Original Article

Comparing National-level English Proficiencies and the Necessity of a Critical Eye

Junichiro Takeno and Paul Moritoshi

Department of International Liberal Arts, Faculty of International Liberal Arts, Chugokugakuen University, Niwase 83, Kita-ku, Okayama City, 701-0197, Japan

Individuals, companies, governmental agencies and other organisations sometimes find it useful to be able to compare the English proficiency of one country with another, but to do so they must rely on the use of a very limited number of data resources which were not intended for that purpose and for which they are not necessarily suitable. Using the Education First (EF) English Proficiency Index and the European Commission's Special Eurobarometer 386 for illustrative purposes, this paper demonstrates why such use of these measures is invalid and why researchers and writers should be more critical of the data and results they use in support of their academic work. Ultimately, it is concluded that no reliable, valid and practical instrument or methodology yet exists by which data from a sufficiently representative sample of each country's population could be gathered, and that therefore comparisons of one country's English proficiency with another is not currently possible. However, the big data provided by EF and the European Commission can be useful in illustrating tendencies.

Keywords: Education First English Proficiency Index, Special European Barometer 386, English-speaking population, European English speakers, Language proficiency

Introduction

It can sometimes be useful, or just interesting, to compare one country with another on some variable. Which has the most rainfall per year or the tallest building. These can be measured objectively with little or no controversy. It is more difficult though to compare countries on issues such as which makes the best cars, has the best transportation network or the best healthcare system. These issues are much more complex, not only because there are many more variables to account for, or because those variables interplay, but also because measuring those variables objectively is more problematic.

Such is the problem when it comes to comparing one country's proficiency in English, or any other language for that matter, with another's, or even knowing how many people speak a particular language in various countries well enough to use it. Certainly, it would be useful for a company to have such comparative data when needing to make an informed decision as to where to site their next international office or manufacturing plant. It would be tempting in such cases to turn to the very small number of resources available that might be used to make such comparisons, with the assumption that they will give valid guidance. However, the first objective of this paper is to provide a note of caution, to encourage users of such information to

Corresponding author: Paul Moritoshi

Department of International Liberal Arts, Faculty of International Liberal Arts, Chugokugakuen University, Niwase 83, Kita-ku, Okayama City, 701-0197, Japan

Tel: +81 86 293 1100 Fax: +81 86 293 2941

E-mail: paulmoritoshi@cjc.ac.jp

evaluate exactly what such resources do (and do not) show, in order to better gauge the extent to which one can rely on them for any given purpose. The second is to give a clearer perspective on the problems involved in attempting to compare one country with another in terms of their respective population's proficiency in any given language or other language-related metrics.

In particular, and only by way of example, we examine two resources: the Education First English Proficiency Index (EF EPI), which is an annual report released by Education First (EF), and the Special Eurobarometer 386 (SE386) (European Commission, 2012) which shows the percentage of people in each country who self-report being able to use English.

Comparing Countries' English Proficiencies

The Japanese are often said to be particularly bad at English, especially when compared to other countries, with Clark (2000, 2009) and Tolbert (2000) being only a very few examples. Even if Japanese do have, on average, lower English proficiency than other countries, and it is not necessarily conceded that they do, there could be several linguistic, political, social or pedagogic reasons to explain it. Among these is that, as shown by Chiswick & Miller (2005), the 'distance' between Japanese and English, i.e., the extent to which they differ grammatically, lexically, phonologically and in their writing systems, is larger than for almost any other language pairing, making it much more difficult for Japanese to learn English than, for example Chinese people, because Chinese is grammatically more akin to English than is Japanese.

Furthermore, it is virtually impossible for the population of any country to know how they compare with that of another because it is almost impossible to measure. There are simply too many variables to account for, few if any of which can be controlled for when researching national populations. Also, national populations are too large. Short of testing a country's entire population, one would have to rely on using a representative sample, of sufficiently large size that the inferential statistics would yield valid results. Even deciding what constitutes a

representative sample at the national level would get bogged down in issues of demographics. Finally, those demographics are, to a greater or lesser extent fluid, changing over time, so that even if a valid comparison could be obtained, it could be out-of-date within a few years.

There is also the fact that countries start English education at different ages, with disparate amounts of contact time allocated to teach it, and that different approaches, methodologies and materials are used within and between countries, so it becomes clear that it is impossible to discern any reasonable, meaningful conclusions from any measures that might be used. However, despite these many concerns, there are still those who seek to compare the English proficiency of certain national populations using what few instruments and 'measures' exist. Two of these are the English Proficiency Index, now published annually by Education First and the Special Eurobarometer 386, published by the European Commission.

Education First English Proficiency Index

Education First English Proficiency Index categorises participating countries into one of five geographic regions: Europe, Asia, Latin America, Africa and the Middle East, though for the comparative and illustrative purposes of this paper, only results relating to the European region from 2016 are used (Education First, 2016) and are shown in Appendix 1 (Marian, 2016).

For a country to be included in the index, its sample must have more than 400 test-takers of either the Education First Standardised English Test (EF SET) or the EF placement test prior to starting an English course at an EF school. These tests consist of reading and listening comprehension. The countries within each region are then ranked highest to lowest based on the test-takers' average score for each country. The most recent index (Education First, 2019) collated more than 2,300,000 test-takers' scores.

However, there are numerous methodological flaws inherent within the scores used to compile the index. Firstly, as Reedy (2000) and Marian (2016) point out, and as Education First (2019) itself concedes, the samples used for each

country are "self-selected and not guaranteed to be representative" and that "[T]his could skew scores lower or higher than those of the general population". In other words, the scores do not represent the country as a whole and *therefore should not be used to do so*. Also, they are an amalgam of scores from two different, albeit perhaps similar tests.

Furthermore, the EF SET only measures reading and listening comprehension. However, it is erroneous to assume that reading and listening proficiency are a reliable indicator of writing and speaking proficiency, or language proficiency in general. Spolsky's (1989) "Condition 9" (p. 17) points out that receptive skills develop earlier and to a higher level than productive skills. The EF EPI scores can therefore mislead users to think that a nation's general English proficiency is higher than it really is. Finally, Takeno & Moritoshi (2017) point out that the EF SET is not a supervised test, though EF point out that test-takers do not benefit from cheating since it is only for self-evaluation purposes.

It should be stressed however that EF are not acting in bad faith and are not necessarily trying to misrepresent what their indices show, or what users can or cannot use them for. Rather, the problem is that some users apply the indices' rankings inappropriately.

The Special Eurobarometer 386

The Special Eurobarometer 386 (SE386) (European Commission, 2012) surveyed people in 28 European countries to determine which language(s) they spoke. The data were then used to calculate the percentage of people in each country's population who could use English to hold a conversation. Marian (2014) has presented these results as a map, provided in Appendix 2.

Marian (2014) however points out that these percentages are calculated based on the average of *self-reported* data in each country, and that such self-reporting is notoriously inaccurate and therefore unreliable.

We are not suggesting that the European Commission is attempting to mislead anyone, but those who wish to compare the English proficiency of one country with another could misuse this report to make unfair and invalid comparisons, especially given the scarcity of other resources by which to do so.

Correlating the EF EPI and SE386

When each participating country's EF EPI scores for the indices published between 2011 and 2015 (Education First, 2011, 2012, 2013, 2014, 2015) were correlated with its corresponding SE386 percentage, the correlations shown in Table 1 below were observed.

These results show that for the participating countries in the years included in this analysis, the EF EPI scores and SE386 percentages are impressively and consistently very highly correlated. However, again, one must take care to understand what these results really mean and what they really show, based on what the data they are derived from really represent. Only then can one gauge the meaningfulness and usefulness of the results and to what extent, if any, they can be relied upon for any given application. Are these results really that impressive? All they really mean, *at best*, is that as the number of people in a country who can use English to communicate at some level increases, the higher the average proficiency of the test-takers from that country, but this is hardly surprising. Indeed, it would seem to be a natural, logical correlation, given the self-selected and self-reported nature of the data.

There is of course also the issue of whether it is even valid to correlate the SE386 data, released in 2012, with the EF EPI scores from other years. This highlights the need to be selective in the data used in analyses.

In actuality, there currently exists no valid, reliable and practical instrument or methodology by which a sufficiently large and representative dataset can be collected to compare one country's population with another with respect to their proficiency in any language. However, we do concede that the big data provided by EF, the European Commission and others in this field can be helpful in confirming some trends or tendencies.

Table 1. Coefficients for EF EPI scores when correlated with SE386 data

Country	EF EPI Survey (Edition)					SE386 % (N=18)
	2007~2009 (2011) (N=15)	2009~2011 (2012) (N=15)	2012 (2013) (N=18)	2013 (2014) (N=19)	2014 (2015) (N=21)	
Netherlands	67.93	66.32	66.19	68.99	70.58	90
Denmark	66.58	67.96	65.15	69.3	70.05	86
Sweden	66.26	68.91	68.69	67.8	70.94	86
Austria	58.58	62.14	62.66	63.21	61.97	73
Finland	61.25	64.37	62.63	64.4	65.32	70
Slovenia	*	*	60.19	60.6	64.97	59
Estonia	*	*	65.55	61.39	63.73	56
Germany	56.64	60.07	58.47	60.89	61.83	56
Luxembourg	*	*	*	*	63.45	56
Latvia	*	*	57.66	59.43	57.16	46
France	53.16	54.28	50.53	52.69	51.84	39
Belgium	57.23	62.46	58.74	61.21	59.13	38
Lithuania	*	*	*	*	55.08	38
Italy	49.05	54.01	50.97	52.8	54.02	34
Poland	54.62	59.08	62.25	64.26	62.95	33
Romania	*	*	*	58.63	59.69	31
Czech Republic	51.31	58.9	54.4	57.42	59.01	27
Portugal	53.62	55.39	57.52	56.83	60.61	27
Slovakia	50.64	56.62	54.58	55.96	56.34	26
Spain	49.01	55.89	53.51	57.18	56.8	22
Hungary	50.8	60.39	60.41	58.55	57.9	20
Correlation coefficient	**0.947**	**0.841**	**0.757**	**0.819**	**0.82**	

＊ Non-participation

Conclusion

This paper has outlined two measures, the Education First English Proficiency Index, and the Special Eurobarometer 386, and demonstrated how they can be used and misused. These measures were selected for no other reason than that they often are, or could be misused to make statements which go far beyond what their respective data represent or what their respective results really show.

Their results were correlated to show how what at first appears to be an impressive result, is merely what one would naturally, logically expect. It was concluded that while such big data resources can be used to confirm some tendencies and *might* be useful in making general comparisons or showing overarching trends, they can only be useful *up to a point*, beyond which their use is inappropriate and invalid. It was therefore recommended that one views the data and results of such resources critically and uses them advisedly during analyses in order not to overstretch what the resource can accurately, validly and reliably do.

Finally, until a valid, reliable and practical instrument and methodology are available by which a representative dataset can be compiled, it is of limited use to make national-level comparisons of language proficiency.
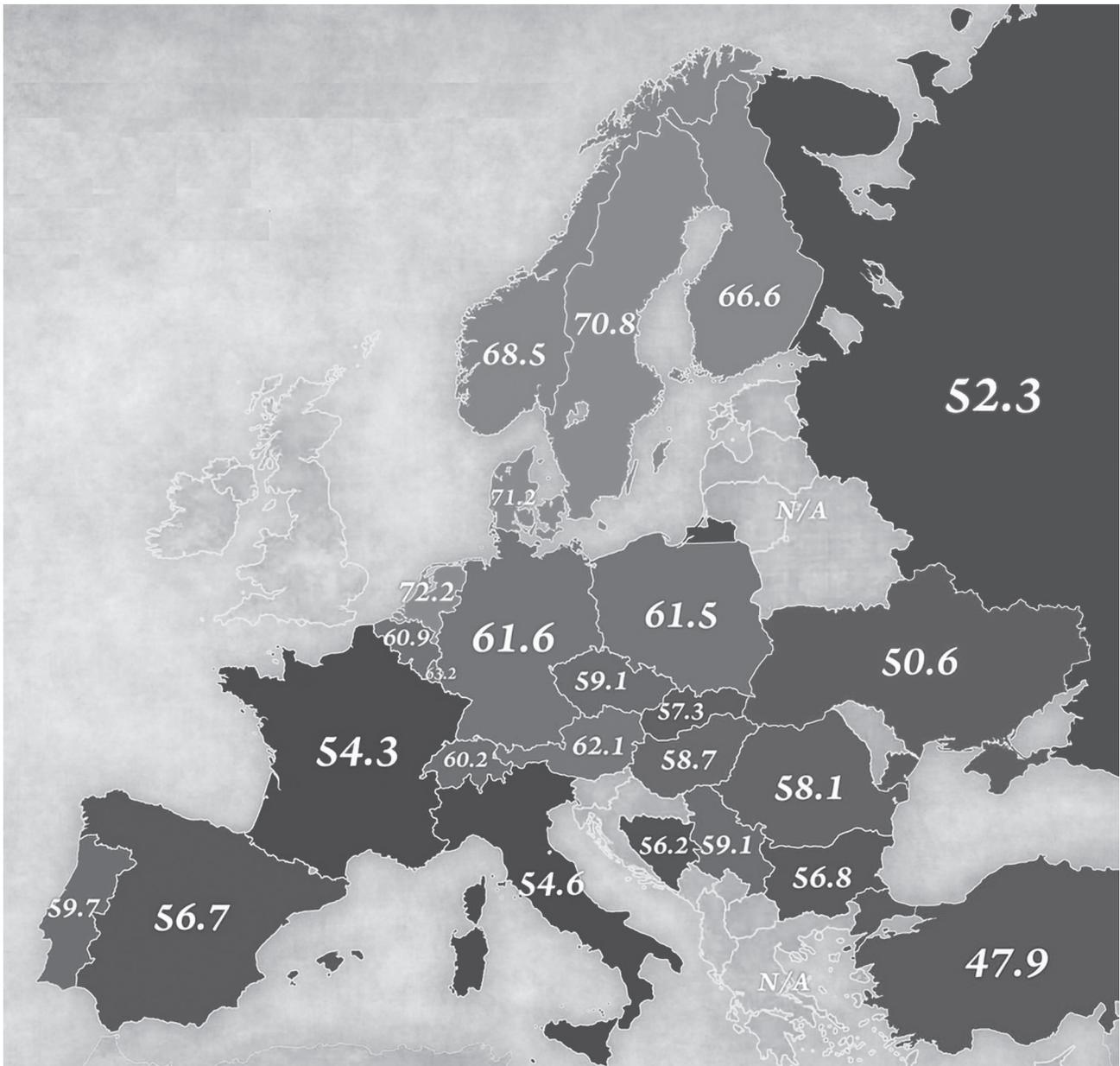
References

Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development*, *26*(1), 1-11.

Clark, G. (2000, January 30). Why Taro can't speak English? *The Japan Times Online*. Retrieved from http://www.japantimes.co.jp/opinion/2000/01/30/commentary/why-taro-cant-speak-english/

Clark, G. (2009, February 5). What's wrong with the way English is taught in Japan? *The Japan Times Online*. Retrieved from http://www.japantimes.co.jp/opinion/2009/02/05/commentary/whats-wrong-with-the-way-english-is-taught-in-japan/#.VAgG02P-KSo

Education First. (2011). *EF English Proficiency Index (2011)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

Education First. (2012). *EF English Proficiency Index (2012)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

Education First. (2013). *EF English Proficiency Index (2013)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

Education First. (2014). *EF English Proficiency Index (2014)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

Education First. (2015). *EF English Proficiency Index (2015)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

Education First. (2019). *EF English Proficiency Index (2019)*. EF Education Ltd. Retrieved from https://www.ef.com/wwen/epi/downloads/

European Commission. (2012). *Special Eurobarometer 386: Europeans and their languages*. European Commission. Retrieved from https://ec.europa.eu/commfrontoffice/publicopinion/archives/eb_special_399_380_en.htm

Marian, J. (2014). *Map of the percentage of people speaking English in the EU by country*. Retrieved from https://jakubmarian.com/map-of-the-percentage-of-people-speaking-english-in-the-eu-by-country/

Marian, J. (2016). *Why the EF EPI rankings are not what you think*. Retrieved from https://jakubmarian.com/why-the-ef-epi-rankings-are-not-what-you-think/

Reedy, S. M. (2000). TOEFL scores in Japan: Much ado about nothing. *The Language Teacher*, *24*(5), 9-11.

Spolsky, B. (1989). *Conditions for second language learning*. Oxford University Press.

Takeno, J. & Moritoshi, P. (2018). Re-examining the English Proficiency Level of Japanese EFL Learners. *Chugokugakuen Journal*, *17*, 35-39.

Tolbert, K. (2000, January 29). English is the talk of Japan. *The Washington Post*, p. A13.

Appendix 1. Map of EF EPI Scores for the European Region

(https://jakubmarian.com/wp–content/uploads/2016/12/ef-epi-europe.jpg)

Appendix 2. Map of the Percentage of National Populations Who Can Use English

(https://jakubmarian.com/wp-content/uploads/2014/05/conversation-english-eurobarometer.jpg)