Cognate Words in Mehri and Hadhrami Arabic

Hassan Obeid Alfadly*　　　　**Khaled Awadh Bin Mukhashin****

Abstract

The lexicon is one important source of information to establish genealogical relations between languages. This paper is an attempt to describe the lexical similarities between Mehri and Hadhrami Arabic and to show the extent of relatedness between them, a very little explored and described topic. The researchers are native speakers of Hadhrami Arabic and they paid many field visits to the area where Mehri is spoken. They used the Swadesh list to elicit their data from more than 20 Mehri informants and from Johnston's (1987) dictionary "The Mehri Lexicon and English-Mehri Word-list". The researchers employed lexicostatistical techniques to analyse their data and they found out that Mehri and Hadhrmi Arabic have so many cognate words. This finding confirms Watson (2011) claims that Arabic may not have replaced all the ancient languages in the South-Western Arabian Peninsula and that dialects of Arabic in this area including Hadhrami Arabic are tinged, to a greater or lesser degree, with substrate features of the Pre-Islamic Ancient and Modern South Arabian languages.

Introduction:

Historically speaking, the Semitic language family from which both of Arabic and Mehri descend belong to a larger family of languages called Afro-Asiatic or Hamito-Semitic that includes Semitic, Egyptian, Cushitic, Omotic, Berber and Chadic (Rubin, 2010).

The Semitic language family is divided into two main branches (West Semitic and East Semitic).The West Semitic branch is divided into three branches including Central Semitic, Ethiopian and Modern south Arabian languages (henceforth MSAL). Though Arabic and Mehri belong to the West Semitic, Arabic descends from the Central Semitic and Mehri from (MSAL) which consists of two branches; the Western branch that includes Mehri, Harsusi Bathari and the Eastern branch which includes Jibbali and Soqotri. Hobyot is a mixture of both branches (ibid).
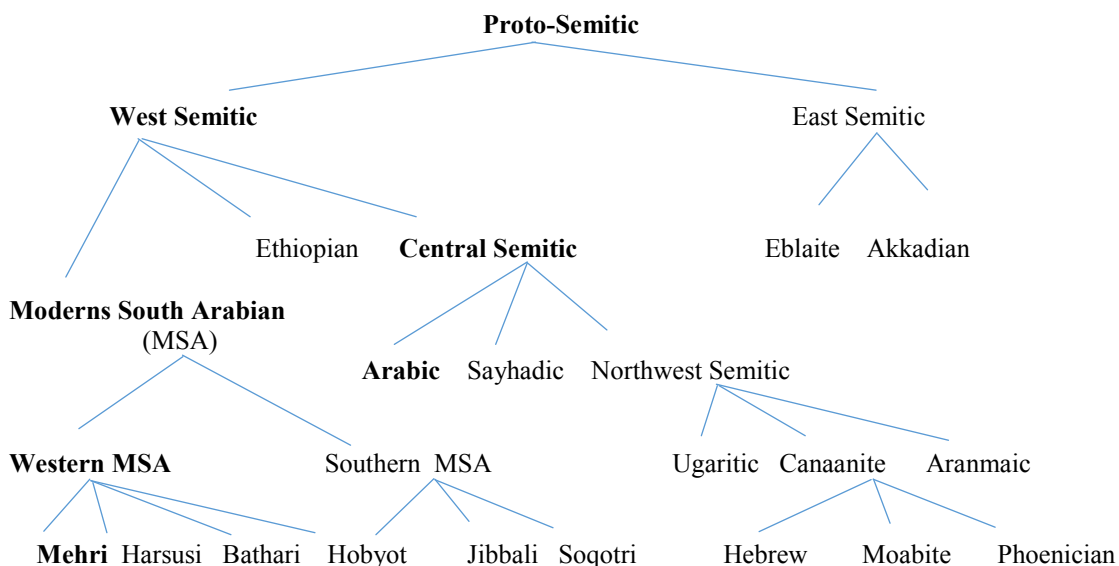

Figure No.1: Semitic Language Family

* Professor at Department of English,College of Arts,
Hadhramout University.
** Associate Professor at Department of English College
of Arts, Hadhramout University.

The Mehri Language:

According to Alrowsa (2014), the word Mehri as a term covers two things: the name of the language itself and the name of the tribe speaking that language. Mehri constitutes the most widely spoken language among all of its Pre-Islamic MSAL sisters, with a population that exceeds 135,764 as indicated in Yemen 2004 census and (SIL, 2000). Simeone-Senelle (2010), Alrowsa (2014) and Al-Qumairi (2015) state that approximately more than 100,000 people speak Mehri in Southeast Yemen, the western part of Dhofar in Oman, and the southern part of Saudi Arabia. The Mehri natives live along the southern borders of the Arabian Peninsula in an isolated area located between the eastern part of Yemen and the western mountains of Oman, Mehri extends from the Arabian Sea in the south to Thamud, on the border of *AlRub' al-Khali*, 'the Empty Quarter' (Alrowsa, 2014) in the north. The western boarder of the Mehri area is Alreideh, Alshehr and Almukala in Hadhramout Governorate. Mehris in Yemen are semi-nomads who breed camels and goats. The sedentary population is settled on the coast. Here, the activities are sea-oriented: trading, shipping, fishing. Some people are employees in public services in the main coastal villages (Ibid).
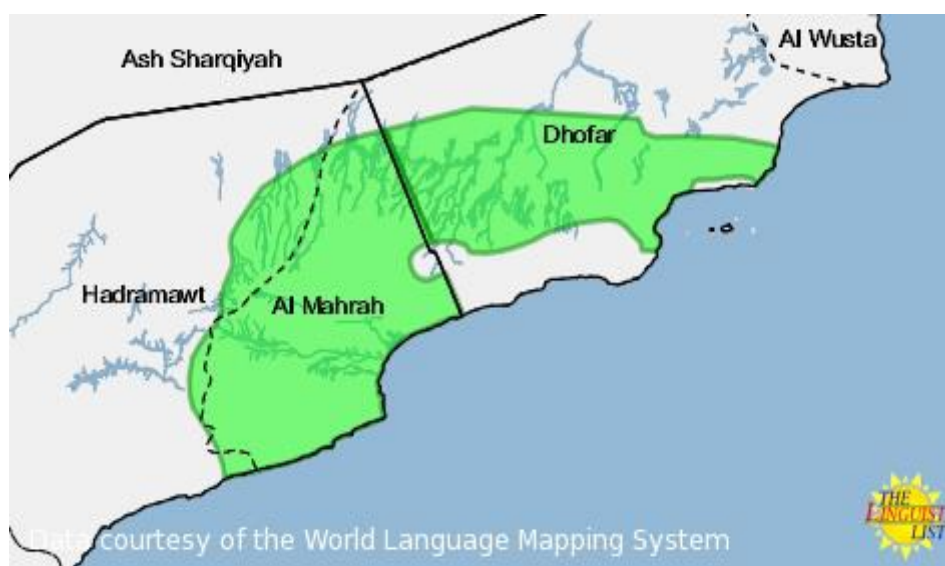

Figure 2. Hadhramout & Mehra locations in Yemen Taken from Alrowsa (2014)

The Hadhrami Arabic:

Hadhrami Arabic (henceforth HA), is a variety of Arabic spoken by Hadhrami people (*Ḥaḍārima*) living in Hadhramaut east of Yemen. It is also spoken by many emigrants, who migrated from Hadhramaut to the Horn of Africa (Somalia and Eritrea), East Africa (Comoros, Zanzibar, Kenya, Tanzania, and Mozambique), Southeast Asia (Indonesia, Malaysia, Brunei and Singapore) and, recently, to the other Arab states of the Persian Gulf (Al-Saqqaf,2006). Hadhramout dialectology consists of two main dialects: Coastal Hadhrami Arabic and valley Hadhrami Arabic dialect (Ibid). This study concentrates on comparing the HA, which is spoken along the Arabian Sea coast from Mukalla up to the border of Seihoot in Mehra. This dialect shares a lot of lexical items with MSAL including Mehri.

The Statement of the Problem:

Arabic dialects in Yemen are spoken as a mother tongue in almost all parts of the country, except in the eastern province of Mahra and on the island of Soqotra, where the MSALs (Mehri and Soqotri) are the native languages of the inhabitants**.** Watson (2011) states that it is traditionally assumed that after the spread of Islam in Arabia, Arabic overwhelmed the original ancient language of the peninsula leaving the language situation in the south-western Arabian Peninsula as one in which dialects of Arabic are tinged, to a greater or lesser degree, with substrate features of the ancient South Arabian languages.

So far no study has been conducted to investigate these traces of the ancient languages in the Arabic dialects especially those spoken in

a geographically-adjacent area to Mehra such as Hadhramout. This study, which aims at investigating the lexical similarities between HA and Mehri, is hoped to shed light on these traces and to bridge the gap in the literature.

Objectives of the Study:
This paper attempts to achieve the following objectives:
1- to find out the lexical similarities between Mehri and HA.
2- to shed light on the extent of relatedness between Mehri and HA.

Literature Review:
Theoretical Literature:
The possibility of applying mathematical and statistical techniques to language study has attracted the imagination of many linguists from the past in the 1950s and 1960s. The quantitative linguistic theory was developed in the work of Morris Swadesh (1952) who believed that statistical data could be used to provide accurate subgroups among related languages. He counted the cognate lexical forms shared by specific languages to determine to what extent the languages were related calling his subgrouping technique "lexicostatistics.

Lexicostatistics has been widely used to detect hypothetical genetic relations among languages (McMahon and McMahon, 2005). Lexicostatistics refers to the statistical manipulation of lexical materials for historical inferences. It compares languages for phylogenetic affinity based on proportion of cognates in a standard basic vocabulary list. A set of word forms from the compared languages is assembled based on a list of basi*c vocabulary*. In practice, linguists usually conduct basic word assembly based on small-scale meaning lists such as the two widely-adopted lists called the Swadesh lists. They compile 100 (Swadesh, 1955) or 200 (Swadesh, 1952). Lexical cognates are identified based on recurrent sound correspondences. Cognates and recurrent sound correspondences give strong evidence of a common origin of languages. These recurrent sound correspondences usually occur in vocabulary of languages that have phylogenetic relations (or systematically borrowed words in languages that have a history of deep contact; Hoijer, 1956; Bergsland and Vogt, 1962).

Though the application of lexicostatistical techniques, linguists managed to quantify the rate and extent of lexical change in languages over time and then could determine the genetic relatedness of the words. The lexicostatistical methods helped the linguist to organize language subgroups around percentage values of relatedness (e.g., languages with a determined relatedness of 81-100% are "dialects of a language," languages with a determined relatedness of 36-81 % are "languages of a family," and so forth. The higher the value, the more closely the languages are related.

Table 1. Linguistic Subgrouping Based on Shared Cognate Percentage (SCP) (Shosted, 2000:14)

Level of subgrouping	SCP in core vocabulary
Dialects of a language	81-100%
Languages of a family	36-81%
Families of a stock	12-36%
Stocks of a micro phylum	4-12%
Microphyla of a meso phylum	1-4%
Mesophyla of a macrophylum	0-1%

Regarding the data used to determine this relatedness among languages, Swadesh (1955) states that there is a distinction between "core" and "peripheral" vocabularies. Crowley (1997) states that core vocabulary consists of forms like "tooth" and "tongue" which are common to the human experience and therefore show up almost universally in the world's languages. However, Peripheral vocabulary forms, like "antelope" and" cilantro" are not found in every language.

These peripheral forms are directly related to the physical environment and material culture of each unique linguistic community.

Previous studies:

Johnstone (1987) published the first dictionary of Mehri. This contribution is regarded as a landmark in the study of Mehri (Alrowsa, 2014). Nakano (1986) collects 2000 words from Mehri, Jibbali and Soqotri and arranged them according to certain semantic fields. So this book is a kind of dictionary of Soqotri, Mehri, Jibbali and English in which the meaning of each Mehri, Jibbali and Soqotri word is given in English. Some of these words are used in sentences.

The most important studies on Hadhrami Arabic are those conducted in the previous century by European scholars, such as Landberg (1901) and Hein & Müller (1909). Landberg'(1901) as cited in Al-Saqqaf (2006 b) provided detailed descriptive linguistic study with reference to the Hadhrami folklore. Hein Müller's work is parallel texts in Mehri, Hadramī and German. Al-Saqaaf (2006 a) conducted a study investigating the linguistics of loan words in Hadhrami Arabic (HA). He found out that HA includes loan words borrowed from the languages to which Hadhramis immigrated such as Swahili, Malay and Urdu.

Method:

The study adopts the previously described approach known as lexicostatistics. The term lexicostatistics is used here in its narrow sense, to apply to the statistical manipulation of lexical similarities for making inferences related to the genetic relationship between HA and Mehri. The researchers who are native speakers of HA conducted field visits to the area where Mehri is spoken. They used the Swadesh list in addition to Altoma's (1969) list to elicit their data from more than 20 Mehri informants and from Johnston's (1987) dictionary "The Mehri Lexicon and English-Mehri Word-list". After that, they specified which Mehri and Hadhrami Arabic words are cognate and which are not and counted the percentage of these cognates.

Findings and discussion:

A good number of shared items are found in both HA and Mehri language within Swadesh list and in the other lists. It has been found that HA and Mehri have an SCP of 30% in the Swadesh list. This means that they are of a common "stock" as shown in Table 1 above. The results show that these cognates in HA and Mehri are of three groups depending on their pronunciation and semantic content. The first group includes a large number of cognates that are phonetically and semantically identical and are not found in Standard Arabic.

Table 2: Cognates with Almost the Same phonetic Shape and Meaning Exclusively Used in Mehri and HA

Mehri	HA	Gloss
/ bəχs /	/bəχs/	pain
/ dəh /	/ dəh /	this
/ ʃərħ /	/ ʃərħ/	dance
/ ħəs /	/ ħəs /	mind
/mənħær/	/mənħær/	chest
/ χɔːr /	/ χɔːr /	valley mouth
/ bʌtʼ/	/ bʌtʼ/	split
/ mæʕsəf/	/ mæʕsəf /	a kind of fish
/ dəfɒr /	/ dfər /	to push
/ɹirɔːb/	/ɹræb/	skin case
//ħɒrəm/	/ ħərəm /	path

The high percentage of such cognates in HA and Mehri shows that HA has preserved so many words from the Pre-Islamic Old South Arabian Languages. a group of four closely related extinct languages used to spoken in the far southern portion of the Arabian Peninsula particularly from Haḍramitic, which was spoken in Hadhramout where HA is spoken now. It also proves that HA has preserved cognate words from MSALs to which Mehri belongs. The location of HA helped it to keep these old cognate words. It is spoken in the coast of

Hadramout, which is on one hand adjacent to the location of the Pre-Islamic MSALs, and on the other hand it is naturally isolated and protected by high mountains and arid plains that separate it from the countryside of Hadhramout (the valley) which was vulnerable to Arabic since the appearance of Islam. This confirms Watson's (2011) claim that some Yemeni Arabic dialects are tinged, to a greater or lesser degree, with substrate features of the ancient South Arabian languages. It also supports her view that Arabic may not have replaced all the ancient languages of the Peninsula, and it provides new evidence that supports her expectation that we may be witnessing the rediscovery of descendants of the ancient languages.

The second group includes Mehri and HA cognate words that have the same consonantal roots and the same meaning but they have slightly different phonetic shapes. This slight difference in their pronunciation is partially attributed to the female gender suffixes / -t / in Mehri and / -h / in HA, for example the word "woman" in Mehri is/ /ħərmeit/ whereas in HA it is / ħərməh/ . In some cognate words such as in /ʔnbɔːtħ / and /ʔnbtˊəh "lay", the difference is caused by the different verb conjugations of Mehri and HA. Another reason for these difference is attributed to the fact that Mehri has lost the voiced pharyngeal consonant / ʕ / and it has been replaced by the laryngeal glottal consonant / ʔ / (Johnstone, 1975). However, HA is still preserving both consonants, consequently, most of the HA cognate words that begin with / ʕ / have Mehri equivalent cognates that begin with / ʔ / for example the HA /ʕdʒem/ 'to be dumb" has the Mehri equivalent /ʔtədʒɒm/.The following table includes more examples.

Table 3: Mehri and HA Cognate Words with the Same Root, Meaning but Slight Different Phonetic Shape

Mehri	HA	Gloss
/ħərmeit/	/ ħərməh /	women
/məzħ-æt/	/məzħ-æh/	'jock'
/ ɹeħleit/	/ɹeħleh /	jug made of clay
/rəħmeit/	/rəħməh/	rain
/mχjəţˊ/	/mχɔːţˊ/	needle
/slɔːb/	/sləb/	personal weapons
/ ʔəri /	/ ʕərri /	cat
/χəlɔːwək /	/χləg/	cloth
/χɔːʃəʕ/	/χʃəʕ/	press
/ʔnbɔːtəħ/	/ʔnbtˊəħ/	lay

The results also show a third group of cognate words which are found in Standard Arabic, HA and Mehri. The roots and meanings of these cognate words are alike but they are phonetically different. This difference lies in the fact that both HA and Mehri allow consonant clusters while Standard Arabic does not. The following example illustrates this difference. The word "devil" in Standard Arabic is pronounced / ʔibIiːs / but it is /bIiːs/ in both Mehri and HA. In some cases the Standard Arabic cognate differ even semantically from those in HA and Mehri, for example the cognate /frχ/ in HA and Mehri means" illegal son or bastard" and its equivalent in Standard Arabic is /fərəχ/ and it means "the young of a bird".

Table 4: Cognate Words in Mehri, HA and Standard Arabic

Mehri	HA	Standard Arabic	Gloss
bIi:s	bIi:s	ʔibIi:s	devil
/χerf/	/χerf/	/χəri:f/	autumn
/frχ/	/frχ/	/fərəχ/	illegal son/ young bird
/ ʕʤə m/	/ ʕʤə m/	/ ʕəʤə m/	dumb / in Arabic non-Arabs
/ʕeigr/	/ʕger/	/ʕəger/	to be cowardly / in Arabic to kill an animal
/ft´n /	/ ft´n /	/ fət´in /	remember / in Arabic to know
/gədr/	/gədr/	/gedir/	pot
/ ʕɒgI /	/ ʕɒgI /	/ʕəgəI /	acquire one's full Intelligence

Conclusion:

Mehri and HA have so many Cognate words. These cognate words are of three types. The first type includes cognate words that do not exist in Standard Arabic at all. They have the same phonetic shape and meaning and are exclusively used in Mehri and HA. The second type of these Mehri and HA cognate words includes cognate words that have the same roots and meanings but they are slightly different in their phonetic shape. These slight phonetic and orthographic differences may be attributed to the different female suffixes and verb conjugations in Mehri and HA. Another reason behind these differences is the nonexistence of the voiced pharyngeal consonant / ʕ / in Mehri. The third type of cognate words in Mehri and HA are words found in Standard Arabic, Mehri and HA. However, those cognate words found in Arabic do not have the same phonetic shape and meaning of their equivalent cognate words in Mehri and HA.

The large number of cognate words in Mehri and HA reveals a genetic relation between them. It also shows clearly that HA still preserve remnants of the MSA lexicon due to the fact that it is spoken in an area near the area of MSALs. The geographical distance and natural barriers isolated HA and protected it from the influence of the dominant Arabic. Further synchronic research studies are needed to discover more linguistic aspects shared by MSA languages and HA.

References:

1- Al-Qumairi, Saeed Saad Najadan (2015). "A minimalist analysis of the animal coding system in Mehri language within probegoal matching approach". International Academic Journal of Social Science 1(2): 177–191.

2- Alrowsa, W.A, (2014). Question Formation in Mehri. Unpublished thesis, University of Florida, USA.

3- AL-Saqqaf. A.H.(2006a). "The Linguistics of Loanwords in Hadramı Arabic", in The International Journal of Bilingual Education and Bilingualism Vol. 9, No. 1 pp.75-93.

4- AL-Saqqaf. A.H.(2006b). "Co-referential devices in Hadrami Arabic". In Zeitschrift für Arabische Linguistik, No. 46 (2006), pp. 75-93

5- Altoma, S.(1969). The Problem of Diglossia in Arabic. USA Harvard University Press.

6- Bergsland K., Vogt H. (1962). "On the Validity of Glottochronology". Curr. Anthrop. 3, 115–153. 10.1086/200264 [CrossRef]

7- Hein, W and Müller, H. (1909). Mehri und Hadrami Texte gesammelt im Jahre 1902 in Gischen, Wien: Alfred Holder.

8- Hoijer H. (1956)."Lexicostatistics: a critique". Language 32, 49–60. 10.2307/410652

9- Johnstone, T.M. (1975). "The Modern South Arabian Languages". Afroasiatic Linguistics 1/5:93-121 [1-29

10- Johnstone, T.M. (1987). "Mehri Lexicon and English-Mehri Word –List". London: School of Oriental and African Studies.

11- McMahon A., McMahon R. (2005). Language Classification by Numbers. Oxford: Oxford University Press.

12- Marie-Claude Simeone-Senelle (2010). Mehri and Hobyot Spoken in Oman and in Yemen. La communication en elle-m^eme a ete presentee sur support power-point lors du Symposium Oman-Yeme.

13- Nakano, A., (1986). Comparative vocabulary of southern Arabic Mehri, Jibbali and
Soqotri. Tokyo: Tokyo Publishing & Printing Co.

14- Rubin, Aaron (2010). "Mehri Language of Oman". Studies in Semitic Languages and Linguistics, Volume 58 :. Leiden, NLD: Brill.

15- Shosted, Ryan (2000) "Romani Roots: A Lexicostatistical Analysis Of Romani, Hindustani, and Czech," Deseret Language and Linguistic Society Symposium: Vol. 26: Iss. 1, Article 3.

16- Swadesh M. (1952). Lexicostatistic Dating of Prehistoric Ethnic Contacts. Proc. Am. Philos. Soc. 96, 452–463.

17- Swadesh M. (1955). Towards Greater Accuracy in Lexicostatistic Dating. Int. J. Am. Linguist. 21, 121–137. 10.1086/464321 [CrossRef]

18- Watson, J C.E. (2011). South Arabian and Yemeni Dialects in Salford Working Papers in Linguistics and Applied Linguistics, Ed. Gerry Howley, volume 1.

19- Wardaugh, Ronald (1993). An Introduction to Sociolinguistics. Oxford: Blackwell.

الكلمات ذات الاصل الواحد في اللغة المهرية ولهجة حضرموت العربية

حسن عبيد الغضلي   خالد عوض بن مخاشن

الملخص

يعد المعجم (المفردات) أحد أهم المصادر التي يعتمد عليها في دراسة التقارب بين اللغات وقد عُنِيَتْ هذه الدراسة ببحث التشابه المعجمي بين اللغة المهرية ولهجة حضرموت ؛ العربية لمعرفة مدى العلاقة والتقارب بين تلك اللغة وهذه اللهجة. وقد اختير هذا الموضوع لقلة مثل هذه الدراسات ، وربما لعدمه في هذا المجال. قام الباحثان بزيارات ميدانية لمحافظة المهرة لجمع معلومات البحث واختارا عشرين شخصاً من تلك المحافظة كي يستقيا معلوماتهما منهم: كما استفاد الباحثان من المعجم المهري الذي نشره الباحث جونستون عام 1987م. جمع الباحثان قائمة من الكلمات الأساسية المهرية وفقا لقائمة كلمات سوادش الأساسية التي أعدها عالم اللغة سوادش لدراسة التقارب بين اللغات ، كما استخدما تكنيكاً معجمياً إحصائياً لكشف إمكانية ومدى ذلك التقارب المعجمي. أكدت نتائج هذه الدراسة أن هناك تقارباً وتشابهاً معجمياً بين اللغة المهرية ولهجة حضرموت العربية ، وهذا يؤكد ما ذهب إليه الباحث واتسن (2011) الذي أوضح أن اللغة العربية لم تستطع أن تمحو تماما بقايا اللغات القديمة من جنوب غرب الجزيرة العربية وأن اللهجات العربية في تلك المنطقة والتي تشمل لهجة حضرموت العربية تحمل بقايا من آثار تلك اللغات العتيقة التي سادت في تلك المناطق قبل ظهور الإسلام بعصور