Between Veneration for the Text and Vernacularization of the Spoken Form in Quran Recitation

Zakaria Ahmad Abuhamdia
Middle East University,Amman, Jordan

Abstract
An argument is made in this paper attempting to show that the purportedly intended sanctification accorded to reciting the Quran by adherence to certain phonetic coarticulation operations goes against the noble intention itself. This principle of Tajweed rules requires that anticipatory (i.e., regressive) consonant assimilation (should) apply where its phonetic conditions obtain. The paper argues that this process results in making the oral product of such recitation come closer to vernacular speech than to standard speech. The argument of the paper extrapolates from the attested universal use of such coarticulation assimilation in vernacular or casual speech in languages of the world.
Keywords: Quran, recitation, formal pronunciation, veneration, casual speech, vernacular, phonetic processes of connected speech

1.   Introduction

The recitation of Holy Scriptures is universal; followers of known religions around the world give special attention to the written word as it is pronounced. Such universality has been observed and attested throughout documented history. (Graham, 1987: 65 – 66) To most Muslims, the recitation of the Quran, whether from a seen text or by heart, must be, and in fact is, done in Arabic irrespective of the native language of the reciters. The recitation cannot be done from any text of interpretation (or translation as some prefer to call it) in any other language. Unlike many other cases, Quran recitation involves yet an added dimension; readers are encouraged and urged to aim for the perfection of recitation in compliance with rules of recitation, Tajweed, which were standardized between the fifth and the ninth century A.H. Four of the works on recitation have become the classics on the subject. The writers of these classics are: Al-Dany (died 444 AH), Al-Qutruby (died 461 AH), Al-Shatiby (died 590 AH), and Ibn Al-Jazary (died 833 AH). The principles laid down by these authors have remained the same to a great extent without any major change, addition, or reduction, let alone challenge. (Al-Jamal, 2011) Most Muslim authorities consider ignoring the phonetic principles of Tajweed a sin, relying on a judgment by Al-Jazary: [*Adhering to Tajweed principles is a must, and those who do not do so commit a sin.*] Therefore, scores and scores of writers have annotated the works, simplified them, or produced an abridgement of one or more of them. The practice of learning Tajweed is a school subject in many countries. Radio and television stations air programs on Tajweed; some are interactive, like talk shows, such as that of Dr. Ayman Rushdi Sweid. The popular recitation styles are those of Haffs, Warsh, and Qaloon, although the seven (or to some four) styles are also common in some countries and institutions of religious studies. The learners and practitioners of recitation are de facto Muslim and include both native and non-native speakers of Arabic. Local, regional, and international competitions are periodically held in many countries, with large financial awards. Internet websites run in the hundreds, some with audio components.

In the oral rendition of a text, any text, pronunciation of words and larger units forms the channel of the act. Therefore, the production of sounds in isolation and the production of word canonical forms are addressed in the description and the teaching-learning process of recitation. This phonetic/phonological aspect is the subject of the next section of this paper.

2. The Production of Speech Sounds: the citation and contextual forms

When speech sounds are described in terms of their articulatory features distinguishing each one from the others, the individual sound in isolation is the immediate target of the effort. This specification of features forms the foundation for the study of the sound patterns in every language. However, sounds are only the building blocks of words and longer expressions which carry meaning. In their sequential occurrence, sounds influence each other in various ways and to different degrees. The influence could be in a forward or a backward direction (perseverative or anticipatory, also called progressive or regressive, respectively). (Ladefoged, 1993: 109) The other possibility of mutual influence (as in palatalization) may also occur. (Also called 'coalescent assimilation', Robertson and Stanton, 2005: 5)

Coarticulation assimilation has been explained from the perspective of saving articulatory effort by the speaker. In the economy of effort view, the speaker would want to deliver the message to the listener by expending the least effort of articulation; that is, imperfectly producing the canonical (i.e., ideal) form of the words or expressions. On the other hand, the listeners, on their part, would also want to expend the least effort in the identification of the sounds in the speaker's message, and therefore expect them to be in their canonical forms.

Both orientations aim at 'economy of effort'. The outcome is a balance between the two orientations. The direction of convergence (i.e., more towards careful speech or careful listening) depends on factors such as the formality of the context (Shockey, 2003: 267 - 270)), the relative status of each one of the interlocutors with respect to each other, and the topic of the message: (from most attention to least attention) minimal pairs, word lists, reading, formal/interview, and casual/vernacular speech (Labov, 1966: 240) Other researchers, such as Trudgill, suggested and used four styles: (from most attention to least attention) word list, reading passage, formal style, and casual speech. (Trudgill, 2000: 87) (Some specialist in religious cantillation e.g., Graham, 1987, suggested that while reading a holy scripture, a person is addressing the source of the scripture, a deity, and therefore the style should be formal rather than relaxed.(Page 100)

This assimilation process has also been explained by reference to the overlap in time and space of gestures that involve the same articulator or sometimes different articulators. (Browman and Goldstein, 1987: 7; Farnetani, 1999)

A third position explains the process on the basis of the difference in speed between commands from the speech-motor area in the brain and the response of speech muscles to the commands. (Guenther, 1995) To Agwuele, et al., (2008), it is the rate of sound production that can lead to sound feature spreading and sound overlap. Data from X-ray palatography, fMRI and other instrumental techniques show images of speech articulators and gestures in overlap and/or modified positions from canonical form positions. (Browman and Goldstein, 198; Omesh, et al., 2011) In addition to the above, , Zharkova adds that speech rate is a function of speech-motor control which can make coarticulation possible in adult speech, although the exact age at which maturity of motor control for assimilation options has not been conclusively established. (2011: 121)

3. The (Phonetic) Consequences of Coarticulation Assimilation

The phonetic output of coarticulation assimilation in 'fluent speech', 'rapid speech', or 'normal speech' (Robertson and Stanton, 2005), 'connected speech' (Farnetani, 1999; Ingram, 1989; Pavlik, 2009; McDonald, 2013), 'informal speech' or 'casual speech' (Labov, 1966; Zwickey, 1972 (also colloquial speech); Browman and Goldstein, 1987; Ernestus, 2000; Pearman, 2007; Pham, 2007 (also casual and colloquial); Torreira and Ernestus, 2011 (also spontaneous speech), 'vernacular speech' (Johnson, 2004 also 'conversational English), or 'spontaneous speech' (Binnenpoorta, et al., 2004; Dilley and Pitt, 2007; Brouwer, et al., 2012; Shockey, 2003) differs from the canonical forms of the words that undergo the phonetic process. Below is a sample from English (Robertson and Stanton, 2005)

Good + boy = Goobboy

White + paper = whyppaper

Hand + bag = hambag (/d/ deletion also)

Can + buy = cammbuy

Saint + Paul = samepaul

Green + park = greempark

Did + you = didgeyou (palatalization or coalescent assimilation)

4. Coarticulation Assimilation in Tajweed

Relevant to this process in Quran recitation are the following consonantal environments. They are all examples of partial / incomplete anticipatory assimilation of place and/or manner or of complete / total assimilation. As listed below, the first sound is the syllable or word coda and the second sound is the onset of the immediately following syllable or word.

(Alshaatiby, 24 and ff.; Ben Aljazary, 144, 153, 155, 157-158)

As can be seen from the International Phonetic Alphabet 2005 revised chart, reproduced below, the sounds which are described in the output as 'phonetically unmotivated' (the nasal labial approximant and the alveolar or alveopalatal coronal nasal) do not appear in the chart and, therefore, I have not used a symbol for each one but used a word qualification for each; however, the statement below the chart does not rule them out as being impossible articulations; rather it alerts the reader or user that only the shaded areas denote articulations that are judged impossible.

5. Phonetic environment Phonetic Output

b + m	m + m (delayed release, long) / irkamma3anaa/
m + b	(phonetically unmotivated bilabial approximant) m + b (does not apply to the Tajweed tradition in Iraq.
n + m	m + m (delayed release or long m /mimmamma3ak/
n + b	m + b (m here being an approximant, phonetically unmotivated, as indicated above)
n + w	w + w (total assimilation; long w; but does not apply to these instances: *Sinwaan, qinwaan, noon walqalam,* and *yaaseen walquran)* miwwaal
n + f	labio-dental n + f
n + th (of *this* in English)	approximant dental n + *th (this)*
n + *th* (of *thin* in English)	approximant dental n + *th* (*thin*)
n + TH (Arabic emphatic *th* of English *th* of *thus*)	approximant *n* + TH
n + t	unmotivated approximant n + t
n + d	unmotivated approximant n +d
n + (emphatic) T	unmotivated approximant n + T
n + D (of English *dawn*)	unmotivated approximant n + D Arabic emphatic
n + s	approximant n + s
n + S (of English *saw*)	approximant n+ (emphatic) S
n + z	approximant n + z
n + sh	approximant n + sh
n + dg	approximant n + dg
n + l	l + l (delayed release or long) (see above) ?allaw
n + r	r + r (delayed release or long or a long trill (does not apply to *man raaq*) mirrab
n + y	y + y (total assimilation) (does not apply to *dunya* and *bunyaan*)

	mayya3mal
n + k	velar n + k
n + q	uvular N + q
l + r	r + r (continuat r) (does not apply to *bal raana)* waqurrabbi

In the description of the sound system of Arabic of his time in the second century AH, Sibawayh (140 AH – 180 AH) in his 4-volume work *al-kitab*   , the data apparently came from the (tribal) dialects at the time. Dialects are represented in the illustrative examples given. All the examples in the Tajweed literature reflect phonetic processes that characterized one or the other of the dialects of the time. His work preceded the earliest of the classics on Tajweed and other works on Arabic phonology by about two hundred years. Some uncommon Tajweed styles have one or more of the following coarticulation assimilation types from his work.

In volume 4, page 437, the double process of vowel deletion and the germination of the bilabial nasal /m/ in the two words *ni3ma +ma* producing the word comes from the dialect of Huthayl. In that dialect, the first word in the phrase is /ni3ima/ but not /ni3ma/; with word-final vowel deletion (/ni3im/) the phonetic environment for the germination of the two m's is thus created allowing for /ni3imma/. In dialects which had the form of the first word as /ni3m/ as in *ni3ma*, the process of germination would produce a non-existent form /ni3mma*/. On page 448, Sibawayh says that the cross-word sequence /b/ + /f/ in a phrase such as /ithhab + fee/ may be regressively assimilated to become /ithhaffee/ . Two pages later, he cites an example of (de)voicing assimilation of the voiced pharyngeal fricative /3/ fricative (from the dialect of Banu Tameem) when it is followed by its voiceless counterpart /7/ /ma3/ + /hom/ (with them) produce /ma77om/. Furthermore, and two pages later (456), Sibawayh does not favor assimilating the lateral /l/ to a following voiceless alveolar voiceless plosive /t/; that is /hal/ + /taraa/ to remain as is and not to become /hattaraa/.

Similarly, page 458, he finds it acceptable to assimilate the same lateral /l/ to a following alveopalatal /sh/ in /hal shay?on/ to get /hashshay?on/ . Again, page 459, he accepts the assimilation of /l/ to a following /t/ or the voiceless /th/. The above are in some uncommon styles of Tajweed. However, Sibawayh does refer to what can be preferred, dispreferred, or odd in speech with such processes. His scale of judgment words range from good to ugly . Dialects then, now, and forever are perceived and received in different ways by others; some are considered 'the best', others 'the most beautiful', and others 'stigmatized, and so forth.

6. Research Findings on Coarticulation Assimilation

The literature on vernacular or casual speech (also by any of the other labels listed above) spans a few decades. In the past two decades, more research focused on the perception of casual speech (including studies on native and non-native subjects, e.g., Pearman, 2007) and on Automatic Speech Recognition (ASR) and Automatic Phonetic Transcription (APT) in computer software programming.

The ASR studies have shown the relative difficulty in the machine perception of casual speech in contrast with the perception of standard speech due to the higher frequency of assimilation. In one study, on Dutch (reported in Binnenpoorta, 2004), the threshold for recognizing variation in of careful speech pronunciation was 2.5 times higher than that in casual speech. This is in agreement with the findings of other investigations surveyed by that author. Using language data from Japanese, the study conducted by Nakamura and colleagues (2008) comparing 'read' texts, (e.g., news broadcasts and prepared texts) and spontaneous speech shows that a 90% recognition rate of 'read' style but a significantly lower rate of recognition for spontaneous speech. Matthies and coauthors (2001) also report that speech shows more coarticulation processes than 'normal' speech, i.e., careful, speech. The higher rate of coarticulation negatively affects the identification of forms when the assimilation is strong. (Gaskell and Snoeren, 2008)

Other ASR and APT studies for speech-to-text software (Binnenpoorta, et al., 2004 on Dutch; Nakamura, et al., 2008 on Japanese; Lindblom, 1983; Omesh, et al., 2011 by digitalizing data for software recognition) obtained results showing that spontaneous speech recognition or the transcription of spontaneous speech falls much below those of recognizing, or transcribing the reading style speech.

Investigations of the phenomenon on language education express the teachers' low tolerance of assimilation in the reading class (Elbow, 2006; Shockey, 2003) Carpenter (2010) reports on the relative difficulty of learning an unnatural phonological process as opposed to the relative ease of learning a natural process. A process is considered natural on the basis of its being attested in many languages.

Natural processes as indicators of casual speech (or the relaxed form of language) have been reported in many languages: Zwickey (1972) on English and Welsh, the work of the Cambridge Casual Speech group from 1986 – 1994 (as reported by Shockey, 2003: 76-77), Ingram (1989) on Australian English, Kohler (1990)

on German, Ernestus (2000) on Dutch, Barry and Andreas (2001) and Gaskell (2001) on several languages, Hsiar (2007) on Malaysian Cantonese, Torreira and Ernestus (2011) on French, Pham (2007) on Vietnamese, Pearman (2007) on English and Catalan each as a first and second language, and Barden (2009) on speech production. The list can still go further, but this sample suffices. This process, among others which do not concern the thesis of this paper, is universal but it varies in quantity and type across languages.

7. Conclusion

The great emphasis placed on the sanctity of Quran by Muslims is the raison d'etre of Tajweed. The standardization of Tajweed in the period between the fifth and ninth centuries AH has remained virtually beyond debate, despite phenomenal advances in the fields relevant to it in phonetics, phonology, and pronunciation. Unquestionably, the Quran text is holy to Muslims. Furthermore, reciting it is a form of worshipping, a serious performance but, it need not be overemphasized that recitation is not a casual speech act. Nonetheless, Muslim scholars of Quran reading continue to give the same detailed descriptions of the phonetics of recitation as they were given by their predecessors. However, the rules concerning consonantal coarticulation assimilation are representations of vernacular speech of Arabic of previous times, some of which continue to be carried into Arabic speech of the present time. Whereas the intention of upholding this coarticulation principle is meant to serve a lofty holy purpose, its implementation goes against the grain of the formality of reading standard Arabic; it, instead, with good intention and, literally, in good faith, lowers the level of attention and care in pronunciation given to standard Arabic.

List of Sources

Agwuele, Augustine, et al. 2008. 'The effect of speaking rate on consonant-vowel coarticulation' *Phonetica* V. 65 (4): 194 – 209

Barden, Katharine. 2009. 'Speech production 1'. *Linguistics at Cambridge, Paper 9*

Barry, William and Bistra Andreeva. 2001. 'Cross language similarities and differences in spoken speech patterns'. *Journal of the International Phonetic Association.* V. 31 (1): 51 – 66

Binnenpoorta, Diana, et al. 2004. 'Multi-word expressions in spoken language: an explanatory study on pronunciation variation'. *Computer Speech and Language.* V. 19: 433 – 449

Brouwer, Susanne, et al. 2012. 'Speech reductions change the dynamics of competition during spoken word recognition'. *Language and Cognitive Processes.* V. 27 (4): 539 – 571

Browman, Catherine P. and Lois Goldstein. 1987. 'Tiers in articulatory phonology with some implications for casual speech'. *Haskins Laboratories Status Report on Speech Research*

Carpenter, Angela C. 2010. 'Natural phonological processes.' *Phonology.* V. 27: 345 - 392

Dilley, Laura and Mark A. Pitt. 2007. 'A study of regressive place assimilation and its implication for spoken word recognition'. *Journal of the Acoustic Society of America.* 122: 2340 – 2353

Elbow, Peter. 2006. 'Vernacular literacy.' *English Department Publication Series.* University of Massachusetts, Amherst.

Ernestus, Mirjam. 2000. 'Voice assimilation and segment reduction in casual Dutch.' Unpublished PhD dissertation. Amsterdam: University of Amsterdam.

Farnetani, Edda. 1999. 'Coarticulation and connected speech processes' In *The Handbook of Phonetic Sciences,* edited by J.W. Hardcastle, John Laver and Fiona E. Gibbon. Oxford: Blackwell. Pp. 371 – 404

Gaskell, M. Gareth. 2001. 'Phonological variation: its consequences for the word recognition system'. *Language and Cognitive Processes.* V. 16 (5-6): 723 – 729

------- and Natalie D. Snoeren. 2008. 'The impact of strong assimilation on the perception of connected speech.' *Journal of Experimental Psychology: Human Perception and Performance.* V. 34 (6) 1632 - 1647

Graham, William A. 1987. *Beyond the Written Word: Oral Aspects of Scripture in the History of Religion.* Cambridge: Cambridge University Press.

Guenther, Frank H. 1995. 'Speech sound acquisition, co-articulation, and rate effects in a neural network model of speech production'. *Psychological Review.* V. 102 (3): 594 – 621

Hsiar, Ong Yin. 2007. 'Phonological elision in Malaysian Cantonese casual speech.' Unpublished M.A. Thesis: National University of Singapore.

Ingram, J. C. C. 1989. 'Connected speech processes in Australian English'. *Australian Journal of Linguistics* V. 9 (1): 21 – 49

Johnson, Keith. 2004. 'Massive reduction in conversational American English'. In *Proceedings of the Workshop on Spoken Speech: data and analysis,* edited by K. Yoneyama and K. Maekawa. Tokyo: National International Institute for Japanese Language. Pp. 29 – 54

Kohler, K. J. 1990. 'Segment reduction in connected speech in German. In *Speech Production and Speech Modelling,* edited by W. J. Hardcastle and A. Marchal. The Hague: Kluwer Academic. Pp. 69 – 92

Labov, William. 1966. *The Social Stratification of English in New York City.* Washington, D.C. : Center for

Applied Linguistics

Ladefoged, Peter. 1993. *A Course in Phonetics, 3ʳᵈ Edition.* New York: Harcourt Brace College Publishers

Lindblom, B. 1983. 'Economy of speech gestures'. In *The Production of Speech,* edited by P. F. MacNeilage. New York: Springer. Pp. 217 – 245

McDonald, Hancock. 2013 'The word blender: Teachers' notes'.

www.hancockmcdonald.com

Accessed online January 17, 2013

Matthies, Melani, et al. 2001. 'Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language, and Hearing Research.* V. 44: 340 – 353

Nakamura, Masanoba, et al. 2008. 'Differences between acoustic characteristics of spontaneous speech and read speech and their effects on speech recognition performance.' *Computer Speech and Language.* V. 22: 171 – 184

Omesh, Wadhawani, et al. 2011. 'Recognition of vernacular language speech for discrete words using linear predictive coding technique.' *International Journal of Soft Computing and Engineerin*g. V. 1 (5)

Pavlik, Radoslav. 2009. 'A typology of assimilation'. *SKASE Journal of Theoretical Linguistics.* V. 6 (1): 2 - 26

Pearman, Andrea. 2007. 'Native and non-native perception of casual speech'. Unpublished PhD dissertation. Universitat de Autonoma de Barcelona

Pham, Andrea Hoa. 2007. Cliticization in casual spoken Vietnamese.' *Cahiers de linguistique-Asie Orientalle. V. 36: 219 – 244)*

Robertson, Callum and Alan Stanton. 2005. 'BBC learning English: talk about English', pp. 1 – 6.

http://bbclearnenglish.com

Accessed online January 26, 2013

Shockey, Linda. 2003. *Sound Patterns of Spoken English.* Oxford: Blackwell

Torreira, Francisco and Mirjam Ernestus. 2011. 'Vowel elision in casual French.' *Journal of Phonetics.*V. 39: 50 – 58

Trudgill, Peter. 2000. *Sociolinguistics: an introduction to language and society, Fourth Edition.* London: Penguin Books

Zharkova, Natalie. 2011. 'Coarticulation as an indicator of speech motor development in children: an ultrasound study'. *Motor Control.* V. 15: 118 – 140

Zwicky, Arnold. 1972. 'On casual speech'. *Chicago Linguistic Society.* V. 8: 607 – 615