

Is Iraqw an easy language to learn?

Bert van Pinxteren

1 Introduction

Considering Maarten Mous' extensive contributions to the study of Iraqw, this volume would not be complete without an article on Iraqw. Maarten Mous is the world's foremost authority on the language, having written a grammar (1992) and a dictionary (2002, co-authored with Roland Kießling and Martha Qorro) and a number of other works, and still being involved in collecting Iraqw materials. Maarten has certainly learned Iraqw, even though, by his own admission, he does not speak it very well¹ - but that may be a case of misplaced modesty.

Learning to speak Iraqw, even if not spoken so well, can safely be considered an admirable achievement. In 2013, Tyler Schnoebelen tried to find out which languages in the world are the weirdest - operationalised as being the most different from all other languages. He used the World Atlas of Language Structures as a basis and compared languages using 21 structural features.² On Schnoebelen's list, Iraqw proudly takes up 7th position, roundly beating other weird languages such as German (in 10th place) and Dutch (in 12th).

Of course, all these classifications, for whatever they are worth, are potentially valuable only for the hypothetical visitor from Mars, who when visiting Earth may want to try first to learn the language that is most typical for us humans (which, according to Schnoebelen, would in fact be Hindi). But for human beings, learning a language as an infant is easiest, and learning additional languages becomes progressively more difficult with age. So, for an Iraqw child, there is nothing weird or difficult about Iraqw – it seems the most natural thing in the world. Therefore, in principle, the question of whether or not Iraqw (or any other language) is easy or difficult depends on who you ask.

However, ease or difficulty of language learning is not only a matter that is of relevance to individuals; it also has a sociolinguistic relevance, especially when considered in an African context. In some countries, such as Tanzania, an African language is used in all of primary education. In most others, African languages are used only for the first few years of primary education,

1 http://www.let.leidenuniv.nl/forum/04_4/personalia/1.htm (Accessed 15 October 2019).

2 The original blogpost is no longer online, but can still be found through the web archive: <https://web.archive.org/web/20130809075053/http://idibon.com/the-weirdest-languages/> (Accessed 15 October 2019).

leading to an ‘early exit’ model (Heugh, 2011). In almost all of Africa, a former colonial language is used as medium of instruction in secondary and higher education. How sustainable is that model with rising enrolment figures? It might be relevant to consider whether or not and at what point a switch to a language that would be easier to learn may become necessary.³

But is it possible to be more precise in assessing ease or difficulty of language learning? In order to discuss that, Section 2 takes a brief look at the concept of ‘language’. Section 3 looks at levels of language learning and section four at the ease or difficulty of learning a language to a specific level. In that section, I propose a new way to approximate the ease or difficulty of language learning for a specific combination of languages. In order to do that, I propose an innovative way of using the Automated Similarity Judgement Program (ASJP) and its associated database.⁴ I do that by benchmarking it to a US government scheme for determining ease or difficulty of language learning for specific language pairs. In Section 5, I argue why there is a wider relevance to such a scheme: in Tanzania as well as in other parts of the world, educational systems are increasingly entrusted with the task of teaching large sections of the population a second or third language. However, what can we expect an educational system to achieve? I advance the idea that it is relevant to weigh knowledge about the ease of language learning when setting expectations of and defining policies for second or third language teaching. Section 6 then returns to Iraqw and to the Tanzanian educational situation and leads to a number of recommendations that are relevant for the Tanzanian situation. The last brief Section 7, offers a summary and some last conclusions.

2 What is a ‘language’?

Before going to the heart of the matter, it is first useful to clarify my position on what I mean when I write about ‘language’. In the literature, there are different opinions.

One approach is taken for example by Blommaert (2013) and by Lüpke and Storch (2014). They criticize the use of the term ‘language’ as if languages were clearly demarcated, countable and bounded objects. Instead, they look at how people actually speak and how they use language in different settings and circumstances. Instead of languages, they prefer to talk about ‘languoids’ or ‘registers’ people use when speaking. Seen in that light, because everybody speaks slightly differently and how individuals speak also varies depending on the circumstances, it makes no sense to enumerate languages. Applied to Iraqw, ‘learning the language’ is an impossibility – what is possible at best is to master a register that allows for a given level of communication with given

³ For a fuller discussion of this, see Van Pinxteren (2018).

⁴ <https://asjp.clld.org/>

other speakers of the family of registers that together perhaps belong to an Iraqw family of languoids or registers.

For the purposes of this contribution, I will adopt a different position. I do not think it is necessary or particularly helpful to stop using the term ‘language’. Even though everybody uses language in a slightly different manner, communication is still possible between those who speak, for example, Iraqw, and Iraqw speakers cannot use their language to communicate with non-speakers. That being said, the borders can be fuzzy: at the borders between two language areas, it may not always be straightforward to decide which dialect belongs to which language or indeed whether or not a dialect should be considered in fact a different language. So, although the borders may be fuzzy, there are still borders.

What interests me is the position of languages as social phenomena – as instruments of power. Consider the metaphor of road building: in olden times, roads (or paths) emerged naturally, as a result of people walking from A to B along a similar route. But little by little, roads became the preserve of engineers and planners, from the army routes in Roman times to the highways of modern times. In the same way, languages originated by people talking to one another, but gradually evolved into complex social constructions, planned, maintained and built out using elaborate mechanisms. Thus, in France, the ‘Alliance Française’ is an institution specifically set up to promote and protect the French language.⁵ Internationally, the *Organisation Internationale de la Francophonie* serves the same purpose.⁶ In the English-speaking world, hosts of style guides and armies of editors work tirelessly to keep English standardized and understandable for ever larger numbers of people. For Iraqw, there exist only few written materials: a grammar and dictionary as mentioned above, parts of the Bible,⁷ a story book (Mous and Sanka, 2017) – and that is just about it. Iraqw is not taught in schools and is not a study topic at any institution of higher learning in Tanzania. This means that Iraqw as a language is closer to the paths of old than to a modern road, to use that metaphor.

3 What does ‘speaking a language’ actually mean?

If we wonder whether or not Iraqw is an easy language to learn, we also need to be specific about the goal: when can we be satisfied that a language has indeed been learned? When do we say that a person ‘speaks’ a language? Is

⁵ <https://www.fondation-alliancefr.org/?cat=536> (Accessed 22 October 2019).

⁶ <https://www.culture.gouv.fr/Sites-thematiques/Langue-francaise-et-langues-de-France/Politiques-de-la-langue/Multilinguisme/Francophonie> (Accessed 26 April 2020): ‘les francophones peuvent s’appuyer sur un dispositif institutionnel voué à promouvoir la langue française’ – ‘Francophones can rely on an institutional mechanism dedicated to promoting the French language’ (author’s translation).

⁷ <https://live.bible.is/bible/IRKBST/MAT/1> and https://archive.org/details/rosetta_project_irk_gen-1/mode/2up (Accessed 26 April 2016).

somebody able to order a beer in a specific language a speaker of that language? Is speaking enough, or do we also look at reading and writing skills? What level of proficiency do we aim for? Fortunately, there is a possibility for being a bit more precise here and there are three main systems for doing that.

In the U.S. there are two, related scales for assessing language proficiency: the guidelines of the American Council on the Teaching of Foreign Languages (ACTFL)⁸ and the scale of the Interagency Language Roundtable (ILR). The ILR scale has five broad levels, ranging from zero to five.⁹ In Europe, the Council of Europe has produced the Common European Framework of Reference (CEFR) scale, which has six levels.¹⁰ Attempts have been made to relate the U.S. and the European scales to one another, but these alignments are approximate at best. To be so proficient in a language that one could follow tertiary education¹¹ in that language, a level corresponding to at least the CEFR B2 level is considered to be necessary. This level stands for 'upper intermediate'. Above it, there are still two more levels, the C1 and the C2 level.

Unfortunately, in statements about the number of speakers of a certain language in Africa or in African countries, the level is often not specified. But it is possible to say a bit more about this for the francophone part of Africa. The figures of who speaks a language are the most comparable for the African 'francophone' nations. Albaugh (2014: 221) gives the average of 18% of French-speakers. These figures are somewhat comparable, because Albaugh is able to use data collected by the *Organisation Internationale de la Francophonie* (OIF).

The OIF uses what it calls a common-sense definition of a francophone in Maurer (2015: 3): 'a person able to express him/herself in French, no matter what his/her level may be or his/her mastery of other competences such as writing or reading.'¹² It is not immediately obvious how to map this on to the CEFR levels, but it clearly includes many more people than those who are at the B2 level, which stands for an ability to 'interact with a degree of fluency

8 <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012> (Accessed 20 July 2019).

9 <https://www.govtilr.org/> (Accessed 20 July 2019).

10 <https://www.coe.int/en/web/common-european-framework-reference-languages/home> (Accessed 20 July 2019).

11 Tertiary or higher education comprises levels 5 through 8 as defined in the UNESCO International Standard Classification of Education (ISCED) scheme as developed by UNESCO (2012). Note that in this scheme, 'tertiary education' refers to more than what is commonly understood as university education: it also includes education for example by polytechnics at the higher vocational level.

12 Author translation. Original: 'Revenons donc au sens commun, qui entent par « francophone » une personne capable de s'exprimer en français, qu'elle que soit son niveau ou sa maîtrise d'autres compétences comme l'écriture ou la lecture.' This definition is different from for example the Collins English definition of a Francophone: 'A **Francophone** is someone who speaks French, especially someone who speaks it as their first language.' <https://www.collinsdictionary.com/dictionary/english/francophone> (Accessed 16 October 2019).

and spontaneity that makes regular interaction with native speakers quite possible without strain for either party'.¹³ Probably, because one of the interests of the OIF is to support the importance of the French language, it uses a wide definition of who is a francophone, which in all likelihood would include also those who are at the A2 level: those able to 'describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.' It is illustrative to note Maurer's assessment that in 2007, 62.4% of teachers in Francophone Mauritania had level A2 or below, a level Maurer considers to be insufficient for teaching in that language (Maurer 2015: 31). Based on this discussion, it is my assumption that the percentage of Francophones in the francophone nations of Africa who can speak the language at B2 level is not more than 9%, and those who have C2 level form an even smaller group.

This discussion of how I use the word 'language' and what is meant by 'speaking a language' enables a closer look at the issue of easy and difficult languages.

4 Which languages are easy to learn – and for who?

How can we know if a language is 'easy'? Most people who have tried to learn more than one additional language have some idea of how difficult or easy it is to do that. But *Ethnologue*¹⁴ discerns more than 7,000 languages the world over. Is there any way of approximating which languages are more easy or more difficult for L1 speakers of any particular language?

4.1 Measuring

The literature usually approaches this question from the point of view of the individual learner: it takes a hypothetical learner and asks what factors influence how and how quickly and to what extent he or she can learn a new language – see, for example Klein (1986). Factors that are mentioned include motivation, aptitude, but also the similarities or differences in sounds, grammar and vocabulary between the languages the learner already knows and the new language. From a sociological point of view, and related to Africa, we see case studies of individual countries, summarized for example in Skattum (2018). However, the question of what ease or difficulty of language learning means for large groups of learners and for an education system has not been asked in the literature in that way. Yet, this is a question of key relevance for Africa, where populations are supposed to be taught in a language that most learners do not speak from birth.

13 <https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale> (Accessed 7 September 2019).

14 <http://www.ethnologue.com>

Common sense suggests to start from the principle that languages that are close to one another are easier to learn and to be taught in formal education than languages that are very different from one another. In other words, the **distance** between any two languages can be taken as an indicative or rough measure for how easy or difficult it may be to learn another language for a speaker of a given language or to teach the new language to large groups of learners.

One way of approximating this is by looking at language families: speakers of Germanic languages may find it easier to learn other Germanic languages than, for example, Bantu or Cushitic languages. But not all language families are the same: some are internally more diverse than others and even within language families, some languages may be more similar to one another than others. Can we somehow get some form of measurement that serves as a more precise indicator of how similar or different languages are from one another?

The measure of linguistic distance as an indicator of ease or difficulty of language learning has advantages, as will be shown further down, but it also has limitations. A first limitation is that the relationship is not necessarily bidirectional: it may be easier for somebody who speaks language A to learn language B than the other way around. This can happen for example if language A has more sounds (phonemes) than language B and there are no sounds in language B that do not also occur in language A. In that case, speakers of language B will have to familiarize themselves with the new sounds that language A has, but speakers of language A do not have that problem if they want to learn language B. The same can be true for the grammar of a language: if language A has a more difficult or strict grammar than language B, it may be easier for speakers of language A to learn language B than vice versa.

A second limitation is that there are various ways of measuring the distance between languages, all of them with their own problems and imperfections. Ginsburgh and Weber (2016) give a useful overview of ways that have been found of measuring linguistic distance.

One way they describe is by comparing languages in terms of their relatedness to a common ancestor, starting from the idea that there once existed one language and that all existing languages have branched off from that common root. By counting the number of 'branchings', the distance between languages can be computed, in the same way that family distances are traced through the distance from a common ancestor. This is often called cladistics (Ginsburgh and Weber 2016: 142).

Another way of measuring distance between languages is through lexicostatistical methods. These methods are based on comparing the common roots of words in the vocabularies of various languages. These comparisons are based on a limited list of words that are assumed to exist in almost all languages with the same meaning. The most famous of these lists is the one developed by the American linguist Swadesh, last published in

1971. Ginsburgh and Weber (2016: 148) suggest using the method first proposed by Levenshtein in 1966 as a way of using these word lists for comparing distances between languages. This is done by computing the number of changes that need to be made to turn one word (such as the English word ‘night’, but spelled phonetically) into its equivalent in another language (such as ‘nuit’ in French, also spelled phonetically).

The cladistic method has the advantage of taking into account more than just vocabulary. However, it relies on a classification of language families that is imprecise at best, often itself relying on the comparison of limited wordlists, and therefore gives only very rough results (Ginsburgh and Weber 2016: 149).

The most precise and most comprehensive tool for computing language distance currently available uses a lexicostatistical method with a simplified 40-item word list derived from the Swadesh list and using Levenshtein distances. This method was developed in an experimental way and its results were compared with the expert knowledge of relevant linguists and refined based on their feedback. This has led to the Automated Similarity Judgment Program (ASJP) and its associated database, started in 2008.¹⁵ It currently contains word lists from 5,067 languages and is able to compute the degree of similarity between any pair of these languages.¹⁶

4.2 Benchmarking the ASJP database

Without further work it is unclear what the ASJP distances mean in terms of ‘ease’ or difficulty of learning a language. In order for them to have practical relevance for predicting ease of language learning, it is necessary to benchmark them against a schema for language learning and to see if there is any relationship between the ASJP scores and such a schema. The schema I propose to use is one that has been developed by the US Government.¹⁷ This schema gives a list of language pairs (mostly involving English as one member of the pair), giving for each pair the number of weeks of full-time formal instruction needed for a talented native speaker to reach the IRL S-3/L-3 proficiency level in a given other language. The S3/R3 level is equal to basic ‘vocational’ proficiency, roughly equivalent to the CEFR C1 level.

15 <https://asjp.cld.org/> (Accessed 23 July 2019). The ASJP is the product of a considerable amount of research; for more information, their Wiki page is a good starting point: https://en.wikipedia.org/wiki/Automated_Similarity_Judgment_Program.

16 ASJP website, 23 July 2019. Note that SIL is the registrar for an ISO norm that tries to list all of the world’s languages, ISO 639-3; this can be seen as listing all the more than 7,000 currently discerned languages of the world. See <https://iso639-3.sil.org/about> for more information.

17 <https://2009-2017.state.gov/documents/organization/247092.pdf> (Accessed 21 July 2019). There is also a (less elaborate) British schema: <http://www.baylanguages.com/language-scale> (Accessed 29 April 2020).

In some cases, it also gives the number of weeks needed to give a student who already speaks a certain language the same level in a related language.

The mapping looks like this:

Table 1
Ease of language learning.

Language pair	Weeks	US classification	ASJP score	My classification
Czech – Slovak	10-12	Closely related	32	Very easy
Bulgarian – Macedonian	10-12	Closely related	32	Very easy
Indonesian – Malay	10-12	Closely related	15	Very easy
Lao – Thai	14-18	Related	53	Very easy
Portuguese – Spanish	14-18	Related	68	Easy
Dutch – German	18-22		49	Very easy
Bulgarian – Serbo-Croatian	30-36		48	Very easy
English – Dutch	24	Cat I	61	Easy
English – Italian	24	Cat I	90	Medium
English – French	30	Cat I	92	Medium
English – German	36	Cat II	69	Medium
English – Haitian Creole	36	Cat II	94	Medium
English – Swahili	36	Cat II	97	Difficult
English – Amharic	44	Cat III (hard)	96	Difficult
English – Hausa	44	Cat III (hard)	98	Difficult
English – Somali	44	Cat III (hard)	103	Very difficult
English – Japanese	88	Cat IV (super hard)	98	Difficult
English – Korean	88	Cat IV (super hard)	99	Difficult

English Mandarin	-	88	Cat IV (super hard)	102	Very difficult
---------------------	---	----	------------------------	-----	----------------

As is clear from the table, the US Government-based classification and my classification based on ASJP scores do not provide an exact match, but they are still reasonably close. The difference between the two systems is never more than one adjacent category. What is also clear is that the ASJP scores do not form a scale with equal distances between points: at the higher end of the scale, the difficulty level increases faster than at the lower end of the scale.

I propose the following classification:

Table 2
Categories of ease of language learning.

ASJP distance score	Category
< 60	Very easy
≥ 60, < 90	Easy
≥ 90, ≤ 95	Medium
> 95, < 100	Difficult
≥ 100	Very difficult

This benchmarking and categorization has the advantage that it leads to an approximate assessment of the ease or difficulty of learning a language for any language pair in the ASJP database. It can therefore be used in order to make a rough assessment of the equitableness and inclusivity of a given language regime. This can be done without any knowledge of the actual languages. However, it is not more than a rough assessment: any assessment of this type would have to be validated against the expert knowledge of local speakers and learners of the languages involved.

Table 3
ASJP database output for Dutch-like languages (figures rounded to the nearest whole).¹⁸

2 SYNONYMS, AT LEAST 28 WORDS		DUTCH		ENGLISH		FRANS_VLAAMS		FRISIAN WESTERN		GRONINGS		LIMBURGISH		SALLANDS		TWEENTS		WEST_VLAAMS		ZEEUWS	
LOANWORDS EXCLUDED	LDND	BRABANTIC	DUTCH	ENGLISH	FRANS_VLAAMS	FRISIAN WESTERN	GRONINGS	LIMBURGISH	SALLANDS	TWEENTS	WEST_VLAAMS	ZEEUWS									
BRABANTIC	0																				
DUTCH	44	0																			
ENGLISH	66	61	0																		
FRANS_VLAAMS	43	38	61	0																	
FRISIAN WESTERN	57	52	67	59	0																
GRONINGS	58	38	71	58	56	0															
LIMBURGISH	54	46	66	57	58	44	0														
SALLANDS	61	35	69	50	62	40	54	0													
TWEENTS	63	47	63	59	64	48	37	0													
WESTVLAAMS	50	43	64	29	65	55	48	60	0												
ZEEUWS	47	37	67	27	61	51	56	48	59	34	0										
BRABANTIC	DUTCH	ENGLISH	FRANS_VLAAMS	FRISIAN WESTERN	GRONINGS	LIMBURGISH	SALLANDS	TWEENTS	WEST_VLAAMS	ZEEUWS											

18 LDND in these tables stands for the Normalized Levenshtein Distance divided between two languages, as defined by Bakker et al (2009).

To illustrate the power of this approach, let's consider the example of the 10 Dutch-like languages contained in the ASJP database for the Netherlands (plus English). As can be seen from Table 3 above, standard Dutch is very easy to learn for all speakers of Dutch-like languages (whereas English is marginally more difficult, falling into the 'easy' category). It makes sense that in the Netherlands, Dutch is used as a common language of instruction. This may also help to explain why the Dutch are often praised for their generally good command of the English language – it is an easy language for them to learn.

4.3 Easy and difficult languages – does it matter?

It is good to recall that formal education in a specific language always involves learning in a medium that is different from the spoken word: it involves learning an 'intellectualized' language,¹⁹ a form of language that will always be in some degree different from the language spoken at home. For some children, this is too difficult in any language: according to a meta-analysis by McKenzie et al (2016), about 1% of all children are estimated to suffer from intellectual disability (although there is quite a bit of uncertainty about this figure). Intellectual disability means that these children either cannot learn to read or write or if so, only to a very limited level. This means that around 99% of all children can acquire basic reading and writing skills. The number of people who are or can become 'literate' in the sense that they can no longer be considered 'functionally illiterate' is smaller. Functional illiterates, according to Schlechty (2004: 7) are those people who have reading and writing skills that are inadequate 'to manage daily living and employment tasks that require reading skills beyond a basic level'. According to the UK National Literacy Trust, one out of six (16.4%) of all adults in England are functionally illiterate.²⁰ In 2019, the World Bank launched a new indicator, dubbed 'learning poverty': 'Learning poverty means being unable to read and understand a simple text by age 10' (World Bank 2019: 6). Even in high-income countries, there is a percentage of children in this category – fewer than 10%; so even with the best education available, not all children can learn to read and understand a simple text.

The inconvenient truth is that learning abilities are not equally divided over the population: some people are more intelligent than others. Tests have been calibrated so that the average IQ is 100 – 50% of all children are supposed to have 'average' intelligence. At the upper extreme, just over 2% of the population score 130 or above. So, some children learn more quickly than others.

19 Prah (2017: 216) quotes the definition of Sibayan from 1999: an intellectualised language is a 'language which can be used for educating a person in any field of knowledge from kindergarten to the university and beyond'

20 <https://literacytrust.org.uk/parents-and-families/adult-literacy/> (Accessed 8 October 2019).

For language learning, it is important to note that IQ is not one-dimensional: there are different, although interrelated, forms of intelligence. Li (2016) has shown that language aptitude is a valid construct. This construct is related to, but independent of general intelligence. What this means is that some children may be good at language, but hopeless in math. For others, it may be the other way around. Then also, of course, some children are good at both.

For Africa, these simple facts have tremendous consequences, although they are usually overlooked. For the top end of the intelligence scale, historical experience has shown that in a way, it does not matter what the language of instruction is. In Europe, elite education for a long time was in Latin or in other languages other than the mother tongue. In India, Sanskrit has been used as a medium of instruction for centuries. All over the Arab world, the classical Arabic used for instruction is very different from the spoken languages. And in Africa, some great intellectuals emerged, in spite of the use of colonial languages.

For the bottom end of the intelligence scale, language of instruction does matter. Almost everybody is able to master some words in a foreign language. However, learning a foreign language to a level high enough to be able to profit from more and more advanced instruction in that language takes time and effort. Lower language aptitude means that more effort is required to reach the same level. Under colonial education systems, this was not considered relevant: education was aimed at selecting those most talented and reached only a minor proportion of the population. Therefore, under colonial education systems, it was not necessary to worry about the language of instruction – any language of convenience could serve equally well. However, if education is supposed to reach larger portions of the population, then the ease of learning the medium of instruction does become relevant: educational systems that use a medium that is close to the mother tongue of the learners will always be more efficient than those that do not.

The question becomes, therefore, what one can realistically expect an educational system to achieve. In theory, given infinite resources, it might be possible to teach almost anybody almost anything. However, in practice resources are never infinite. The question what to expect of an educational system can be broken down into two questions:

- A) What percentage of the population can an education system educate to a reasonable level of proficiency in a foreign language?
- B) What percentage of the population can an education system provide with secondary or higher education?

If the percentage under A is greater than that under B, there will be no problem in providing secondary or higher education in a foreign language. If, on the other hand, the percentage under B is greater than that under A, foreign-language secondary or higher education will not be an option for everybody.

This theoretical question has never been asked, in part probably because for most parts of the world, it is not relevant; most developed countries have developed tertiary education systems that offer at least parts of the curriculum in a language that is close to the mother tongue of most learners in that country. Another reason why it has not been asked is probably because finding the answer to the question may not be straightforward. Yet, for Africa, a continent that uniformly relies on foreign languages for secondary and higher education, this is a key question. It could be that at the moment, there are not enough places in higher education to accommodate everybody who has the required language level. But in the future, the situation may be reversed. In any case, it is clear that from a sociological and economic point of view, those educational systems that provide education using a medium of instruction that is close to the mother tongues of the learners will be the most efficient.

What does this mean for Tanzania and for Iraqw?

5 Is Iraw an easy language to learn – and for whom?

Now we have come to a point where it is possible to give an approximate but still meaningful answer to the question of how easy or difficult it is to learn Iraqw, using the benchmarked ASJP output for a number of relevant languages, as contained in Table 4 below.

A first thing that is worth noting is that the data suggests that Iraqw is indeed a very difficult language to learn for Dutch L1 speakers. However, this is in fact true of any language spoken in Tanzania (and vice-versa); Swahili is not any easier for Dutch L1 speakers than Iraqw. And considering the relative difficulty to learn Iraqw for a Swahili L1 speaker, having a command of Swahili did not make it much easier for Maarten Mous to then learn Iraqw.

A second point to note is how standard Swahili appears in fact to be a logical choice as language of instruction in Tanzania. It is not the only choice that would have been possible from a theoretical point of view: Sukuma, for example, would be just as easy. Nevertheless, for speakers of one of the Tanzanian Bantu languages, Swahili is either very easy or at least easy to learn (and much easier than English). This point may seem obvious, but it is not. Thus, a respected linguist such as Blommaert (2014: 63) erroneously follows Khamisi (1974) by assessing that the consequence of the nationwide adoption of Kiswahili was “that most children in non-urban inland areas, where Swahili was not the mother tongue of the population, were faced with a ‘hidden’ language barrier when they entered primary school: that of Swahili. For them, Swahili was as foreign a language as English”. This point could be read in two ways. One could say that the sentence postulates a dichotomy between mother tongues and all other languages – the ‘foreign’ languages. In such a reading, English and Swahili are both equal in the sense that they fall in the category of ‘foreign’ languages. One could also say that there is a distinction of degree: some languages could be more ‘foreign’ (or more difficult to learn) than others. My benchmarking of ASJP distances suggests that this second reading is more appropriate: some languages are a good deal more ‘foreign’. For the great majority of Tanzanians (though less so for the Iraqw), Swahili is far easier to learn than English.²¹ In terms of the U.S. schema discussed above, it is the difference between 15 weeks of full-time instruction for a linguistically gifted person and 80-plus weeks for such a person.²²

Africa is replete with failed language policies: either they were never implemented at all, staying letters on paper (Bamgbose 2000), or they were implemented but failed and later repealed. Chaudenson (2006) discusses several examples. There are also nuanced cases. Thus, Altinyelken et al (2014) discuss the example of Uganda, where the policy of using English as medium of instruction in urban areas leads to a lower status of indigenous languages

21 This also has consequences for either using English or Kiswahili in secondary and tertiary education: most Tanzanians will feel far more comfortable using Kiswahili, rather than English. This also means that teaching in Kiswahili would be more cost-effective than teaching in English. For an interesting perspective on this by an education practitioner, see <https://www.thecitizen.co.tz/magazine/success/1843788-4009890-ygpdwpz/index.html> (Accessed 8 August 2019).

22 These figures serve to illustrate the order of magnitude we should be thinking about when comparing between easy or difficult to learn language pairs. Precise values for specific language pairs and specific countries and educational systems can only be established through further research.

and undermines the policy of using them as medium of instruction outside of cities. If implementing a language policy in Africa is so problematic, why could Swahili 'stick' in Tanzania? Blommaert (2014) shows how the promotion of Swahili was coupled with the state ideology of 'Ujamaa'. This ideology, he argues, was based on overly simplistic and utopian ideas of pan-African values, based on an idealized communal village life which supposedly was the cradle for 'African Socialism' (Blommaert 2014: 15). He points to the contradiction between these utopian ideas and the actual situation on the ground: the actually existing cultural differences within the country were seen as potentially divisive, were branded as backwards and were consciously ignored or played down, not studied and not used as the basis for developing the 'Ujamaa' ideology (Blommaert 2014: 32). The idea, therefore, was to build a new Tanzanian nation and citizenry, with the Ujamaa ideology and the Swahili language as unifying factors. Blommaert shows how this project failed: the idea of a new Tanzanian nation built around Ujamaa greatly underestimated the 'cultural resilience' of the 'common man' (Blommaert 2014: 40). However, establishing Swahili as the dominant language in the country did not fail: it was a big success and now seems unchallengeable (Blommaert 2014: 148). Why? Why was it possible to establish Swahili in this way, but not to build a culturally homogeneous nation? Blommaert does not provide an answer; his analysis stops at this point. Topan (2008: 264) lists a variety of factors explaining the choice for and success of Swahili: the existing caravan routes; the use of Swahili by the Germans and the British; the use by missionaries; and lastly, the role of the first President of Tanzania, Julius Nyerere, himself a proud teacher. These are all factors that explain the initial choice for Swahili, but they do not fully explain its continued success. The fact that Swahili is easy for most Tanzanians and therefore a rational choice does, at least in part; the ASJP benchmarking exercise carried out above allows us to see this.

The borders of Tanzania are not linguistic borders, though. What is true for the Tanzanian languages in fact also holds for many languages of the Narrow Bantu language family, as Table 5 below shows.

Table 5
ASJP database output for selected Bantu languages.

2 SYNONYMS, AT LEAST 28 WORDS		LINGALA		LUBA		SOTHO SOUTHERN		SWAHILI		SHONA		ZULU	
LOANWORDS EXCLUDED													
LDND													
	GIKUYU	KOONGO	LINGALA	LUBA	SOTHO SOUTHERN	SWAHILI	SHONA	ZULU					
GIKUYU		0											
KOONGO		93											
LINGALA		87		0									
LUBA		86		80	0								
SOTHO_SOUTHERN		96		83	83	0							
SWAHILI		90		80	69	83	0						
SHONA		89		90	77	83	73	0					
ZULU		96		89	82	75	79	82	0				
	GIKUYU	KOONGO	LINGALA	LUBA	SOTHO SOUTHERN	SWAHILI	SHONA	ZULU					

Iraqw, however, is Cushitic, not a Bantu language. Therefore, Iraqw speakers are at a comparative disadvantage compared to Bantu speakers when they have to learn Swahili: for them, Swahili is not easy. It is of medium difficulty – that means it is still easier than English.

A third thing worth noting from Table 4 is that the non-Bantu languages spoken around the Iraqw area are all very different from one another; it is not much easier for an Iraqw speaker to learn Datooga than it would be to learn English.

It could be that the pattern that we have seen for the Narrow Bantu language family also applies to the Cushitic language family, to which Iraqw belongs. In other words, would it help the Iraqw if one of the other Cushitic language could be used as language of instruction in the same way Swahili is used for Bantu-language speakers? This appears not to be the case, as shown in Table 6: Iraqw is similar to Alagwa and Burunge, but not to any of the other major Cushitic languages. The East Cushitic languages²³ do share similarities, but this does not extend to Iraqw.

23 Glottolog distinguishes 35 East Cushitic languages. The table shows the East Cushitic languages Gedeo, Komso, Western Oromo, Rendille and Sidamo. A full list is at <https://glottolog.org/resource/languoid/id/east2699>.

Table 6
Iraqw and selected other Cushitic languages.

2 SYNONYMS, AT LEAST 28 WORDS									
LOANWORDS EXCLUDED									
LDND	XAMTANGA	GEDEO	SIDAMO	KOMSO	RENDILLE	W OROMO	ALAGWA	BURUNGE	IRAQW
XAMTANGA	0								
GEDEO	97	0							
SIDAMO	98	42	0						
KOMSO	96	80	85	0					
RENDILLE	99	91	92	83	0				
W_OROMO	100	79	88	71	85	0			
ALAGWA	96	95	91	94	99	96	0		
BURUNGE	98	96	93	98	100	95	60	0	
IRAQW	95	94	92	91	98	96	53	70	0
	XAMTANGA	GEDEO	SIDAMO	KOMSO	RENDILLE	W OROMO	ALAGWA	BURUNGE	IRAQW

Again, the benchmarked ASJP database gives only a rough indication of the ease or difficulty of learning for specific language pairs. However, this could serve as a starting point for a discussion on equitable language policies involving linguistic experts. This knowledge is more than just a fun fact: it demonstrates not only the sound basis for Swahili as a national language in Tanzania, it also demonstrates that this in fact creates a significant disadvantage for monolingual speakers of non-Bantu languages, such as Iraqw. If Tanzania is interested in providing equal educational opportunities to all its citizens, then at least two policy changes would seem to be necessary:

- 1) It is to be expected that educational results of those native speakers of the minority languages listed above who are not bilingual from birth will fall below those of the Bantu speakers.²⁴ This is wasteful in terms of talent and resources. This waste could be reduced by introducing primary education in a relatively small number of related non-Bantu local languages, at least for the larger population groups, such as the Iraqw. Again, this should be done in a gradual and well-planned way.
- 2) This also means that the study of languages spoken in Tanzania besides Swahili should be taken up in the country, as recommended by Muzale and Rugemalira (2008). Students should be encouraged to study a Tanzanian language besides Swahili and language departments should be set up for other Tanzanian languages, to start with the minority languages. It is telling that the Swedish-funded 'Languages of Tanzania' project that ran until 2008 was located in the 'Department of Foreign Languages and Linguistics' of the University of Dar-Es-Salaam.

6 Conclusions

Asking an apparently simple question such as whether Iraqw is an easy language to learn has led us to a not-so simple answer. It was necessary to look at what we mean by the word 'language' and at what we mean when we talk of 'speaking a language'. Those questions were answerable (more or less) using existing terminology: for 'language' a common-sense approach could be used, for 'speaking' a reference to existing systems for assessing language proficiency was possible. However, for determining which languages are 'easy' or 'difficult' to learn and for whom, no ready answer is available in the

²⁴ Because Tanzania publishes its exam results at the school and district levels, it might be possible to substantiate this through research. However, this means controlling for such factors as bilingualism, linguistic mix in schools, general socio-economic factors and school-related factors such as the ratio of (qualified) teachers to students and the enrolment ratio in the district. A quick comparison of 2019 primary school results shows an average score for three districts that have at least a large proportion of non-Bantu speakers that is 89% of the score of three districts with mostly Bantu speakers. (The comparison is between Bunda, Hanang and Kondoa districts versus Ikungi, Manyoni and Singida Rural districts). See <https://www.necta.go.tz/> for information on the exam results.

existing literature. This article proposes an innovative way of approaching the problem by making use of the ASJP database, a database which helps to approximate the linguistic distance between any two languages. The outputs from the database could be benchmarked to a schema taken from the U.S. government. This new approach has made it possible to at least give approximate information on how difficult or easy it is for a speaker of a language 'A' to learn a given other language 'B'. It has enabled us to examine all-too-easy generalisations, such as quoted in Blommaert (2014), of the type that English and Swahili are equally difficult for most Tanzanians. I have argued that weighting the relative proximity of the languages used in instruction and the mother tongues is relevant, because learning an 'easy' language requires far less effort than learning a 'difficult' language. This matter becomes more relevant if more and more students are enrolled in formal education. Education for more than only the elite takes more effort, because language abilities are not distributed evenly over the population. Given a finite amount of resources, any educational system will perform better if the medium of instruction is easier to learn for students.

Coming back to the original question posed in this article, then: the answer is clearly 'yes': Iraqw is indeed a difficult language to learn, at least for everybody except for those who speak a handful of other South Cushitic languages. This realization should not be without consequences: I have ended this contribution with two related policy recommendations that both seem rational from the point of view of providing equal access to education to all Tanzanian citizens: Iraqw should be introduced as a language of instruction in primary education; and the study of Iraqw and other Tanzanian languages deserves to be taken up in Tanzania itself.

Acknowledgements

I am grateful to the two reviewers, Dr Andrew Harvey and Dr Jenneke van der Wal, for their constructive comments.

References

- Albaugh, Ericka. 2014. *State-Building and Multilingual Education in Africa*. New York: Cambridge University Press.
- Altinyelken, Hülya, Sara Moorcroft & Hilde van der Draai. 2014. The dilemmas and complexities of implementing language-in-education policies: Perspectives from urban and rural contexts in Uganda. *International Journal of Educational Development* (36). 90-99.
- Bakker, Dik, André Müller, Viveka Velupillai, Søren Wichmann, Cecil H. Brown, Pamela Brown, Dmitry Egorov, Robert Mailhammer, Anthony Grant & Eric W. Holman. 2009. Adding typology to lexicostatistics: A

combined approach to language classification. *Linguistic Typology* (13). 167-179.

Bamgbose, Ayo. 2000. Language planning in West Africa. *International Journal of the Sociology of Language* (141). 101-117.

Blommaert, Jan. 2013. Citizenship, language and superdiversity: Towards complexity. *Journal of Language, Identity & Education* 12(3). 193-196.

Blommaert, Jan. 2014. *State ideology and language in Tanzania* (2nd edn) Edinburgh (UK): Edinburgh University Press.

Chaudenson, Robert. 2006. *Éducation et langues - Français, créoles, langues africaines* [Education and languages – French, creoles, African languages]. Paris: l'Harmattan.

Dryer, Matthew & Martin Haspelmath, Martin (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
<http://wals.info> (Accessed 22 October 2019).

Ginsburgh, Victor & Shlomo Weber. 2016. Linguistic distances and ethnolinguistic fractionalization and disenfranchisement indices. In Victor Ginsburgh & Shlomo Weber (eds.), *The Palgrave Handbook of Economics and Language*, 137-173. Houndsmills: Palgrave Macmillan.

Heugh, Kathleen. 2011. *Theory and practice – language education models in Africa: research, design, decision-making and outcomes*. Hamburg: Unesco Institute for Lifelong Learning.

International Standard Classification of Education - ISCED 2011. UNESCO Institute for Statistics, Montreal, 2012.
<http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf> (Accessed 22 February 2017).

Khamisi, Abdu M. 1974. Swahili as a national language. In Gabriel Ruhumbika (ed.), *Towards Ujamaa: Twenty Years of TANU Leadership*, 288-308. Kampala: East African Literature Bureau.

Klein, Wolfgang. 1986. *Second language acquisition*. Cambridge: Cambridge University Press.

Li, Shaofeng. 2016. The construct validity of language aptitude – A meta-analysis. *Studies in Second Language Acquisition* (38), 801-842.

Lüpke, Friederike & Anne Storch. 2014. *Repertoires and choices in African languages*. Boston: De Gruyter Mouton.

Maurer, Bruno (ed.). 2015. *Mesurer la francophonie et identifier les francophones - Inventaire critique des sources et des méthodes*. [Measuring Francophonie and identifying the francophones – critical inventory of sources and methods.] Paris: Éditions des archives contemporaines.

- McKenzie, Katherine, Meagan Milton, Glenys Smith & H el ene Ouellette-Kuntz. 2016. Systematic review of the prevalence and incidence of intellectual disabilities: Current trends and issues. *Current Developmental Disorders Reports* (2016)3. 104-115.
- Mous, Maarten. 1993. *A grammar of Iraqw*. Hamburg (Germany): Buske Verlag.
- Mous, Maarten, Martha Qorro & Roland Kie bling. 2002. *Iraqw-English dictionary: With an English and a thesaurus index*. Cologne: R udiger K ppe.
- Mous, Maarten & Safari Sanka. 2009. *Geeso  Duqangw na hadithi nyingine za K iraq w*. [Geeso Duqangw and other Iraqw stories]. 2nd corrected edition with Basilisa Hhao 2017. Nairobi: Kijabe Press.
- Muzale, Henry & Josephat Rugemalira. 2008. Researching and documenting the languages of Tanzania. *Language Documentation & Conservation* (2)1. 68-108.
- No author. 2017. *RISE Tanzania Research Programme Overview January 2017*. Washington DC: Georgetown University Initiative on Innovation, Development and Evaluation (guide).
- Pinxteren, Bert van. 2018. What do enrolment data say about education? In Anneke Breedveld & Jan Jansen (eds.), *Education for Life in Africa*, 35-50. Leiden: African Studies Centre.
- Prah, Kwesi. 2017. The intellectualisation of African languages for Higher Education. *Alternation* (24)2. 215-225.
- Schlechty, Philip. 2001. *Shaking up the schoolhouse: How to support and sustain educational innovation*. Hoboken: Wiley.
- Skattum, Ingse. 2018. Language of instruction in anglophone, francophone and lusophone Africa: An overview. In Anneke Breedveld & Jan Jansen (eds.), *Education for Life in Africa*, 67-81. Leiden: African Studies Centre.
- Topan, Farouk. 2008. Tanzania: The development of Swahili as a national and official language. In Andrew Simpson (ed.), *Language and National Identity in Africa*, 252-266. Oxford: Oxford University Press.
- Wichmann, S oren, Eric Holman & Cecil Brown (eds.). 2018. *The ASJP Database (version 18)*.
- World Bank. 2019. *Ending learning poverty: What will it take?* Washington DC: World Bank.