

Does literacy no longer need an institution to remain sustainable? Some reflections on the impact of texting and messaging

Maik Gibson, SIL International

The future of many, or indeed most, of the world's languages is endangered; this is not a matter of debate. What is genuinely hard to predict is the trajectory of each language - which ones will be passed on to the next generation, and to the one after that. There is agreement on some general factors involved in language maintenance, shift and death: attitudes both internal and external; changes in lifestyle; contact and relationship with other groups; patterns of multilingualism etc. But how these factors play out in individual cases does not leave us with predictive power of what the precise future of each language is.

A related question is what the literacy status of a language is – some language varieties are used exclusively for oral communication, while others are used for writing, and for a variety of purposes. Just as there are social factors that influence the transmission of a language from one generation to another, so there are other factors which help the continued use of a language in a written form. The central question of this paper is whether the conditions which sustain literacy have changed with the advent of writing on digital platforms – more specifically, whether digital writing means that social institutions are no longer strictly necessary for the maintenance of community's use of the written form of the language.

On this topic of digital writing, Professor Kornai has written what has proved to be the seminal work: his 2013 paper *Digital Language Death*. In it he demonstrates that most languages have failed to ascend digitally, that is, to become vehicles of written digital expression, whether on a computer or phone. One of the key questions for linguists is whether digital writing presents an opportunity for minority languages to be used in new sociolinguistic domain (Fasold 1984:60), or whether the digital environment is another factor in the decrease of prestige, and shift to another language. It seems clear that texting, messaging and the internet in general are something of both an opportunity and a danger to minority languages, depending on a wide variety of factors.

One of Kornai's key observations is that for digital language use to be vital, it must involve "active use in a broad variety of two-way contexts such as social networks, business/commerce, live literature, etc" (2013:3). With this comment he rightly dismisses the presence of dictionaries and the like as sufficient conditions of digital vitality - interaction is the key, rather than the mere existence of the documentation, as important and laudable task as it is. We can build on this observation by recalling Abercrombie's (1963:14) comment that "writing is a device developed for recording prose, not conversation" - digital writing is, on the other hand, often conversational, as it shares with conversation the in-the-moment interactivity that other forms of writing, fixed on paper, do not. As such, we can make the case that messaging has opened the door to genuinely conversational writing. And it seems to be the case that in writing, conversation makes the use of any vernacular more likely - whether an endangered, minoritised, or non-standard, variety of language.

Conversation is where most minority languages survive - in discussions at home, in informal gatherings, in the fields or bars. In multilingual societies one often finds languages of higher prestige used in education, formal writing and so on, but in general the vernacular or the non-standard is the appropriate variety to use when the communication is interactive. These conversations disappear from record the moment after utterance (interestingly enough, a practice shadowed by some messaging apps such as Snapchat). When used for in-the-moment person-to-person interaction, digital writing can mirror these spoken practices, and thus we see an impetus around the world to use non-standard or vernacular forms of writing, from textspeak (Crystal 2008) in English, to a multilingual situation where speakers write in a language normally reserved for speech. Digital writing is a more natural domain for the writing of otherwise rarely written languages. It seems reasonable to suggest that this use may then, at the very least, serve as a model for other writing, showing that writing the vernacular is a possibility.

We also note that the internet can increase exposure to other languages, and in some cases may be part of the hastening of a shift to a language which is perceived as having richer benefits (Karan 2011). But in this paper we have decided to focus on opportunities rather than the threats, which can look after themselves.

Just as it is easier to work on preventing a village from falling into a river rather than reconstructing it after it has fallen in ('a stitch in time saves nine'), so the most effective strategies of language preservation are those where the language is still vital, rather than cases where the language has already mainly been lost. The analogy applies also to digital engagement with minority languages - it is more likely to take root where the language has not yet started to fall into the river, that is, it is still used within all generations of the community in daily spoken communication.

In order to be able to address the question of whether digital media have introduced a new opportunity for communities to use their languages in written form, we introduce parts of the two frameworks that help raise the issue of what constitutes sustainable, vital use of language in written form, EGIDS and Kornai's measure of digital vitality.

EGIDS

In order to be able to talk of language endangerment and revitalisation with greater precision, we introduce a part of EGIDS (Lewis and Simons 2010, 2016) – an extension of Fishman's (1991) *Graded Intergenerational Disruption Scale*. Unlike the UNESCO framework of language vitality (Brenzinger et al. 2003) which views a language as vulnerable if it is not used in education, EGIDS' central question is whether "The language is used for face-to-face communication by all generations and the situation is sustainable" (Lewis and Simons 2016: 80). If this is the case, the language use is classed as *Vigorous*, and assigned at the very least a Level 6a (the lower the number, the higher the level of vitality). This is distinguished from Level 6b, *Threatened*, defined as "The language is used for face-to-face communication within all generations, but it is losing users" (ibid:81), which is the case when not all children within the language community are acquiring the communal language. For EGIDS the primary question for the future of the language is whether it is currently being transmitted orally to children, rather than questions of use in literacy or education, important as

these can be. It is only once it is established that the language is at a minimum level of 6a, *Vigorous*, that other factors such as use of the language in writing are considered. Overall a language which does have written presence, but is losing speakers, is regarded as being less vital than one where oral transmission continues, but without literacy.

The next two levels, above 6a, do take into account the status of literacy in the language, and are:

- 5, *Developing*, “The language is in vigorous use, with literature in a standardized form being used by some, though this is not yet widespread or sustainable”
- 4 *Educational*, “The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.” (Simons & Lewis 2016:80)

Level 4, *Educational*, shares with Level 6a the fact that it is a sustainable level – in contrast to 6b, where the loss of some speakers, will, without effective intervention, lead to a further degradation of the vitality status of the language. At Level 4, as long as an educational system is being sustained (with not just funding for the schools, but also some system for training teachers), the level is expected to be stable. In most cases the institution will have government backing, but other possibilities for sustainability exist, such as religious institutions, community organisations, or NGOs. In these cases, the level will be stable only as long as the institution sees the value of continuing the education in this language.

Level 5, *Developing*, is a natural stage on any transition from a language used purely for oral expression to one used in a sustainable educational system. There needs to be some development of literacy in a language used only orally before it can be used in education, such as developing a writing system, agreeing the vocabulary to be used, and developing materials. And a body of users (e.g. of those who will teach the language) must also develop. However, an insight of the EGIDS model is that this level, without a sustainable institution, is not stable. If there is activity at this level which does not continue, the language will naturally slip back to level 6a; even to sustain this level needs constant activism. This is an important observation, which can guard against the false optimism of developing a writing system which a few people use, under the assumption that these practices will spread themselves. In most cases, unless there is a champion within the community who then is key in establishing an institution, literacy will not be more widely adopted.

A diagram showing the relevant part of the ‘language mountain’ is shown below. The full diagram of the whole EGIDS scale is shown in Simons and Lewis (2016:116). A flat area represents an EGIDS level which is sustainable. In the diagram we see that this is the case for levels 4 and 6a. Level 6b is a situation which is inherently unstable – without communal intervention, the language is likely to move eventually to a level where the youngest generation of speakers has reached adulthood (Level 7). Moving up to 6a would require substantial effort, signified by the steep gradient. In the same way, Level 5 is seen as a level where without institutional support in teaching the language, the language will most likely revert to use only in oral domains. We argue here in this paper that the gradient is perhaps not so steep at Level 5, if a community starts writing its language in digital media.

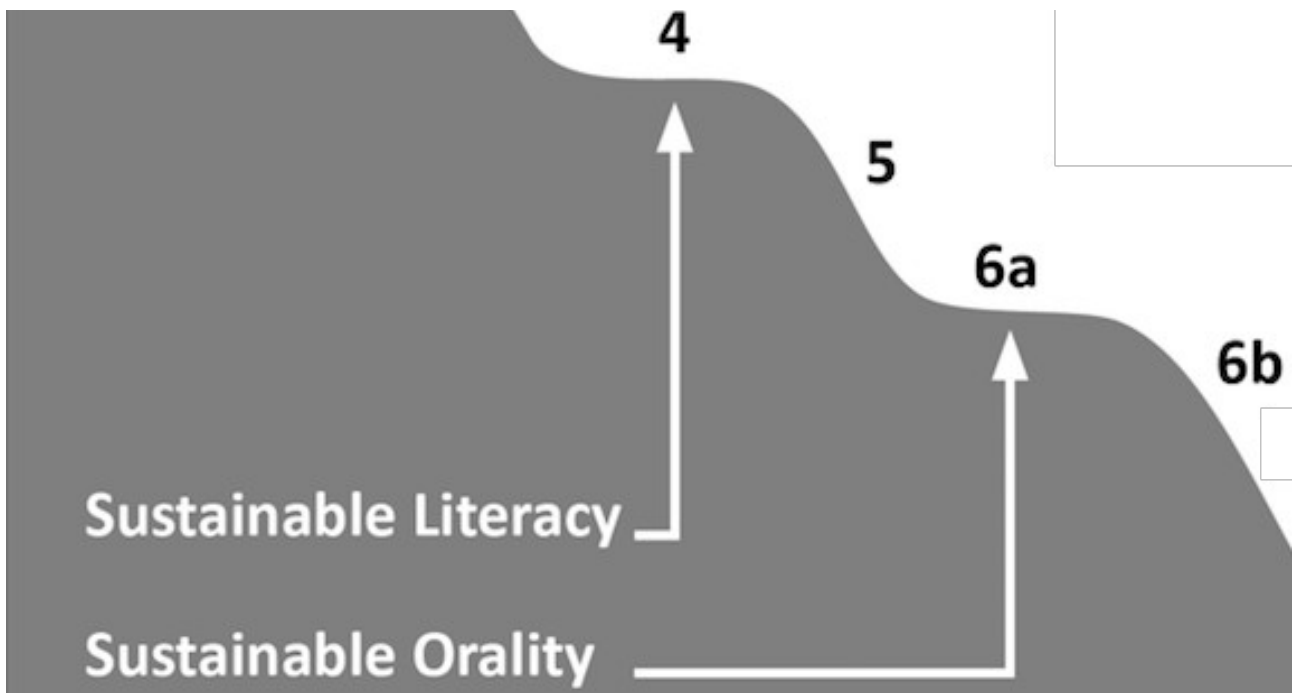


Table 1: Levels 4-6 in the EGIDS, adapted from Lewis and Simons (2016:116)¹

Measuring Digital Vitality

EGIDS was not designed with digital writing in mind, though in their later work, Lewis and Simons (2016:195) do mention that “mobile telephones, text messaging, and the worldwide web, can also be identified as venues for local-language use and sharing of identity-re-enforcing knowledge”. Kornai (2013) is to some extent an attempt to apply the principles behind EGIDS to digital writing, backed up with empirical evidence of the digital presence of different languages obtained through web crawling. The resulting scale is much simpler than EGIDS, of which we have presented only a part.

Kornai presents four categories of digital vitality. I refer the reader to his work for a fuller account.

<i>Thriving</i>	T
<i>Vital</i>	V
<i>Historical/Heritage</i>	H
<i>Still</i>	S

Thriving and *Still* are somewhat self-explanatory; the key distinction for our purposes in this paper is the difference between *Vital* and *Heritage*. In both cases the language is present on the internet in some form, but with the *Vital* category it is used between speakers for two-way communication, as opposed to *Heritage* (or *Historical*), in which case the “language is not used by native speakers (L1) for communication in the digital world” (2013:2). At the *Heritage* level there may be

¹ Used by permission, © SIL International, (Lewis and Simons, *Sustainable Language Use*, 2016) further redistribution prohibited without permission

documentation, or even Wikipedia entries, or a dictionary, but the two-way interaction, also key for Level 6a *Vigorous* in EGIDS, is absent. Further categories of digital vitality, *Emergent* and *Latent*, are proposed by Gibson (2015, 2016), primarily to help identify which activities may be useful to enable digital ascent, in much the same way as the *Sustainable Use Model* (Lewis and Simons 2016) is focused on helping communities identify useful activities for securing the future of their language. EGIDS, which it builds upon, is primarily a tool of description and analysis.

It is evident that if a language is classified as being *Vital* in digital vitality, then it will be normally at the very least at EGIDS Level 4, i.e. supported by a sustainable institution. In fact, for the *Vital* category, Kornai (2013:9) states “This group contains about two thirds (66%) of the EGIDS 1 languages; less than half (46%) of EGIDS 2; 13% of EGIDS 3; 8% of EGIDS 4; 2% of EGIDS 5; and less than 1% of all higher classes.” We see that in fact most languages at the lower end of sustainable literacy are not judged to be digitally vital – some may well be borderline between vital and still. But the presence of even some languages below EGIDS Level 4 at the digital *Vital* level demonstrates that these language communities have found ways to use the language in written communication with each other, without a sustainable institution. In cases where this use is marginal (e.g. only a small portion of the community uses the language in written communication), and EGIDS level 5 would seem appropriate. EGIDS figures listed in Simons and Fennig (2017) are not the final word, and are subject to change in the presence of new data – that Kornai has found digital presence of a language at the *Vital* level is surely sufficient to recategorise these languages as at 5 at the very least, with the possible exceptions of where the digital usage is not matched by oral communal usage.

Some Examples of Digital Writing

We now report on three examples of digital writing in language varieties which do not have any significant institutional support. In each case we give but a brief example, and discuss the context in which this use is found. The examples give but a small flavour of this kind of interaction, and are here to demonstrate that digital writing can exist where there is no institution which teaches people how to read and write the language variety, or decrees which is the correct way of writing it. Rather than giving detailed documentation of multiple examples, or rates of use, we demonstrate usage, that is, however, not isolated, but reflective of broader trends.

For people to be able to write to each other effectively, there do need to be some conventions – an agreement on which script or scripts are appropriate, for example, and which letters are appropriate to represent which sounds. Total agreement on these issues is not necessary – just as in oral communication, we can effectively communicate where there is a difference in dialect, but there are also limits to effective communication where divergence increases. In each of the cases the language users have learnt to read and write in other language varieties, and have transferred this skill to a variety not taught at school. However, the development of orthographic conventions – normally the role of an institution, is arrived at by some level of negotiation, whether overt or developmental.

Under Kornai’s paradigm these three cases can be described as *Vital*, with evidence of two-way

communication going on between community members. In two of the three cases there is a question over whether the target variety is in fact an independent language, but in some ways this is also a moot question – we will see that conventions for writing a language variety have come into play without significant intervention from an institution – the community of web users can be said to have taken over this function more often performed by an institution. The definition of what constitute separate languages is notoriously difficult and controversial (e.g. see Simons & Fennig (2017)² on *The Problem of Language Identification*), so we have chosen to remain non-dogmatic here – the issue is broadly similar, whether the variety is defined as a separate language, or as a non-standard variety of another language. We will start by looking at the case which is indisputably a separate language, Rangi of Tanzania.

Rangi, (ISO 639-3: LAG), is listed in Simons and Fennig (2017) as at 6a on EGIDS, and is spoken by over 400,000 people, but not used in the education system. There are some Biblical and other texts – the Bible uses barred vowels in addition to the typical Latin five vowels, the result of careful phonological analysis being applied to an orthography. There is however no agreed standard orthography, but the most available model is that of the Bible. As for all of Tanzania, the media of instruction in school are Swahili and English. However, there is a Facebook page³, which is for community members to use the language. The page does not function exclusively in Rangi, with Swahili also present, in line with the multilingual practices of the community. It is important to note that the page is focused on the use of the language by the community, and as such is less of strong example of vital digital use than the other two cases below, where the name of the language variety is not on the page – the interaction in those cases has no tinge of language activism – the message is clearly of more importance than the medium. In the case of Rangi, unlike the two that follow, we are unable to confirm that what is observed on a Facebook page is reflective of broader communal practices commonly found elsewhere.

I will quote a couple of texts on the site, merely to demonstrate usage related to the barred vowels of the biblical text. The first example (dated 5 December 2017), contains the following text: “Kei si kwa mirimo yaanyu tũkũ, sa mũũntũ yoyoosi adiire kwiivaa kipeembe”. The reader will note the use of barred vowels - this is a Bible verse, which is then repeated in both Swahili and English. However other posts do not demonstrate these barred vowels (e.g. a post from 30 December 2017 has “Kalarira saka eeh mukulu mikate yosi jei na ndii wuuuu kibirya mpia sana”). We can assume that the Bible passage was copied and pasted from a soft version of the Bible. However, other online use has eschewed the symbols that are not used in Swahili (the barred vowels), and which are also more of a challenge to find on a phone or computer keyboard. This is a tendency which seems quite common when community members write minoritised languages – making do with simpler solutions, and not necessarily making an effort to reflect a standard form of the language. This demonstrates that even where there is a model (in this case the Bible) where phonemes are orthographically distinguished, speakers do not always adopt these solutions, often adopting solutions which are easier to implement, or conform more closely to orthographic models already available. In this case, both factors might be at play in the lack of uptake of the barred vowels.

² <https://www.ethnologue.com/about/problem-language-identification>

³ <https://www.facebook.com/groups/TuluusikeKilaangi/>

Overall, the language may have reached the *Vital* (or at the very least a *Borderline*) level in the digital sphere, through the presence of the Facebook group, and this should also qualify the language to be deemed at Level 5 on EGIDS, and that this level may be sustainable. How much deliberate effort is being made by those promoting the written use of the language is not clear. There is less evidence of more widespread use of the language than in the following two examples. The conventions followed seem to be as close a match to Swahili as possible.

Tunisian Arabic (ISO 639-3: AEB, also listed as part of the Arabic macrolanguage ARA), is spoken by over 11 million people (Simons and Fennig 2017). Formal writing in Tunisia is generally either in Standard Arabic or French, and there is no official standard for Tunisian Arabic. However many advertising slogans are written in the variety, and there has been an increased use of the oral variety in broadcast media over the last twenty years. In general the spoken norms of the capital Tunis are followed in such writing. Children are not taught to read and write in this variety, despite its being the home variety of over 99% of the population. It is listed at level 3 on EGIDS, as it functions as the vehicular language outside formal domains, even if without any formal status, or a standardised variety.

I will present just one piece of data which is demonstrative of how the variety is used in writing: “N3ichou elwahn w msad9inou”. This is found on the public page of a Tunisian Facebook user, written by another user in response to a post. Such usage is extremely common on Facebook, not restricted to pages which mention Tunisia or Tunisian Arabic. Here we find Tunisian Arabic written in an adapted form of Latin script, with numbers added to this to represent sounds for which there is no good Latin equivalent. In this case ‘3’ stands for the voiced pharyngeal fricative [ʕ], which in Arabic is written ع, which looks like a laterally inverted 3. Likewise ‘9’ represents a voiceless uvular plosive /q/, which is written with the Arabic letter ق. Again, the reader can see the similarity. In Tunisia, normal Latin transcription of Arabic names (for example on road signs) uses ‘k’ for this sound, which is ambiguous as it also represents /k/. This system, which can also use 2, 5 and 7 as letters, is sometimes known as Arabizi (e.g. Darwish 2013). This comment, which means “We are living the illusion and believing it”, is followed by a comment from another user, written in Standard Arabic, and in Arabic script. Many other comments are written in Tunisian Arabic, but in Arabic script. The conventions for writing Arabizi (or the Arabic chat alphabet) are not taught in school, but learnt through interaction with others online. The form represented is also clearly non-standard, representing Tunisian speech.

We are not claiming that Tunisian Arabic is a separate language from Standard Arabic, as that is not how most Tunisians view their language practices. The innovative choices that are made draw on general education, including learning to read and write in French, but do not reflect any practice taught at any formal institution: specifically writing the Tunisian form of Arabic; writing it in Latin characters; and using numbers to represent certain sounds. These conventions are however shared across the community of Tunisian web users. While certain conventions exist, we cannot talk of an enforced standard, but rather a few parameters which are shared and aid effective written digital communication. A more detailed coverage of this phenomenon would show a great amount of variation in written forms – we cannot say that there is a standardised variety of written Tunisian

Arabic, but there are nevertheless some shared practices which have been adopted by a wide community. So we can say here that the institutions of Tunisia are what have taught people to read and write – in Arabic and French, but that there are some orthographic conventions which have developed (not just in Tunisia, but throughout the Arab World) through communal use. These conventions are normally promoted by an institution, but here the institution, if any, is that of the body of the internet users. This reflects spoken norms, in that what is deemed appropriate for many spoken language varieties is not decided by an institution, but by an aggregate of the speakers themselves. An example is that the Tunis variety of Tunisian Arabic functions as a *de facto* prestige variety of Tunisian Arabic, without any institutional support (Gibson 2002).

Our final example is **Sheng**, an urban variety of Swahili spoken in Nairobi featuring much lexical innovation and language mixing, or translanguaging. Some scholars (e.g Bosire 2006, Rudd 2008) in fact claim that it is an independent language – in any case, this question of the status of Sheng does not impact our judgement as to whether users of digital media are able to establish some level of orthographic conventions among themselves without intervention by an official body. Sheng does not have an ISO 639-3 code, as it has not met the criteria for inclusion as a separate language from Swahili. Sheng usage in written communication is also found widely on platforms such as Facebook, and has a lot of use in written advertising slogans, a trait it shares with Tunisian Arabic. In such cases it is sometimes referred to as *Kenyanese* (Erastus 2013), to avoid the sometime negative associations of the term *Sheng*.

The example given here is perhaps not of quite the same nature as those for Rangi and Tunisian Arabic, as it is taken from the Facebook page of ‘DJ Boyie’, a fictional character who is part of the Shujaaz.FM⁴ multimedia platform. An example sentence is “Nadai ku’get the best hustlas ii mwaka”, which demonstrates some of the features of Sheng, in that uses innovative lexis (*dai* for ‘to want’) and English expressions (in this case a whole noun phrase). Note the spelling of *hustlas*, presumably influenced by African American norms also found in the ‘z’ of the *Shujaaz* (‘heroes’) in the platform’s name. We also have the English root ‘get’ used with a Swahili infinitive marker, with the convention of using an apostrophe to separate the English root from the Swahili affixation. And then we have “ii mwaka” ‘this year’, where *ii* is the Swahili *hii* ‘this’ with the *h* dropped, and the standard Swahili order of the modifier and head inverted. *Shujaaz.FM* has also used hyphens (such as in *ku-come* ‘to come’) to separate Swahili affixes from English roots – in this case, the hyphen helps with recognising that the pronunciation of the root should follow English and not Swahili pronunciation rules. We also find no marking from DJ Boyie “nlibuy” *I bought*, along with a contributor using “imeneglect” *It has neglected*. However users of the website do not seem to have adopted either the convention of the hyphen or apostrophe. Instead strategies include leaving a space between the Swahili morphology and English root “wana stay” *they stay*, “nisha cheki” *I already checked*.

While we might consider that Shujaaz.FM is some sort of institution, even its conventions do not yet seem to have been adopted. But written communication proceeds nonetheless. Again we see that the educational institutions of Kenya have effectively introduced literacy in Swahili and English,

4 <https://www.facebook.com/DJBoyie/>

and this knowledge has been used in writing this non-standard, mixed, variety. But when confronted with cases which are somewhat alien to either language – using an English verbal root with Swahili verbal morphology, users resort to a variety of strategies, the dominant ones seeming to be either splitting the morphology from the root or combining them with no marking. The solution primarily followed by *Shujaaz.FM*, using an apostrophe (or a hyphen) to mark the boundary between forms originating in different languages, despite its rationale, does not seem to have been adopted by the broader community. Thus, as with Rangi and Tunisian Arabic, the usage of the broader community has not been fixed by an institution, but is negotiated among the community, thus circumventing one of the functions of the institution.

Is Digital Writing Creating a New Kind of Sustainability for Literacy?

Kornai's paradigm of digital endangerment and death has driven us to ask questions which bring us back to evaluate more closely EGIDS, and the effect that digital writing has on the broader context of the use of a language variety in writing. In each of the cases we have seen, there is no institutional support for the conventions that speakers appear to be using. As previously noted, these conventions are loose, with much room for latitude, but the very fact that people continue to use these forms of writing shows that effective communication is taking place – otherwise it would not be sustained. The case for this claim is the strongest in the case of Tunisian Arabic – using Arabizi is a widespread phenomenon, seen even in advertising in Lebanon. Interestingly, all three cases are to some extent a challenge to Kornai's (2013:4) statement that "languages without mature writing systems are unlikely to digitally ascend". In each case we note that the writing system is not mature, as it is not standardised, and conventions are still being negotiated, akin to the negotiation and valorisation of social norms for speech. However, it would seem sensible to concede that the challenges for digital ascent of non-standardised varieties might well be greater than those of languages with an accepted standard version, due to extended use.

Now, in no case am I arguing that there is no institutional support for writing – we can safely assume that all the people using these varieties have had an education, but one which taught them how to write another language or languages. Conventions for those languages were learned, sometimes fully, sometimes imperfectly perhaps, and some of those conventions have been drawn into their literacy practices when writing other varieties. There is obviously a feeling that at least one of the scripts that has been learnt is an appropriate vehicle for writing something other, at least on social media. While Rangi exhibits some greater phonological complexity than Swahili, with a seven- rather than five-vowel system, as well as distinctive tone, this has not been an insurmountable barrier to effective writing in a system more closely aligned to Swahili phonology. However, where tone bears a heavier load, and the language of education does not use it, we can imagine greater challenges in successful community-led orthography construction. Even though Tunisians often use Latin script (alongside Arabic script) to write their vernacular, the system also references Arabic script in a way that formal transcription standards do not – in fact their use shows the deficiencies of the Latin transcription system used for signage in Tunisia.

We note that the formal education system has its role to play in the implementation of vernacular writing, probably along with some level of linguistic similarity between what is learnt at school and

what is spoken at home or on the street (and all our cases have been at the less challenging end of the spectrum in this respect). We suggest that at the very least, where these conditions have been met, literacy may take hold without the involvement of a sustainable institution. Again, we have demonstrated this only for limited domains, with limited examples, and have not yet demonstrated that this use effectively spreads to more formal domains – there is much room for further study. But we would like to ask the question, in the environment of widespread digital literacy, whether incipient use of a language for writing might now in some cases be sustainable *without* an institution behind it. And we ask the question with an idea of the answer – that, yes, there is evidence of digital writing practices establishing themselves without systematic institutional support. We might expect the use of the vernacular in writing to be even more widespread in private messaging, including texting, but we have more of a challenge in observing this, as opposed to the semi-public domain of Facebook which have used for this paper. Therefore Level 5 in EGIDS should no longer be so hard to climb up in the language mountain as the picture shown above.

Having addressed a question of theory, we are driven to ask what difference this might make in practice. If, as seems to be the case, digital spaces are places where orthographic innovations can take root, then there are implications for those of us concerned with applying our linguistic knowledge for the benefit of minority ethnolinguistic communities. We may wish to work with them in identifying which activities are the most likely to be effective in helping them meet their own goals for the uses of their language. From the evidence presented here, it would seem that messaging using mobile phones will be a fruitful, or perhaps essential, step for languages which are at EGIDS levels 6a or 5 – more considerations for effective engagement are suggested in Gibson (2015, 2016). Digital literacy seems to present an opportunity for further written development of varieties which employ it; how effective this may be is yet to be seen.

References

Abercrombie, David. (1963). *Problems and Principles in Language Study*. 2nd edition. London: Longman.

Bosire, Mokaya. (2006). “Hybrid Languages: The Case of Sheng”. In *Selected Proceedings of the 36th Annual Conference on African Linguistics*, ed. Olaoba F. Arasanyin and Michael A. Pemberton, 185-193. Somerville, MA: Cascadilla Proceedings Project.
<http://www.lingref.com/cpp/acal/36/paper1423.pdf>

Brenzinger, Matthias, Akira Yamamoto, Noriko Aikawa, Dimitri Koundioubu, Anahit Minasyan, Arianne Dwyer, Colette Grinevald, Michael Krauss, Osahito Miyaoka, Osamu Sakiyama, Rieks Smeets, and Ofelia Zepeda. (2003). *Language vitality and endangerment*. Paris: UNESCO Ad Hoc Expert Group Meeting on Endangered Languages.
http://www.unesco.org/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf

Crystal, David. (2008). *Txtng: the gr8 db8*. Oxford: Oxford University Press

Darwish, Kareem. (2013). *Arabizi Detection and Conversion to Arabic*.
<http://arxiv.org/pdf/1306.6755.pdf>

Erastus, Kanana Fridah. (2013). “Manipulation of Sheng in advertisements and awareness campaigns in selected businesses in Kenya”. Paper presented at African Urban and Youth Language Conference, Cape Town, July 2013

Fishman, J. A., (1991). *Reversing language shift*, Clevedon, UK, Multilingual Matters Ltd.

Gibson, Maik (2002). “Levelling in Tunisian Arabic: towards a new standard?” in Aleya Rouchdy (Editor) *Language Contact and Language Conflict in Arabic* (Curzon Studies in Arabic Linguistics). New York: RoutledgeCurzon. 24-40.

Gibson, Maik (2015). “A Framework for Measuring the Presence of Minority Languages in Cyberspace”. In *Linguistic and Cultural Diversity in Cyberspace. Proceedings of the 3rd International Conference* (Yakutsk, Russian Federation, 30 June – 3 July, 2014). – Moscow: Interregional Library Cooperation Centre. 61-70.
http://www.ifapcom.ru/files/2015/khanty/yak_mling_2015.pdf

Gibson, Maik (2016). “Assessing Digital Vitality: Analytical and Activist Approaches”. In *Proceedings of the LREC 2016 Workshop “CCURL 2016 – Towards an Alliance for Digital Language Diversity”*, Claudia Soria et al. (eds.). 46-51. http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-CCURL2016_Proceedings.pdf#page=53

Karan, Mark E. (2011). Understanding and forecasting ethnolinguistic vitality. *Journal of Multilingual and Multicultural Development* 32(2):137–149.

Kornai, András. (2013). Digital Language Death. *PLoS ONE* 8(10): e77056.
doi:10.1371/journal.pone.0077056
<http://www.plosone.org/article/info:doi/10.1371/journal.pone.0077056>.

Lewis, M. Paul and Gary Simons. (2010). Assessing endangerment: Expanding Fishman’s GIDS. *Revue Roumaine de Linguistique* 55 (2):103-120. <http://www.lingv.ro/RRL%202010%20art01Lewis.pdf>.

Lewis, M. Paul and Gary Simons. (2016). *Sustaining Language Use: Perspectives on community-based language development*. Dallas, Texas: SIL International.
<https://leanpub.com/sustaininglanguageuse>

Rudd, Philip. (2008). “Sheng: the mixed language of Nairobi”. PhD thesis, Ball State University

Simons, Gary F. and Charles D. Fennig (eds.). (2017). *Ethnologue: Languages of the World*, Twentieth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>