Bonny Sands[†], *Andrew Harvey*[‡], *Maarten Mous*[§], *Mauro Tosco*[††]
[†]Northern Arizona University; bonny.sands@nau.edu
[‡]University of Bayreuth; andrew.harvey@uni-bayreuth.de
[§]Centre for Linguistics, Leiden University; m.mous@hum.leidenuniv.nl
[††]University of Turin; mauro.tosco@unito.it

Why Hadza is (probably) not Afroasiatic: a discussion of Militarev's "Hadza as Afrasian?"

Problems with the lexical evidence used to posit Hadza as an Afroasiatic language are discussed. The failure to identify certain loanwords and the misanalysis of certain Hadza morphemes are problems rectified by having greater familiarity with Hadza and its linguistic contact history. Other problems are more general to the specific methodology employed. The overly wide semantic range of meanings often allowed in establishing form-meaning sets increases the likelihood of chance resemblances. The use of certain words that are likely onomatopoeic also reduces the impact of the proposed cognate sets. Ultimately, it is the lack of regular, repeated sound correspondences between Hadza and Afroasiatic that makes the proposal of their familial relationship unconvincing.

Keywords: language isolates; Hadza language; Khoisan languages; Cushitic languages; Afroasiatic languages; Chadic languages; long-distance relationship.

1. Introduction

There have been three main linguistic genetic classifications of Hadza that have been proposed: 1) linguistic isolate, 2) Khoisan language, 3) Afroasiatic language. There is currently a mainstream acceptance among Africanists in labeling Hadza as an isolate (Sands 1998, Güldeman 2014, Starostin 2012). Honken (2013: 20) states that "Hadza is currently regarded as an isolate by nearly everyone", though the Khoisan classification for Hadza continues to be used by some researchers (e.g. Chebanne 2017, Elderkin 2014, Ehret 2013a). Suggestions that Hadza has affinities to Afroasiatic have been made over time and include: Tucker (1966, 1967), Elderkin (1982), Starostin (2008) and now Militarev (this volume).

In this paper, we review the evidence proposed by Militarev (this volume) but in the interest of space do not include a critique of the similarities noted by Tucker (1966, 1967) and Elderkin (1982). We argue that the arguments put forth by Militarev are not sufficiently convincing proof of a linguistic genetic relationship between Hadza and Afroasiatic. We fail to see evidence of regular, repeated sound correspondences; nor do we see a convincing number of grammatical morphemes whose similarities suggest relatedness.

We provide examples of alternative analyses of many of the comparison sets, drawing upon our extensive first-hand experience with Hadza (Sands, Harvey) and Cushitic languages (Tosco, Mous, Harvey) as well as our knowledge of Bantu and Nilotic languages. First, our knowledge of Hadza allows us to show that certain Hadza terms are incompatible with the Proto Afroasiatic (AA) forms that Militarev has connected them to. In many cases, we provide alternative citation forms and in some cases we disagree with his morphological analysis. Note that our transcriptions fail to mark tone in many cases though tone is contrastive in the language[1].

[1] Jeremy Coburn (Indiana University) is currently researching tone in Hadza and we (Harvey & Sands) hope to collaborate with him in producing a fully tone-marked Hadza dictionary (cf. Coburn et al. forthcoming). As

Second, our knowledge of languages surrounding Hadza often allows us to identify loan sources that provide better etymologies than the AA forms suggested.

Our critique also discusses the sound correspondences and semantic correspondences proffered by Militarev, arguing that these do not meet the standard of evidence necessary to conclude that Hadza *must* be an Afroasiatic language. Problems of morphological analysis, semantic laxity and sound correspondences cannot be evaluated entirely separately because these factors all interact. In many instances, only a portion of a Hadza root, or only a portion of a reconstructed AA root are directly comparable yet no analysis is proffered to explain the differences in the roots. We analyze many of these cases as simply being due to chance resemblance. Furthermore, we think that Militarev has underestimated the role of contact in contributing to sets of similar words.

Given these methodological challenges of chance resemblances it is vital to restrict ourselves in how to go about establishing cognate sets. One sound restriction would be to compare only to solid reconstructed families. Militarev uses the impressive and extensive Afroasiatic reconstructions in an openly accessible database (available at the Tower of Babel website: https://starlingdb.org/). Reconstructions by others at various levels of families and subfamilies are not used nor cited in the article. In this contribution, we cannot assess the reconstructions from his database and we can only work under the assumption that the cited reconstructions are solid. Restricting the cognate sets to only reconstructed forms in (sub)families would reduce the danger of chance resemblance. The amount of cognate sets would be reduced although presumably it would be possible to propose new (sub)family reconstructions for some of the sets.

2. Problems with the Data

2.1. Citation Forms

A multilateral comparison involving hundreds of languages necessitates the use of sources that the author does not have first-hand experience with. This is not to say that such comparisons should not be done, rather, care should be taken to use the most recent and authoritative sources on each language; sources cited should withstand scrutiny from specialists on those individual languages.

For Hadza, Militarev uses a draft dictionary compiled by Kirk Miller (Miller et al. 2021); we have this same draft dictionary but it has a 2013 date and has not been revised to 2021. Miller et al.'s orthographic system has not been explained by Militarev, though he does present phonetic transcriptions in some cases. Miller et al. represent ejectives with symbols for voiced consonants (e.g. *pedla* 'white' (*petl'a-* in Sands 2012) (more properly cited as **/pecɻ̓a-/**); *jjowa* ['tʃoa] 'gecko') which might be distracting for readers.

Forms in this paper are modified from Miller et al. (2013) draft dictionary. We provide phonemic transcriptions based on the system used in Coburn, et al. (2023). In some cases, translations are provided based on our own Hadza research. We also refer to Peterson (2013) because his plant and animal identifications have been made in consultation with biologists.

Just as the Hadza data cited leaves something to be desired, so too, the Afroasiatic sources cited by Militarev are not always the most recent or authoritative. For a large part, the data on individual languages are from old sources. For instance, Militarev primarily sources Somali

tone does not enter into the discussion of Afroasiatic comparanda, we will not discuss it further here, apart from noting that it would need to be accounted for in any historical reconstruction.

data from Reinisch (1902) rather than use Zorc & Osman (1993) & *Dizionario Somalo-Italiano*. For Beja data, Reinisch (1895) is used rather than Wedekind et al. (2007).

It is not always possible for us to evaluate the data because the original sources (and therefore, the phonological interpretations) are not always cited. For instance, a language name Darasa is cited, without reference to the original source. This is an older, derogatory name for the language that has been consistently referred to as Gedeo (or Gede'o) since the 1970s. We consider Wedekind's Gedeo Dictionary (2008) and/or Hudson's Gedeo-English Index (1989: 229–265) of his *Highland East Cushitic dictionary* to be more authoritative sources for this language.

More generally, it is evident that there are degrees of reliability using primary sources. In particular, one needs to be hesitant to use words collected from languages that were no longer comfortably spoken and for which the speakers had already shifted to another language. This is the situation for the Cushitic languages Aasáx, Qwadza, Yaaku. Unfortunately, adding to lack of reliability, the earliest collections for these languages were done by non-linguists.

2.2. Alternative Morphological Analyses

There are a number of cases where words are cited without sufficient morphological analysis. We present some cases where we disagree with Militarev on the identification of the morpheme(s) in question.

#62 'not' *ʼukuwa* (more properly cited as /ʔuk^{h}u-/ 'to not exist') ~ AA **(ʔa-)k^{w}ay-* 'not'

Militarev glosses this form as 'not', whereas Miller (2013: 37–38) lists the forms [ʔuk^{h}u] as 'to not exist', and [ʔuk^{h}uu̯a] as 'to lack' or 'to not have'. Not only are the semantics of Hadza *ʼukuwa* different from the characterisation given in Militarev, but *ʼukuwa* is morphologically complex, formed of /ʔuk^{h}u/, plus *=a*, a form of the copula. As such, the form to be used in comparison would be /ʔuk^{h}u/, which seems rather more different to the forms offered by Militarev here (Proto-Afroasiatic **(ʔa-)k^{w}ay-* 'not', and the **kway* forms of Proto-Chadic and Proto-Omotic.

It should also be added that the Iraqw suffix offered by Militarev here (written in his work as *-Vkä*, where the diaeresis on the *a* is unexplained), probably refers to the verbal negative suffix *-Vká*. This can in turn be internally reconstructed in Iraqw to the lexical verb *kaáhh* [ká:ħ] 'to be absent', and itself cannot be given as cognate to the Proto-Cushitic reconstructed prefix **ka-* offered here.

#96 'what?' *akwiʔa* and #98 'who?' *akwaza* (more properly cited as /ʔaku-/ 'which?')

Militarev presents the forms *akwiʼa* 'what?' and *akwaza* 'who' as morphologically simple, when, in fact, both are morphologically complex. The form *akwiʼa* [ʔak^{w}iʔa] is formed of the interrogative *hadza aku* [ɦadza ʔaku] 'person', and *ihia* ['ʔiɦiʔa] 'thing', (cf. Miller 2013: 8) and could just as easily be translated as 'which thing?'. The form *akwaza* [ʔak^{w}adʑa] is formed of the same interrogative *aku*, and *haza* 'person', and could just as easily be translated as 'which person?'. As such, the form to be used in both comparisons would be *aku*.

This does not seem to challenge the forms offered as cognates (e.g. Proto Afroasiatic **k^{w}ay-* 'what? who?'), but does weaken the overall argument – after all, the cognacy here rests on one consonant, [k].

#34 'good' *muta-na, mta-na* (more properly cited as /mtana ~ mutana/ 'fine, good')

While Miller (2013:177) presents this form as *mut=ana*[2], it is unclear as to why Militarev presents this form as multimorphemic *muta-na* in this way. At any rate, the form *mutana* (also

[2] The diacritic after the stop indicates that it has been confirmed as being unaspirated.

commonly heard as *mtana*) provides no evidence that it is (or was at any point in Hadza) a multimorphemic form. We analyze this root as a Bantu loan, discussed below in Section 3.

2.3. Evidence from Internal Reconstruction

It appears to be the case that some body part words in Hadza feature a fossilized form of the possessive enclitic =*kwa* (the 1st person singular possessive in contemporary Hadza).

This may possibly be the motivation for the presentation of #**56 ‘mouth’** ***awani-ka*** **[ʔawani-ka]**, where the *-ka* is parsed from the rest of the word and disregarded in the offered cognates (e.g. the Proto-Afroasiatic reconstruction **ʕawin-* ‘tongue; (part of) mouth’). This word is more properly cited as /ʔawanika-ko/. If *ka* is a separate morpheme, it would probably derive from enclitic =*kwa*.

Where this fossilized =*kwa* seems to have been taken into account for ‘mouth’, Militarev seems to have missed it in both #31 ‘foot, leg’ *aphukwa* [ʔap^{h}uk^{w}a] or *uphukwa* [ʔup^{h}uk^{w}a], as well as in #37 ‘hand’ *ukhwa-* [ʔuk^{wh}a-] (and ‘fingers’ with the plural suffix *-bii* [biʔi], *ukhwabii* [ʔuk^{wh}a-biʔi]). For ‘foot, leg’, the form to be used in comparisons should probably be something like /ʔap^{h}u/ or /ʔup^{h}u/, which differs rather greatly from the cognates offered. For ‘hand’ (and ‘fingers’), the form to be used in comparisons should probably be something like /ʔuɦu/ – again, entirely different from the cognates offered (e.g. the reconstructed Proto-Afroasiatic **ḳawiʕ-* ‘claw, fingernail, hand’).

2.4. Imprecise Translations

‘rhino’ ***losho*** (more properly cited as /loːʃo/ ‘sighting/hunting name for the rhinoceros’ (when it is stooped and ready to charge)) ~ AA **warŝ-* ‘rhinoceros’

Militarev glosses this form as ‘rhinoceros’, but Miller (2013: 158) gives this form not as ‘rhinoceros’, but as the sighting or hunting name for the rhinoceros (the more typical name of the rhinoceros in Hadza is *tlhakate* /cʎ̝̊^{h}akate/). This distinction is crucial here, as these special hunting names of animals in Hadza are not nouns, but verbs (see also Blench 2013a). In this case, the source verb is /loʃo/ ‘to stoop’. This would be consistent with other animal hunting names, which often characterize the animal in some way (the hunting name for ‘ostrich’ comes from a verb meaning ‘to puff up’, for example, and the hunting name for ‘baboon’ comes from ‘thirst’ (in reference to a baboon’s concave stomach)). In this case, the verb ‘to stoop’ would seem to evoke the posture of the rhinoceros, especially its stance when about to charge. The verbal nature of /loʃo/ is further reinforced in that this term has no plural form. As such, any comparison of the (verbal) hunting name for ‘rhinoceros’ /loʃo/ with (nominal) words for ‘rhinoceros’ in Afroasiatic (e.g. reconstructed Proto-Afroasiatic **warŝ-*) is misguided.

#32 ‘to be many, to be plenty, to be full’ ***furu-ne*** (more properly cited as /fuɾune/ ‘to be many, to be plenty, to be a lot’) ~ **ʕVpVr-* ‘full’

The basic meaning of the Hadza root is ‘to be many’. The meaning ‘to be full’ is a secondary one, and would not be used to refer to a full cup of water or a full belly (the verb /ǀˀoʔa/ ‘to be full’ would be used in these cases). It appears that the Afroasiatic terms compared all mean ‘be full’ and not ‘be many’, but it is difficult to confirm this as the AA **ʕVpVr-* ‘full’ root does not appear in our searches of the StarLing Database. Unless it can be shown that a reconstructed AA form shares the basic meaning ‘to be many’, we do not think the roots can be considered to be related.

There are additional issues with the comparison. Militarev presents this form as *furu-ne*. Phonologically, it is important to note that Hadza does not feature the phoneme [r], and the form actually has a tap/flap (thus: [fuɾune]). Morphologically, the word *furune* (or *furuni*) is

monomorphemic in Hadza, and though Miller offers the Iraqw and Proto-West Rift form *xooroo* 'crowd, community, people in group' for comparison (Miller 2013: 83–84), there is no evidence that the word can or ever could be analyzed as some morpheme *furu* and some morpheme *-ne* in Hadza. As such, the form to be used in comparison ought to be *furune* [fuɾune] (or the alternative form *furuni* [fuɾuni]). This makes comparisons with forms offered by Militarev (such as the reconstructed Proto-Afroasiatic **ʕVpVr-* 'full') rather less convincing.

The Hadza and Afroasiatic forms are not directly comparable because no analysis is presented explaining why the initial syllable is missing in the Hadza form. An AA *ʕ corresponds to a (medial) Hadza glottal stop in #1 'all'. A medial AA *ʕ (AA **ma-ʕi(n)ʒ-* 'k. of bovid') appears to be deleted in Hadza *minza* (more properly cited as /mínza/ [míndza] 'reedbuck'). No explanation is given as to why there might be different reflexes of AA *ʕ in Hadza.

In addition to this, it has been argued that the voiceless labiodental fricative [f] is a recent addition to the Hadza consonant inventory, and that most if not all Hadza words containing [f] can be shown to be borrowings or mimetic (Harvey 2021). If this were the case, any comparison with Afroasiatic forms would be inapplicable.

2.5. Ma'a Data

There are languages that one should use with utmost care in building cognate sets. A prima example is Ma'á / Mbugu. This language is known as a mixed language. Mous (2003) uncovers how this "strange case of Mbugu", as Goodman (1971) has put it, came about. Briefly, the language is a Bantu language, with an extra parallel lexicon. This parallel lexicon has partly been built consciously in an attempt to re-create a Cushitic language that was lost due to shift. The extra lexicon shows words from two different Cushitic sources, from Maasai, from Taita Bantu and manipulated words from the regular lexicon (Pare). Taking words from Ma'á / Mbugu requires that one knows from which of the two lexica it comes and, if from the mixed parallel lexicon, from which source. An additional challenge is that even the basic Pare lexicon contains borrowings from Cushitic as all Bantu languages in the area do. No evidence from Ma'á should be used to build a cognate set on; it can only be taken as additional evidence.

3. Lexical Sets Better Analyzed as Loans

A number of Militarev's lexical sets arguing for an Afroasiatic membership of Hadza are better analyzed as instances of loans from different languages and language families (especially Bantu, Nilotic and West Rift Cushitic). As such, they are certainly interesting for the history of Hadza, but they can hardly be taken as proof of genetic relationship. We expect that there may be more cases of loans that we cannot yet confirm due to currently available, limited resources on languages such as Ihanzu (Bantu) and Datooga (Nilotic), and the timing of the loans is not just of a recent date.

3.1. Bantu

#16 *dza* 'come'

Hadza *dza-* < Ihanzu *dza*

The Hadza term may be straightforwardly interpreted as a loan from the neighboring Bantu language Ihanzu (Harvey 2019). The concept 'come' does colexify with 'go' in CLICS[3] (Rzymski et al. (2019), so it is possible that an Afroasiatic term meaning 'go' might be related to a term meaning 'come', yet the borrowing analysis is, in our opinion, the most likely scenario.

There is a different Hadza root *botʃʰo* 'come' used in imperative forms that has no connection to Afroasiatic (that we are aware of), though it can be plausibly linked to a similar irregular root used in the imperative in Bantu languages (c.f. Ihanzu *nzuu* 'come!', or Swahili *njoo* 'come!').

Section 1.3.2. 'large male lion' *mondo* (more properly cited as /móndo/)

Miller (2013: 173) notes that this form is ultimately from a Bantu source. **mondo* 'tiger-cat, serval' (Bastin et al. 2002: 6665) is a form with distribution across northeast Bantu. This seems to be more plausible than the cognates offered by Militarev.

Section 2 'epilepsy' *nkoro-ko* (more properly cited as /ŋkʰoɾo-ko/)

The word for 'epilepsy' in the neighbouring Bantu language Ihanzu is *nkólo* (Harvey 2019). Cf. also **kódò* 'heart', with distribution across northeast Bantu (Bastin et al. 2002: 1889).

#34 'good' *muta-na, mta-na* (more properly cited as /mtana ~ mutana/ 'fine, good')

As mentioned above, there is no synchronic (or diachronic) evidence for the morpheme breaks in this word proposed by Militarev. From a semantic perspective, Militarev provides the gloss for *mutana* as 'good' (adjectival), whereas Miller (*ibid.*) provides the gloss as both 'good' (adjectival) and 'fine' (adverbial). In fact, in the author's work (Griscom and Harvey 2020), the use of *mutana* in an adjectival sense is absent, and only the adverbial sense is recorded. The root used for 'good' in an adjectival sense is *nube- ~ nubi-*.

Phonetically, a preponderance of words in Hadza beginning with prenasalized stops have been shown to be borrowings from Bantu languages. Despite Militarev's assertion to the contrary (fn. 15), it is highly likely that *mtana* or *mutana* also has a Bantu source.

Hadza *minza* ~ AA **ma-ʕi(n)ʒ-* 'k. of bovid' (more properly cited as /mínza/ [míndza] 'reedbuck')

Even if we were to accept the dubious semantic correspondences between a wide-range of cloven-hoofed ungulates: 'cow', 'he-goat', 'female topi'[3], etc. and Hadza *mínza* 'reedbuck', there is a very strong likelihood that the Hadza form is a loan, as is demonstrably the case for many words which contain a prenasalized obstruent (cf. Elderkin 1978). Another word for 'reedbuck' /ndofeda/ (Peterson 2012) is also likely a loanword for the same reason.

Hadza *ndama* 'calf'

This is from Cushitic (Nurse and Hinnebush 1993: 304), also borrowed into Bantu, and Hadza must have gotten it from Bantu due to the prenasalized stop; cf. *ndama* 'calf' Sukuma, Swahili.

3.2. Nilotic

Hadza *gondera* 'dog' < *gudeeda* 'dog' in Datooga.

This is clearly a loan from Datooga *gùdé:d* (Griscom 2019: 93). Note that the Hadza word is not the most commonly used word for dog, /‖ʰaʔano/, nor is it the second most commonly used word /tiʰiŋgi/.

3.3. West Rift (Southern Cushitic)

#80 Hadza *ntsa-ko* 'star' ~ AA **(t/wV-n)ciʕ(ciʕ-)*

Militarev gives Iraqw *cacēʕ* as evidence for the Hadza form to be Afroasiatic. The Iraqw root is *ts'atsa'* (m) 'stars' (Mous et al. 2002); PWR **ts'its'aʕu* (m) 'stars, starlit sky' (Kießling &

[3] To this list might be added Melo (Omotic) *mintsá* 'buffalo' (Seibert & Caudwell 2002) and Shabo (isolate) *mijad* 'buffalo' and *mindʒa* 'female cow' (Kibebe 2015: 443)..

Mous 2003) but an even closer cognate would be Proto West Rift *tsaaʔa-s* (causative verb) 'shine, shed light'. The initial nasal in Hadza reminds us of a Bantu class 9 noun class prefix but there is no Bantu source in the vicinity with a similar word for 'star'. It is possible that this root came into Hadza from Cushitic via a Bantu intermediary.

Hadza *slaa* [ŝaʔa] (more properly cited as /ɬaʔa/) 'to love' ~ AA *ŝaʔ/w- 'want, wish, love, like'

This root is indeed attested in Proto West Rift *ɬaʔ* 'live, like, want' (Kießling & Mous 2003) and we agree with Militarev that borrowing from West Rift cannot be ruled out. The suggested cognates with Iraqw work equally well with Proto West Rift (Kießling & Mous 2003) and there is very little current lexical transfer between Iraqw and Hadza. One should keep in mind that it was only a few decades ago that the Iraqw moved into the vicinity of the Hadzabe. It is not unlikely, however, that the speakers of Pre-Proto West Rift were in contact with Hadza during their movements into Tanzania.

4. Problems with Semantic Correspondences

Semantic correspondences for alleged cognates are sometimes baffling. For instance, set #56 includes meanings as diverse as: 'mouth', 'tongue', 'cheek' and 'muscle in the neck'. These are all compared without further justification and we find these semantic correspondences to be rather implausible. By allowing such wide semantic comparisons, Militarev increases the likelihood that chance resemblances occur.

We also call into question the semantic correspondences in set #74 which include diverse meanings: 'to sit', 'to stand' and 'to sleep'. These are all quite different things. While the correspondence between 'hand' and 'finger' may be accepted, it still looks rare (cf. #37). The correspondences in the second part of the paper between different types of animals are also concerning.

We would recommend that the types of allowable semantic comparisons be limited to the prototypical types of semantic change mentioned in introductory textbooks on historical linguistics or historical semantics (e.g. widening, narrowing, metonymy, etc.). It is not sufficient to propose cognate sets based on two glosses occuring in the same semantic field. We do not find it terribly plausible that a term for a gazelle or a ram somehow semantically changes meaning to allow reference to a buffalo, or that terms for herons or cranes might somehow end up referring to ostriches.

#1 is a typical example in many ways. We doubt that the correspondence between 'all' (Hadza) and 'one' (all the rest) is semantically plausible.

#1 Hadza *waʔi* 'all' (more properly cited as /waʔi-/ 'all of X')

The Hadza root *waʔi-* is not a noun and must take an affix.

The concept 'all' may colexify with 'every' or 'many' but does not colexify with 'one' in any language in the CLICS[3] Database of Cross-Linguistics Colexifications (Rzymski et al. 2019, accessed 27 June, 2023)

#56 *awani-ka* 'mouth' (more properly cited as /ʔawanika-ko/ 'mouth') ~ AA *ʕawin- 'tongue; (part of) mouth'

As discussed above in Section 2.3, we do not agree with the segmentation of the word into a root ***awani-***[4]. Besides our disagreement over the morphological segmentation, there is an is-

[4] Miller internally reconstructs **ʔawa* 'mouth' for Hadza based on /ʔawanika/ 'lower lip', /ʔawanika-ko/ 'mouth, beak' and notes the similarity to Proto West Rift **ʔafa* 'mouth'. It is possible to internally reconstruct ***ʔawa-ni-k(w)a** (along with ***ʔawa-ti** 'upper lip') but not ***awani**.

sue with the equivalence of the semantics. The concept 'mouth' colexifies with a number of different concepts in CLICS[3] (Rzymski et al. (2019) including: 'lip', 'tooth', 'beak', 'language', 'edge', 'door', 'face', 'chin', 'word', 'mother', 'jaw', 'hole', 'throat'. We are dubious that the concepts 'mouth' and 'tongue' (or 'inside of cheek') might be connected through semantic change from a common meaning.

Hadza *geweda-ko* 'dikdik' (more properly cited as /géwéda-ko/ 'dik-dik')

The Hadza word is compared to Afroasiatic words for various ungulates including 'greater kudu', 'buffalo' and 'duiker'. Dik-diks are tiny antelopes, very visually and functionally distinct from the kudu and the buffalo, one of the most dangerous animals on the continent. Although it is not impossible for a word to come to refer to different ungulates over time, we are concerned that the great semantic laxity in such a comparison *greatly* increases the likelihood that a chance resemblance may have occurred. The AA form **guday* provides a match for two of the three consonants in the Hadza root, but readers are left to imagine for themselves how the two roots might be related through a series of sound changes. The sound changes that would be necessary to connect the roots seem to us to be implausible and not supported by other correspondence sets as a regular, repeated correspondence. The shape of the Hadza root (especially the final *da* syllable) immediately bring to mind the possibility that the word may be a Southern Nilotic loan[5]. The relatively poor documentation of Datooga lects in the present day makes it impossible to conclude that the Hadza word could not be a loan. Although not proveable at this time, we feel this alternative hypothesis is no more unlikely than the one set forth by Militarev.

Hadza *garaani* 'heron' ~ AA 'ostrich' (more properly cited as /gaɾaʔani-ko/ [galaʔani-ko] 'black-headed heron')

There is very little to semantically connect a heron with an ostrich other than the fact that both birds have relatively long legs.

Hadza *gaga* 'grasshopper' (more properly cited as /gagá/)

There is very little to semantically connect a grasshopper with a spider or flea. This word is best analyzed as a loan from the Bantu language Ihanzu; the Ihanzu word for 'grasshopper' is *ngaga* (Harvey 2019).

5. Onomatopoeia

In this section, we discuss some sets that we feel are better analyzed as being due to onomatopoeia.

#6 Hadza *tʰítʰí, tĩ̀ti* 'bird' (more properly cited as /ti:ti-/ 'bird' (generic term for birds excluding ostriches)) ~ AA **diʔ(diʔ)* 'k. of bird'

#30 Hadza *pururu* 'fly' (v.) (more properly cited as /puɾuɾu-/ 'to fly off') ~ **AA *pir- 'to fly'**

Militarev compares the Hadza form *garaani* 'heron' with AA roots **gawir-* and **garay-* 'heron, crane, ostrich'. We note that the Kanuri (Saharan) word *gəragə́ra* 'heron' (Cyffer 1994: 88) nicely matches the first CVC of the Hadza, presumably completely due to chance or to onomatopoeia. The Afroasiatic roots also only match an initial voiced velar, low vowel and liquid.

[5] The Hadza root resembles Datooga *géewèedà* 'shoe' (Roland Kießling, p.c. 23 June 2023). Though the semantics do not match, they are one of the most abundant antelopes in the area and the leather is quite likely used in making shoes.

The Hadza is specifically *Ardea melanocephala* 'black-headed heron' /gala?ani-ko/ (Peterson 2012: 2018), the term used when referring to the bird's habit of swallowing snakes (or, the name *!omolo* may be used when referring to its long neck, Peterson, *op. cit.*). Hadza does not contrast r/l but it does contrast VV and V?V. None of the proposed cognates explain the lack of a consonant corresponding to the Hadza glottal stop in this word. If the Hadza root is bimorphemic, it would be possible for the glottal to be part of a second root (*gala-*?ani-) but we have no internal evidence for this. The proposed correspondence set is unsatisfactory both from the standpoint of the phonological correspondences (which ignore half of the Hadza word and the correct spelling of the Hadza) and from the semantic correspondence (which equates herons and ostriches). If any two types of birds might be compared, then it is unsurprising to find at least one root with a voiced velar and a liquid in one of over a hundred possible languages. The fact that two different AA roots have been proposed is unsatisfactory as well, as it introduces an indeterminacy in the historical scenario being proposed.

6. Evaluation of Selected Sound Correspondences

In this section, we discuss correspondences involving a subset of Hadza consonants.

6.1. Glottal stop

The presence of a glottal stop is not always noted in the orthographic representations cited by Militarev though it is shown in the phonetic transcriptions (when provided). Glottal stops are not merely phonetic in Hadza, but phonemic.

Glottal stops in Hadza roots have been compared to AA *? in several sets:

#41 'I' *ono* [ʔono, ʔono-ko] (more properly cited as /ʔono-/) ~ AA **ʔani/u* 'I'

#31 'foot' -2 *asenako* ['ʔasenako] (more properly cited as /ʔasena-ko/) ~ AA **ʔa-sin-* 'foot, leg'

#9 'blood' *átʰaʔmá-, átàmaˆ* [ʔatʰama] (more properly cited as /ʔátʰamá/) ~ AA **(ʔa-)dam-* 'blood'

#64 'person' *unu* [ʔunu] (more properly cited as /ʔunu/) ~ AA **ʔa/inay-(n)-* 'man, person (also elder kin)'

#76 'sleep' *ʔase* (also 'lie') (more properly cited as /ʔase/) ~ AA **sayʔ-* / **ʔays-* 'sit, sleep, rest'

#71 'say' -1 *ʔî* (<*ʔiy?) (more properly cited as /ʔi:/) ~ AA **ya-* / **ʔiy-* 'say'

#79 'stand' *ikha-* (also 'to stop') [ʔíkʰà] (more properly cited as /ʔíkʰà/) ~ AA **ḳaʔ/w* / **ʔVḳ(k)-* 'rise, be high, stand (up)'

#82 *isho-ko* (isho 'sunlight') [ʔiʃoko] (more properly cited as /ʔiʃo-ko/) ~ AA **ʔa/is-* 'sun'

At first glance, these sets might appear to be a sufficient collection to demonstrate a regular, repeated sound correspondence. But, it is necessary for all of the segments in roots to regularly correspond, and not just a single segment. First, however, note that #9 'blood' and #31 'foot' cannot be considered independent of each other as they appear to reconstruct to a single morpheme **ʔa-* in Afroasiatic. These correspondences are dubious, however, since there is no evidence that these syllables are separate morphemes in Hadza. The corresponding segments in #41 'I' are not demonstrated to be regular and there are difficulties with comparing the generally restricted set of consonants that occur in pronominal systems with one another; the small number of consonants that tend to be used greatly increases the likelihood of a chance resemblance (Gordon 1995). Although we do not examine vowel correspondences in this paper, those shown in these sets do not appear to be regular. We do analyze voiceless stops and

fricatives and cannot consider the correspondences shown here involving these to be regular. If only one segment in a word has a regular correspondence, the entire word cannot be presumed to be cognate.

We question the first impression that the /ʔ/~ *ʔ correspondence is regular.

There is no glottal stop comparable to AA *ʔ in:

#33 'give' *kwe-* ~ AA *kaʔ- / *ʔaVk-

No conditioning environment for this is discussed, nor is an explanation given to account for the metathesis.

There also does not appear to be a segment in AA corresponding to the glottal stop in:

#24 'egg' *usle-ko* [ˈʔuɬeko] (more properly cited as /ʔúɬe-ko/) ~ Chadic **ŝay(ŝay)*

#31 'foot' -1 (also 'leg') *a/uphukwa* [ʔa/ʔup^{h}uk^{w}a]~ Cush. S. **fanḳ^{w}-*; Kera *káma-a* 'foot, leg' (met.), etc.

There are a few cases where a Hadza glottal appears to correspond to a pharyngeal in AA *ʕ:

#56 *awani-ka* 'mouth' (more properly cited as /ʔawanika-ko/ 'mouth') ~ AA **ʕawin-* 'tongue; (part of) mouth'

#52 'many' *ʔaso* (more properly cited as /ʔáso-/) ~ AA **w/yasaʕ*

#1 'all' *waʔi* ~ AA **waʕ* 'one'

#45 'know' *tethaʼo* [tet^{h}aʔo] (also 'understand') (more properly cited as /tet^{h}aʔo/) ~ AA **(y)daʕ-* 'know'

63 'one' *itchâme* [ʔitʃ^{h}aame] (more properly cited as /ʔitʃ^{h}aːme/) ~ AA **ʕist(-an)-* 'one'

Since Hadza does not have a pharyngeal fricative, it might be presumed that such a sound may have merged with the glottal stop. Also, the correspondence between Hadza ʔ ~ AA *ʕ in #52 is only allowable if one presumes metathesis, yet no explanation is given to motivate the metathesis. Furthermore, this analysis would not explain why other AA roots with *ʕ have no corresponding glottal stop in Hadza, e.g.:

#23 'eat' *seme, simi* ~ AA **suʕVm-*

#32 'to be many, to be plenty, to be full' *furu-ne* (more properly cited as /fuɾune/ 'to be many, to be plenty, to be a lot') ~ **ʕVpVr-* 'full'

#80 'star' *ntsa-ko* ~ AA **(t/wV-n)ciʕ(ciʕ-)*

To summarize, we see irregular rather than regular, repeated correspondence sets involving the Hadza glottal stop. We do not see sufficient evidence to hint at a linguistic genetic relationship between Hadza and Afroasiatic.

6.2. Voiceless Stops

The distinction between aspirated and unaspirated stops (and other obstruents) in Hadza is contrastive and regular sound correspondences should be established involving each type of consonant.

Hadza /t^{h}/ is compared with AA *d:

#9 'blood' *átʰaʔmá-, átàmaˆ* [ʔat^{h}ama] (more properly cited as /ʔát^{h}amá/) ~ AA **(ʔa-)dam-* 'blood'

#45 'know' *tethaʼo* [tet^{h}aʔo] (also 'understand') (more properly cited as /tet^{h}aʔo/) ~ AA **(y)daʕ-* 'know'

But #49 shows /t^{h}/ is compared with AA *t*, with no explanation proposed for this seeming irregularity:

#49 'long' *thas* [t^{h}as-] (more properly cited as /t^{h}as-/) ~ AA **tays-* 'long'

Hadza /t/ is also compared with AA *d:

#6 'bird' *t^{h}ít^{h}í, tîti* (more properly cited as /ti:ti-/ 'bird' (generic term for birds excluding ostriches)) ~ AA **di?(di?)* 'k. of bird'

Hadza /t^{h}/ and /t/ are compared with AA *d yet there is no comparable collapse (or alternation) of voiceless (aspirated and unaspirated) labial and velar phonemes with their voiced counterparts as shown by forms such as:

————————p
#30 'fly' (v.) *pururu* (more properly cited as /puɾuɾu-/ 'to fly off') ~ AA **pir-* 'to fly'
#91 'two' pi^{h}e, pie-(be) ~ AA **(hV)pV(p)-* / **H/yV(m)p/b-* (< **hVṗ-*?) 'two'
————————p^{h}

————————k
#33 'give' *kwe-* ~ AA **ka?-* / **?aVk-*
#92 'walk' *haka* (ɦaka 'go', *ʔetlhikwa* 'walk') (more properly cited as /ɦaka/ 'to go', /ʔic$\underset{\circ}{\Lambda}^{h}$íkwa/ 'to walk, to go, to leave')~ AA **k^{w}VH-* / **HVk^{w}-* 'go, walk'
#67 'road' *yeke* (more properly cited as /jeke/ 'path') ~ AA **kaw/y(k)-* / **yVhk-* / **hVwk-* 'go, walk, come; road'
————————k^{h}
#62 'not' *'ukuwa* (more properly cited as /ʔuk^{h}u-/ 'to not exist') ~ AA **(ʔa-)k^{w}ay-* 'not'
#79 'stand' *ikha-* (also 'to stop') [ʔík^{h}à] (more properly cited as /ʔík^{h}à/) ~ AA **ḳaʔ/w* / **ʔVḳ(k)-* 'rise, be high, stand (up)'

We find it implausible that the stop series would not pattern together in a similar way. A sufficient number of regular, repeated sound correspondences have not been presented for the sets involving pulmonic stops to be convincing.

Since it is difficult to find repeated correspondences with any particular consonant, we might instead compare consonants grouped into natural classes in order to detect correspondence patterns that might otherwise fall below the level of significance. However, in grouping together the voiceless plosives, we see that voiceless unaspirated and unaspirated velars and labials pattern differently with respect to coronals; the coronals are posited to be cognate with voiced stops in AA while the other stops are posited to correspond to voiceless plosives. Not only is there no explanation for the contrast between aspirated and unaspirated voiceless stops in Hadza, there is no explanation posited by Militarev to explain why the natural class of voiceless stops does not pattern together. Grouping consonants together into natural classes is one way to test the phonological naturalness of the sound correspondences proposed.

6.3. Fricatives

The only fricative with more than a few proposed cognate sets is /s/. The other sets involving fricative cannot be said to show regular, repeated sound correspondences. (The sets involving /ɬ/ at least have one repetition without a contradictory sound correspondence, but only one of these ('to love') reconstructs to AA and it is almost certainly a loan from West Rift (South Cushitic)).

————————f

#32 'to be many, to be plenty, to be full' ***furu-ne*** (more properly cited as /fuɾune/ 'to be many, to be plenty, to be a lot') ~ *ʕVpVr- 'full'

————————ɬ

#24 'egg' ***usle-ko*** **['ʔuɬeko]** (more properly cited as /ʔúɬe-ko/) ~ Chadic *ŝay(ŝay)

Hadza *slaa* [ŝaʔa] (more properly cited as /ɬaʔa/) 'to love' ~ AA *ŝaʔ/w- 'want, wish, love, like'

————————ʃ

#82 ***isho-ko*** **(isho 'sunlight') [ʔiʃoko]** (more properly cited as /ʔiʃo-ko/) ~ AA *ʔa/is- 'sun'

Hadza *dushu* 'distended, big stomach' (more properly cited as /dúʃu/[6] 'distended stomach (from malnutrition)' also /dúʃu-ko/ 'mole cricket') ~ AA *dVs(-Vm/n)- 'fat belly'

Hadza *shububu-bi* 'lungs' (more properly cited as /ʃububú-/) ~ AA *ci/anṗ- 'lungs'

'rhino' ***losho*** more properly cited as /loːʃo/ 'sighting/hunting name for the rhinoceros' (when it is stooped and ready to charge)) ~ AA *warŝ- 'rhinoceros'

————————z ~ dz[7]

Hadza *uzame-ko* [ʔudzameko] **'spotted (laughing) hyena' (**more properly cited as /ʔudzame-ko/ 'spotted hyena') ~ Chadic *zVm- 'lion'

Hadza /s/ corresponding to AA *s is one of the few consonantal correspondences with more than 2 examples. However, in many of these cases, there are elements in the AA forms which have no correspondence whatsoever in the Hadza forms. For instance, there is no explanation for the missing second syllable in Hadza *sa* 'to rain' compared to AA **sawiʔ*- 'rain', or of the difference in order of segments in Hadza *ʔase* 'sleep' compared to AA **sayʔ-/ʔays*- 'sit, sleep, rest'.

#23 Hadza 'eat' ***seme-*** compared to AA **suʕVm*- 'eat'

#31.2 Hadza [ʔasena-ko] 'toes' compared to AA **ʔa-sin*- 'foot, leg'

#49 Hadza *thas-* [tʰas-] 'long' compared to AA **tays* 'long'

#52 Hadza *ʔaso* 'many' compared to AA **w/yasaʕ*- 'big, many'

#65 Hadza *sa* 'to rain' compared to AA **sawiʔ*- 'rain'

#76 Hadza *ʔase* 'sleep' compared to AA **sayʔ-/ʔays*- 'sit, sleep, rest'

#93 Hadza 'to warm *siɦi-* (oneself by a fire)' compared to Semitic **šVḫan*- '(be) warm, hot; warm oneself'

Hadza *biso-ko* 'wildebeest' compared to AA **bus*- 'goat, bushbuck'

Hadza *ʔakwisiti-ko* 'sinew that runs along the spine and neck' compared to AA **kac/sw*- 'back with shoulders'

Militarev's correspondence sets are curated to some extent. For instance, **#17 Hadza *misi* 'die'** is not scored even though it is similar to S. Bauchi **mis*- 'die' and AA **mwt*, because the S. Bauchi form cannot be shown to be a regular reflex of the AA form, and the *s~t* correspondence between Hadza and AA would also be irregular.

Despite Militarev's attempts to show regular, repeated sound correspondences, we would need to see an explanation of the vowel and approximant correspondences in these same words to make sense of these as related forms.

[6] This word is a rare instance of a clickless word that appears to be synchronically related to another word in Hadza with a click, i.e. /!ʰuʃuː/ 'have a big/stretched stomach' (Miller 2013). Also suspiciously similar is: /!ʔuʃu-ko/ 'navel and/or umbilical cord', 'fat critter with visible innards' (such as some toads or crickets) (Miller 2013).

[7] Note that although we prefer to consider that this root has a /dz/ phoneme, there is no contrast between /dz/ and /z/ in the language so it is possible to analyze this as a fricative.

6.4. Central Ejective Affricates

Hadza distinguishes phonemes /tʃ/ and /ts'/ and there is no synchronic evidence or evidence from internal reconstruction to connect the two phonemes. Militarev compares Hadza /tʃ/ to AA *ʦ̣ and *ʧ̌, and Hadza /ts'/ is compared to AA *ʦ̣, *ʧ̌ *and* *ʨ̂. There is no explanation proffered for the irregular correspondences. Looking at the sets below, we see other issues with the comparanda. Evidence proposed based on central ejective affricates does not support the hypothesis that Hadza is an Afroasiatic language.

In the cognate set #21 'ear' /tʃ/ is connected to the Hadza phoneme /ts'/ in set #46 'leaf' (compared to AA *ḫ/ḥaʦ̣- 'leaf; ear'). The Afroasiatic Database (AADB) provides two Proto-Omotic reconstructions for 'leaf': *Hyaʦ̣- (supported by forms Chara: *yēʧ̌a*; East Mao (Diddesa): *yaʦ̣-*) and *Hwayš- (supported by: Koyra (Badditu, Amuru): *wāše*; Ganjule: *waše*; Janjero (Yamma, Yemsa): *t-oša*; Gimirra (Benesho, She): She *aisi*). Neither of these reconstructions matches the Ometo reconstruction cited by Militarev: *Hayʦ̣-. Although there are a number of similar-looking roots in Ometo languages, these do not contain ejectives: *haytta* (Wolayta), *hayða* (Gofa, Dawro/Kullo), *hayð* (C'ancha) 'leaf' (Alemayehu Abebe 2002: 8). Note that in many of these languages, the same word may also refer to 'ear': *haytta* (Wolayta), *hayða* (Gofa, Gamo, Dawro/Kullo), *hayð* (C'ancha) (*op. cit.*: 7). A correspondence in form and meaning between '(h)ear'/'leaf' cannot be presumed to be evidence of common inheritance. Not only can the connection between 'leaf' and 'ear' occur in languages unrelated to Afroasiatic, as shown in Figure 1, but colexifications may spread within a linguistic area, and be more indicative of geography than inheritance (Urban 2012, Campbell et al. 1986).

Language	Family	Form
Dera	Afro-Asiatic	kumo
Rasawa	Lakes Plain	ura
Diyari	Pama-Nyungan	tharlpa
Ngamini	Pama-Nyungan	talpa
Wulguru	Pama-Nyungan	bina
Yandruwandha	Pama-Nyungan	thalpa

Figure 1. Colexifications for 'leaf' and 'ear' in CLICS[3] Database of Cross-Linguistics Colexifications (Rzymski et al. 2019, accessed 27 June, 2023)

The Hadza consonant /tʃ/ is proposed by Militarev to have a repeated correspondence with AA *ʧ̌ in two words. In the word *tʃowa* 'gecko', which is compared to the AA root *ʕiʧ̌ay-, no explanation is given to explain why the Hadza word does not have a syllable preceding the /tʃ/ to correspond to AA *ʧ̌. There is a similar problem with the other Hadza word containing this affricate: *watʃo* 'slender mongoose' (Peterson 2012: 216). The Hadza word is said by Militarev to correspond to AA *ʧ̌ in words with varied meanings ('rat', 'mouse', 'weasel', 'mongoose', 'jerboa'). Hadza speakers distinguish between five species of mongoose (Peterson 2012: 216) and there is no sense in which these carnivores would be referred to as if they were mice or rats. Furthermore, there is an initial syllable in most of the Afroasiatic forms (**ʔV(n)ʧ̌aw-)* that has no correspondence with Hadza.

In set #78 'smoke', the AA *ʧ̌ consonant is connected to Hadza /ts'/ rather than /tʃ/. No explanation for the different correspondence is noted (compared to 'gecko' and 'slender mongoose'), however, the AA form may alternatively be reconstructed with the *ʦ̣ consonant. The

Hadza consonant /ts'/ in the word *ts'eʔa-* 'to shit' is also proposed to correspond to AA **ǰ́*, but the same consonant in 'to smile' is proposed to correspond to **ĵ́* (*ts'ukuts'uku* 'to smile' ~ AA **ĵ́Vḥ/Vḳ* 'to laugh'). No explanation is given for the irregularity of the proposed correspondences. We are not convinced by evidence consisting of a single repeated sound correspondence, especially when a set that violates the proposed regularity has also been proposed.

6.5. Lateral Ejective Affricates

Sets involving lateral ejective affricates do not show regular, repeated sound correspondences though the comparanda here are quite striking and are deserving of some discussion.

Sets #10 'bone' and #97 'white' appear to show a single repetition of a sound correspondence between the Proto-Afroasiatic sound **ĵ́* and Hadza /c̬ɬ̓/. For a sound correspondence to count as evidence of a linguistic genetic relationship, however, a sound correspondence must be both regular and repeated – linguists have not quantified the number of times a correspondence must occur to count as evidence, but certainly, the difference between one occurrence and two occurrences is not statistically significant. Furthermore, set #36 'hair' shows Hadza /c̬ɬ̓/ does not repeat this same correspondence with AA. In this set, /c̬ɬ̓/ corresponds to **ĉ/ŝ* and no explanation is proffered to account for the different correspondence.

#10 'bone' *midla* (Miller et al. 2021), *mitl'a* (Sands 2012: 5) (more properly cited as **/micc̬ɬ̓a-/**) =AA **ma/iĵ́-* 'bone': =Cush. C. **ŋaǰ́-* (<**maǰ́-*); =E.: HEC: Darasa, Burji *miǰ́ǰ́-o*; Yaaku *moǰ́-o*; =S.: Dahalo *miĵ́ĵ́-o*[1] // =Omot. N.: Mao (Sezo) *mãlṭ-é* ◊ AADB 1269.

#97 'white' *pedla* (*petl'a-* in Sands 2012) (more properly cited as **/pecc̬ɬ̓a-/**) =Sem. **ṗayṣ̂-* (AA **ṗayĵ́-*) 'white'.

#36 'hair' *hadle* (more properly cited as **/ɦacc̬ɬ̓e-/**) ~ AA **(Ha-)ĉ/ŝVw-* 'hair, feather': Cush. S. =Ma'a *aŝu* 'hair' *(*ĉ* and *ŝ* both render *ŝ* in Ma'a according to Takács 2011) // Eg. (Pyr.) *šw.t* 'feather' ◊ AADB 1284.

Using Ma'a to establish a sound correspondence is problematic as we explained above. Both Egyptian (and Ma'a) have a rounded segment (*w* and *u*, respectively) that has no regular correspondence with the Hadza front vowel. No other comparison sets with Hadza /c̬ɬ̓/ are proposed by Militarev. Two sets of corresponding sounds is not a high enough number to be considered by us to be a 'regular, repeating sound correspondence'; rather, this is more typical of a chance resemblance or resemblance due to borrowing.

Even though the sound correspondence involving the sets 'bone' and 'white' is striking, we must still ask whether or not one of these forms might be due simply to chance resemblance. Unrelated Kanuri (Saharan) has a form *bûl* 'white' (adj.; Cyffer 1994: 212) which strongly resembles the set, particularly when one considers that Kanuri does not have lateral obstruents.

There is another Hadza root *petla* /pec̬ɬa-/ 'to shine, glitter, gleam' with a pulmonic (not ejected) lateral affricate that is conceivably connected to /pec̬ɬ̓a-/ 'white', though not through any known synchronic alternation.

Even within Afroasiatic, it is unclear which forms the Hadza form(s) should be compared to. Militarev connects Hadza to a Proto-Semitic form **ṗayṣ̂-* rather than to the Proto-Omotic form. A Proto-Omotic form may have yielded reflexes without lateral ejective obstruents, e.g. Yemsa *bìtʃà* 'yellow' (Aklilu Yilma et al. 2002: 26), Melo (Omotic) *bó:ts* 'white' (Siebert & Caudwell 2002); *botta* (Wolayta), *boða* (Gofa, Dawro/Kullo), *buð* (C'ancha, Dorze) (Alemayeh Abebe 2002: 12). Takács (2011: 185–186) connects Proto-Omotic **bōǰ́* to Semitic **byḍ* (Arabic *byḍ* : *bāḍa* 'to grow yellow', *bayyaḍa* 'to make white', etc.) and to Chadic forms such as Mafa-Mada

**baḍ* 'white', all derived from AA **b-ĉ̣* 'white' (which is comparable to the AA form **ṗayĉ̣-* cited by Militarev). If the rounded vowels in Omotic might be due to coloring from the initial labial, it is not out of the question, then, that Hadza might have borrowed a form with an ejective from the same type of source as 'bone'.

Another issue with the sound correspondence proposed for 'bone' and 'white' is that AA **ĉ̣* has also been proposed by Militarev to have other reflexes in Hadza: *ts'ukuts'uku* 'to smile' (Miller et al. 2021: 560) is compared with AA **ĉ̣Vḥ/Vḳ* 'to laugh'. Even disregarding the semantic differences between these two words, there is no explanation of the different sound correspondence between this set and 'bone'/'white' (not to mention different correspondences involving Hadza /ts'/). If a word has a sound correspondence that is irregular and cannot be explained as a conditioned sound change, then it cannot count as evidence for a linguistic genetic relationship.

7. Role of chance not fully evaluated

In this section, we discuss some comparison sets that raise the question of whether the similarities are due to chance rather than common inheritance from Afroasiatic.

#31.2 *asena-ko* ['ʔasenako] 'toes'

The Hadza root /ʔasena/ is compared to AA **ʔa-sin-* 'foot, leg'. Some Chadic forms have the initial syllable, but not the Omotic, Cushitic or Egyptian forms. The paper provides no explanation of the presence or absence of purported *ʔa-* prefix nor of its function in Hadza and Afroasiatic. The second and third consonants of the Hadza root are also reasonably similar to those in /sɔ̃ːnt^{h}/ 'foot' in Northern Koma (Koman, Nilo-Saharan) (Bender 1971). We must stress that even if Hadza *were* Afroasiatic, it is still possible for chance resemblances to occur, particularly when involving relatively high-frequency consonants such as /s/ and /n/.

#32 *furu-ne* 'to be many, to be plenty, to be full' (more properly cited as /fuɾune/ 'to be many, to be plenty, to be a lot')

The first part of the Hadza root is phonologically similar to English 'full', which reminds us that resemblances due purely to chance may certainly occur.

Hadza *uzame-ko* [ʔudzameko] **'spotted (laughing) hyena' (**more properly cited as /ʔudzame-ko/ 'spotted hyena')

The Hadza root is as similar both in terms of phonology and semantics to Kanuri (Saharan) *zazórma* 'leopard' (Cyffer 1994: 106) as it is to any of the Chadic forms proposed (**zVm-* 'lion', etc.). None of the forms attempt to show a correspondence to the initial *ʔu-* syllable in the Hadza form.

It is not difficult to find two words in even a very short wordlist that have a single repeated sound correspondence that are roughly comparable to the correspondence sets that Militarev proposes (i.e. where not all segments or even syllables have corresponding segments or syllables). For instance, in comparing Hadza with the Nigerian language Akye[8] (Benue-Congo, Plateau) (Decker et al. 2021), similar forms involving *k ~ k* include:

[8] This language was randomly chosen for comparative purposes.

Hadza	Akye[9]
yeke- /jeke-/ 'road, path'	*ókée* 'road, path'
/ʔaku/ 'what kind, which'	*āké* 'what?'
/k^{w}a‖^{h}a/ 'to vomit'	*kwɛ̀* 'vomit'
/kuku, k^{w}ak^{w}a/ 'to dig with the hands'	*kɔ̃ŋ* 'dig'

Even more pairs could be listed if the semantic correspondences were stretched. Additional similar forms could be added if we violate the *k~k* pseudo-correspondence. We are definitely not trying to claim that Hadza and Akye might be related; rather, that lexical similarities occur due to chance quite often. This is especially true when the segment inventories and phonotactics of the languages are similar.

8. Discussion

Militarev's proposal of Hadza as an Afroasiatic language (with or without a particularly close connection to Chadic, Cushitic or Omotic) concerns itself with linguistic data but is not an ecologically or historically-situated proposal. By this, we mean that there are no proposed times or locations of a Hadza-Afroasiatic homeland. Although Militarev cites Ongota data to support Afroasiatic etymologies, he makes no reference to Fleming (2006) and the reconstruction of prehistory proposed therein.

Militarev's proposal raises many more questions than it answers. If Hadza is Afroasiatic, then does that mean that Proto-Afroasiatic should not be reconstructed as a language of agriculturalists/pastoralists? If Hadza is more closely connected to Chadic than to other branches of Afroasiatic, then did the Hadza people migrate southwards (e.g. in a scenario similar to that proposed by Blench 2013b)? There is no explanation as to why foragers would migrate (or why non-foragers would have resided further south at a time when there is no archaeological trace of their habitation). It is unclear whether Militarev considers Hadza a separate branch of Afroasiatic (similar to Fleming's (2006) analysis of Onogota), a branch of Chadic, or other. Regardless of which of these scenarios obtains, later contacts with Cushitic should be expected and not treated as independent sources of evidence for Afroasiatic affiliation.

Early loans between Afroasiatic languages and Hadza (pre-dating contact with Proto West Rift / Southern Cushitic) undoubtedly occur. We know that clicks occur in Dahalo, a Cushitic language formerly spoken by hunter-gatherers, even though these consonants have never been reconstructed for Cushitic or Afroasiatic, so the direction of borrowing need not be assumed to be from Afroasiatic into Hadza but the reverse may also have occurred. Hunter-gatherer groups existed throughout Kenya, Tanzania and SW Ethiopia and as people shifted from a pre-existing language of foragers to an Afroasiatic language, they may have retained vocabulary items, or acquired vocabulary items through later, continued contact with foragers. Since obsidian found near Lake Eyasi (where Hadza is spoken) can be traced to Central Kenya (Goldstein 2022), it is not outrageous to suggest that the area in which a Hadza-like language was spoken may once have been much larger, or that Hadza people once traveled more widely, especially following the wildebeest migrations northward. It is also quite likely that some words may be shared by groups that are not directly connected but which were connected through an intermediary language. Certainly, it should be unsurprising to find Wanderwörter shared by languages that today appear to have no contact at all.

[9] Forms cited are those of Kiguna village.

Our alternate analysis is that 1) there are chance resemblances between Hadza and Afroasiatic; 2) there are some loans from Afroasiatic into Hadza, 3) there are later loans from Cushitic into Hadza; 4) there are Hadza (and Sandawe) loans into Cushitic. That borrowing from preexisting groups into newcomer languages occurred is evident in more recent examples of borrowings from Hadza or Sandawe into Cushitic (see Ehret 2013b and Kießling & Mous 2012). Early contact between Hadza and Afroasiatic is not sufficiently explored; even if we entertained the notion that Hadza were Afroasiatic, there would still be a need to distinguish contact from inherited items.

It is hard to critique a paper that does not propose a concrete scenario. For example, #24 'egg' is compared to two different reconstructed Chadic roots, **ŝay(ŝay)-* and **ʔi(n)ŝ-*, and the path between these roots and the Hadza is not made explicit. #31 'foot' is compared to forms in Chadic, with metathesis, and to South Cushitic **fanq^{w}-* (therein **fanḳw-*) – both forms ignoring the initial syllable **ʔa**phukwa or **ʔu**phukwa. There are three roots reconstructed with the meaning 'all' in the Afroasiatic database (AADB) but none match the forms proposed to connect to the Hadza form *wa'i* with that meaning.

In his discussion, Militarev suggests that Hadza may be parallel to Cushitic and Omotic as a sub-branch of South Afrasian but continues to stress the striking similarities with Chadic. The link with Chadic is probably more surprising and indeed tantalizing. It would be interesting to attempt to develop that idea by comparing Hadza to Proto-Chadic. We have the impression that Proto-Chadic is more often than Proto-Cushitic suggested in the proposed Hadza-Afrasian cognate sets. Working with reconstructed languages rather than individual languages in these proposals for cognate sets has the methodological advantage of reducing chance resemblances.

9. Conclusion

In this paper we have critiqued many of the lexical comparison sets proposed by Militarev to provide support for a linguistic genetic relationship between Hadza and Afroasiatic. We argue that some of these sets are better explained as being due to contact with other languages and others may be discounted because they do not accord with a Hadza-internal analysis. Of the sets which remain, none show more than two repeated, regular sound correspondences; this falls under the level of proof needed to secure a label of relatedness. Given that Hadza has many consonants, the number of lexical items with any one consonant or consonant type is relatively low compared to languages with fewer consonants; this means that it is relatively difficult for the language to have retained a large number of words with each consonant over a great time depth. Just as it can be difficult to recognize old, phonologically-assimilated loans from Latin, etc. in Basque and Berber (cf. Trask 1996, Kossmann 2013), so too, is it difficult to recognize early loans in Hadza – particularly since the donor languages are not anywhere near as well-attested as Latin. The challenge to *prove* that Hadza might be related to another language is an attractive but daunting one. We hope that we have outlined some of the types of errors that can befall the intrepid linguist willing to take a chance on exploring such little-known ground.

References

Aklilu Yilma, Ralph Siebert, Kati Siebert. 2002. Sociolinguistic Survey of the Omotic Languages Sheko and Yem. *SIL Electronic Survey Reports (SILESR)*, 2002-053.

Alemayehu Abebe. 2002. Ometo Dialect Pilot Survey Report. *SIL Electronic Survey Reports (SILESR)*, 2002-068.

Bastin, Yvonne, André Coupez, Evariste Mumba, Thilo C. Schadeberg (eds.). 2002. *Bantu Lexical Reconstructions 3*. Tervuren: Royal Museum for Central Africa.

Blench, Roger. 2013a. Linguistic aspects of Hadza interactions with animals. In: Alena Witzlack-Makarevich, Martina Ernszt (eds.). *Khoisan Languages and Linguistics (Proceedings of the 3rd International Symposium July 6-10, 2008, Riezlern/Kleinwalsertal). Quellen zur Khoisan-Forschung / Research in Khoisan Studies, 29*: 101–110. Köln: Rüdiger Köppe.

Blench, Roger. 2013b. Links between Cushitic, Omotic, Chadic and the position of Kujarge. In: Marie-Claude Simeone-Senelle, Martine Vanhove (eds.). *Proceedings of the 5th International Conference on Cushitic and Omotic Languages, Paris, 16-18 April 2008 (Cushitic and Omotic Studies, 3)*: 67–80. Köln: Rüdiger Köppe.

Campbell, Lyle, Terrence Kaufman, Thomas C. Smith-Stark. 1986. Mesoamerica as a linguistic area. *Language* 62(3): 530–570.

Chebanne, Andy. 2017. Common Khoisan: A reflection on their linguistic and cultural relationships. In: Kwesi Kwaa Prah, Lazarus Miti (eds.). *Deconstructing the African Tower of Babel: Between the Harmonisation and Fragmentation of African Language Orthographies (CASAS Book Series, No. 120)*: 95–116. Cape Town: Centre for Advanced Studies of African Society.

Coburn, Jeremy, Bonny Sands, Andrew Harvey, Richard Griscom (forthc.). Tonal Patterns of Hadza nouns. In: *The Proceedings of the 10th World Congress on African Linguistics, Leiden, 7-11 June 2021 (Contemporary African Linguistics)*. Berlin: Language Science Press.

Cyffer, Norbert. 1994. *English-Kanuri dictionary* (*Westafrikanische Studien: Frankfurter Beiträge zur Sprach- und Kulturgeschichte, Bd 3*). Köln: Rüdiger Köppe.

Decker, Ken, Innocent Jonah, John Muniru, Yakubu Danladi, Benard Abraham, David Onoja, Christina Riepe. 2021. A Sociolinguistic Profile of the Akye (Ake) [aik] Language of Nasarawa State, Nigeria. *Journal of Language Survey Reports* 2021(042): 1–51.

Ehret, Christopher. 2013a. Sub-Saharan Africa: Linguistics. In: Immanuel Ness, Peter Bellwood (eds.). *The Encyclopedia of Global Human Migration. Vol. I: Prehistory*: 96–106. Malden, MA / Oxford: Wiley-Blackwell.

Ehret, Christopher. 2013b. The Extinct Khoesan languages in East Africa. In: Rainer Vossen (ed.). *The Khoesan Languages*: 465–479. New York: Routledge.

Elderkin, Edward Derek. 1978. Loans in Hadza: Internal evidence from consonants. *Occasional Paper 3*, Dept. of Foreign Languages, Dar es Salaam. University of Dar es Salaam.

Elderkin, Edward Derek. 1982. On the classification of Hadza. *Sprache und Geschichte in Afrika* 4: 67–82.

Elderkin, Edward Derek. 2014. Clicks, prosodies and Khoisan. In: Tom Güldemann & Anne-Maria Fehn (eds.). *Beyond 'Khoisan': Historical Relations in the Kalahari Basin*: 103–122. Amsterdam: John Benjamins.

Fleming, Harold C. 2006. *Ongota: A decisive language in African prehistory*. (*Aethiopistische Forschungen, 64*). Wiesbaden: Harrassowitz.

Goodman, Morris. 1971. The strange case of Mbugu. In: Dell Hymes (ed.). *Pidginization and Creolization of Languages*: 243–254. Cambridge University Press.

Gordon, Matthew J. 1995. The phonological composition of personal pronouns: implications for genetic hypotheses. *Berkeley Linguistics Society* 21: 117–128.

Griscom, Richard. 2019. *Topics in Asimjeeg Datooga Verbal Morphosyntax*. Ph.D. thesis, University of Oregon.

Griscom, Richard, Andrew Harvey. 2021. *Hadza: an archive of language and cultural material from the Hadzabe people of Eyasi (Arusha, Manyara, Singida, and Simiyu regions, Tanzania). Endangered Languages Archive*. Available online at: hdl.handle.net/2196/82e2b99d-5c62-4210-8903-8dd976337c10.

Güldemann, Tom. 2014. 'Khoisan' linguistic classification today. In: Tom Güldemann & Anne-Maria Fehn (eds.). *Beyond 'Khoisan': Historical Relations in the Kalahari Basin*: 1–41. Amsterdam: John Benjamins.

Harvey, Andrew 2019. *Ihanzu: an archive of language and cultural material from the Ihanzu people of Mkalama (Singida Region, Tanzania)*. London: SOAS, Endangered Languages Archive. Available online at: hdl.handle.net/2196/00-0000-0000-0014-1365-F.

Harvey, Andrew. 2021. *The lack of labiodental in Ihanzu as a result of contact with Hadza*. Talk given at the 10th World Congress of African Linguistics (WOCAL10), Leiden (Online) 10/06/2021. doi: 10.17605/OSF.IO/HNDC5

Honken, Henry. 2013. Genetic relationship: An overview of the evidence. In: Rainer Vossen (ed.). *The Khoesan Languages*: 13–24. New York: Routledge.

Hudson, Grover. 1989. *Highland East Cushitic dictionary*. Hamburg: Helmut Buske.

Kibebe Tsehay Taye. 2015. *Documentation and grammatical description of Chabu*. Addis Ababa: Addis Ababa University doctoral dissertation.

Kießling, Roland, Maarten Mous. 2003. *The lexical reconstruction of West Rift (Southern Cushitic)*. Köln: Rüdiger Köppe Verlag.

Kossmann, Maarten. 2013. *The Arabic Influence on Northern Berber. Studies in Semitic Languages and Linguistics*. Leiden: Brill.

Miller, Kirk (ed.), w. Mariamu Anyawire, G. G. Bala, Bonny Sands. 2013. *A Hadza Lexicon. Mang'ola, Tanzania.* Ms.

Militarev, Alexander. 2023. Hadza as Afrasian? *Journal of Language Relationship* 21(2): ???.

Mous, Maarten, Martha Qorro, Roland Kießling. 2002. *An Iraqw – English Dictionary. Cushitic Language Studies, 15*. Köln: Rüdiger Köppe.

Nurse, Derek, Thomas J. Hinnebusch. 1993. *Swahili and Sabaki, a Linguistic History*. *University of California Publications in Linguistics, 121*. Berkeley: University of California Press.

Peterson, Daudi (w. Richard Baalow & Jon Cox). 2013. *Hadzabe, By the Light of a Million Fires*. Dar es Salaam: Mkuki na Nyota Press.

Reinisch, S. Leo. 1902. *Die Somali-Sprache. Bd II: Wörterbuch Somali-Deutsch, Deutsch-Somali*. Vienna: Alfred Hölder.

Reinisch, S. Leo. 1895. *Wörterbuch der Bedauye-Sprache*. Vienna: Alfred Hölder.

Rzymski, Christoph, Tiago Tresoldi et al. 2019. *The Database of Cross-Linguistic Colexifications: reproducible analysis of cross- linguistic polysemies*. Available online at: clics.clld.org. doi: 10.1038/s41597-019-0341-x

Sands, Bonny. 1998. *Eastern and Southern African Khoisan: Evaluating Claims of Distant Linguistic Relationships*. *Quellen zur Khoisan-Forschung/Research in Khoisan Studies, 14*. Köln: Rüdiger Köppe.

Siebert, Ralph, Simon Caudwell. 2002. Sociolinguistic Survey Report of the Melo (Malo) and Mursi Languages of Ethiopia. *SIL Electronic Survey Reports (SILESR)*, 2002-046.

Starostin, George. 2008. K voprosu o geneticheskoj prinadlezhnosti jazyka hadza [On the issue of the genetic affinity of Hadza]. In: V. F. Vydrin (ed.). *Afrikanskiy sbornik* 2007: 262–278. Saint Petersburg: Nauka.

Starostin, Georgij S. 2013. *Jazyki Afriki. Opyt postrojenija leksikostatisticheskoj klassifikatsiji. T. I. Metodologija. Kojsanskije jazyki* [Languages of Africa: an attempt at a lexicostatistical classification. Volume 1: Methodology; Khoisan languages]. Moscow: Jazyki slav'anskoj kultury.

Takács, Gábor. 2011. Omotic lexicon in its Afro-Asiatic Setting I: Omotic *b- with dentals, sibilants, and velars. In: Luca Busetto, Roberto Sottile, Livia Tonelli, Mauro Tosco (eds.). *"He bitaney lagge": Studies on Language and African Linguistics in Honour of Marcello Lamberti*: 183–200. Milano: Qu.A.S.A.R.

Trask, Robert Lawrence. 1996. *The History of Basque*. London: Routledge.

Urban, Matthias. 2012. *Analyzability and Semantic Associations in Referring Expressions: A Study in Comparative Lexicology*. Ph.D. thesis, Leiden University.

Wedekind, Klaus. 2008. *Gedeo Dictionary, Revision of 1978, Listing by Semantic Categories*. Available online at: www.kwedekind.de/Eingang1/Biblio1.htm.

Wedekind, Klaus, Charlotte Wedekind, Abuzeinab Musa. 2007. *A Learner's Grammar of Beja (East Sudan). Grammar, Texts and Vocabulary (Beja-English and English-Beja). (Afrikawissenschaftliche Lehrbücher Volume 17).* Köln: Rüdiger Köppe.

Zorc, R. David, Madina M. Osman. 1993. *Somali-English Dictionary with English Index, 3rd edition*. Wheaton MD: Dunwoody Press.

Бонни Сэндс, Маартен Моус, Мауро Тоско, Эндрю Харви. Почему хадза (скорее всего) не афразийский язык: ответ на статью А. Ю. Милитарева «Хадза — афразийский язык?»

В настоящем ответе на статью А. Милитарева обсуждаются проблемы, связанные с использованием лексических данных для обоснования афразийского происхождения языка хадза. Некоторые из них, как, например, некорректный морфологический анализ ряда сложных форм и неопознанные заимствования из соседних языков, легко устранимы при более глубоком изучении внутренней структуры и лингвистических контактов хадза. Другие скорее связаны с выбранной методикой исследования; так, широкий разброс семантики сравниваемых слов увеличивает вероятность случайных сопоставлений, равно как и включение в сравнение слов, носящих звукоподражательный характер. Главной проблемой, не позволяющей признать гипотезу доказательной, тем не менее, следует считать отсутствие регулярных и рекуррентных звуковых соответствий между хадза и афразийскими языками.

Ключевые слова: языковые изоляты; хадза язык; койсанские языки; кушитские языки; афразийские языки; чадские языки; дальнее родство языков.