

Phonetic Rules in Arabic Script

Mansour M. Al-ghamdi*, Husni Al-Muhtasib and Moustafa Elshafei**

King Abdul Aziz City for Sciences and Technology, Riyadh

King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

(Received 8 January 2002; accepted for publication 17 November 2002)

Abstract. Conversion of Arabic script into phonetic rules is one of the major obstacles facing the researchers on Arabic text-to-speech systems and speech recognition. Although Arabic is one of the oldest languages that its sounds and phonological rules were extensively studied and documented (more than 12 centuries ago), these valuable studies need to be compiled from scattered literatures and formulated in a modern mathematical framework.

The objective of this paper is to present to the interested researchers a mathematical formulation for a comprehensive set of these rules.

قوانين الفونولوجيا العربية

منصور بن محمد الغامدي*، حسني المختسب و مصطفى الشافعي**

*مدينة الملك عبدالعزيز للعلوم والتقنية، الرياض

** جامعة الملك فهد للبترول والمعادن، الظهران

(قدّم للنشر في ٠٨/٠١/٢٠٠٢م؛ وقبل للنشر في ١٧/١١/٢٠٠٢م)

ملخص البحث. يجد الباحثون العاملون في مجال النطق الآلي والتعرف الآلي على الكلام العربي صعوبة لحد ما في التعامل مع النص العربي وذلك من حيث تحويل رموزه المكتوبة إلى رموز صوتية. ورغم أن العربية تكاد تكون أقدم لغة حُدّدت أصواتها ودون نظامها وقوانينها الفونولوجية (قبل أكثر من ١٢ قرناً) إلا أنه لم توضع هذه الخصائص والقوانين الفونولوجية بشكل يمكن للعلمين في مجال الحوسبة الاستفادة منها بطريقة مباشرة. ولهذا فإن هذه الورقة تعيد طرحها من جديد ولكن بطريقة ميسرة وتبينها بالأمثلة حتى يمكن توظيفها في مجالي التعرف الآلي على الكلام والنطق الآلي.

١. المقدمة

مع الطفرة التقنية التي يعيشها العالم في هذا العصر وخاصة الإلكترونية منها أصبح تحويل النص اللغوي المكتوب إلى منطوق text-to-speech مطلباً مهماً ليس لتسهيل التواصل بين المتحدثين فحسب ولكن لخدمة أعمال أخرى لا تقل عنها أهمية كالتواصل عبر شبكات الحاسوب بين المعاقين وتحويل الكتب والدوريات والجرائد والنصوص المكتوبة

إلا أن تنفيذ مشروع للناطق الآلي العربي ليس بالأمر اليسير، مثله في ذلك مثل تنفيذ نفس النظام لأية لغة أخرى ولو أن الخصائص اللسانية العربية من أقلها صعوبة عند مقارنتها بتلك في اللغات الأخرى من حيث التعامل مع الحاسب الآلي وبرامجه. فالعربية لغة خوارزمية من حيث النحو والصرف والاشتقاق والكتابة، على الأقل. بمعنى أنه يمكن وضع قواعد رياضية تحدد السلوك اللساني في العربية. فعلى سبيل المثال، لا نجد قانوناً يمكن أن يفسر نُطق ع في الإنجليزية في الكلمات: laugh, through, good, geography. بينما لكل حرف في العربية رمز صوتي لا يشذ عنه إلا لوجود قانون محدد كما هو موضح في القوانين التي سنأتي على ذكرها في هذه الورقة. وبالنسبة للنص العربي فإنه لا بد من المرور بالمراحل التالية عند الرغبة في تحويله من مكتوب إلى منطوق:

١. تشكيل النص

أي وضع علامات التشكيل diacritics (ُ ُ ُ ُ) ^(١) على الأحرف العربية لتتوفر جميع الرموز الكتابية للأصوات اللغوية العربية grapheme والتي تتكون من الحروف وعلامات التشكيل. ولقد اعتاد العرب على قراءة النص العربي دون تشكيل وذلك لأن المعرفة اللسانية linguistic knowledge عندهم تقوم بوضع علامات التشكيل تلقائياً وبدون أن نشعر بذلك. والأعجب من هذا أن العرب كانوا قبل عصر التدوين يكتبون ليس بدون تشكيل فحسب ولكن بدون نقاط أيضاً فكلمة مثل "بعيد" يمكن إذا حُذفت نقاطها أن تعني: "تعيد، تعيد، يعيد، يعند..." مما يزيد الأمر صعوبة لدى القارئ للوصول للكلمة المقصودة. أما الآن فإن النص العربي يحمل النقاط لكنه في معظم الأحيان لا يكون مشكلاً. ولكن المعرفة اللسانية لدى العربي المعاصر كفيلاً بتشكيل النص غير المشكل بسرعة وفعالية فائقتين مما يجعلنا لا نحس بها ولا تُبطئ من عملية القراءة. وعند التمعن في هذه العملية أو محاولة محاكاتها حاسوبياً فإننا نجد أن النص غير المشكل يحتاج لعمليات لسانية معقدة منها

(١) نود الإشارة إلى أن السكون ليس رمزا لصوت معين وإنما هو رمز يدل على أن الصامت الواقع عليه متبوعاً بالصامت التالي دون وجود صائت بينهما. لذا فإن وجوده تكرر لا ضرورة له، فهو لا يمكن أن يرد مرادف لصائت على صامت واحد ولا يأتي إلا في غياب الصائت. فغياب الصائت يدل على أن الصامت غير متبوع بصائت، ومن ثم فوروده من عدمه سيان.

كما تشتمل هذه المرحلة على تحويل الرموز والأرقام والاختصارات والكلمات الشاذة في العربية إلى نصوص مشكّلة ينطبق عليها ما ينطبق على غيرها من النصوص الأخرى. فعلى سبيل المثال، يحول الرقم ٢٣ إلى "ثلاثة وَعَشْرُونَ"، وتحول الكلمة "هذا" إلى /هاذا/ والقائمة التالية تسرد الاستثناءات والاختصارات:

- إله = تزداد ألفا فتتطق /إلاه/ (١).
- أنا. = اختصار لـ "أنبأنا"
- اه. = اختصار لـ "انتهى"
- أولئك = تحذف الواو وتزداد ألفا فتتطق /الألك/.
- أولو = تحذف الواو الأولى فتتطق /ألو/.
- أولي = تحذف الواو فتتطق /ألي/.
- بسم = أصلها "باسم" إلا أنها تنطق /بسم/ ولا ترد كتابتها "بسم" إلا في البسمة، أما فيما سواها فإنها تكتب "باسم"، ولهذا تحذف الألف في هذه الحالة وتحل محلها كسرة /بسم/.
- تحذف الألف في نهاية الفعل المنتهي بواو الجماعة. مثل : رموا، كتبوا، عملوا.
- ثنا. = اختصار لـ "حدثنا"
- الخ. = اختصار لـ "إلى آخره"
- داود = تنطق /داوود/.
- ذلك = تزداد ألفا فتتطق /ذالك/.
- الذي = تضعف اللام فتتطق /الذي/. وكذلك بالنسبة للكلمتين: "الذين" و"التي".
- الرحمن = تزداد ألفا فتتطق /الرحمان/.
- رحه. = اختصار لـ "رحمه الله"
- رضه. = اختصار لـ "رضي الله عنه"

(١) في هذه القائمة، استخدمت علامتا الاقتباس "" للدلالة على أن النص الوارد بينهما هو النص كما يكتب، واستخدم الخطان المائلان // للدلالة على أن النص الوارد بينهما هو حروف للدلالة على الأصوات المنطوقة وهذا لا يتفق دائما مع المكتوب.

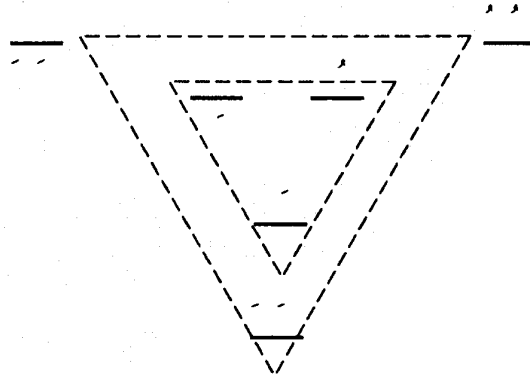
- السموات = تزداد ألفا فتنتطق /السموات/.
- ص = اختصار لـ "صلى الله عليه وسلم".
- طه = تزداد ألفان لتصبح /طاها/
- عمرو = تنطق /عمرُ/.
- لكن = تزداد ألفا فتنتطق /لاكن/.
- الله = تُزداد ألفا فتنتطق /الله/.
- اللهم = تزداد ألفا فتنتطق /اللاهم/.
- نا. = اختصار لـ "أخبرنا".
- ياسين = تزداد ألفا فتصبح /ياسين/.
- هؤلاء = تزداد ألفا فتنتطق /هاؤلاء/.
- هذا = تزداد ألفا فتنتطق /هاذا/.
- هذان = تزداد ألفا فتنتطق /هذان/.
- هذه = تزداد ألفا فتنتطق /هاذه/.
- هذين = تزداد ألفا فتنتطق /هاذين/.

٢. تحويل الحرف العربي والحركة التي عليه إلى رموز صوتية فونيمات (phonemes) والقصد من الرموز هنا هو تعريف نطق كل حرف بحيث يمكن تحديد خصائصه الأكوستية acoustic characteristics ومن ثم توليدها (الجدول رقم ١) والشكل رقم (١)). وهذه المرحلة تشتمل على نطق الحروف والرموز العربية، بما فيها التشكيل في جميع المواقع.

٣. وضع القواعد الفونولوجية التحويلية (Transformational phonological rules) ويتم في هذه المرحلة تحويل الفونيمات إلى ألو فونات allophones ، حيث تتولى القواعد الفونولوجية توضيح تأثير الأصوات (الفونيمات) على بعضها البعض وتحديد الصوت الذي يتم إخراجه في نهاية المطاف. ويندرج تحت هذه القواعد تأثيرات الأصوات المتجاورة وما يتجاوزها إلى الأصوات التي بعدها.

الجدول رقم (١). أصوات العربية الفصحى المعاصرة ، الأصوات المجهورة تقع على يمين العمود بينما تقع المهموسة على يساره
 (* تعني مفخم شديد ، ** تعني مفخم زخم) [١].

	Bilabial	Labiodental	Interdental	Alveodental	Alveopalatal	Palatal	Velar	Lab-velar	Uvular	Pharyngeal	Glottal
Nasal	m			n							
Stop	p			t			k		q		ʔ
Stop											
Emphatic Stop*				tʰ							
Fricative		f	θ	s	ʃ				x	ħ	ʕ
Emphatic Fricative**			θʰ	sʰ							
Affricate											
Glide						j		w			
Lateral					l						
Glottal											
Trill				r							



الشكل رقم (١). الصوائت العربية الستة: ثلاثة منها قصيرة والأخرى طويلة. ويقابلها في الرموز

$$\begin{aligned} /a:/ &= / \text{---} / & /u/ &= / \text{---} / & /i/ &= / \text{---} / & /a/ &= / \text{---} / \\ /u:/ &= / \text{---} / & /i:/ &= / \text{---} / & & & & \end{aligned}$$

٤. ربط الألفونات ببعضها

بحيث يتوفر الانتقال الأكوستي (coarticulation) بينها، فالإشارة الأكوستية لا تنتقل فجأة من صوت لغوي ما إلى صوت آخر، إذ لا بد من فترة للانتقال بين الأصوات اللغوية. وهذا يعطي الكلام خاصية الوضوح بحيث يمكن للسامع فهم ما يُقال.

٥. تحسين الصوت

حتى يكون قريبا قدر ما أمكن من الصوت اللغوي الصادر من جهاز الصوت البشري وذلك بإضافة الخصائص الفوققطعية (prosody/suprasegmental) كأمداد مقاطع الصوت duration وشدته amplitude وتردده الأساس frequency (F₀) fundamental. وهذا يعطي الكلام خاصيته الطبيعية بحيث يرتاح له السامع ولا يشعر بأنه كلام آلي.

القوانين الفونولوجية

لكل مرحلة من المراحل السابقة تحدياتها وطرق معالجتها، أما هذه الورقة فتركز على المرحلة الثانية والثالثة وهي تحويل الرموز الكتابية graphemes إلى فونيمات phonemes ثم إلى ألفونات allophones. وفي البداية لا بد من تحويل النص المكتوب إلى وحدات صوتية مستقلة - فونيمات. بعد ذلك تُطبق القوانين الفونولوجية phonological rules واحدا تلو الآخر متى ما توفرت متطلبات القانون. وقواعد هذه القوانين شائعة في مجال التحليل الصوتي للقواعد الصوتية في اللغة. وتحتوي على رموز متعارف عليها أيضا [٢]. وقوانين الفونولوجيا العربية ليست جديدة، فقد كتب عنها اللغويون العرب منذ القرن الثاني الهجري وما تبعه [٣، ٤]. وكتب عنها المعاصرون أيضا. إلا إنها إما أن تكون اختصارا لقليل منها [٥، ٦]، أو إنها ليست مجموعة بطريقة ميسرة للعاملين في مجال الحوسبة حتى يمكن الاستفادة منها مباشرة [٧-٩]. فهي مكتوبة على شكل نصوص تحتاج لوقت وجهد ومتخصص في اللسانيات لاستخلاص كل قانون منها لتقديمه للمتخصصين في البرمجة والحوسبة الذين غالبا ما يضطرون للجوء إلى عملية الاستثناءات والدخول في المستويات اللغوية الأخرى كالنحو والصرف والدلالة لتطبيق ما يجدونه في كتب التراث وما يشابهها من الكتب المعاصرة [١١، ١٠]. بينما يمكن الاعتماد المباشر على التسلسل الصوتي وما يعرف بالنظام والقواعد الفونولوجية. وهذه الورقة تقدم عرضا رياضيا لهذه الخصائص مستخلصة من الكتب اللغوية التي كُتبت عن العربية منذ القرن الثاني الهجري ومقدم لها أمثلة وتطبيقات بطريقة ميسرة ومتسلسلة. ولتيسير فهم القوانين الفونولوجية فإن الجدول رقم (٢) يفسر الرموز المستخدمة فيها.

الجدول رقم (٢). الرموز المستخدمة في القوانين الفونولوجية في هذه الورقة

الرمز	الدلالة
∅	لا شيء
←	ويعني يصبح أو يتحول إلى...
١×	بداية العبارة
×١	نهاية العبارة
⊗	في أي مكان ولكن ليس في بداية العبارة

تابع الجدول رقم (٢)

الرمز	الدلالة
١#	بداية الكلمة
#١	نهاية الكلمة
—	موضع الصوت أو الحرف الذي ينطبق عليه القانون
/	تستخدم للفصل بين عملية التحول التي تتم قبل هذا الرمز ومكان وجود الصوت قبل التحويل
C _n	سلسلة من الأصوات غير محددة العدد
C	استخدم هذا الرمز للدلالة على صامت consonant وسبب استخدامه لثلا يكون هناك ليس مع أصوات العربية
V	أي من الصوائت، القصير منها والطويل
v	صائت قصير short vowel (— ، — ، —)
v:	صائت طويل long vowel (— ، — ، —) ^(٣)
""	علامة الاقتباس، وتستخدمان للدلالة على أن الرمز الذي بينهما هو المستعمل في كتابة اللغة
()	يمكن لما هو داخل القوس أن يرد أو لا يرد
[]	كل صوت داخل القوس يقابله صوت آخر مماثل له في الترتيب داخل قوس أو أقواس آخر من نفس النوع في نفس القانون
{ }	ينطبق القانون على كل واحد من مكونات القوس منفردا بنفسه عن غيره من المكونات الأخرى داخل نفس القوس
ح	مُفَحَّم
س	مُرْسَل
ص	قصير
ع	مُسْتَعْل
غ	غُنَّة
ق	مُرَقَّق

(٣) رموز الصوائت المستخدمة هنا مقتبسة من [١].

بعد التعرف على الرموز المستخدمة في هذه الورقة نسرد القوانين الفونولوجية مرتبة حسب أولوية تطبيقها. هذا يعني أن القانون الأول يأتي قبل بقية القوانين وأن الثالث يأتي بعد الأول والثاني وقبل بقية القوانين الأخرى وهكذا.

$$١. \bar{C} \leftarrow CC^{(٤)}$$

$$٢. "ا" \leftarrow \emptyset / \otimes CC^{(٥)}$$

$$\leftarrow \text{ء} \text{ / } \text{—} \times \text{ ل}^{(٦)}$$

$$\leftarrow \text{ء} \text{ / } \text{—} \times \text{ CC}^{(٧)}$$

$$\leftarrow \text{ء} \text{ / } \text{—} \times \text{ ل} \text{ غير "ل"}^{(٨)}$$

(٤) هذا القانون يتعلق بتحويل الصوامت المشددة إلى صامتتين متماثلتين، كما في "تَبَّتْ" التي تُتَنَقَّضُ /ث — ب ب — ت /.

(٥) يبين هذا القانون أن الألف يصبح لا شيء (لا يُنطق) إذا كان متبوعاً بصامتتين متتاليتين وليس في بداية العبارة. مثل "بالمدرسة معلم" فإننا ننطقها /ب ل م — د ر — س — ت — م — ع — ل ل — م /، وكما في "فاكتب" في العبارة "إذا عدت إلى المنزل فاكتب درسك" فننطق /ف — ك — ت — ب /.

(٦) إذا كان "ا" جزء من "ال" التعريف في بداية العبارة فإننا ننطقه /ء — / مثل "القلم أمضى" فإننا ننطقها /ء — ل — ق — ل — م — م — م — ض — /.

(٧) إذا كان "ا" جزء من فعل، ثالث حرف فيه مضموم ضمة أصلية، وكان في بداية العبارة فإننا ننطقه /ء — / مثل "اركض، ادعوا، استهزئ" فإننا ننطقها /ء — ر — ك — ض /، /ء — د — ع — /، /ء — س — ت — ه — ز — /، على التوالي.

(٨) إذا كان "ا" لا ينطبق عليه القانونين السابقين (ليس جزء من ال التعريف أو من فعل ثالث حرف فيه مضموم) ويقع في بداية العبارة فإننا ننطقه /ء — / مثل "ابن، امرؤ، اسم" فإننا ننطقها /ء — ب — ن — ، /ء — م — ر — /، /ء — س — م /، على التوالي. يلاحظ أن هذا القانون والقانونين السابقين له يعتمدان على معرفة نوع الكلمة (اسم معرف بأل، فعل) وليس على التسلسل الفونولوجي. فلا يمكن وضع قانون لكلمة مثل "البس" التي من الصعب معرفة ما إذا كانت ال هنا جزء من الكلمة أم ال التعريف بمجرد النظر إلى التسلسل الفونولوجي. ومن ثم فإن القوانين الفونولوجية لوحدها لا تكفي في هذه القوانين الثلاثة فقط. وهذا بعكس بقية القوانين في هذه الورقة التي تعتمد كلياً على التسلسل الفونولوجي.

- ← م / في أي مكان آخر^(٩)
٣. "ى" ← م / #١^(١٠)
٤. "ة" ← هـ / ×١
- ت / فيما عدا ذلك^(١١)
٥. "ل" ← ∅ / CC^(١٢)
٦. "[أ، ء، ∅]" ← ∅ / ×١
- "[أ، ء، ∅]" ← ن / في أي مكان آخر^(١٣)
٧. "آ" ← ء / ^(١٤)

(٩) كما في "عالم" فإننا نطقها /ع ل م / أو /ع ل م /، يلاحظ هنا أن هناك حركتين بعد /ع/ الأولى الفتحة الموجودة على العين أصلاً والثانية ناتجة عن تحول الألف إلى فتحة أخرى في هذا القانون، فتدمجان لتصبحا صائت واحد طويل. ويمكن إضافة قانون تابع لهذا القانون يحول الصائتين المتجاورين المتمثلين إلى صائت واحد طويل: ٧٧ ← ٧٧ .

(١٠) تأتي "ى" في نهاية الكلمة ويسبقها دائماً فتحة قصيرة وتُتطقان فتحة طويلة كما في "رمى" /ر م /.

(١١) القانون هذا والقانون السابق له يخصان التاء المربوطة التي تُتطق دائماً تاء إلا إذا أتت في نهاية العبارة فإنها تتحول إلى هاء.

(١٢) هذا قانون حذف اللام الشمسية كما في كلمة "النور" التي عند تجزئتها كتابتها فإنها تظهر: "ل ن ن م ن" وور ، وعند تطبيق القانون تصبح كما تُتطق /ء ن ن ن /ر.

(١٣) هذا القانون وسابقه يقومان بحذف التتوين في نهاية العبارة ويحولانه إلى حركة ونون في أي مكان آخر كما في "صيراً" حيث تُتطق /ص ب ر ن / وفي الكلمة "علم" حيث تُتطق /ع ل م ن / وفي الكلمة "شيء" حيث تُتطق /ش ي ء ن /. هذا القانون يسبق أي قانون آخر تكون /ن / أحد مكوناته، فينطبق على التتوين ما ينطبق على /ن /. يلاحظ في هذا القانون أن التتوين والألف كلاهما يتحولان إلى نون داخل العبارة، أما الألف المتبوعة بتتوين فلا تحذف في نهاية العبارة وإنما يحذف التتوين فقط كما في "رُزقت علماً" فتتطق /ر ز ق ت ع ل م /.

(١٤) يتحول حرف المد "آ" بموجب هذا القانون إلى / ء م / كما في "آية" و "الأفة" حيث تتطقان: /ء م ي م هـ /، /ء ل ء ف هـ /، على التوالي.

٨. [أ، { ئ، إ، و }، و] ← [ءَ ، ءِ ، ءُ] ، [ءَ ، ءِ ، ءُ] (١٥)
٩. [و ، ي] ← [و ، ي] (١٦)
١٠. ∅ ← [و ، ي] / [و ، ي] ، [و ، ي] / [و ، ي] ، [و ، ي] ، [و ، ي] CC - C (١٧)
١١. ن ← م / م - { م ، ب } (١٨)
- ← [ر ، ل] / [ر ، ل] (١٩)
- ← [و ، ي] / [و ، ي] ، [و ، ي] (٢٠)

- (١٥) يقوم هذا القانون بفك الحروف التي تحمل صوت الهمزة إلى مكوناتها. فتصبح كما تنطق أي همزة متبوعة بحركة كما في "مَلَائِكَةٌ" إذ تتحول إلى /م ل ءَ ءِ ءُ / كَ ءَ ءُ / .
- (١٦) يتيح هذا القانون تحويل الواو المسبوقة بضمة والياء المسبوقة بكسرة إلى ضمة طويلة وكسرة طويلة، على التوالي، كما في "سوق"، "عيد" اللتان تنطقان /س ءُ ق/ ، /ع ءِ د/، على التوالي.
- (١٧) عند تواجد ثلاثة صوامت متتالية فيقوم هذا القانون بإدخال صائت قصير بعد الصامت الأول يكون هذا الصائت /ـ ، ءُ ، / إذا كان الصائت الذي قبل الصوامت الثلاثة /ـ ، ءُ ، /، على التوالي. كما في "مِنَ الْقَمَرِ"، "انْتَمَ الْقَادَةُ"، "مَنْ الْقَادِمِ" فإننا ننطقها على التوالي: /م ـ ن ـ ل ق / م ـ ر /، /ء ـ ن ت ـ م ـ ل ق / د ـ هـ /، /م ـ ن ـ ل ق / د ـ م / .
- (١٨) هذان قانونا الانقلاب والإدغام الكامل في الميم حيث تصبح النون ميمًا في كلتا الحالتين، كما في: "عنبر" حيث تُنطق /ع ـ م ب ـ ر/، و "من ماء" حيث تنطق /م ـ م م ـ ع/ .
- (١٩) هذا إدغام كامل تتحول فيه النون إلى الصوت الذي يليها كما في "من رأى" و "من له" حيث تُنطقان /م ـ ر ر ـ ع /، /م ـ ل ل ـ هـ /، على التوالي.
- (٢٠) هذا ما يُعرف بالإدغام بغنة حيث تبقى خاصية الغنة، أما موقع اللسان في الفم فإنه يكون في موقع الصوت الذي يليه كما في "من وراء" و "من يعمل" حيث تُنطقان /م ـ وْ و ـ ر ع /، /م ـ يْ يْ ع م ـ ل/، على التوالي، ولا يكون هذا الإدغام إلا بين الكلمات أما إذا تسلسلت هذه الحروف في كلمة واحدة فإنه لا يكون إدغاما كما في "دنيا" ، "عنون" فإنهما تنطقان /د ـ ن ي ـ ع /، /ع ـ ن و ـ ن/، على التوالي.

← [ق غ ، ك غ ، ج غ ، ش غ ، ض غ ، ص غ ، س غ ، ز غ ، ط غ ، د غ ، ت غ ،
ظ غ ، ذ غ ، ث غ ، ف غ] / — [ق ، ك ، ج ، ش ، ض ، ص ، س ، ز ، ط ،
د ، ت ، ظ ، ذ ، ث ، ف] (٢١)

١٢. " [ذ ، ت ، ت] ← [ظ ، ط ، د] / — [ظ ، ط ، د] (٢٢)

" د " ← ت / { C_n ١ # — C_n ١ # ، # ١ C_n # ، # ١ ق — — ١ # ١ # ت }

١٣. " ل " ← ر / — ر (٢٣)

ل ل ← ل غ ل غ / { ١ × — ، — (C) ، — (C) } — (اللام هنا في
لفظ الجلالة فقط) (٢٤)

(٢١) وهذا شبيه بالإدغام السابق وهو ما يُعرف عند اللغويين بالإخفاء، وبعد تدقيق النظر فيه وفي القانون السابق تبين أن الشبه بينهما كبير وهو أن يتحول النون إلى صوت أنفي مخرجه هو مخرج الحرف الذي يليه كما في "من جاء" فإننا ننتطقها /م — ج غ ج — ع/. وهذا القانون ينطبق على الأصوات المذكورة فيه سواء أتت بين الكلمات أو داخل الكلمة الواحدة كما في "منشار" التي تنطق /م — ش غ ش — ر/.

(٢٢) يُعرف هذا بالإدغام التام حيث يصبح الصوت مطابقاً تماماً للصوت الذي يليه كما في "إذ ظلموا" حيث تُنتطق /ء — ظ — ل — م — /، وكما في "قالت طائفة" فتنتطق /ق — ل — ط — ط — ع — ف — هـ/، وكما في "أجيبك دعوتكما" فتنتطق /ء — ج — ب — د — د — ع — ت — ك — م — /، على التوالي. وينطبق نفس القانون على "د" في الجزء الثاني من هذا القانون إلا أنه لا بد أن يكون هو والحرف الذي يليه داخل كلمة واحدة كما في "أرذت" تُنتطق /ء — ر — ت — ت — / وعندما يكونا في كلمتين فيُشترط في هذه الحالة أن تكون الكلمة السابقة "قد" كما في "قد تعلم" التي تُنتطق /ق — ت — ت — ع — ل — م/.

(٢٣) كما في "قل ربي" إذ تُنتطق /ق — ر — ر — ب — /.

(٢٤) هذا قانون الترخيم والترقيق للام. فاللام دائماً مرققة إلا تلك في لفظ الجلالة "الله" أو "اللهم" فإنها تكون مفخمة إذا جاء اللفظ في بداية العبارة أو كان مسبوقة بـ /—/ أو /—/، أما عدا ذلك فإن اللام تكون مرققة، كما في المثالين التاليين: "إن الله" "بالله"، حيث تُنتطقان: /ء — ن — ل غ ل غ — هـ/ ، /ب — ل غ ل غ — هـ/، على التوالي. يلاحظ أنه في هذا القانون فقط استخدم الرمز — للدلالة على لفظ الجلالة وليس على الحرف الذي هو موضوع القانون هنا وهو "ل".

ل ← ل ق / (فيما عدا ذلك، سواء كانت اللام في لفظ الجلالة أم في غيرها من الكلمات)

١٤. "ر" ← ر ق / - (ر) - ، - C (٢٥)

← ر ح / فيما عدا ذلك

١٥. " [ص ق ، ط] ← [ص ك ، ع ت] / - [ك ، ت] (٢٦)

١٦. V ← V ح / { ص ، ض ، ط ، ظ ، ق ، غ ، خ } - (٢٧)

١٧. v ← v ص / - { #١ ، CC } (٢٨)

١٨. [v :] ← [v] / CC - (٢٩)

١٩. v ← ∅ / - ×١ (٣٠)

(٢٥) هذا قانون الترقيق والتخفيف للراء، فالراء مرققة إذا كانت مكسورة أو كانت ساكنة مسبوقة بكسرة، ومفخمة فيما عدا ذلك، كما في المثالين "فرعون" و "مريم" حيث تنطقان /ف - ر ق ع - ون / ، /م - ر ح ي - م /، على التوالي. هذا القانون يحل إشكالية الراء المضعفة فيطبق مرتين كما في "الرياح" حيث ترقق الراء الثانية /ء - ر ر ح ي - اح / ، والمرة الثانية عندما ترقق الراء الأولى /ء - ر ر ق - ي اح /.

(٢٦) يُعرف هذا بالإدغام الناقص حيث يكون مخرج القاف هو مخرج الطاء ومخرج الكاف هو مخرج التاء، إلا أن الصائت السابق لهما يكون مُستعليا كما في "تَخَلَّقَكُمْ" حيث تُنطق /ن - خ ل - ك - ك - م / وفي "أحطت" حيث تُنطق /ء - ح - ع - ت /.

(٢٧) تصبح الصوائت مفخمة إذا أتت بعد الأصوات المطبقة /ط ظ ص ض / أو بعد أصوات الاستعلاء /غ خ ق / أو بعد الأصوات المفخمة مثل: /ل ح /، /ف ن /، /ب /، /ن س /، /ء ل ل ح ل ع ه /، /ط ع ر /، على التوالي.

(٢٨) يبين هذا القانون أن الصائت القصير يصبح أقصر في أمده نسبيا عندما يكون في نهاية كلمة متبوعة بكلمة أخرى أو عندما يكون متبوعا بصائتين متتاليتين، فالضمة في كلمة "كتب" /ك - ت - ب / أطول منها في تلك التي في الكلمات التالية: "علم بلادي" /ع - ل - م - ب - ل - د - /، "مقبِل" /م - ق - ب - ل /، الفقرة الأولى من القانون ذكرها سيبويه في كتابه الكتاب أما الثانية فهي معروفة عبر اللغات بما فيها العربية وقد تم قياسها في العربية أكوستيا [١٢].

(٢٩) تصبح الصوائت الطويلة قصيرة إذا أتت بعدها صائتين متتاليتين، لأن المقطع العربي لا يجيز هذا التسلسل الفونولوجي، كما في "رمى الرجل" فتتطق /ر - م - ر - ر - ج - ل /، وكما في الليل فتتطق /ب - ل - ل - ي - ل /، وهذا بعكس لو لم يكن في التسلسل صائتين متتاليتين، كما في "رمى عمر" و "قي بيتنا" حيث تنطق /ر - م - ع - م - ر /، /ف - ب - ي - ت - ن - /، على التوالي.

(٣٠) هذا القانون يتولى حذف الصوائت القصيرة في نهاية العبارة كما في: "يُجْتَهَدُ التَّلْمِذُ"، حيث تنطق /ي - ج - ت - ه - د - ت - ل - م - ذ /.

٢٠. "ء، ق، ك، ط، ت، د، ب" ← "ء، ق، ك، ط، ت، د، ب" / — ×١ (٣١)

كان هذا سردا لجميع القوانين الفونولوجية العربية مستقلة عن المستويات اللغوية الأخرى كالنحو والصرف والدلالة، وجميعها تعتمد كليا على تسلسل الأصوات عدا الألف - وليس الهمزة - التي تقع في بداية العبارة - وليس داخلها أو في نهايتها، ففي هذه الحالة فقط لا بد من معرفة حالتين هما: ما إذا كان الألف جزء من أَل التعريف أو أنه جزء من فعل ثالث حرف فيه متبوع بضمّة: ففي الأولى يصبح همزة وفتحة وفي الثانية يصبح همزة وضمّة. وإذا علمنا أن أصل الجملة في العربية هو الجملة الفعلية فهذا يعني أن وجود الأسماء في بداية الجملة ومن ثم المعرفة بأل هو قليل الوجود في النصوص العربية. وإذا وزدت فإن اللبس بينها وبين الأفعال التي تبدأ بألف قليل كذلك، ذلك لأن الأفعال التي تبدأ بألف غالبا لا يتبعها لام أما الأفعال التي تبدأ بألف ولام فقليلة العدد مثل: "العب"، "اليس"، "الطم"، "الكم"، "المز"، "الشم"، "الجم". ومن ثم فإنه يمكن في هذه الحالة - حالة الألف في بداية الجملة - عمل قائمة بالأفعال التي يبدأ جذرها بلام وتكون متعدية حتى يمكن تحويلها إلى أمر ومن ثم تبدأ بألف ولام، كما في الأمثلة السابقة وفصلها عن الألف الذي هو جزء من أَل التعريف. كما يمكن عمل قائمة بالكلمات التي تتحول فيها الألف إلى همزة وكسرة كما في: "ابن"، "امرؤ"، "اسم" وهي قليلة أيضا، للتفريق بينها وبين الألف التي تكون جزء من فعل كما في: "اكتب". وبهذا يمكن تجاوز هذه الحالة الاستثنائية الوحيدة

(٣١) يبين هذا القانون أن الأصوات الشديدة stops ينطلق بها عضو/عضوي النطق عند الوقوف عليها في نهاية العبارة فتكون مرسلة released، وهذا ما يعرف بالقلقلة عند الوقف على الأصوات الشديدة المجهورة voiced stops والنفخ aspiration عند الوقف على الأصوات الشديدة المهموسة voiceless stops. فيُقلقل /د/ في العبارة "مررت بأحمد"، ويتبع /ك/ نفخ في الجملة "قرأت كتابك". ويذهب بعض اللغويين إلى أن هذا القانون ينطبق على الأصوات المذكورة عندما تكون ساكنة - ليست متبوعة بصائت - حتى وإن كانت داخل العبارة أو داخل الكلمة إلا أن جمهور اللغويين يرون بتطبيقه فقط عند الوقف عليه (أي في نهاية العبارة).

للقواعد الفونولوجية باعتماد قائمة بالمفردات الشاذة وهي كما ذكرنا ليست طويلة ، أو بتوظيف محلل صرفي يقوم بتحديد نوع الألف والكلمة اللتين تقعان في بداية العبارة.

أمثلة تطبيقية

بعد سرد القوانين الفونولوجية لا بد من عرض أمثلة لكيفية تطبيقها عند تحويل النص المكتوب إلى رموز صوتية تحاكي العملية النطقية عند العربي بحيث يكون الناتج في نهاية المطاف هو نتاج تطبيق جميع القوانين الفونولوجية التي لها في العبارة المستهدفة مجال للتطبيق. ويلاحظ أنه لا بد أن تكون العبارة في الأصل تحمل كافة رموز التشكيل عدا السكون الذي ليس له دلالة صوتية كما سبق ذكره. وفيما يلي بعض الأمثلة ومراحل تطبيق القوانين الفونولوجية عليها. ويظهر خط تحت الرموز التي طُبِقَ عليها القانون وتغيرت لتوضيح موضع تطبيق القانون:

• مثال (١). "رَأَيْتُ مُعَلِّمًا فِي الْفَصْلِ"

١. الفونيمات مكتوبة كما هي إلا أنها مجزأة:

رَ أ ي ت م م ع ل ل م أ ف ي ا ل ف ص ل
/

٢. يطبق القانون الأول الخاص بفك التضعيف:

رَ أ ي ت م م ع ل ل ل م أ ف ي ا ل ف ص ل
ل

٣. يُطبق القانون الثاني الخاص بحذف الألف من أل التعريف:

رَ أ ي ت م م ع ل ل ل م أ ف ي ل ف ص ل
/

٤. يطبق القانون السادس الخاص بالتونين:

رَ أ ي ت م م ع ل ل ل م ن ف ي ل ف ص ل
ل

٥. يُطبق القانون الثامن الخاص بالهمزة:

ارَاءَ ي ت م ع ل ل م ن ف ي ل ف
ص ل /

٦. يُطبق القانون التاسع الخاص بالصوائت الطويلة:

ارَاءَ ي ت م ع ل ل م ن ف ل ف
ص ل /

٧. يُطبق القانون الحادي عشر الخاص بالنون:

ارَاءَ ي ت م ع ل ل م ن ف ي ل ف
ص ل /

٨. يُطبق القانون الثالث عشر الخاص بترقيق اللام:

ارَاءَ ي ت م ع ل ل م ن ف ي ل ف
ف ص ل /

٩. يُطبق القانون الرابع عشر الخاص بتفخيم الراء:

ارخَاءَ ي ت م ع ل ل م ن ف ي ل ف
ف ص ل /

١٠. يُطبق القانون السابع عشر الخاص بتقصير الصوائت القصيرة:

ارخَاءَ ي ت م ع ص ل ل م ن ف ي ل ف
ل ف ص ل /

١١. يُطبق القانون الثامن عشر الخاص بتحويل الصوائت الطويلة إلى قصيرة:

ارخَاءَ ي ت م ع ص ل ل م ن ف ي ل ف
ل ف ص ل /

١٢. يُطبق القانون التاسع عشر الخاص بحذف آخر صائت في العبارة:

ارِخَاءَ صَيِّتِ مُمُعَ صِلَلِ مَمَّصَ فِغْفِ
لِفَصَلِ

● مثال (٢) . تَتَسَبَّبُ الْغَازَاتُ الْمُنْبَعِثَةُ مِنَ الْمَصْنَعِ فِي تَلَوُّثِ الْهَوَاءِ.

١. الفونيمات مكتوبة كما هي إلا إنها مجزأة:

ا ت س ب ب ب ل غ ا ز ا ت ا م
ن ب ع ث ة م ن ا ل م ص ا ن ع
ف ي ت ل و و ث ل ه و و ا ا

٢. يطبق القانون الأول الخاص بفك التضعيف:

ا ت س ب ب ب ل غ ا ز ا ت ا م
ن ب ع ث ة م ن ا ل م ص ا ن ع
ف ي ت ل و و ث ل ه و و ا ا

٣. يطبق القانون الثاني الخاص بالألف:

ا ت س ب ب ب ل غ ا ز ا ت ا م
ن ب ع ث ة م ن ل م ص ا ن ع
ف ي ت ل و و ث ل ه و و ا ا

٤. يطبق القانون الرابع الخاص بالتاء المربوطة:

ا ت س ب ب ب ل غ ا ز ا ت ا م
ن ب ع ث ت م ن ل م ص ا ن ع
ف ي ت ل و و ث ل ه و و ا ا

٥. يطبق القانون التاسع الخاص بتحويل الصوائت الطويلة:

ا ت س ب ب ب ل غ ا ز ا ت ا م
ن ب ع ث ت م ن ل م ص ا ن ع
ف ي ت ل و و ث ل ه و و ا ا

٦. يطبق القانون العاشر الخاص بإدخال صائت في حالة وجود ثلاثة صوامت:

ا ت س ب ب ب ل غ ز ت ل م
 ن ب ع ث ت م ن ل م ص ن
 ع ف ت ل و و ث ل ه و و

٧. يطبق القانون الحادي عشر الخاص بالنون:

ا ت س ب ب ب ل غ ز ت ل م
 م ب ع ث ت م ن ل م ص ن
 ع ف ت ل و و ث ل ه و و

٨. يطبق القانون الثالث عشر الخاص بترقيق اللام:

ا ت س ب ب ب ل غ ز ت ل م
 م ب ع ث ت م ن ل م ص ن
 ع ف ت ل و و ث ل ه و و

٩. يطبق القانون السادس عشر الخاص بتفخيم الصوائت:

ا ت س ب ب ب ل غ ز ت ل م
 م ب ع ث ت م ن ل م ص ن
 ع ف ت ل و و ث ل ه و و

١٠. يطبق القانون السابع عشر الخاص بتقصير الصوائت القصيرة:

ا ت س ب ب ب ل غ ز ت
 م ل م ب ع ث ت م ن ل م ص ل م
 ص ل م ن ع ف ت ل و و ث
 ص ل ه و و

٥. يطبق القانون الخامس الخاص بحذف اللام الشمسية:

/تَ دَ فَ عَ كَ ثَ يَ رُ مَ نَ نَ نَ نَ نَ سَ
لَ لَ وُ صُ وُ لَ إَ لَ سَ سَ رَ رَ بَ /

٦. يطبق القانون السادس الخاص بالتنوين:

/تَ دَ فَ عَ كَ ثَ يَ رُ مَ نَ نَ نَ نَ نَ سَ
سَ لَ لَ وُ صُ وُ لَ إَ لَ سَ سَ سَ رَ
بَ /

٧. يطبق القانون الثامن الخاص بالهمزة:

/تَ دَ فَ عَ كَ ثَ يَ رُ مَ نَ نَ نَ نَ نَ سَ
سَ لَ لَ وُ صُ وُ لَ إَ لَ سَ سَ سَ رَ
بَ /

٨. يطبق القانون التاسع الخاص بالصوائت الطويلة:

/تَ دَ فَ عَ كَ ثَ يَ رُ مَ نَ نَ نَ نَ نَ Sَ
سَ لَ لَ وُ صُ وُ لَ إَ لَ Sَ Sَ Sَ Rَ
بَ /

٩. يطبق القانون العاشر الخاص بإدخال صائت قصير بين ثلاثة صوامت متتالية:

/تَ دَ فَ عَ كَ ثَ يَ رُ Mَ Nَ Nَ Nَ Nَ Sَ
سَ لَ Lَ وُ صُ وُ Lَ إَ Lَ Sَ Sَ Sَ Rَ
بَ /

١٠. يطبق القانون الحادي عشر الخاص بالنون:

/تَ دَ فَ عَ كَ ثَ يَ Rُ Mَ Mَ Nَ Nَ Nَ Nَ Sَ
سَ Lَ Lَ وُ صُ وُ Lَ إَ Lَ Sَ Sَ Sَ Rَ
بَ /

١١. يطبق القانون الثالث عشر الخاص بترقيق اللام:

ا ت د ف ع ك ث ر م م ن ن
 س ل ل و ل ص ل ل ء ل ل س س
 ر ب /

١٢. يطبق القانون الرابع عشر الخاص بتفخيم الراء:

ا ت د ف ع ك ث ر م م ن ن
 ن س ل ل و ل ص ل ل ء ل ل س س
 ر ب /

١٣. يطبق القانون السادس عشر الخاص بتفخيم الصوائت:

ا ت د ف ع ك ث ر م م ن ن
 ن س ل ل و ل ص ل ل ء ل ل س س
 س ر ب /

١٤. يطبق القانون السابع عشر الخاص بتقصير الصوائت القصيرة:

ا ت د ف ع ك ث ر م م ن
 ص ن ن س ل ل و ل ص ل ل ء ل ل س
 ل ل س س ر ب /

١٥. يطبق القانون الثامن عشر الخاص بتحويل الصوائت الطويلة إلى أخرى

قصيرة:

ا ت د ف ع ك ث ر م م ن
 ص ن ن س ل ل و ل ص ل ل ء ل ل س
 ل ل س س ر ب /

١٦. يطبق القانون التاسع عشر الخاص بحذف الصائت القصير في نهاية العبارة:

ا ت د ك ف ع ص ك ث ر ح م م ن
 ص ن ن س ل ل و ص ل خ ل ص ء
 ل ل س س ر ح ب

١٧. يطبق القانون العشرون الخاص بالأصوات المرسلة:

ا ت د ك ف ع ص ك ث ر ح م م ن
 ص ن ن س ل ل و ص ل خ ل ص ء
 ل ل س س ر ح ب

الخاتمة

بينت هذه الورقة كيف يمكن تحويل النص العربي المكتوب والمشكّل من رموز كتابية graphemes إلى رموز صوتية phonetic symbols بتطبيق القوانين الفونولوجية حتى يمكن توظيفها لتوليد كلام منطوق أيا كانت الطريقة المستخدمة في ذلك: التوليد بالتسلسل أو التوليد بالقانون أو التوليد بالأنموذج [١٠]، ويمكن كذلك توظيفها في نظم التعرف الآلي على الكلام العربي. ويلاحظ أنه في البداية لا بد من البحث في النص عن الكلمات الخارجة عن قواعد الكتابة في العربية والمسرودة في مقدمة هذه الورقة وتحويلها إلى رموز صوتية، ثم تُطبق بعد ذلك القوانين المذكورة في هذه الورقة مرتبة حسب أولويتها في التطبيق. بمعنى آخر أن نبدأ بالقانون الأول فالثاني... وهكذا. وجميع القوانين هنا مستقلة عن المستويات اللغوية الأخرى كالنحو والصرف والدلالة وتقوم كلياً على تسلسل الأصوات عدا الألف الواقع في بداية العبارة.

شكر وتقدير

يود الباحثون أن يتقدموا بالشكر والتقدير لمدينة الملك عبد العزيز للعلوم والتقنية

وذلك لدعمها للمشروع رقم: ١٨-١٢ والذي من نتائجه هذه الورقة

المراجع

- [١] الغامدي، منصور بن محمد. *الصوتيات العربية*. الرياض، المملكة العربية السعودية: مكتبة التوبة ١٤٢١هـ.
- [٢] Lass, Roger . *Phonology: an Introduction to Basic Concepts*., London: Cambridge University Press, 1984.
- [٣] سيويو، أبو بشر عمرو بن عثمان (ت ١٨٠هـ). *الكتاب* ، تحقيق عبد السلام محمد هارون. القاهرة : دار الكتاب العربي، ١٩٨٦م.
- [٤] ابن جني، أبو الفتح عثمان (ت ٣٩٢هـ). *سر صناعة الإعراب* ، تحقيق مصطفى السقا وآخرين . مصر : مكتبة مصطفى الخليلي، ١٣٧٤هـ .
- [٥] حركات، مصطفى. *الصوتيات والفونولوجيا* . بيروت : المكتبة العصرية ، ١٤١٨هـ .
- [٦] بدري، الدكتور كمال إبراهيم . *علم اللغة المبرمج* . الرياض : مطابع جامعة الملك سعود، ١٤٠٨هـ.
- [٧] الحمد، غانم قدوري. *الدراسات الصوتية عند علماء التجويد* . بغداد : مطبعة الخلود ، ١٤٠٦هـ.
- [٨] الغول، محمد بن شحاده. *بغية عباد الرحمن لتحقيق تجويد القرآن* . الدمام : دار بن القيم ، ١٤١٠هـ.
- [٩] Nassir, Abdulmunim abdulmir. "Sibawayh the Phonologist: A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in His Treatise Al Kitab." *unpublished Ph. D. thesis*, University of York, England, 1985.
- [١٠] Al-Muhtasib, Husni; Elshafei, Moustafa and Al-ghamdi, Mansour. "Techniques for High Quality Arabic Speech Synthesis." *The Third KFUP Workshop on Information & Computer Science*, (2000), 73-82.
- [١١] Fathi, H. A., and others. "A Real-Time Arabic Text-to-speech System Based on Parallel Processing." *J. King Saud University*, Vol. 7,(1995), 49-67.
- [١٢] Hussain, A. A. "An Experimental Investigation of Some Aspects of the Sound System of the Gulf Arabic Dialect with Special Reference to Duration." *unpublished Ph. D. thesis*, Essex University, England, 1985.

Phonetic Rules in Arabic Script

Mansour M. Al-ghamdi*, Husni Al-Muhtasib and Moustafa Elshafei**

*King Abdul Aziz City for Sciences and Technology, Riyadh
King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia*

(Received 8 January 2002; accepted for publication 17 November 2002)

Abstract. Conversion of Arabic script into phonetic rules is one of the major obstacles facing the researchers on Arabic text-to-speech systems and speech recognition. Although Arabic is one of the oldest languages that its sounds and phonological rules were extensively studied and documented (more than 12 centuries ago), these valuable studies need to be compiled from scattered literatures and formulated in a modern mathematical framework.

The objective of this paper is to present to the interested researchers a mathematical formulation for a comprehensive set of these rules.

