MATHEMATICAL FORMULAS FOR GENERATING SYLLABLES USED IN ARABIC SPEECH SYNTHESIS

Chegrani Lamari, Guerti Mhania, Boudraa Bachir

Algeria

ABSTRACT

This study aims to create speech of the Arabic language requires in this function is concatenation support, we can say that the presence of mathematical forms can ensure the generation of many sequences. This research aims to design a system that helps illiterate people in Arab nations learn spoken language in all its components (verbs, nouns, and simple particles). Further, with just 84 sub-syllables we can generate speech at many levels of complexity (syllable, word, sentence or text)depending on the differing strategies designed.

KEYWORDS

Arabic speaker, generative scheme, Arabic language, referential syllable, short vowels, long vowels.

1. INTRODUCTION

Humans primarily communicate through speech, which is the production of word sounds. In the last few decades, research on speech synthesis focused on making people with disabilities understand recognizable synthesizers and automation has made it possible to produce mechanized computers capable of producing understandable synthetic speech. The quest for a talking machine started in the eighteen century, and although modern speech synthesizers have made significant progress in their quality of articulation, they still lack the issues pertaining to the quality and naturalness of sound produced. Such research makes speech synthesis a crucial area of development in primary languages including Arabic. Because of various applications, the speech synthesizer or Text-To-Speech (TTS) technology is one of the most revolutionary inventions today. It can be found in human-computer interfaces in multi mediasoftwares and apparatuses for device usage, is used for reading text of emails or SMS on mobile phones, in computers for reading emails and text documents, etc. It is handy for most blind people as a basic reading machine, gives opportunities for communication to deaf and speech disabled people who are not familiar with sign language, and can be used in many educational activities such as speech therapy and teaching pronunciation for foreign languages.With over 437 million speakers around the globe, Arabic ranks as the most spoken language in the world, belonging to 25 countries [1].Alongside this, the language also holds religious significance for more than 1.6 billion Muslims[2].

Moreover, the Arab world has an excess of 5 million visually impaired people[3]. It is important that an automated accurate and simple TTS system is developed for the Arabic language so that these people can communicate using SMS, email, webpages, etc. A speech synthesizer consists of a Digital Signal Processing (DSP) module and a text processing unit, which are its main two portions. The former does two critical functions. First, it converts “unformatted” text containing symbols such as numbers and acronyms to its more readable form. This procedure is also referred to as text normalization. Then, it provides the text’s data in another symbolic form to the DSP or synthesizer, which changes the symbols into speech. Different means and various languages have been synthesized by many researchers to produce speech. In 1987 NETtalk, Sejnowsky and Rosenberg built a neural network learning to speak English text. This system was developed using numerous parallel network systems capable of capturing various regularities and exceptions of English pronunciation and so facilitated transforming English text strings into phonetic representations[4]. Karaali et al. [5]developed a rule-based system that consists of two neural networks. The time-delay neural network convertsthe phonetic representation of speech into the acoustic representation andgenerates speech. The second neuralnetworkisappliedfor output timingcontrol.Concatenation-based speech synthesis uses the speech inventory bychoosingunitsandalgorithmsforjoiningthemalong with some kind of processing, such as smoothing the bordersbetween the concatenated parts.[6].

Telecommunications serviced at present time is the primary market for such techniques. These types of services exemplify scenarios that use speech synthesis as an important means for a computer system to transmit information to the user.Our work centers on the development of speech synthesis system based on symmetry technique in generating formant transition patterns using a dictionary of 84 different acoustic units consisting only of [CV]-type sound units (Consonant-Vowel).Based on mathematical models, speech sequences are generated into small, medium or large unit sequence varieties such as syllable sequence, word sequence and sentence sequence.The work of this article reports the process of text-to-sound sequence conversion, first from small to exponentially large sequences of sounds.Thus, for example in creating a sentence sequence, we first construct a syllable sequence, followed by a word sequence, and lastly a sentence sequence; with smooth increments in quality at every transition upward from syllable to sentence.

2. SYLLABLES OF THE ARABIC LANGUAGE

The investigation of the operation of the Arabic word can be illustrated by the written components of sentences in our corpus. The Arab grammarians have undertaken several studies concerning this language which is utilized by a fourth of the world's population [3]. A different aspect of this study is we will be focused on the analyses of Arabic syllables that have been categorized into six types as follows: CV, CVV, CVC, CVVC, CVCC and CVVCC which occurred only sporadically at the end of sentences (finalizing with a double consonant preceded by a long vowel).

The objective is to establish a method for creating every possible form of syllables, from one syllable that, for our purposes we have labelled reference syllable. The selection of reference syllable is contingent on a number of criteria for the method to yield satisfactory results of speech synthesis on the basis of its rules. The selection criteria of reference syllable are influenced by the language's characteristics; and it is determined from the speech sequence dependent on the successional states of consonants and vowels. In the Arabic language the beginning of a sequence is always consonant followed by a short or long vowel (CV or CVV). Likewise, the sequential state of consonants (C...C), can have a maximum of two consonants. In this study, to develop mathematical models, we need to first give the variables and constants of our data. The main distinction between variable and constant is the number which are equal to 3 for vowels and equal to 28 for consonants. It is preferable to consider consonants as variables and vowels as constants. We take the upper (maximum) for variable and the lower (minimum) for constant. From the data, we find that the reference syllable will be CVCC.

The following expressions can give an idea about the generation of secondary syllables from the reference syllable $CnV_nC_{nn}C_{nnn}$:

- $CV = \sum((C,V,C,C)\times(C,V,0,0))$
- $CVC = \sum((C,V,C,C)\times(C,V,C,0))$
- $CVV = \sum((C,V,C,C)\times(C,V,V,0))$
- $CVCC = \sum((C,V,C,C)\times(C,V,C,C))$
- $CVVC = \sum((C,V,C,C)\times(C,V,V,C))$
- $CVVCC = \sum((C,V,C,C)\times(C,V,V,C),C)$

Nevertheless, this method of lexical form generation has flaws relating to the usage of variables, in cases where one has switched the value of (C) with a vowel value and using (C) without support of concatenation to generate CVVCC .well, the value of (C) cannot be that of a value of (V);for this purpose, the last concatenated expressions of long vowels will need to be reformulated as follows:

- $CV = \sum((C,V,C,C)\times(C,V,0,0))$
- $CVC = \sum((C,V,C,C)\times(C,V,C,0))$
- $CVV = \sum((C,V,C,C)\times(C,V,0,0)(C,V,C,C)\times(0,0,V,0))$
- $CVCC = \sum((C,V,C,C)\times(C,V,C,C))$
- $CVVC = \sum((C,V,C,C)\times(C,V,0,0),(C,V,C,C)\times(0,0,V,C))$
- $CVVCC = \sum((C,V,C,C)\times(C,V,0,0),(C,V,C,C)\times(0,V,C,C))$

3. SYNTHESIS OF SPEECH

The process of speech synthesis is presented in the Arabic language through the created templates shown in table (1), we can extract syllables representing the conditions of the short vowels, and their complements to construct the conditions of the long vowels:

Table 1: vowel generation using syllables of the Arabic language

	Syllables			**able complements**		
Syllable to generate	*CV*	*CVC*	*CVCC*			*VCC*
values of C(s)	*CV*00	*CVC*0	*CVCC*0*V*	00	0*VC*0	0*VCC*
vector of generation	(C,0,0)	(C,C,0)	(C,C,C)	,0)	,0)	,C)

According to the values of the vowels in the syllables and the complements of the syllables, we can classify the results into three categories as indicated in the table (2):

Table 2: vowels (Fat'ha , Dhamma and Kasra) using syllables of the Arabic language

	ables			**Syllable complements**		
Syllable to generate	CV	CVC	CVCC	V	VC	VCC
Fat'ha case	Ca00	CaC0	CaCC	0a00	0aC0	0aCC

Dhamma case	Cu00	CuC0	CuCC	0u00	0uC0	0uCC
Kasra case	Ci00	CiC0	CiCC	0i00	0iC0	0iCC

The schemes corresponding to the units are given in table (3):

Table 3: Generation of syllable patterns in the Arabic language.

	Syllables			**Syllable complements**		
Syllable to generate	*CV*	*CVC*	*CVCC*	*V*	*VC*	*VCC*
Fat'ha case	Fa00	Faʕ 0	FaʕL	0a00	0aʕ 0	0aʕ L
Case of Dhamma	Fu00	Fuʕ 0	Fuʕ L	0u00	0uʕ 0	0uʕ L
Kasra case	Fi00	Fiʕ 0	Fiʕ L	0i00	0iʕ 0	0iʕ L
vector of generation corresponds to schema	(F,0,0)	(F,ʕ,0)	(F,ʕ,L)	(0,0,0)	(0,ʕ,0)	(0,ʕ,L)

4. GENERATION OF SYLLABLE SCHEMES WITH SHORT VOWELS

In this case, we have three cases which are:

A. Cas de [CV]

For the example of the coordination tool (huruuf al-ʕatˤ f) $[wa]$, its scheme is (Faʕ L) with a value generation vector (*w,0,0*), and the schema can be reduced to (Fa00).$(w, 0{,}0)$

B. Case of [CCV]

For the example of Nasb's tool (huruuf an-Nasb) $[lan]$, its scheme is Faʕ L with a value generation vector (l,n,0) and the scheme can be reduced to (Faʕ 0).

C. Case of [CVCC]

For the two examples [radd] and [dʒurĦ] which have respectively two identical consonants and two different ones, their respective schemes are FaʕL and FuʕL with value generation vectors (r,d,d) and (dʒ,r,Ħ) respectively.

5. GENERATION OF SYLLABLE SCHEMAS WITH LONG VOWELS

For this we can define three different cases:

A. Case of [CVV]

- For the case of two identical vowels, which is the case of long vowels, as in the example of the Nasb tool (HelBarf an-Nasb)/maa/, its schema consists of two schemas faʕlfaʕl with two vectors generating values (m,0,0) and (0,0,0). The new scheme resulting from this concatenation is as follows:

$$sch\grave{e}me(maa) = sch\grave{e}me(\sum(CV, V))$$

We have:

$$CV = \sum(CVCC) \times (C, V, 0{,}0)$$

And:

$$V = \sum (CVCC) \times (0, V, 0,0)$$

So, we will have:

$$(CV, V) = \sum (CVCC) \times (C, V, 0,0), \sum (CVCC) \times (0, V, 0,0)$$
$$(CV, V) = \sum ((CVCC) \times (C, V, 0,0), (CVCC) \times (0, V, 0,0))$$

For our example:

$$sch\grave{e}me(maa) = sch\grave{e}me[\sum ((CVCC) \times (m, a, 0,0), (CVCC) \times (0, a, 0,0))]$$
$$sch\grave{e}me(maa) = \sum ((Fa\text{ʕ}L) \times (m, 0,0), (Fa\text{ʕ}L) \times (0,0,0))$$

The result will be:

$$sch\grave{e}me(maa) = \sum (Fa \times (m), a) = Faa \times (m)$$

So the schema is [Faa] with a generation value: F = m.f = m

- For the case of two different vowels of *CVV*, which is the case of diphthongs, we have two types in the Arabic language: [aw] (or [au]) and [aj] (or [ai]). In the same way, we will have the following results:

  - For the syllable [Cau]: the scheme is either [Fau] orfaw .
    - Example: scheme of [ʔ aw] isfau with: (F = ʔ).
  - For the syllable [Cai]: the schema isfai] or [Faj].
    - Example: scheme of [kai] iskaj [fai] with: (f = k)

B. Case of [CVVC]

- For the case of two identical vowels, which is the case of long vowels, the example of the defective verb (Fiʕ L Naaqisˁ) [kaan], its schema is composed of twofaʕlfaʕl schemas faʕ lfaʕ l with two value generation vectors: (k,0,0)(0,0,0)(k, 0,0)(0,0,0). The new scheme resulting from this concatenation is as follows:

$$sch\grave{e}me(kaan) = sch\grave{e}me(\sum (CV, VC))$$

We have:

$$CV = \sum (CVCC) \times (C, V, 0,0)$$

and:

$$VC = \sum (CVCC) \times (0, V, C, 0)$$

So, we will have:

$$(CV, VC) = \sum (CVCC) \times (C, V, 0,0), \sum (CVCC) \times (0, V, C, 0)$$

$$(CV, VC) = \sum((CVCC) \times (C, V, 0,0), (CVCC) \times (0, V, C, 0))$$

For our example:

$$sch\grave{e}me(kaan) = sch\grave{e}me[\sum((CVCC) \times (k, a, 0,0), (CVCC) \times (0, a, n, 0))]$$

$$sch\grave{e}me(maa) = \sum((Fa\text{ʕ}L) \times (k, 0,0), (Fa\text{ʕ}L) \times (0, n, 0))$$

The result will be:

$$sch\grave{e}me(maa) = \sum(Fa \times (k), a\text{ʕ} \times (n) = Faa\text{ʕ} \times (k, n)$$

So the schema is (faaʕ 0) with a value generation vector: (F,ʕ) = (k,n).faaʕ

- For the case of two different vowels, which is the case of diphthongs, we have two types of diphthongs in the Arabic language [aw] (or [aw]) and [aj](or [ai]). If we do the same, we will have the following results:

  - For [CauCCauC]: the scheme is: [Fauʕ] or elsefawʕ].
    - Example: the schemefauq is [Fauʕ] with $G_{RCN:}(\mathrm{f}, \text{ʕ}) = (\mathrm{f}, \mathrm{q})$
  - For [CaiC]: the scheme is either:faiʕor].fajʕ
    - Example: Eastkajf withfaiʕ G*RCN scheme*$_{:}(\mathrm{f}, \text{ʕ}) = (\mathrm{k}, \mathrm{f})$

Note: for the case of diphthongs, it will be better to take waw and yaʔ as consonants and their schema will befaʕl with two vectors of generation of values (C,w,C) and ((C, y, C)) respectively, to avoid heavy forms of the Arabic language.

C. Case of [CVVCC]

For the example [Ħ$aarr$] which has two identical consonants, its schema is composed in the same way as the previous one, i.e., of two schemas (faʕlfaʕl) with two vectors of generation of values: (X,0,0) and (0,r,r). The new scheme resulting from this concatenation is as follows:

$$sch\grave{e}me\ (\text{Ħ}aarr) = sch\grave{e}me(\sum(CV, VCC))$$

We have:

$$\mathrm{CV} = \sum(\mathrm{CVCC}) \times (\mathrm{C}, \mathrm{V}, 0,0)$$

And:

$$VCC = \sum(CVCC) \times (0, V, C, C)$$

So, we will have:

$$(\mathrm{CV}, \mathrm{VCC}) = \sum (\mathrm{CVCC}) \times (\mathrm{C}, \mathrm{V}, 0{,}0)\,, \sum (\mathrm{CVCC}) \times (0, \mathrm{V}, \mathrm{C}, \mathrm{C})$$

$$(\mathrm{CV}, \mathrm{VCC}) = \sum ((\mathrm{CVCC}) \times (\mathrm{C}, \mathrm{V}, 0{,}0)\,, (\mathrm{CVCC}) \times (0, \mathrm{V}, \mathrm{C}, \mathrm{C}))$$

For our example:

$$sch\grave{e}me(\mathrm{H}aarr) = sch\grave{e}me[\sum ((\mathrm{CVCC}) \times (\mathrm{H}, \mathrm{a}, 0{,}0)\,, (\mathrm{CVCC}) \times (0, \mathrm{a}, \mathrm{r}, \mathrm{r}))]$$

$$sch\grave{e}me(\mathrm{H}aarr) = \sum ((\mathrm{Fa\text{ʕ}L}) \times (\mathrm{H}, 0{,}0)\,, (\mathrm{Fa\text{ʕ}L}) \times (0, \mathrm{r}, \mathrm{r}))$$

If we reduce the null value boxes, the result will be:

$$sch\grave{e}me(\mathrm{H}aarr) = \sum (\mathrm{Fa} \times (\mathrm{H}), \mathrm{a\text{ʕ}L} \times (\mathrm{r}, \mathrm{r}) = \mathrm{Faa\text{ʕ}L} \times (\mathrm{H}, \mathrm{r}, \mathrm{r})$$

So the scheme is Faaʕ L with a value generation vector:(X, r,r).

6. Generation of Patterns of a Sentence

Utilizing the information in the table, we can develop the frameworkfor automatic synthesis of the Arabic speech from any text. In this case we assume the entry is a sentence consisting of three parts:FiʕL+FaaʕiL+maFʕuuL.

We have chosen a simple phrase which is: [qara′ a قرأ معلم درسا muʕallimundarsan]or

Table 4: generation of the patterns of a sentence [qara′ a muʕallimundarsan]

Phrase in Arabic	قرأ معلم درسا								
Phonetic writing	qara′ a muʕallimundarsan								
Segmentation into lexical units	qara′ a			muʕallimun				Darsan	
Segmentation into syllables and Complements of Syllables	qara′ a			muʕallimun				Darsan	
	Qar	a′	a	muʕ	all	im	one	Dars	An
Syllables and Complements of Syllables	CaCaC a			CuCaCCiCuC				CaCCaC	
ableScheme and SyllableComplements	Faʕ a			FuʕaʕLiʕuʕ				FaʕLaʕ	
Generation of the general scheme of the sentence (expression maths)	Faʕ a			FuʕaʕLiʕuʕ				FaʕLaʕ	
Stage of synthesis	(F,ʕ,ʕ)=(q,r,′)			(F,ʕ,ʕ,L,ʕ,ʕ)=(m,ʕ,l,l,m,n)				ʕ,L,ʕ)=(d,r,s,n)	

In this instance, we obtained in the speech synthesis stage, speech concatenation expressions with variables expressed by (F,ʕ,L) which are automatically substituted by the values of the generation vector that only contains the consonants extracted in the first phase (sequence selector) from phonetic writing.

7. Null Values in a Syllable

The process of synthesizing a speech-generating system of Arabic language as proposed in this method wouldrequire various points which have valuezero. These points may beconsidered aspoints of discontinuity produced during the synthesis of an Arabic word, which will be heard as silence or pauses in the sound that corresponds to the lexical unit. This discontinuity affects the meaning ofthe phrase (the semantics) thatweareaiming for.In thefollowing section, we attempt to demonstrate the points ofdiscontinuity corresponding to the null bridges of the value associated with eachconnectingvector, with the timing or these points provided to understand if this method is sufficiently capable and usable in synthesizing Arabic speech.

8. CALCULATIONS OF THE TIMES OF THE VOIDS

The method proposed in this section is that every syllable in Arabic language is represented basedon the CVCC reference syllable.This indicates that the sequence is constructed by concatenation of a limited number of CVCC reference syllables, this number is equal to the number of vowels belonging to the sonorized sentence. In case the sentence has 10 vowels for instance, the form generation of this sequence applieson a concatenation of 10 CVCC reference syllables,each represented by its respective scheme according to the vowels.This schemecan be, as mentionedabove of the types of shemes:FaʕL for the syllable CaCC, andFuʕL for the syllable CuCC, and FiʕL for syllable CiCC. Referringtothepreviousentence (qaraʹ a muʕallimundarsan), it included 9 vowels,Itispossible to generate the schema of this sentence from a schema representation of 9 CVCC reference syllables as follows:

Indication of vowels, vector of values:

$$VV_v = \left[a \quad a \quad a \quad u \quad a \quad i \quad u \quad a \quad a \right]$$

Syllable vector:

$$V_v = \left[CaCC \quad CaCC \quad CaCC \quad CuCC \quad CaCC \quad CiCC \quad CuCC \quad CaCC \quad CaCC \right]$$

Consonants indication: vector of values:

$$VV_c = \left[q \quad r \quad ' \quad m \quad \varsigma \quad l \quad l \quad m \quad n \quad d \quad r \quad s \quad n \right]$$

Table 5: calculation of the time of null values

<table>
<tr><td>hrase in Arabic</td><td colspan="9">قَرَأَ مُعَلِمٌ درسًا</td></tr>
<tr><td>honetic writing</td><td colspan="9">qaraʹ a muʕallimundarsan</td></tr>
<tr><td>lexical units</td><td colspan="3">qaraʹ a</td><td colspan="4">muʕallimun</td><td colspan="2">Darsan</td></tr>
<tr><td>and V coding</td><td colspan="3">CaCaCa</td><td colspan="4">CuCaCCiCuC</td><td colspan="2">CaCCaC</td></tr>
<tr><td>epresentation in CVCC</td><td>CaCC</td><td>CaCC</td><td>CaCC</td><td>CuCC</td><td>CaCC</td><td>CiCC</td><td>CuCC</td><td>aCC</td><td>CaCC</td></tr>
<tr><td>yllable scheme</td><td>faʕ l</td><td>faʕ l</td><td>faʕ l</td><td>fuʕ l</td><td>faʕ l</td><td>fiʕ l</td><td>fuʕ l</td><td>faʕ l</td><td>faʕ l</td></tr>
<tr><td>CVCC Coding</td><td colspan="3">CaCaCa</td><td colspan="4">CuCaCCiCuC</td><td colspan="2">CaCCaC</td></tr>
<tr><td>Zero values</td><td>CaC0</td><td>0aC0</td><td>0a00</td><td>CuC0</td><td>0aCC</td><td>0iC0</td><td>0uC0</td><td>aCC</td><td>0aC0</td></tr>
<tr><td>onsonant coding</td><td colspan="3">CaC00aC00a00</td><td colspan="4">CuC00aCC0iC00uC0</td><td colspan="2">CaCC0aC0</td></tr>
<tr><td>Decoding</td><td colspan="3">qar00aʹ 00a00</td><td colspan="4">muʕ 00all0im00un0</td><td colspan="2">dars0an0</td></tr>
<tr><td>eneration scheme</td><td colspan="3">faʕlfaʕlfaʕ l</td><td colspan="4">fuʕlfaʕlfiʕ l</td><td colspan="2">faʕlfaʕ l</td></tr>
<tr><td>nsertion of null values</td><td colspan="3">faʕ 00aʕ 00a00</td><td colspan="4">fuʕ 00aʕ l0iʕ 00 0</td><td colspan="2">faʕ l0aʕ 0</td></tr>
<tr><td>Remarks</td><td colspan="3">* ʕ]2pt(0)[a</td><td colspan="4">* ʕ]2pt(0)[a
* l]pt(0)[i
* ʕ]2pt(0)[u</td><td colspan="2">* l]pt(0)[a</td></tr>
</table>

Time of silence	* $\frac{2}{11025}s < 0.2ms$	* $\frac{2}{11025}s < 0.2ms$ * $\frac{1}{11025}s < 0.1ms$	* $\frac{1}{11025}s < 0.1ms$
general scheme	Faʕ a	FuʕaʕLiʕuʕ	FaʕLaʕ
VVc	(F,ʕ,ʕ,F,ʕ,ʕ,L,ʕ,ʕ,F,ʕ,L,ʕ)=(q,r,′,m,ʕ,l,l,m,n,d,r,s,n)		

9. COMPARISON WITH EXIST METHOD

We compare the results of Arabic speech synthesis using the generative grapheme technique to produce [CVh] with the results of linear smoothing for a hybrid composition system using a polyphems method [8]. The important result that can be concluded from this comparison is the overall percentage of the regeneration method compared to the schemes method, which is approximately 97.9%, with the correction of only two ambiguous sounds in the fourth and seventh sentences, i.e., only 22 listening tests. In contrast, the results of the polyphems method use the correction and modification results for all sentences three times for each sentence, i.e., 60 listening tests. This yields the previous results[7]; the overall average is estimated at 90.9%, with a 7% difference from the overall speech recognition rate of the schemes method. Moreover, the values of this method are almost equal with values over 92%, which explains the effectiveness, efficiency, discrimination, preference and stability of the recognition rate of different sentences of the schemes method.

10. CONCLUSION

The identification rate of the constructed Arabic speech sequences largely depends on the quality of the sound snippets used. The generative form that carries formant transitions, a rather advantageous technique to automatic synthesis of speech within Arabic language, is based on the mirror technique that generates symmetric syllables of those opposites of type [CV]. These syllables are certainly carrying formant transitions to the left and right of a vowel which we have referred to as dependent on the left and right, information contained in the continuity and intelligibility section of the synthesized speech, that is the region between the consonant and the vowels carrying this information acoustically. It is the variation of the values of the frequencies F1 and F2, named formants. It exists in the sound units of type [CV] of various possible cases, (28x3), that is, 28 consonants combined with three different Arabic vowels (Fat'ha, Dhamma, and Kasra). So this is an extremely important stage. The corresponding vowels to the emphatic phonemes can be used in the case of realization of long vowels [CVV], we do segmentation in several qualities:

- of type]VV] to generate long vowels of closed syllables [CVV];
- of type]VV[to generate long vowels for open syllables, this is the case of generation of syllables of quality [CVVC and CVVCC];
- and of type]V] for the generation of closed syllables with short vowel [CV], this is the case of an emphatic consonant with short closed vowel (example: [bɑHr]).

As we have seen, these segments are identical in both the emphatic and ordinary cases. The possibility of generating all Arabic syllables symmetrically while maintaining formant transitions (mirror technique) and the total number of Arabic consonants to generate the cases of the two successive consonants (CVCC, CVVCC), or all Arabic sequences, is made possible by this technique, which enables us to synthesize speech using only [CV] type units.

REFERENCES

[1] "Liste des pays ayant l'arabe pour langue officielle," in Langue internationale ou mondiale, https://creativecommons.org/licenses/by-sa/3.0/deed.fr, Ed., ed. Etats-unis: organisation de bienfaisance régie par le paragraphe 501(c)(3) du code fiscal des États-Unis., 2022.

[2] B. J. Grim and B. Hsu, "Estimating the global Muslim population: Size and distribution of the world's Muslim population," Interdisciplinary Journal of Research on Religion, vol. 7, 2011.

[3] A. B. Kain and J. P. v. Santen, "A speech model of acoustic inventories based on asynchronous interpolation," in Eighth European Conference on Speech Communication and Technology, 2003.

[4] W. H. Organization, "Country focus: annual report 2008," World Health Organization2008.

[5] O. Karaali, G. Corrigan, and I. Gerson, "Speech synthesis with neural networks," in World Congress on Neural Networks, San Diego, 1996, pp. 45-50.

[6] A. Indumathi and E. Chandra, "Survey on speech synthesis," Signal Processing: An International Journal (SPIJ), vol. 6, no. 5, p. 140, 2012.

[7] L. Chegrani, G. Mhania, and B. Bachir, "The symmetric technique of formant transition generation for use in speech synthesis in Arabic," International Journal of Information Technology, vol. 17, no. 2, pp. 1235-1245, 2025/03/01 2025.

[8] T. Saidane, M. Zrigui, and M. B. Ahmed, "Un système de synthèse de la parole arabe par concaténation de polyphèmes: Les résultats de l’utilisation d’un lissage linéaire," in 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, Tunis, https://aclanthology.org/2005.jeptalnrecital-recitalcourt.11., 2005.