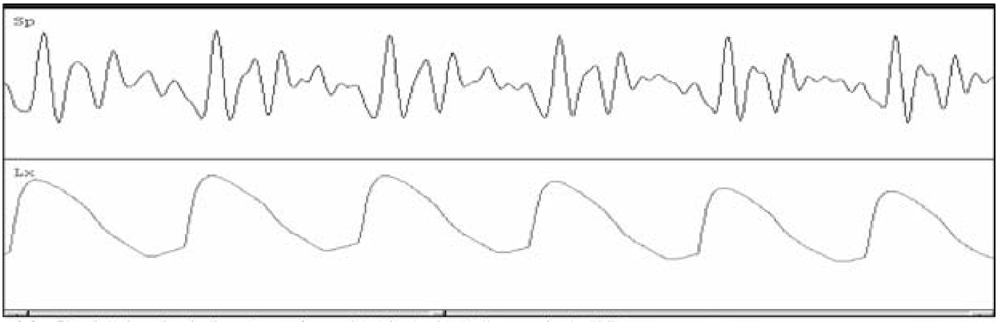


(a) A further 'looser' token showing the characteristic drop in F0 and amplitude.
(b) Spectrum of (a) at the dashed line showing positive spectral tilt h1–h2 of +10.2dB, and h1–h5 (first F1 harmonic) of +24.5dB.

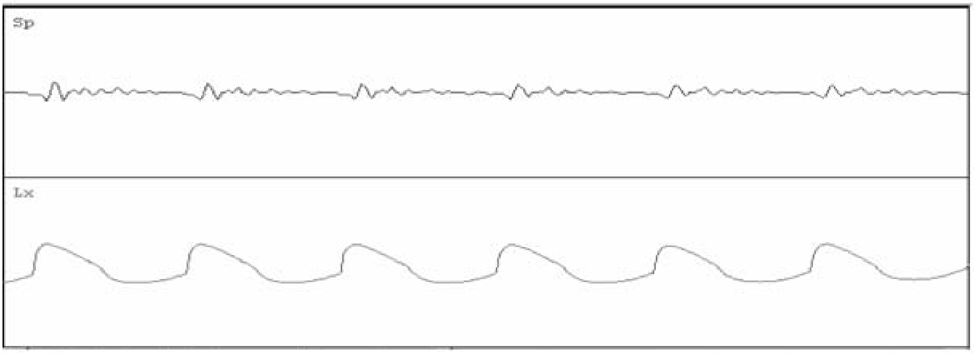
Figure 13 A tight approximant realisation of geminate *'ayn* in the word /waʕʕad/ by a male speaker from Cairo. Laryngograms of this token are shown in figure 14.

epilaryngeal vestibule. A greater degree of laryngeal tension is more consistent with the adjustments necessary to cause the bandpass filter effect, but this is not the only evidence in favour of whispery voice. Breathy voice tends to be accompanied by lowering of the larynx (Laver 1980: 31) whereas the informant who provided this token reports that her larynx rises, which the present author confirmed for subsequent productions by this speaker. As already discussed in section 2.4, larynx raising is an almost categorical feature of realisations of *'ayn*. Spectral evidence also supports a whispery voice interpretation. The first harmonic in breathy voice is expected to be stronger than higher harmonics due to a general lax musculature in the larynx and also to a maintained aperture between the arytenoid cartilages (Hanson, Stevens, Kuo, Chen & Slifka 2001: 453) which is thought to be more likely in female speakers than in male speakers (Hanson et al. 2001: 460). Whispery voice also has a maintained arytenoidal aperture (Laver 1980: 134) but not a lax musculature. We can probably attribute to the arytenoidal gap the fact that the lower harmonics are not as radically attenuated in figure 11 compared to the examples in figures 4, 5 and 6, and the fact that the first harmonic is stronger than the second and third (see figure 11(b)), thus giving a negative spectral tilt (Hanson et al. 2001: 459). Combined with the higher pitch of this female's voice, and the greater inter-harmonic intervals resulting from it, the whispery voice phonation sounds mellower and less strained than the male productions of the tight approximant, and may be a gender marker. A tendency for female Arabic speakers to use less laryngeal tension in realisations of the emphatic /t^{ʕ}/ is reported, for example, by Khattab, Al-Tamimi & Heselwood (2006: 155f.) in Jordanian Arabic on spectrographic evidence. Further evidence from Jordanian Arabic is provided in nasoendoscopic images in Heselwood & Al-Tamimi (2006) which show less epiglottal retraction in a female production of emphatic /s^{ʕ}/ compared with a male production, suggesting a weaker engagement of the laryngeal sphincter mechanism. For productions of /ʕ/, the male speaker's epiglottis flattens out more against the rear pharyngeal wall resulting in a slit-like aperture, also suggestive of greater tension in the aryepiglottic sphincter than is the case with the female speaker.

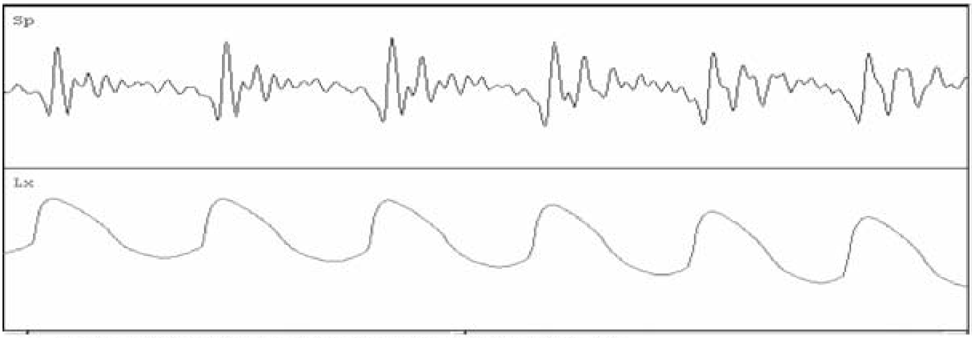
A token of a geminate *'ayn* with some similar properties to that in figures 11 and 12 but produced by a male speaker from Cairo, Egypt, is shown in figures 13 and 14.



(a) Six glottal cycles in the stressed vowel [a] just prior to the geminate [ʕʕ].
Average Fx = 127 Hz. Average Qx = 52%. Lx waveform typical of modal voice.



(b) Six glottal cycles in the centre of the geminate [ʕʕ].
Average Fx = 108 Hz. Average Qx = 43%. Lx waveform shows shorter excursions and longer open phase.
Acoustic waveform shows very low amplitude pulses.



(c) Six glottal cycles in the unstressed vowel [a] just after the geminate [ʕʕ].
Average Fx = 134 Hz. Average Qx = 51%. Waveform typical of modal voice.

Figure 14 Panels showing acoustic (top) and Lx (bottom) waveforms comparing geminate [ʕʕ] with preceding and following vowels. Male speaker, Cairo. Same token as in figure 13.

Although his H1–H2 spectral tilt is positive, the Qx value decreases for the *‘ayn* compared to the adjacent vowels as it does for the female speaker but not the other males. There is a growing body of sociolinguistic evidence to suggest that some groups of Cairo males exhibit some of the phonetic characteristics traditionally associated with female speech (Royal 1985,

Table 1 Harmonics in bold within 1.25 ERBs and predicted not to be separately processed by the auditory filters in the tokens represented in figures 5, 4c and 9, respectively. (Harmonics higher than no. 8 are also not separately processed.)

F0 and speaker	Harmonic number	Frequency	ERB number	ERB separation value
90 Hz	4	360	8.78	
	5	450	10.12	1.34
Male, Amman, Jordan	**6**	**540**	**11.26**	**1.14**
	7	**630**	**12.29**	**1.03**
	8	**720**	**13.22**	**0.93**
110 Hz	4	440	9.97	
Male, Irbid, Jordan	5	550	11.38	1.41
	6	**660**	**12.61**	**1.23**
	7	**770**	**13.70**	**1.09**
	8	**880**	**14.67**	**0.97**
211 Hz	4	844	14.36	
Female, Qatar	5	1055	16.03	1.67
	6	1266	17.44	1.41
	7	**1477**	**18.67**	**1.23**
	8	**1688**	**19.75**	**1.08**

Haeri 1996) and this may be another example of it; although it may, of course, simply be idiosyncratic.

3.4 Auditory processing

The bandpass filter effect shaping the output spectrum of the tight approximant presents the ear of the listener with a signal comprising just a few prominent harmonics in the mid-frequency region. On the basis of reported experimental results, only the lowest five, and possibly the lowest eight, harmonics of a complex tone are held to be separately resolved by the basilar membrane (Moore 1997: 100). Moore & Ohgushi (1993), cited in Moore (1997: 102), found that only those harmonics which are more than 1.25 ERBs apart are resolved separately. ERB width increases as the frequency of a given harmonic increases so that more harmonics are separately resolved the higher the fundamental frequency of the complex tone. The female production in figures 11 and 12 may owe its mellower quality in part to this phenomenon as well as to its whispery voice phonation. It can be seen in table 1 that more of the lower harmonics are separately resolved in the female production than in the two male productions given for comparison. The table gives the ERB numbers for the harmonics using the formula provided by Hayward (2000: 142):

$$\text{Number of ERBs} = 21.4 \log_{10}((0.00437^{*} f) + 1)$$

Where harmonic separation is greater than 1.25 ERBs, each harmonic is processed by a separate auditory filter. Higher harmonics where the separation is less than 1.25 ERBs, however, inflict interference on their neighbours because two or more harmonics are processed by the same auditory filter (Moore 1997: 204). Generally, the third, fourth and fifth harmonics determine pitch perception, though there is some inter-listener variation (Moore 1997: 198f.). However, according to the temporal theory of pitch perception, if these lower components are missing or substantially weakened, as they are in the tight approximant variant, then the pitch sensation results from the interference pattern created by the harmonics vying for the attention of the same filter. This pattern is a kind of 'beating' (Howard & Angus 2001: 130) at a frequency equal to the difference between the harmonics in question, which is equal to the

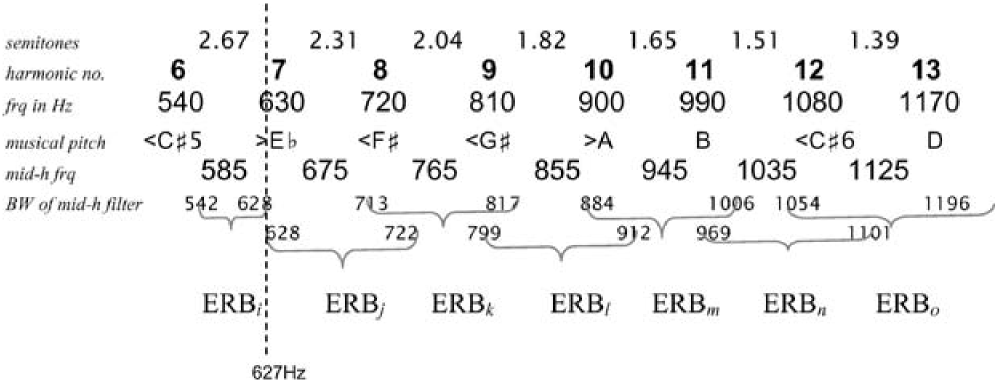


Figure 15 Auditory critical band filtering of a tight approximant *‘ayn* with F0 = 90 Hz. Filters to the right of the dashed line capture two adjacent harmonics, creating interference difference-frequencies equal to F0.

F0. The ‘beating’ is too fast to be heard as beating if it is above about 12–15 Hz (Howard & Angus 2001: 75), but it does affect the timbre of the sound (Plomp 1976: 97, Moore 1997: 188). (The situation is analogous to voicing where the individual glottal pulses follow in succession too quickly to be separately ‘heard out’.) Pitch perceived in this way is known as ‘residue’ pitch (Moore 1997: 188) and has ‘an impure, sharp tone quality’ (Moore 1997: 194). Howard & Angus (2001: 75) describe the auditory effect as ‘roughness’, the term also used by Plomp (1976: 97), which may explain Gairdner’s ‘growl’ mentioned above. If we take an example of a tight approximant *‘ayn* having an Lx value of 90 Hz, and harmonics 6–13 band-passed, we can calculate the centre-frequency (*cf*) of that auditory filter whose bandwidth is 90 Hz by:

$$\frac{-(93.39{*}10^{-3}) \pm \sqrt{(93.39{*}10^{-3})^2 - ((24.92{*}10-6)\times(28.5-90))}}{12.46{*}10^{-6}}$$

(Howard & Angus 2001: 126). This gives us $cf = 627$ Hz. Any mid-harmonic filter with $cf > 627$ Hz will therefore have a bandwidth greater than 90 Hz and will capture two adjacent harmonics which will interfere in the filter and produce an output difference-frequency, or ‘beat’, of 90 Hz. The bandwidths of the mid-harmonic filters are wider the higher the *cf* value and are obtained by the ERB equation (cf. Howard & Angus 2001: 76):

$$\mathrm{ERB} = ((6.23{*}10^{-6}{*}cf^2) + (93.39{*}10^{-3}{*}cf) + 28.52)\,\mathrm{Hz}$$

Figure 15 shows how this sound is thought to be processed by the auditory critical band filters.

More recent research suggests that interference patterns are not necessary for the perception of residue pitch (Moore 1997: 199–202). Instead, it is proposed that the intervals of time between firings of neurones are compared and the reciprocal of the most frequently occurring time interval is the perceived pitch. Table 2 shows the firing intervals in response to harmonics 6–13 of a complex tone with a fundamental of 90 Hz. Neurones responding to each harmonic come closest to firing simultaneously at a time interval almost, but not quite, equal to the time period of the fundamental, 11.11 ms. For each harmonic this happens when the firing number is the same as the harmonic number or an exact multiple of it, i.e. they coincide most closely on the seventh, fourteenth, twenty-first, etc. firings in response to the seventh harmonic, and the eighth, sixteenth, twenty fourth, etc. firings in response to the eighth harmonic, and so on. Because there is variation around the F0 time period, the

Table 2 Neural responses to eight non-resolvable harmonics with F0 = 90 Hz. The responses coincide most closely at time intervals almost equal to the time period of F0 (in bold). The F0 time period is 11.11 ms, the same as harmonic no. 11.

	Harmonic number, frequency, and firing times after onset							
Firing number	6–540 Hz	7–630	8–720	9–810	10–900	11–990	12–1080	13–1170
1	1.85 ms	1.59	1.39	1.23	1.11	1.01	0.93	0.85
2	3.70	3.18	2.78	2.46	2.22	2.02	1.86	1.70
3	5.55	4.77	4.17	3.69	3.33	3.03	2.79	2.55
4	7.40	6.36	5.56	4.92	4.44	4.04	3.72	3.40
5	9.25	7.95	6.95	6.15	5.55	5.05	4.65	4.25
6	**11.10**	9.54	8.34	7.38	6.66	6.06	5.58	5.10
7	12.95	**11.13**	9.73	8.61	7.77	7.07	6.51	5.95
8	14.80	12.72	**11.12**	9.84	8.88	8.08	7.44	6.80
9		14.31	12.51	**11.07**	9.99	9.09	8.37	7.65
10			13.90	12.30	**11.10**	10.10	9.30	8.50
11				13.53	12.21	**11.11**	10.23	9.35
12					13.32	12.12	**11.16**	10.20
13						13.13	12.09	**11.05**
14							13.02	11.90
15								12.75
f of closest common interval	90.09 Hz	89.85	89.93	90.33	90.09	90.00	89.61	90.50
average *f*	90.05 Hz							

perception of pitch is less clear than if the pitch were perceived from the processing of the lower separately resolvable harmonics (Moore 1997: 206), as would be the case with the adjacent vowels. The switching into and out of this different way of obtaining pitch information, and consequently in and out of a ‘blurring’ of the pitch percept with its associated difference of timbre, may define the auditory quality of the tight approximant allophone of *‘ayn*, or at least contribute substantially to it. An appreciation of the responses of the auditory system to the tight approximant variant helps us understand the contributions made by the several component adjustments brought about by contraction of the laryngeal sphincter mechanism to the auditory quality of the sound through the attenuation of the harmonics below and above numbers 6–13 or thereabouts. Without considering the end product of articulatory and acoustic activity, i.e. the ‘sound-as-perceived’, we might not attain the same level of insight into why the sound is made the way it is.

3.5 Perception of *‘ayn* as a stop

We must return, however, to tokens which exhibit the acoustic characteristics of tight approximants but which may be perceived as stops. It was noted above how these present a problem of classification. As instrumental phonetics becomes more sophisticated, phonetic theory must find ways to deal with discrepancies between properties of sounds in different phonetic domains. Sounds which strike the listener as stops but which instrumental records show have uninterrupted acoustic output and no complete articulatory occlusion are qualitatively different from, on the one hand, sounds perceived as stops which *do* exhibit interrupted acoustic output, and, on the other hand, sounds which are *not* perceived as stops. Decisions have to be made how the sounds in question should be symbolised. Should the symbol reflect the listener’s experience or the instrumental records? On the grounds that instrumental records speak for themselves, IPA symbols can most usefully be used to express what such records cannot express, i.e. what the sound sounds like, intra- and inter-subjective

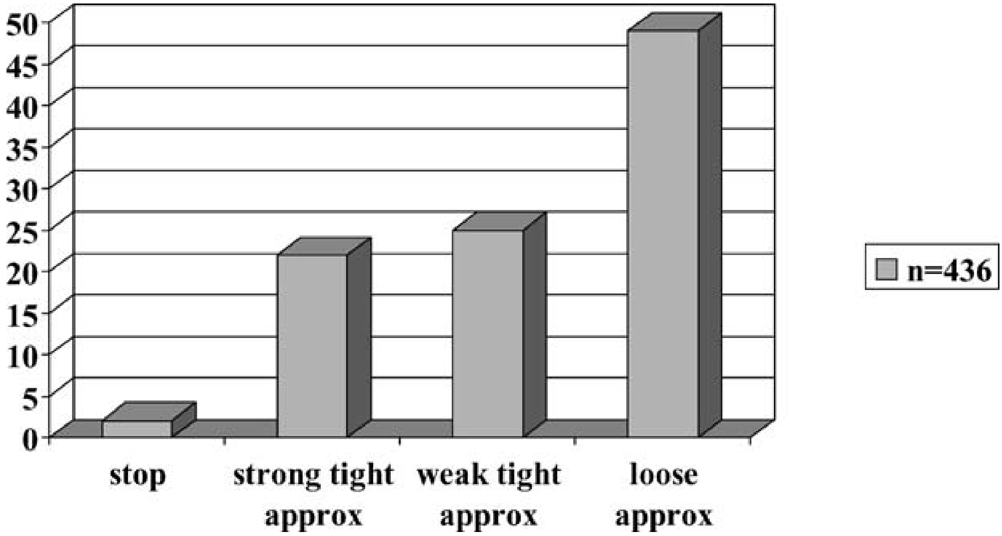


Figure 16 Percentages of realisation types from a sample of 436 tokens of *‘ayn*.

uncertainties notwithstanding. That is to say, if it sounds like a stop, then a stop symbol should be used, and the sound should be classified auditorily as a type of stop which shows continuous acoustic activity rather than as a type of continuant that sounds like a stop, although the traditional link between IPA symbols and articulatory properties may mean the latter is intuitively more satisfactory for phoneticians.

Only a subset, then, of tokens classifiable instrumentally as tight approximants have the full auditory effect of a ‘blurred’ pitch sensation and altered timbre described in the previous section. However, those tokens which are perceived as stops, and possibly even those actually produced as stops, may induce this auditory effect at their edges such that what is perceived is a stop bounded by a tight approximant onset and/or offset, e.g. [ʕ̝ʡʕ̝]. The example in figure 3 seems to display this kind of transitional edge pattern. These edge phenomena might be what acoustically and auditorily distinguish some realisations of *‘ayn* from realisations of *hamza*, the glottal stop.

4 Prevalence and distribution of the tight approximant variant

Concerning the prevalence of this variant, the 436 recorded *‘ayn* tokens were sorted into four categories on acoustic criteria: stopped realisations, realisations in which the tight approximant resonance characteristics are strongly present, those where they are less strongly present, and those where they are not present at all, i.e. ‘normal’ approximants rather than ‘tight’ ones (labelled ‘loose’ on figures 16 and 17) and thus lacking the bandpass filtering effect of the tight approximants. The distinction between ‘strong’ and ‘weak’ tight approximants is an auditory-impressionistic one which cannot claim to be objective in the absence of correlations between perceptual judgments and acoustic or articulatory parameter values, but it is useful to recognise that there are different points on the continuum from very tight approximant through to very loose approximant forming part of a larger continuum of pharyngealisation (Esling 1999: 369); this could be pursued in the future through perception tasks using controlled synthetic stimuli in which the role of spectral tilt steepness as an indicator of strength of ‘tightness’ might usefully be explored. Figure 16 shows the percentages of each type in the

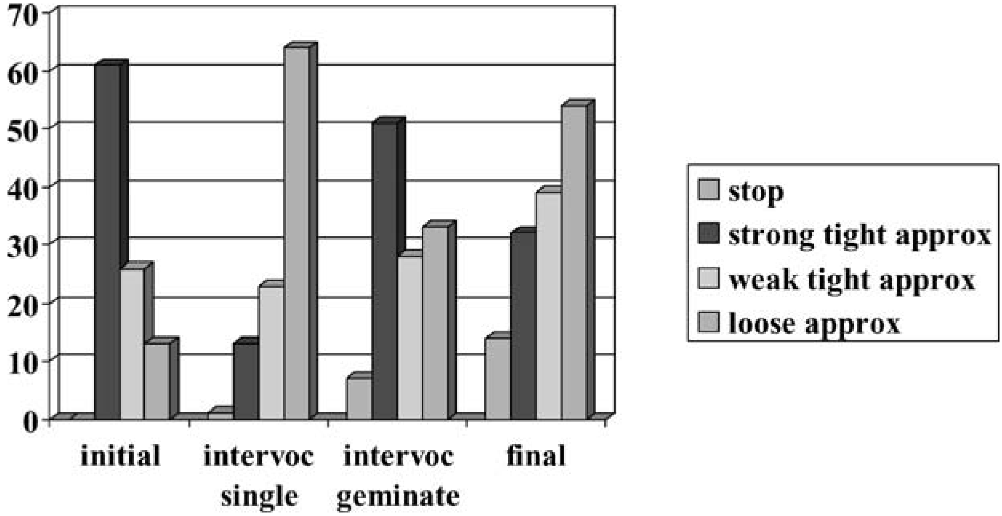


Figure 17 Percentages of phonotactic distribution of realisation types in the sample of 436 tokens of *'ayn*.

sample. Taking the strong and weak tokens together, tight approximants account for nearly half (47%) of all tokens, persuasively suggesting that this variant is in widespread usage and should be recognised as a common allophone of the Arabic *'ayn*. All the speakers in the sample produced the tight approximant allophone in either its strong or weak form, or both, indicating that it occurs in all the regions represented in the sample, at least in the styles of speech employed.

The categories were then correlated with four phonological contexts: initial prevocalic singleton, intervocalic singleton, intervocalic geminate, and final post-vocalic singleton. The results are presented in figure 17. Over half of initial singleton (61%) and intervocalic geminate (51%) occurrences of *'ayn* were realised as tight approximants, and nearly a third (32%) of word-form final occurrences. Only thirteen percent of intervocalic singletons were realised this way, the most common realisation being a 'looser' open-approximation approximant (64%). The results for intervocalic singletons are hardly surprising given that such contexts are known to induce consonantal lenition (Lavoie 2001: 161) and geminates are known to resist lenition more than singletons (Lass 1984: 181f.). The tight approximant variant is most common in initial prevocalic position, a context generally favouring consonantal strength (Vennemann 1988: 13f., Lavoie 2001: 161). Many of these tokens, particularly those in utterance-initial position (see figure 18), begin with a stop release but have not here been classified as stops because the focus of interest is on the occurrence of tight approximation and its associated spectrum. The relative frequencies of occurrence of the four types in post-vocalic word-form final position are hard to comment on. This was the most likely context for stop allophones particularly in utterance-final position. In utterance-medial contexts weak tight approximants and loose approximants predominate (see figure 19). Whether final contexts generally favour consonantal strength or not is hard to evaluate. Vennemann (1988: 24–27) formulates a somewhat contradictory 'coda law' which allows for lenition processes turning obstruents into approximants (e.g. Spanish *cautivo* from Latin *captivus* 'captive'), and for processes traditionally seen as fortition (e.g. final obstruent devoicing). The Arabic *'ayn* would provide a good test case for predictions about which contexts favour which manners of articulation because, as reviewed in section 2.2, all manners of articulation can occur. The examples in figures 18 and 19 indicate the importance of taking context beyond syllable boundaries and word-form boundaries into account as well as illustrating the potential difficulties

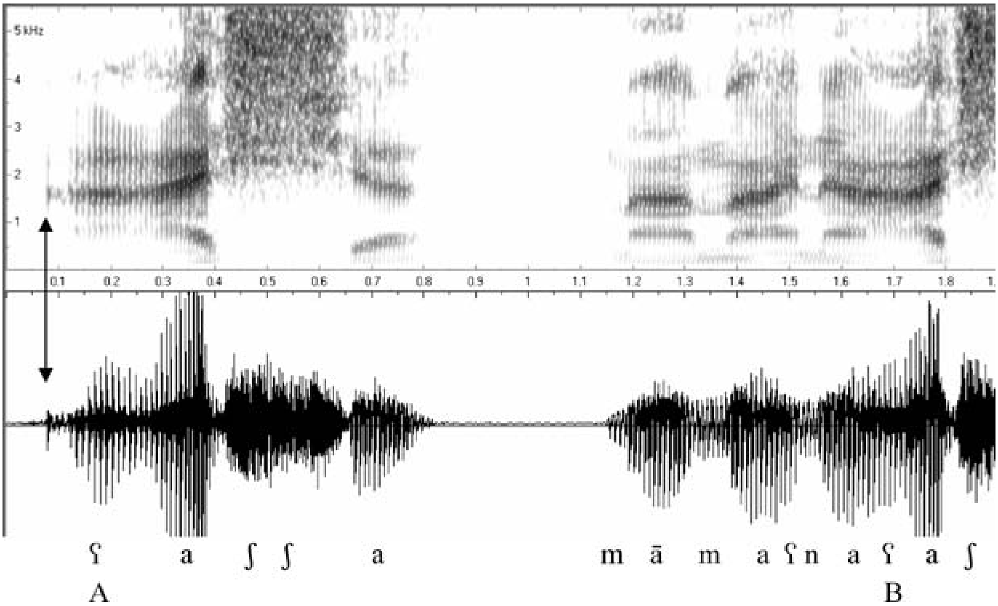


Figure 18 Utterance initial (A) and utterance-medial (B) realisations of word-form initial *‘ayn* in /ʕaʃʃa/ by a male speaker from Kuwait. The utterance-initial realisation has a stop release (indicated by the arrow) while the utterance-medial realisation does not.

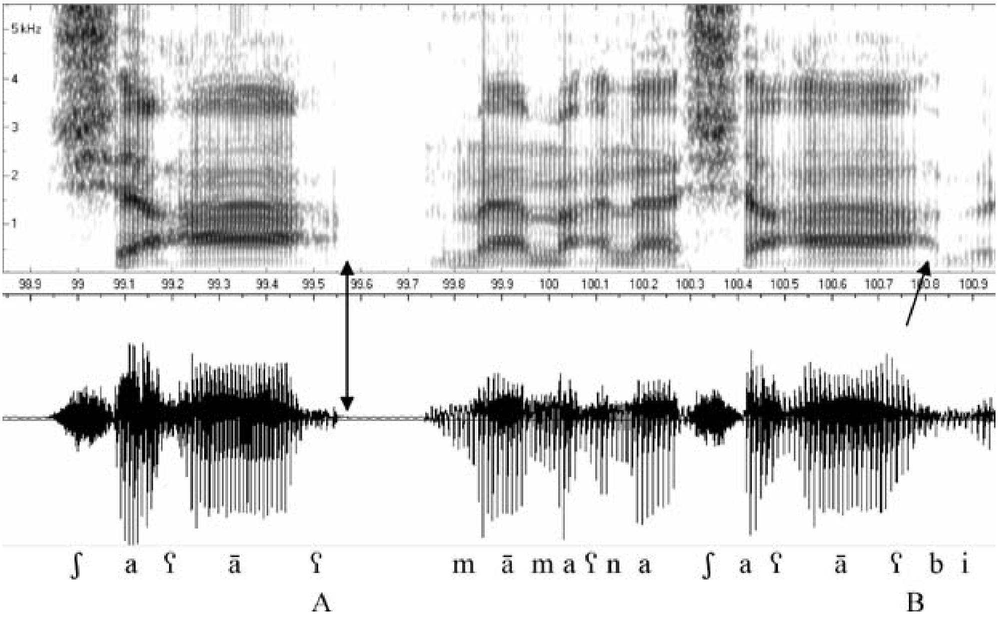


Figure 19 Utterance-final (A) and utterance-medial (B) realisations of word-form final *‘ayn* in /ʃaʕʕ/ by a male speaker from Gaza. The utterance-final realisation terminates in flat formant transitions and a stop closure (at the arrow) while the utterance-medial realisation is an approximant with negative formant transitions (arrowed) leading into the following /b/.

inherent in classifying individual tokens of *‘ayn* as either one manner of articulation or another.

5 Conclusion

Tight approximant realisations of Arabic *‘ayn* are quite common cross-dialectally in certain phonological contexts, especially geminates and initial singletons. They vary on a continuum from strongly to weakly articulated, which can be seen in the acoustic spectrum in terms of strength of bandpass filtering, and on Lx waveforms in terms of closed quotient values and waveshape. The variation may have sociolinguistic significance, particularly in relation to gender, but this is speculative and needs further investigation.

Lx waveform variation indicates that the tight approximant can have a source spectrum produced by creak, harsh voice, pressed voice or whispery voice, depending on degree of general tension and compression in the pharynx and larynx. These different phonation types are associated with different parts of the continuum from strong to weak tight approximants. Some realisations at the strong end of the continuum can be stops from an auditory-perceptual point of view.

Evidence suggests that the tight approximant variant may have developed diachronically from full occlusion insofar as the degree of compression of the laryngopharyngeal spaces can be explained as a reductive articulatory change operating on the kind of articulation seen in the X-ray picture in figure 2. Full occlusion still occurs in the more conservative dialects and in more formal styles of speech. It would be problematic and a mistake, however, to think of full occlusion as the ‘target’ manner of articulation of *‘ayn*. Individual speakers may well have different ‘targets’ for different phonological contexts and different styles of speech, and the targets might be quite flexible incorporating more than one manner of articulation. Geminate tokens provide the best opportunity to study the characteristics of the tight approximant. By looking in considerable detail at articulatory, acoustic and auditory properties of this allophonic type in geminates it has been possible to see how events in these three phonetic domains are related in the chain of cause-and-effect linking speaker and listener, and how they combine to create a phonetically intriguing sound which may occur not only in varieties of Arabic, but also in some other unrelated languages that make use of pharyngeal sounds, e.g. Nuuchahnulth and !Xóõ. The study is also an example of why the auditory processing of sounds should not be neglected if we are to better understand why sounds are the way they are and appreciate how they work in spoken language as perceived objects, or parts of perceived objects.

Taking Laufer’s suggestion that the IPA [ʕ] symbol should be used to denote an approximant, the class of tight approximants can be generally symbolised using the IPA ‘raised’ diacritic as [ʕ̝], with further diacritical modifications to indicate phonation type. Strength of articulation can be denoted by using the ExtIPA diacritic for strong articulation [͈] (Duckworth, Allen, Hardcastle & Ball 1990).

Acknowledgements

The research reported in this paper was completed and written up during a sabbatical semester granted by the School of Modern Languages and Cultures, University of Leeds, in 2004–05. I am grateful to all the informants who gave some of their time to have their speech recorded, also to Feda Al-Tamimi for organising some of the recording sessions. Thanks are also due to the three anonymous reviewers who made several helpful observations and suggestions for improvement, and to John Esling whose insights into pharyngeal articulation led me to a substantial re-conceptualisation of the analysis. Parts of the study were presented at a Department of Linguistics and Phonetics research seminar at the University of Leeds, 18 November 2003, and at the British Association of Academic Phoneticians’ Colloquium in March 2004 at the University of Cambridge.

Appendix. Words containing *‘ayn* used for data collection

Final vowels were omitted by some speakers as is usual in colloquial varieties.

Initial	Intervocalic singleton	Intervocalic geminate	Final
ˈ**ʕaʃʃaʃa** ‘to settle in’ (pronounced /ˈ**ʕaʃʃa**/ by some informants)	**ʃaˈʕaaʕ** ‘perplexed’	ˈ**ʃaʕʕa** ‘to disperse’	**ʃaˈʕaaʕ** ‘perplexed’
ˈ**ʕalam** ‘a flag’	ˈ**raʕaba** ‘to be afraid’	ˈ**waʕʕad** ‘to make s.o. promise’	ˈ**daaʕ** ‘need’
	ˈ**raʕa** ‘to take care of’	ˈ**daʕʕa** ‘to turn aside, to rebuff’	
	ˈ**naʕi** ‘condolences’		

References

Adamson, B. (1981). A spectrographic analysis of pharyngeal consonants in Sudanese Arabic. *Work in Progress* **3**, 81–96. University of Reading Phonetics Laboratory.

Agius, D. (1992). Morphological alternatives in the Gozitan dialects of Maltese. In Vanhove, M. (ed.), *Materiaux arabes et sudarabiques*, 111–161. Paris: Association Gellas, Université de Paris III.

Al-Ani, S. (1970). *Arabic Phonology*. The Hague: Mouton.

Al-Ani, S. (1978). An acoustical and physiological investigation of the Arabic /ع/. In Al-Ani, S. (ed.), *Readings in Arabic Linguistics*, 89–101. Bloomington, IN: Indiana University Linguistics Club.

Al-Nassir, A. A. (1993). *Sibawayh the Phonologist*. London: Kegan Paul International.

Al-Shareef, J. (2002). *Language Change and Variation in Palestine: A Case Study of Jabalia Refugee Camp*. Ph.D. thesis, University of Leeds.

Al-Tamimi, F. Y. A. (2001). *Phonetic and Phonological Variation in the Speech of Rural Migrants in a Jordanian City*. Ph.D. thesis, University of Leeds.

Berg, J. van den (1955). On the role of the laryngeal ventricle in voice production. *Folia Phoniatrica* **7**, 57–69.

Browman, C. P. & Goldstein, L. M. (1991). Gestural structures: distinctiveness, phonological processes and historical change. In Mattingley, I. & Studdert-Kennedy, M. (eds.), *Modularity and the Motor Theory of Speech Perception*, 313–338. Hillsdale, NJ: Erlbaum.

Bukshaisha, F. A. M. (1985). *An Experimental Phonetic Study of Some Aspects of Qatari Arabic*. Ph.D. thesis, University of Edinburgh.

Butcher, A. & Ahmad, K. (1987). Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic. *Phonetica* **44**, 156–172.

Cantineau, J. (1960). *Études de linguistique arabe*. Paris: Klincksieck.

Carlson, B. F. & Esling, J. H. (2003). Phonetics and physiology of the historical shift of uvulars to pharyngeals in Nuuchahnulth (Nootka). *Journal of the International Phonetic Association* **33**, 183–193.

Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Chiba, T. & Kajiyama, M. (1941). *The Vowel: Its Nature and Structure*. Tokyo: Kaiseikan.

Daher, J. (1998). Gender in linguistic variation: the variable (q) in Damascus Arabic. In Benmamoun, E., Eid, M. & Haeri, N. (eds.), *Perspectives on Arabic Linguistics* **XI**, 183–206. Amsterdam: John Benjamins.

Duckworth, M., Allen, G., Hardcastle, W. J. & Ball, M. J. (1990). Extensions to the International Phonetic Alphabet for the transcription of atypical speech. *Clinical Linguistics & Phonetics* **4**, 273–280.

Elgendy, A. M. S. (2001). *Aspects of Pharyngeal Coarticulation*. Ph.D. thesis, University of Amsterdam.

El-Halees, Y. (1985). The role of F1 in the place-of-articulation distinction in Arabic. *Journal of Phonetics* **13**, 287–298.

El-Saaran, M. H. A. (1951). *A Critical Survey of the Phonetic Observations of the Arab Grammarians*. Ph.D. thesis, University of London.

Esling, J. H. (1996). Pharyngeal consonants and the aryepiglottic sphincter. *Journal of the International Phonetic Association* **26**, 65–88.

Esling, J. H. (1999). The IPA categories 'pharyngeal' and 'epiglottal': laryngoscopic observations of pharyngeal articulations and larynx height. *Language & Speech* **42**, 349–372.

Esling, J. H., Fraser, K. E. & Harris, J. G. (2005). Glottal stop, glottalised resonants, and pharyngeals: a reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics* **33**, 383–410.

Esling, J. H. & Harris, J. G. (2005). States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In Hardcastle, W. J. & Mackenzie Beck, J. (eds.), *A Figure of Speech: A Festschrift for John Laver*, 347–383. Mahwah, NJ: Erlbaum.

Gaber, A. M. (1972). *The Phonology of Verbal Piece in Cairo Egyptian Arabic*. Ph.D. thesis, University of London.

Gairdner, W. H. T. (1925). *The Phonetics of Arabic*. London: Oxford University Press.

Garbell, I. (1958). Remarks on the historical phonology of an East Mediterranean Arabic dialect. *Word* **14**, 303–337.

Gasim, A. (1965). Some aspects of Sudanese colloquial Arabic. *Sudan Notes and Records* **46**, 40–49.

Gerratt, B. R. & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics* **29**, 365–381.

Ghali, M. M. (1983). Pharyngeal articulation. *Bulletin of the School of Oriental and African Studies* **XLVI**, 432–444.

Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic*. Unpublished PhD thesis, University of Texas at Austin.

Gordon, M. & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics* **29**, 383–406.

Greenman, J. (1979). A sketch of the Arabic dialect of the Central Yamani Tihāmah. *Zeitschrift für Arabische Linguistik* **3**, 47–61.

Haeri, N. (1996). *The Sociolinguistic Market of Cairo: Gender, Class and Education*. London: Kegan Paul.

Hanson, H. M., Stevens, K. N., Kuo, H-K. J., Chen, M. Y. & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics* **29**, 451–480.

Harrell, R. S. (1957). *The Phonology of Colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Hayward, K. (2000). *Experimental Phonetics*. London: Longman.

Heselwood, B. (1992). *Extended Axiomatic-Functionalist Phonology: An Exposition with Application to Modern Standard Arabic*. Ph.D. thesis, University of Ulster.

Heselwood, B. (1996). Glottal states and emphasis in Baghdadi and Cairene Arabic: synchronic and diachronic aspects. In Dickins, J. (ed.), *Three Topics in Arabic Phonology*, 20–44. Occasional Paper 53, Centre for Middle Eastern and Islamic Studies, University of Durham. Available online at <http://www.personal.leeds.ac.uk/~lnpbch/Glottal%20states%20and%20emphasis.pdf>.

Heselwood, B. & Al-Tamimi, F. Y. A. (2006). A nasoendoscopic study of epiglottal activity in the laryngeal, pharyngeal and pharyngealised consonants of Jordanian Arabic. British Association of Academic Phoneticians' Colloquium, Queen Margaret University College, Edinburgh, April 10–12, 2006.

Hetzron, R. (1969). Two notes on Semitic laryngeals in East Gurage. *Phonetica* **19**, 69–81.

Hillenbrand, J. M. & Houde, R. A. (1996). Role of F0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech & Hearing Research* **39**, 1182–1190.

Howard, D. M. & Angus, J. (2001). *Acoustics and psychoacoustics* (2nd edn.). Oxford: Focal Press.

Ibrahim, M. H. (1986). Standard and prestige language: a problem in Arabic sociolinguistics. *Anthropoligical Linguistics* **28**, 115–126.

Jacobsen, W. (1969). Origin of the Nootka pharyngeals. *International Journal of American Linguistics* **35**, 125–153.

JAKOBSON, R. (1957). Mufaxxama – The 'emphatic' phonemes in Arabic. In Pulgram E. (ed.), *Studies Presented to Joshua Whatmough*, 105–115. The Hague: Mouton.

KAPLAN, H. M. (1960). *Anatomy and Physiology of Speech*, New York: McGraw-Hill.

KHATTAB, G., AL-TAMIMI, F. Y. A. & HESELWOOD, B. (2006). Acoustic and auditory differences in the /t/–/ṭ/ opposition in male and female speakers of Jordanian Arabic. In Boudela, S. (ed.), *Perspectives on Arabic Linguistics* **XVI**, 131–160. Amsterdam: John Benjamins.

LADEFOGED, P. (2001). *A Course in Phonetics* (4th edn.). Fort Worth, TX: Harcourt College Publishers.

LADEFOGED, P. (2003). *Phonetic data analysis*. Oxford: Blackwell.

LADEFOGED, P. & MADDIESON, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

LARADI, W. J. (1983). *Pharyngealisation in Libyan (Tripoli) Arabic*. Ph.D. thesis, University of Edinburgh.

LASS, R. (1984). *Phonology*. Cambridge: Cambridge University Press.

LAUFER, A. (1996). The common [ʕ] is an approximant and not a fricative. *Journal of the International Phonetic Association* **26**, 113–117.

LAUFER, A. & BAER, T. (1988). The emphatic and pharyngeal sounds in Hebrew and Arabic. *Language & Speech* **24**, 39–61.

LAUFER, A. & CONDAX, I. D. (1979). The epiglottis as an articulator. *Journal of the International Phonetic Association* **9**, 50–56.

LAVER, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

LAVOIE, L.M. (2001). *Consonant Strength*. New York: Garland Publishing.

MACCURTAIN, F. (1982). *Pharyngeal Factors Influencing Voice Quality*. Ph.D. thesis, University of London.

MCCARTHY, J. J. (1991). Guttural phonology. In Comrie, B. & Eid, M. (eds.), *Perspectives on Arabic Linguistics* **III**, 63–91. Amsterdam: John Benjamins.

MCCARTHY, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In Keating, P. A. (ed.), *Phonological Structure and Phonetic Form* (Papers In Laboratory Phonology **III**), 191–233. Cambridge: Cambridge University Press.

MAY, J. G. (1981). Acoustic factors that may contribute to categorical perception. *Language* **24**, 273–285.

MOORE, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (4th edn.). San Diego, CA: Academic Press.

MOORE, B. C. J. & OHGUSHI, K. (1993). Audibility of partials in inharmonic complex tones. *Journal of the Acoustical Society of America* **93**, 452–461.

MOWREY, R. & PAGLIUCA, W. (1995). The reductive character of articulatory evolution. *Rivista de Linguistica* **7**, 37–124.

ODISHO, E. (1973). *The Role of the Rear Section of the Vocal Tract in Baghdadi Arabic*. M.Phil. thesis, University of Leeds.

PADDOCK, H. J. (1970). *An Acoustic and Perceptual Study of Distinctive Features*. Ph.D. thesis, University of London.

PEPINSKY, A. (1942). The laryngeal ventricle considered as an acoustic filter. *Journal of the Acoustical Society of America* **14**, 32–35.

PERKELL, J. S. (1980). Phonetic features and the physiology of speech production. In Butterworth, B. (ed.), *Language Production: Speech and Talk*, 337–372. London: Academic Press.

PETTORINO, M. & GIANNINI, A. (1984). F1 locus and place of articulation. In Van den Broecke, M. P. R. & Cohen, A. (eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences*, 201–204. Dordrecht: Foris Publications.

PIKE, K. L. (1943). *Phonetics*. Ann Arbor, MI: University of Michigan Press.

PLOMP, R. (1976). *Aspects of tone sensation*. London: Academic Press.

RABIN, C. (1951). *Ancient West Arabian*. London: Taylor's Foreign Press.

ROYAL, A-M. (1985). *Male/Female Pharyngealisation Patterns in Cairo Arabic: A Sociolinguistic Study of Two Neighborhoods*. Ph.D. dissertation, University of Texas at Austin.

SHAHEEN, K. (1979). *The Acoustic Analysis of Arabic Speech*. Ph.D. thesis, University of Wales at Bangor.

SHANK, S. & WILSON, I. (2000). Acoustic evidence for ʕ as a glottlised pharyngeal glide in Nuu-chah-nuulth. In Gessner, S. & Oh, S. (eds.), *Proceedings of the 35th International Conference on Salish and Neighbouring Languages*, 185–197. Vancouver: Department of Linguistics, University of British Columbia.

STEVENS, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

SWADESH, M. (1939). Nootka internal syntax. *International Journal of American Linguistics* **9**, 77–102.

TRIGO, L. (1991). On pharynx-larynx interactions. *Phonology* **8**, 113–136.

VENNEMANN, T. (1988). *Preference Laws for Syllable Structure*. Berlin: Mouton de Gruyter.

WATSON, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.