

Licensing Semi-syllables in Northern Yemeni Arabic: Evidence from Bedouin Dialects

Basheer Mufleh and Sami Alquhali

Abstract

In this paper, we investigate the licensing of semi-syllables in Northern Yemeni Arabic Dialects (henceforth, NYAD). According to Kiparsky's classification, there are three types of Arabic dialects; CV dialects, VC dialects and C-dialects. C dialects exist when semi-syllables (unsyllabified consonants, adjoined to the higher prosodic word) are licensed. In this paper, we present data from Northern Yemeni dialects which show that, of all the Yemeni dialects, only Bedouin dialects license semi-syllables and can be classified as C-dialects. The other dialects are CV dialects.

The essential argument for a C dialect is the licensing of –CCC- clusters. Some Yemeni dialects do allow –CCC- clusters and group with C-dialects in licensing semi-syllables. The reason is that the constraint *Reduce* is high ranked in VC and C-dialects, License- μ is lower-ranked and *Strict Layering Hypothesis* is violable in Arabic. The violation of Strict Layering Hypothesis is due to affiliating the unsyllabified mora with the lowest possible superordinate category which results in violating the constraints on *foot size* and then the possible affiliation would be the higher prosodic word which is not restricted to any size considering an affiliation of this type. The analyses here are formulated from an OT and PRAAT perspectives.

Keywords: Semi-syllable, C-dialects, Yemeni Arabic Phonology, Optimality Theory.

1. Introduction

Yemeni Arabic is one of the Arabic varieties spoken by around 25,000,000 million people in Yemen; a country in the southern part of the Arabic peninsula. Yemeni Arabic can be divided roughly into several main dialects, each with its own slightly distinctive vocabulary and phonology. The most well-known dialects are San'ani, Ta'izzi, Adeni, Tihami, Hadhrami, Ibbi, yaffi'ii, Thamari.

In the northern part of Yemen, which we are investigating here, dialects can be classified into Urban and Bedouin dialects. Urban dialects are the ones spoken in cities like Sana'a, Taiz, Ibb, and the main cities of governorates like Amran, AlMahweet and Hajjah. The dialects spoken in the outskirts of these governorates are Bedouin dialects. The dialects spoken in Aljawf governorate, the northern parts of Amran governorate, Saadah Governorate, and Tihamah (a large area of northern Yemen including Alhudaidah and Hajjah governorates) are also classified as Bedouin dialects (henceforth, Bed-NYAD). Both Urban and Bedouin dialects have the common five or six syllabification types, for example; six as in Taiz (Al-Samadi, 2011) and five as in Amrani, (Saif Bareq, 2017). In Kiparsky's (2003) and Watson's (2007) classification, the Yemeni dialects (al-Hudaida, San'ani, Yariimi, Ibbi), Egypt (Cairene, Middle Egyptian dialects), Saudi Arabia (Meccan) are CV dialects.

In this paper, we claim that Bed-NYAD differ from those NYAD in the syllabification types and semi-syllable licensing in particular. They are grouped as C dialects, not CV dialects as previously classified, an argument that is discussed in the paragraphs to follow. Such difference is really interesting and absolutely worth investigating. With the aim of approaching some interesting findings, the researchers here are going to investigate syllable types, syllabification patterns and semi-syllable licensing in NYAD and Bed-NYAD.

2. Data and Methodology

2.1.Data

As mentioned above, –CCC- clusters are proposed to be the main criterion for determining the type of a dialect. Therefore, words where –CCC—clusters are possible are recorded by speakers from NYAD & Bed-NYAD. These words are shown in the table below. For the NYAD, the second author of this paper is a native speaker of Amrani dialect, which is one of the NYAD. He also helps in testing the data we found in previous works on NYAD like Ta’izzi dialect-Samadi (2011) and Sana’ani dialect-Watson (2007). For the data collected for Bed-NYAD, the first author of this paper belongs to one of the Bed-NYAD spoken in the north of Amran. Speakers of other Bed-NYAD like Aljawf, Mareb and Saadah were recorded as well. The device used for recording is Samsung cellphone. No high-tech noise cancelling device is needed because the purpose of recording is simply seeing whether a vowel is present or deleted in the second syllable. This is clearly achieved using the smartphone device.

Table 1: words selected for analysis

NYAD, CV dialect	Bed-NYAD,C-dialect	Meaning
Gul.ta.lih	gul.t.lih	said to him
Yik.ru.mo	yik.r.mo	they give generously
Sir.ta.lih	sir.t.lih	went to him
Saa.hi.bi	saa.h.bi	my friend
ya.la.tak	yaa.l.tak	your aunt
Sa.laa.ma.tak	sa.laa.m.tak	Your safety

2.2. Optimality Theory (OT)

Optimal theory (OT) is one of the recent theories and most powerful methodology in phonology that came to replace the derivational rule-based theory. It is a constraint-based theory which was initiated by Prince and Smolensky (1993). The core of the optimal theory is that the correct surface or outputs of the phonological inputs are determined by the interaction of universal, violable constraints. These correct outputs are called Optimals.

According to Prince and Smolensky (2008), There are three components in OT: GEN (generator), EVAL (evaluator), and CON (constraints). The function of GEN is to generate a universal set of candidates for a given input, one of these candidates is selected as the optimal form which is determined by how well it satisfies the constraints in CON which is the universal set of violable constraints present in the grammars of all languages. The final OT component EVAL evaluates in parallel the set of output candidates according to how well they satisfy the constraint hierarchy and determines which output form is optimal.

John J. McCarthy & Alan S. Prince (1995) discuss two types of constraints: faithfulness and markedness constraints. Faithfulness constraints demand identity between two strings (such as an input and an output), while markedness constraints favor structurally unmarked forms at the expense of

modifying the input. The optimal or actually occurring output forms can violate lower-ranked constraints if such violation secures satisfaction of higher-ranked constraints, which its competitors violate. The winning candidate, then, is the most harmonic form that best satisfies the high-ranked constraints. For more mutual understanding and further details (see Archangeli & Langendoen: 1997. Kager: 1999, and McCarthy: 2002).

2.3. PRAAT

Besides the OT analysis, we run a PRAAT analysis for two sounds; one belongs to the NYAD and the other belongs to Bed-NYAD. The word selected for the PRAAT analysis is Sir.ta.lih (went to him). The presence of these three syllables indicate the absence of semi-syllables but the deletion of the vowel in the second syllable results in a –CCC- cluster (-rtl- sir,ta.lah). Due to constraints of foot size, the (t) cannot join the previous or the following syllable and is forced to join the higher prosodic word resulting in a semi-syllable. The purpose of this analysis is showing practically how CV and C dialects work and using PRAAT to present semi-syllables in a more explicit way. PRAAT is downloaded at www.praat.org . The version used in this paper is PRAAT 6.0.36. See section 9 for the PRAAT analysis.

3. Semi-syllables in Prior Studies

3.1. Kiparsky's Account on Semi-syllable

Kiparsky (2003) has conducted a study on Arabic dialects and classified them into three groups CV, VC and C-dialects. C-dialects include the dialects in North Africa (Morocco and Tunisia), CV dialect exists in Egypt and the Libyan Dessert (with little reference to North Yemen dialect), and VC dialect which includes Syria, Lebanon, Palestine and Iraq.

According to Kiparsky (2003), the characteristics of the three dialects (CV, VC, C) include:

- a. Phrase-final clusters occur unrestrictedly only in CV and C- dialects. They can be broken up by an epenthetic vowel, under conditions that vary according to style and dialect. E.g. / katab-t/ katabit, katabt (CV-dialects), ktabt (C- dialects) _ I wrote _ VC-dialects either permit no -CC clusters (kalib) or permit them only with falling sonority (kalb, but katabit) .Therefore, such –CC clusters that violate sonority sequencing in the coda occur only in CV- and C-dialects.
- b. Phrase-initial onset CC-clusters are allowed in VC- and C-dialects which accordingly allow the pan-Arabic process that deletes high vowels in open syllables to reduce even initial CiC- to CC- .The resulting clusters are often broken up by a prothetic vowel (which in turn is phonetically preceded by a glottal stop, in satisfaction of Arabic’s undominated ranking of Onset) e.g. ?islaah, _ plowshare _ ?iklaab _ dogs _ (silaah, kilaab in CV-dialects) .
- c. Medial –CCC- clusters are broken up as –CCiC in VC- dialects, and as –CCiC- in CV dialects. E.g. Iraqi gilitla, Cairane ?ultilu, Moroccan qaltlu I/you (M) said to him. In other words, Medial –CCC- clusters can be parsed in VC and C-dialects by making the middle consonant a semi-syllable, e.g. /gil.t.la/ -- (word level) (gil).t̩.la – (postlexical) (gi.li)t̩.la in VC dialects, but /ʔul.ta.lu/-- /ʔul.ti.lu/ in CV-dialects.
- d. – Metathesis II of medial CCiC- to –CiCC occurs only in VC-dialects, e.g./yi.ktib.u/ yikitbu they write.CV dialects always retain –CCiC- (yiktibu), and C-dialects simply drop the vowel in the corresponding cases (yiktbu)
- e. High vowel deletion occurs after geminates only in VC- and C-dialects, e.g. /y-kallim-u/ (y)ikal(l)mu _ they talk to someone.All CV dialects retain the vowel, e.g. yikallimu.

Kiparsky argues that the three groups of Arabic dialects mainly differ in licensing-semi-syllables: “A semi-syllable arises where a constraint **License-μ** which requires all moras to be licensed by syllables, is outranked by markedness constraint on the form of syllables and feet”. The direct consequence of the low ranking of the Constraint License-μ is the violability of the *Strict Layering Hypothesis* which requires that every non highest prosodic or metrical element to be in its entirety a constituent of an element belonging to the next higher category on the prosodic hierarchy.

3.2.Watson’s Account

Watson (2007), after studying Kiparsky’s classification, has found other groups which do not confirm to the three sets provided by Kiparsky. She refers to the dialects which combine characteristics of two groups as Cv dialects, with a lower case v. Watson has studied many

Yemeni dialects like (Sana‘ni, al-Hudaida, Yariimi, Yaafii). The finding was that these dialects form an intermediate class, falling between CV and C-dialects. She proposes a fourth typology which is Cv with a lower case v. In Cv dialects, semi-syllables are not permitted at either lexical or post lexical level. Mora sharing is permitted if the syllable rhyme contains a long segment.

Watson (2007) argues that — “a closer look at some of the data and consideration of new data shows that the analysis cannot cope with all syllabification phenomena for all dialects. Dialects that fail to exhibit predicted phenomena most consistently are those in which derived CCC clusters are typically syllabified as CCVC—Kiparsky’s CV dialects. As a result, she extended Kiparsky’s three-way-typology to a four-way-typology, adding Cv with a lower case v, for those dialects that share characteristics of both CV and C- dialects” (p.336). We are not going to argue more about Cv for it is not relevant to this study so we will refer to them as CV, the only purpose for highlighting this study is to prove that a closer look at new data in Northern Yemeni dialects shows that Kiparsky’s analysis can’t cope with all syllabification phenomena for all dialects, especially the Bedouin ones. The data below show that Bedouin dialects are C dialects as they allow semi-syllables.

4. A Rule-Based Account for Semi-syllable in NYAD

NYAD enjoy the following types of syllables: light syllable, heavy syllable and super heavy syllable. Samadi (2011):

Table 2: Types of syllables as set out in Samadi (2011)

Syllable	Light	Heavy	Superheavy
1-	CV : ji: (come)	CVV : maa (water)	CVVC : baaʔ (sold)
2-		CVC : Shil (take)	CVCC : gult (I said)
3-			CVCCC: ma.kuntsh. (I was not)

5. Characteristics of NYAD (CV Dialects)

1. NYAD do not allow phrase initial onset CC- clusters. There is always an epenthetic vowel breaking this cluster. This cluster is, however, permitted in both VC and C- dialects.

Table 3: Initial onset clusters

VC and C-dialect	CV
ʔihmar	Himar
ʔiktaab	Kitab
ʔiklaab	Kilab

2. Some dialects like Bedouin Amrani allow –cc clusters in the coda position (a pattern shared by C and CV dialects) where they are allowed unrestrictedly, They occur in some words and are broken by a vowel in other forms.

Katabt (I wrote) Thakart (I remembered)

Katabit (she wrote) Thakarāt (she remembered)

- c. In CV dialects, the –ccc- clusters are allowed only in negation which takes place word- finally permitting no semi-syllable. Interestingly, in C-dialects, -CCC- cluster does occur in the middle resulting in an unsyllabified consonant adjoined to the prosodic word.

6. Bed-NYAD as a C Dialect

We have argued above that Bed-NYAD is a C dialect. In other words, it allows semi-syllables both at the word level and at the post lexical level. Obviously, most of the above-mentioned data indicates that Bed-NYAD is a C-dialect rather than a CV dialect .However, In almost all the positions where semi-syllable is allowed, there are data that do not confirm to this, they group with CV structures in banning semi-syllables through vowel epenthesis.

The occurrence of this unrestricted licensing of semi-syllable is a result of other dialect-specific phonological processes and rules. For example, medial initial cluster –CCC- is not possible when the first consonant is an affricate.

Yik.t.bu (they write)
Yig.t.lu (they kill)

yiš.ta.ru (they buy)
yidʒ.ma.ʕu (they collect)

Figure (1)

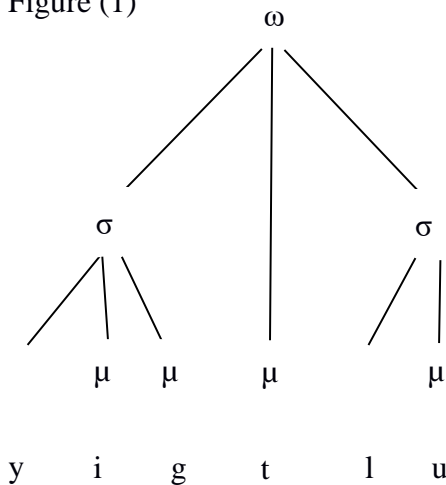
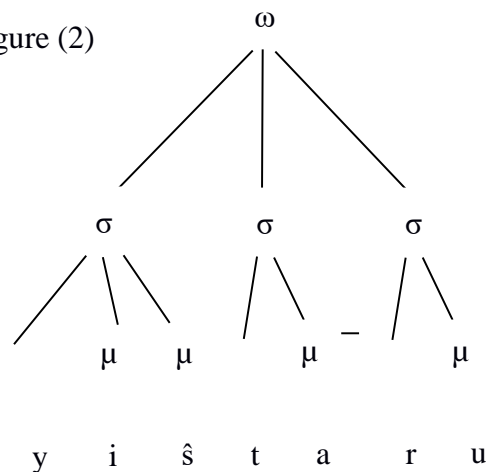


Figure (2)



7. Across-Dialectal Look at Semi-syllables

It has been pointed out that the main criterion for determining whether a dialect is CV, VC, or C-dialect is the treatment of morphologically and phonologically –CCC- clusters. Regarding –CCC- clusters, in CV dialects a vowel is epenthesized to the right of the unsyllabified consonant, as in Sana'ni /gult-lah/ /gul.ta.lah/, in VC dialects the vowel epenthesis takes place to the left of the unsyllabified consonant, as in Iraqi /gilt-la/ /gil-it-la/, in C-dialects no vowel epenthesis takes place.

Table 4: Semi-syllables across dialects

Sana'ni (CV)	Cairane (CV)	Iraqi (VC)	Morrocان (C)	Bed-NYAD ©
Gultalah	gultulu	Gilitla	Gultlu	Gultlih

The following syllabification trees would show the difference between the Sana'ni dialect as CV dialect, Iraqi as VC and Moroccan and Bed-NYAD as C-dialects.

Figure (3): Sana'ani

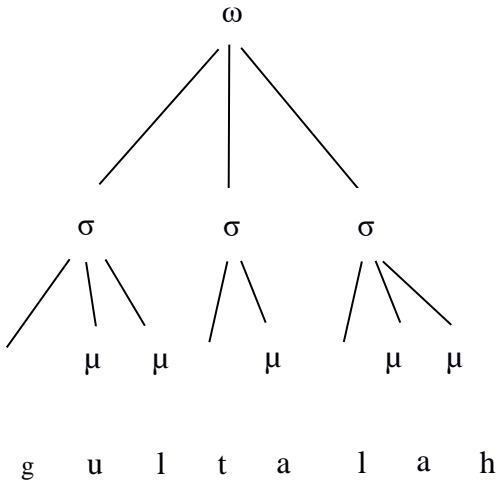


Figure (4): Iraqi

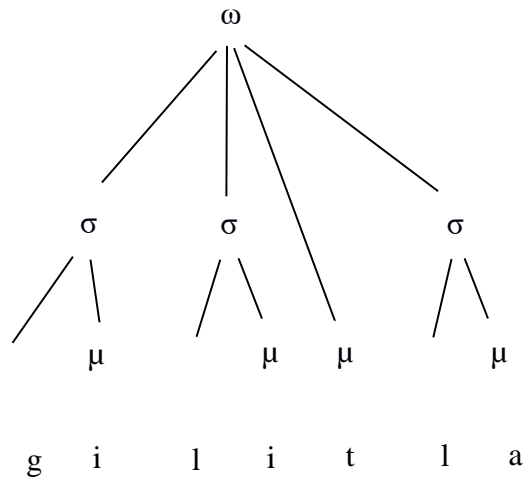


Figure (5): Amrani

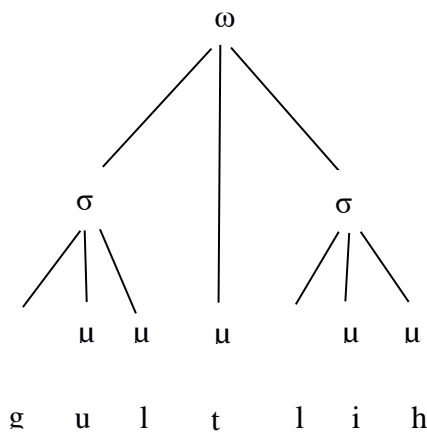
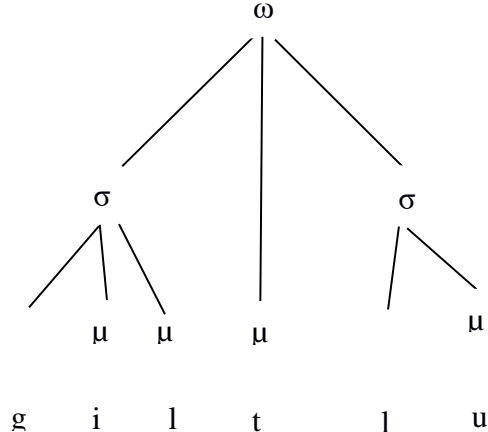


Figure (6): Moroccan



8. An OT account on semi-syllable in Bed-NYAD

8.1.Active Constraints

8.1.1. License- μ

This constraint requires every mora to be licensed by a syllable. Adjunction of a mora to a higher node other than a syllable is a violation of this constraint. For example, in a word like *saa.h μ .bi*, *h* is unsyllabified consonant which is not licensed by a syllable. This unsyllabified consonant is adjoined to the higher prosodic word which violates Licence- μ constraint.

8.1.2. Reduce

This constraint is motivated to delete the moras (high vowels) of the light non-final (CV) syllable. For example, in the word /*fihimna*/. Reduce requires the deletion of the short high vowel following *f* resulting in /*fhimna*/, syllabified as / *f.him.na*/ with *f* as unsyllabified consonant.

8.1.3. Max- μ

This constraint dictates that moras in the input and output to be the same. A deletion of a mora in the output is a violation of this constraint. Since onset represents no mora, a deletion of a mora in the rhyme violates Max- μ .

8.1.4. Dep- μ

This constraint does not allow an epenthesis of a mora. Moras in the input must be the same as moras in the output and the epenthesis of mora is a violation of this constraint.

8.2.Analysis

The following tableau shows the ranking of these constraints in non-Bedouin NYAD. (CV dialects)

Table 5: Tableau of ‘*sahibi*’ in NYAD, CV dialects

CV-dialect	License- μ	Max- μ	Reduce
Input: { <i>saa</i> }. { <i>hi.bi</i> }				
1a. <i>saa.h.bi</i>	*!	*	**	
2a. <i>sah.bi</i>		**	**!	
3a. ► <i>saa.hi.bi</i>			***	

According to the tableau above, candidate (3a) is the optimal because of the high ranking of Licence- μ , no unsyllabified consonant takes place. Since license- μ is high ranked, all moras are licensed by a syllable and no mora is adjoined to the higher prosodic word. It is the high rank of this constraint that does not allow CV dialects to have unsyllabified consonants.

On the contrary, the following tableau shows that Bed-NYAD has the characteristics of C dialects since it gives low ranking for the constraint License- μ .

Table 6: Tableau of ‘sahibi’ in Bed-NYAD, C dialects

C-dialects				
Input: { saa }. { hi.bi }	Reduce	Max- μ	License- μ
1a. ► saa.h.bi	**	*	*	
2a. saah.bi	**	**!		
3a. saa.hi.bi	***!			

As we can see in the tableau, Reduce is high ranked. Reduce requires, as we discussed above, the raising of short low vowels or the deletion of short high vowels. Since Reduce is high ranked, deletion of short vowel takes place resulting in an unsyllabified consonant which can’t be licensed by the syllable. Rather it is adjoined to the highest prosodic word. License- μ which requires every mora to be licensed by a syllable is low-ranked in C-dialects so Semi-syllables take place.

9. PRAAT Analysis of Semi-syllable

The above section shows how semi-syllables behave in Bed-NYAD and how the concepts of CV and C dialects work. It highlights how the deletion of vowels result in a C dialect with the consonant attached to the higher prosodic word due to the inability of attaching it to the syllable. Here, we provide a PRAAT analysis to show practically the deletion of vowels in C dialects leaving the consonant hanging alone. The word we choose for the PRAAT analysis is ‘sir.ta.lih’ (went to him) in two dialects; the Amrani NYAD and Aljawf Bed-NYAD. The former is a CV dialect whereas the latter is a C dialect.

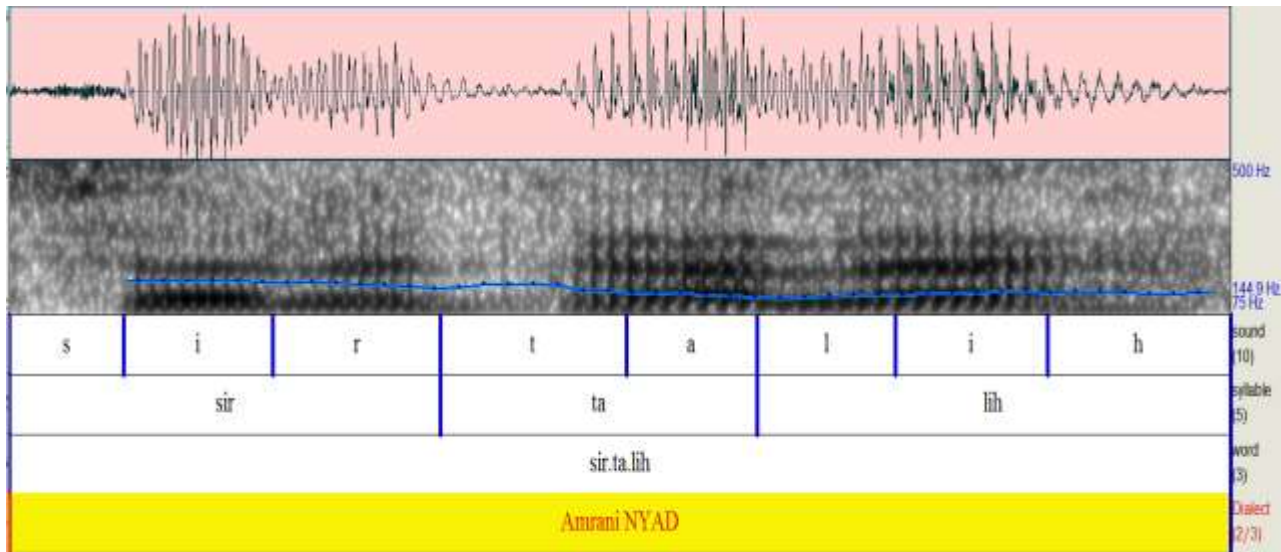


Image 1: Amrani, NYAD, CV Dialect

As the picture above shows, there are three syllables: /sir/, /ta/ and /lih/. There is no vowel deletion in this CV dialect and no medial –CCC- cluster is permitted. All syllables have a nucleus and are licensed as independent syllables. This is the typical syllable structure in all NYAD CV dialects. They tolerate no vowel deletion.

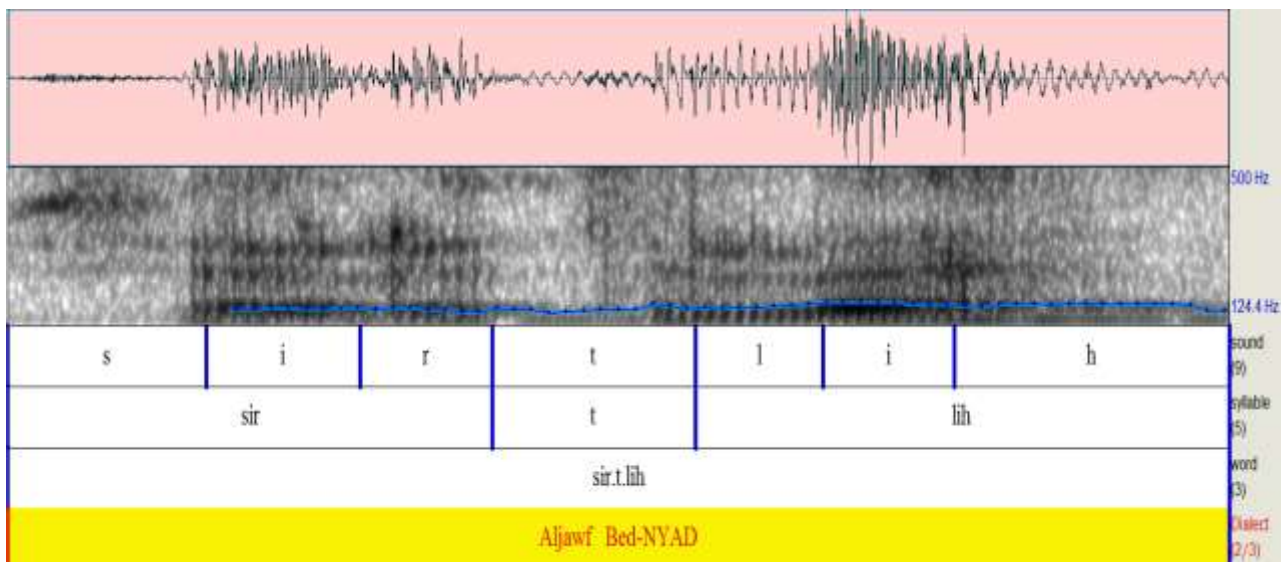


Image 2: Aljawf, Bed-NYAD, C Dialect

In this image, however, the case is different. As you can see the second syllable consists of the /t/ only. The deletion of vowel results in a medial –CCC- cluster. The remaining part of the syllable can join neither the previous nor the following syllable as it violates the foot size constraints. /t/ cannot be part of the syllable because it lacks the nucleus which is the essential part of the syllable. Therefore, it is attached

to the higher prosodic word yielding in what we call a semi-syllable. Semi-syllables like this occur in all the Bed-NYAD.

10. Conclusion

We have argued above that there exist CV and C-dialects in north Yemeni Arabic Dialects. Bedouin dialects are C-dialects which are distinct from many other Yemeni dialects which follow the CV pattern. Both the OT account and the PRAAT analyses show that there exist semi-syllables in Bed-NYAD. The data provided above show that Bed-NYAD allow unsyllabified consonants (moras unaffiliated to syllables, they are adjoined to the higher prosodic word). As –CCC- clusters are proposed to be the main criterion for determining the type of dialect, allowing –CCC- clusters mark either CV or C-dialects and disallowing them indicate CV-dialects, Bed-NYAD do allow – CCC- clusters and group with C-dialects in licensing semi-syllables. The reason for licensing semi-syllables is that the constraint **Reduce** is high ranked in VC and C-dialects. Reduce which requires the deletion of short high vowels creates semi-syllables after deleting the vowel and leaving the consonant unsyllabified. On the contrary, License- μ is lower-ranked and Strict *Layering Hypothesis* is violable in Arabic so semi-syllables exist.

References

- Almashaqba, Bassil Mohammad. (2015). The phonology and morphology of Wadi Ramm Arabic. Diss. University of Salford, 2015.
- Bareq, Saif Mohammed, and Vivek R. Mirgane. (2017) "Syllable Structure of Amrani Yemeni Arabic." *Imperial Journal of Interdisciplinary Research*, 3, 2.
- Caroline, Fery and Ruben, Van De Vijver (2003). *Syllable in Optimality Theory*, University of Potsdam, Germany, Cambridge University Press.
- Damom, M. (2010). *Syllabification and Vowel Deletion in Tihami Yemeni Arabic*, EFL- University.
- Gordon, Raymond G. (2005). *Ethnologue: Languages of the world*. Ed. Barbara F. Grimes. Vol. 15. Dallas, TX: SIL International.
- Kiparsky, Paul (2003). *Syllables and Moras in Arabic*. Stanford University.
- McCarthy, John J., and Alan Prince. (1995). "Faithfulness and reduplicative identity." 249.
- Watson, Janet C. E. (2007). *Syllabification patterns in Arabic dialects: long segments and mora sharing*, University of Salford: Cambridge University Press.
- Prince, Alan, and Paul Smolensky. (2008). *Optimality Theory: Constraint interaction in generative grammar*. John Wiley & Sons.
- Samadi, Sameer (2011). *Syllable Structure of Taizi Yemeni Dialect of Arabic*, EFL University.

Selkirk, 1982] Selkirk, Elisabeth O. 1982. The syllable. Vol. 2. Dordrecht: Foris, pp. 337–385.

Basheer Mufleh (corresponding author)
Research Scholar
Department of Linguistics
University of Kerala
India
bmufleh@gmail.com

Sami Alquhali
Research Scholar
Department of Linguistics
University of Mysore
India
samisofan81@gmail.com